



國立臺灣大學工學院應用力學研究所

碩士論文

Institute of Applied Mechanics

College of Engineering

National Taiwan University

Master's Thesis

車用電動天窗之異音偵測

Anomalous Sound Detection for Automotive Electric
Sunroofs

沈琮育

Tsung-Yu Shen

指導教授：劉佩玲 博士

Advisor: Pei-Ling Liu, Ph.D.

中華民國 114 年 8 月

August, 2025



國立臺灣大學碩士學位論文
口試委員會審定書
MASTER'S THESIS ACCEPTANCE CERTIFICATE
NATIONAL TAIWAN UNIVERSITY

車用電動天窗之異音偵測
Anomalous Sound Detection for Automotive Electric
Sunroofs

本論文係沈琮育 (R12543035) 在國立臺灣大學應用力學研究所完成之碩士學位論文，於民國 114 年 7 月 18 日承下列考試委員審查通過及口試及格，特此證明。

The undersigned, appointed by the Institute of Applied Mechanics on 18th July, 2025 have examined a Master's Thesis entitled above presented by TSUNG-YU, SHEN (student ID: R12543035) candidate and hereby certify that it is worthy of acceptance.

口試委員 Oral examination committee:

劉佩玲 劉佩玲 (指導教授 Advisor)

郭茂坤 郭茂坤

藍政英 藍政英

所長 Director: 陳瑞琳 陳瑞琳

致謝



謹向指導教授劉佩玲老師致上最高敬意，老師無與倫比的豐富經驗和專業知識不斷訓練著我思維上的狹隘與邏輯缺陷，不論是在學科上，研究上，甚或是科學界的知識都收穫滿滿。而老師所教導的更甚的是能夠引領我在未來解決問題的思考能力，並在研究遭遇瓶頸時，適時的引領我們。最後，老師一貫嚴謹治學、實事求是的態度，縱使身體抱恙，仍堅持授業解惑，其執教精神深深感動並激勵著我，成為我日後追求卓越的楷模。衷心感謝老師在學術與人生道路上的悉心栽培。

感謝郭茂坤教授、藍政英博士在百忙之中，仍撥空擔任口試委員，並提供研究、論文撰寫之寶貴建議，使內文更加完善，在此表達真摯的謝意。

本研究首要感謝藍政英藍處長，在您的引介與資源挹注下，本研究得以順利推進，能夠參與貴機構的研究計畫，實屬莫大榮幸。感謝江季達江副理，在每周的討論會議中，給予我實務上的寶貴經驗，提供我不同的思考方向。在此也要感謝從本研究從頭至尾一直教導我的楊宜興楊學長，不論是在我胡說八道或是需要更多訊號，乃至於幫忙我反覆確認了許許多多的異常訊號，都十分有耐心的支持本研究，並持續推進著進度，您在工作上的專業令我感動，研究路上因為您的幫忙而順遂許多。

這本論文後面的意義也代表我的求學生涯到此結束，父母親犧牲了他們的青春歲月，並將時間與辛勞投注在我身上，這筆帳怎麼算都沒有還清的一天，謝謝你們這 25 年來的辛勞，辛苦了！請容許我慢慢地報答，也謝謝我的妹妹多年來的拌嘴。在我多次被課業操到掛急診都是陳語樓小姐在身邊照顧，強大的藥品贊助與陪伴貫穿了我的碩士生涯，真心感謝。最後感謝實驗室的大學長，葉柏涼葉學長，在教書之餘仍每周回到學校與我們一同 Seminar，也會提點寫論文的技巧以及實驗室的奇聞軼事，令研究生涯增添不少輕鬆感。與引領我進入實驗室的大學長承展，勁凱，瑋澤學長，你們的提攜我會銘記於心。

摘要



現行車用天窗品質檢驗多仰賴人員經驗、觸覺或目視判斷，易受主觀因素影響，評估結果缺乏一致性與客觀性。為提升檢測精準度，本研究提出一套以訊號處理為核心的異音偵測方法，建立客觀且可靠之異音指標，並實現檢測流程自動化。

本研究之檢測音訊係由車用電動天窗製造廠提供與標註。該製造廠在品檢程序中，天窗共進行四個動作，分別為天窗玻璃由關閉狀態以水平移動至開啟位置，由開啟位置以水平移動回到關閉狀態，關閉狀態上抬至通風位置，由通風位置回到關閉狀態。本研究所收集之音訊，天窗異音皆出現在天窗玻璃由關閉狀態以水平移動至開啟位置。此動作可細分為 4 個時段，第 1，2 時段可能因滑塊與滑槽搭配不佳產生異音，稱為 NG1、NG2；第 3 時段可能因彈簧未定位好產生異音，稱為 NG3；第 4 時段則可能因擋風板回彈產生異音，稱為 NG4；另外還有橫跨各時段，因拉繩摩擦所產的異音，稱為 NG5。

本研究比較三個常用的音訊時頻分析，包含短時傅立葉轉換(Short-Time Fourier Transform, STFT)、梅爾頻譜(Mel Spectrogram)及梅爾倒頻譜係數(Mel-Frequency Cepstral Coefficients, MFCC)。由於異音發生時，音訊能量會變大，故除了 MFCC，STFT 能量譜與梅爾頻譜皆可觀察到異音所造成的亮帶，其中以 STFT 能量譜的亮帶更為清晰，故本研究採用 STFT 對前述 5 類異音分別進行分析。

為掌握異音的特徵頻帶，我們先觀察各類異音最大聲響樣件之 STFT 亮帶所對應的頻帶，以此頻帶對訊號做帶通濾波，經品保師聆聽後確定此頻帶之聲響即為我們所關注的異音。接著，以帶通濾波之訊號繪製能量歷時曲線訂定異音長度。為精準的擷取實際異音發生區間，我們將掃描窗大小設定為異音長度，對時間域訊號做掃描，並計算所有掃描窗之特徵頻帶能量積分值。而當異音出現在某一掃描窗時，該掃描窗將具備最大能量值，本研究將此能量值定義為異音指標，作為

後續分析異音的依據。

計算出所有訊號之異音指標後，我們觀察到正常(OK)、異常(NG)訊號會出現能量差異，而欲決定一音訊是否存在異音，必須先選擇合理的閾值，本研究採用兩種閾值：第一種是以 NG 訊號中最小的異音指標作為閾值，採用此閾值不會將 NG 訊號誤判為 OK；第二種是以 OK 訊號中最大的異音指標作為閾值，採用此閾值不會將 OK 訊號誤判為 NG。經測試，以第一種閾值做分類，NG1~NG5 之準確率分別達 97.7%、97.3%、98%、100%、100%；以第二種閾值做分類，NG1~NG5 之準確率分別達 99.2%、94.9%、98.4%、100%、100%。證實特徵頻帶能量積分值為有效之異音指標，證實此異音指標之有效性。

由於前述之異音特徵頻帶係由人為觀察 STFT 能量譜決定，為進一步優化異音特徵頻帶，本研究發展出一系統化方式，除能自動搜尋準確率最佳的特徵頻帶，還同時考慮了該頻帶的穩健性。結果顯示，以第一種閾值做分類，NG1~NG5 之準確率分別達 100%、96.9%、100%、100%、100%；以第二種閾值做分類，NG1~NG5 之準確率分別達 100%、99.6%、100%、100%、100%。整體而言，異音偵測的準確率確有提升。

最後，本研究對異音指標進行泛化性測試。發現將此異音指標套用於不同型號之天窗訊號時，檢測時可能產生誤判，這主要是因為不同型號天窗的異音特徵頻帶與本研究之異音特徵頻帶不同。因此對不同型號之天窗進行異音檢測，需以本研究之方法重新分析訊號。

針對特定機種，本研究所提出之異音偵測方法可有效分辨 OK 與 NG 訊號。未來在實務應用上，除了可以發展為自動檢測系統，輔助品保師有效率地執行品保工作，亦可利用各類異音的特徵生成大量的訊號，作為品保人員的訓練素材。

關鍵詞：車用電動天窗、品質控管、訊號處理、時頻分析、異音偵測

Abstract

Current quality inspection of automotive sunroofs heavily relies on human experience, tactile feedback, or visual judgment, which are easily affected by subjective factors and lack consistency and objectivity. To improve detection precision, this study proposes a signal-processing-based anomaly detection method. The approach establishes an objective and reliable acoustic anomaly index and realizes an automated inspection workflow.

The audio signals analyzed in this study were labeled and provided by the cooperating manufacturer. According to their inspection procedure, sunroof operation is divided into four phases: (1) horizontal movement from fully closed to fully open, (2) returning from open to closed, (3) tilting up from the closed position to a ventilation position, and (4) returning from ventilation to fully closed. All abnormal sounds in this study occurred during the first phase. This phase can be further segmented into four time intervals: NG1 and NG2 are associated with poor fitting between sliders and guide rails in intervals 1 and 2; NG3 arises from spring misalignment in interval 3; NG4 results from wind deflector rebound in interval 4; and NG5 spans all intervals, caused by rope friction.

Three time–frequency analysis methods are compared: short-time Fourier transform(STFT), Mel spectrogram, and Mel-Frequency Cepstral Coefficients (MFCC). Since anomalies typically exhibit increased signal energy, both the STFT power spectrogram and Mel spectrogram reveal bright bands corresponding to these events. The STFT power spectrogram provides clearer visual features and is therefore adopted in this study for analyzing the five types of anomalies.

To identify the characteristic frequency bands of each type of anomaly, STFT power spectrograms of the loudest samples were observed to locate bright bands. Then,

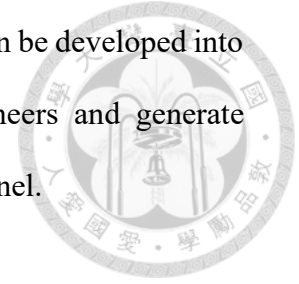
bandpass filters were designed based on the characteristic frequency bands and applied to the audio signals. After auditory confirmation by quality assurance engineers, these bands were verified to represent the target anomalies. The filtered signals were used to construct energy envelopes for estimating anomaly durations. A sliding window, set to the estimated anomaly length, was applied to the time-domain signal to compute the energy within the characteristic band for each window. The window with maximum energy was defined to contain the anomaly, and this energy was defined as the anomaly index.

After computing the anomaly index across all signals, a clear distinction between normal (OK) and abnormal (NG) signals was observed. Two thresholding strategies were employed: (1) the minimum anomaly index among NG signals, ensuring no NG is misclassified as OK; and (2) the maximum anomaly index among OK signals, ensuring no OK is misclassified as NG. Using the first threshold, classification accuracies for NG1–NG5 were 97.7%, 97.3%, 98.0%, 100%, and 100%, respectively; with the second threshold, the accuracies reached 99.2%, 94.9%, 98.4%, 100%, and 100%. These results confirm the effectiveness of the proposed anomaly index.

Since the initial characteristic bands were determined manually, this study further develops a systematic approach to automatically search for optimal frequency bands, considering both classification accuracy and robustness. The optimized bands yielded improved accuracies: 100%, 96.9%, 100%, 100%, and 100% (first threshold), and 100%, 99.6%, 100%, 100%, and 100% (second threshold). The results demonstrate enhanced detection performance.

Finally, a generalization test was conducted using signals from different sunroof models. Misclassification occurred due to model-specific differences in anomaly frequency bands, suggesting that the proposed method must be re-applied for each model. Nevertheless, for a given sunroof type, the method effectively distinguishes OK

and NG signals. In practical applications, the proposed technique can be developed into an automated detection system to assist quality assurance engineers and generate diverse anomaly samples as training data for quality control personnel.



Keywords: automotive electric sunroof, quality control, signal processing, time–frequency analysis, anomalous sound detection.

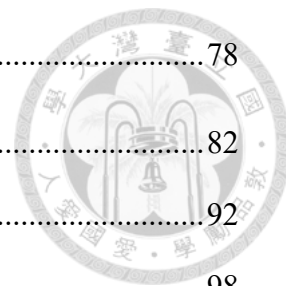
目次



口試委員審定書.....	I
致謝.....	II
摘要.....	III
Abstract.....	V
目次.....	VIII
圖次.....	XI
表次.....	XV
第一章 前言.....	1
1-1 研究背景	1
1-2 文獻回顧	2
1-3 研究動機與目的	4
1-4 本文簡介	5
第二章 訊號量測與初探.....	6
2-1 無響室簡介.....	6
2-2 量測系統	8
2-2-1 實驗架構.....	8
2-2-2 傳聲器	10
2-2-3 前置放大器	14
2-2-4 音源校正器	16
2-2-5 資料擷取器	18
2-2-6 線材	19
2-3 天窗資料介紹	20
2-3-1 天窗機構動作	20
2-3-2 資料集	24

2-3-3 天窗異音肇因紀錄.....	25
2-3-4 天窗異音初探.....	28
第三章 訊號處理方法.....	35
3-1 訊號前處理.....	35
3-1-1 切割訊號.....	35
3-1-2 濾波.....	39
3-2 A 加權濾波器.....	45
3-3 傅立葉轉換.....	47
3-3-1 傅立葉轉換原理.....	47
3-4 短時傅立葉轉換.....	49
3-4-1 短時傅立葉轉換原理.....	49
3-5 梅爾頻譜.....	53
3-6 梅爾倒頻譜係數.....	55
3-7 比較 STFT、Mel spectrogram、MFCC.....	56
第四章 異音特徵擷取與判讀.....	62
4-1 異音的特徵頻帶與時間範圍.....	62
4-1-1 異音之特徵頻帶， $\mathcal{F}_{anomaly}$	62
4-1-2 異音之長度， $D_{anomaly}$	65
4-2 異音判讀.....	68
4-2-1 異音指標.....	68
4-2-2 分類閾值.....	70
4-3 異音頻帶優化.....	71
4-4 頻帶穩健性.....	73
第五章 結果與討論.....	76
5-1 觀察異音特徵頻帶之能量分布.....	76

5-2 自動調整(f_{low}, f_{high})，並計算各種組合之準確率	78
5-3 異音頻帶之穩健性	82
5-4 泛化性	92
第六章 總結與未來展望.....	98
參考資料.....	101



圖次

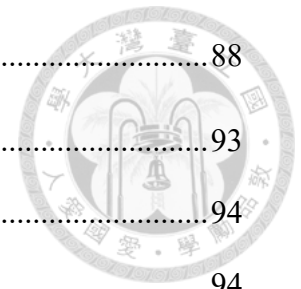


圖 2-1	半無響室門口	6
圖 2-2	半無響室內部(地板之綠色材料為 EPOXY)	7
圖 2-3	量測訊號流程圖	8
圖 2-4	麥克風與天窗放置位置	9
圖 2-5	完整架設圖	9
圖 2-6	麥克風尺寸(MM).....	11
圖 2-7	上曲線為自由音場曲線，下曲線為壓力場曲線之頻率響應。	12
圖 2-8	前置放大器	14
圖 2-9	音源校正器	17
圖 2-10	NI-9239	18
圖 2-11	NI-9239 模組.....	18
圖 2-12	MODEL 070A02	19
圖 2-13	天窗總成示意圖：(A)俯瞰圖；(B)側視圖	21
圖 2-14	天窗動作示意圖	21
圖 2-15	CLOSE-OPEN 動作	22
圖 2-16	OPEN-CLOSE 動作	22
圖 2-17	CLOSE-TILT 動作	23
圖 2-18	TILT-CLOSE 動作	23
圖 2-19	天窗滑軌(淺藍色區塊).....	26
圖 2-20	彈簧機構.....	26
圖 2-21	擋風板機構.....	27
圖 2-22	原始錄製、A 加權濾波之訊號	28
圖 2-23	天窗運行的四個動作	29
圖 2-24	NG1、NG2 之 STFT	31

圖 2-25 NG3 之 STFT.....	32
圖 2-26 NG4 之 STFT.....	33
圖 2-27 NG5 之 STFT.....	34
圖 3-1 A 加權振幅絕對值訊號.....	36
圖 3-2 A 加權振幅絕對值訊號.....	36
圖 3-3 A 加權訊號切割結果.....	37
圖 3-4 原始錄製訊號切割結果.....	37
圖 3-5 切除無效訊號.....	38
圖 3-6 被標註 NG1、2、4 之天窗樣件 STFT.....	40
圖 3-7 被標註 NG1、2、4 之天窗樣件 STFT.....	41
圖 3-8 被標註 NG1、2、4 之天窗樣件 STFT.....	42
圖 3-9 無標註異常之天窗樣件 STFT.....	43
圖 3-10 無標註異常之天窗樣件 STFT.....	44
圖 3-11 A-WEIGHTING CURVE.....	45
圖 3-12 原始錄製、A 加權之時間域訊號.....	46
圖 3-13 原始訊號、A 加權訊號之 FFT.....	46
圖 3-14 矩形窗(RECTANGULAR WINDOW).....	51
圖 3-15 0% OVERLAP 之 FRAME 取用訊號.....	52
圖 3-16 50% OVERLAP 之 FRAME 取用訊號.....	52
圖 3-17 90% OVERLAP 之 FRAME 取用訊號.....	52
圖 3-18 MEL SCALE 與 HERTZ 之對應曲線.....	54
圖 3-19 梅爾三角濾波器組.....	54
圖 3-20 NG1 之三種時頻分析比較.....	57
圖 3-21 NG2 之三種時頻分析比較.....	58
圖 3-22 NG3 之三種時頻分析比較.....	59
圖 3-23 NG4 之三種時頻分析比較.....	60

圖 3-24 NG5 之三種時頻分析比較	61
圖 4-1 NG1 發生時段(0~0.5 SEC)之 STFT	63
圖 4-2 NG2 發生時段(0.5~0.9 SEC)之 STFT	63
圖 4-3 NG3 發生時段(0.9~1.1 SEC)之 STFT	63
圖 4-4 NG4 發生時段(1.5~2 SEC)之 STFT	64
圖 4-5 NG5 發生時段(0.9~5.6 SEC)之 STFT	64
圖 4-6 時間域訊號難以觀察異音之起始~終止點	65
圖 4-7 各異常之能量歷時曲線	66
圖 4-8 掃描訊號示意圖	68
圖 4-9 局部極值示意圖	71
圖 4-10 自動調整 f_{low}, f_{high} 之二維矩陣	72
圖 4-11 頻帶之 $\min_{NG} E_c - \max_{OK} E_c$ 差值示意圖	74
圖 4-12 $Accuracy(f_{low}, f_{high})$	74
圖 4-13 3*3 KERNEL SUMMATION	74
圖 4-14 NG4 之 $Accuracy_{3 \times 3}(f_{low}, f_{high})$	75
圖 4-15 重新挑選之最佳頻帶於 NG4 $Accuracy(f_{low}, f_{high})$ 之位置	75
圖 5-1 各 NG 訊號之能量分布	77
圖 5-2 以 $E_{Threshold} = \min_{NG} E_c$ 之 NG1~NG5 $Accuracy(f_{low}, f_{high})$	80
圖 5-3 以 $E_{Threshold} = \max_{OK} E_c$ 之 NG1~NG5 $Accuracy(f_{low}, f_{high})$	81
圖 5-5 NG1~NG5 最佳頻帶於 $Accuracy(f_{low}, f_{high})$ 位置	85
圖 5-6 NG1~NG5 最佳頻帶於 $Accuracy(f_{low}, f_{high})$ 位置	86
圖 5-7 $Accuracy_{3 \times 3}(f_{low}, f_{high})$	87

圖 5-8 重新挑選之最佳頻帶於 $\text{Accuracy}(f_{low}, f_{high})$ 之位置.....	88
圖 5-9 墨西哥天窗樣件之 STFT.....	93
圖 5-10 墨西哥樣件一之泛用性測試.....	94
圖 5-11 墨西哥樣件二之泛用性測試.....	94
圖 5-12 墨西哥樣件三之泛用性測試.....	95
圖 5-13 墨西哥樣件四之泛用性測試.....	95
圖 5-14 台灣與墨西哥樣件之中高頻異音之 STFT 比較.....	97
圖 5-15 台灣與墨西哥樣件之 NG5 特徵頻帶之 STFT 比較.....	97



表次



表 2-1	電容式麥克風規格	12
表 2-2	前置放大器規格	15
表 2-3	CLOSE-OPEN 動作之 NG1~NG5 發生異常樣件編號與異常樣本數	24
表 2-4	異音肇因紀錄表	25
表 2-5	異音發生時段	30
表 4-1	由 STFT 觀察之異音亮帶(Hz)	62
表 4-2	異音長度(SEC).....	67
表 5-1	不同閾值下之最佳準確率(%)	76
表 5-2	不同閾值下之最高準確率(%)	79
表 5-3	不同閾值下之最大 $\min_{NG} E_c - \max_{OK} E_c$ 差值頻帶(Hz).....	82
表 5-4	最佳頻帶與對應之準確率、 $E_{Threshold}$	89
表 5-5	最佳頻帶與對應之準確率、 $E_{Threshold}$	89
表 5-6	異音之音訊	90
表 5-7	異音之音訊	91
表 6-1	以 $E_{Threshold} = \min_{NG} E_c$ 為分類閾值	99
表 6-2	以 $E_{Threshold} = \max_{OK} E_c$ 為分類閾值	99

第一章 前言



1-1 研究背景

在車用零組件製造業中，品保師除了對天窗產品要進行測試外，也是判定天窗能否出貨的關鍵崗位。天窗之檢測流程與判定標準往往是由公司培訓，並同時在產線上不斷累積經驗，因此傳統的品保作業大多仰賴品保師的個人經驗及主觀判斷，例如以觸覺、視覺、聽覺等人體感官來識別天窗是否存在異常或缺陷。但不同檢驗者在判讀上可能會存在身體感知與評價的主觀差異，導致檢測標準難以統一，容易出現漏檢、誤判，甚至在天窗出貨後由於品質不穩定而面臨退貨風險。

隨著造車工藝日新月異，車用天窗作為車輛外觀與重要功能的零組件，其複雜機構搭配所產生的異常肇因繁多，若需同時判定數種異音不僅顯得耗時，且判定之準確度對品保師更是一大挑戰。

與我們合作的車用電動天窗製造廠也嘗試對整段天窗音訊進行傅立葉轉換等訊號處理方式，然而天窗音訊內包含不同種類之異音，且與背景噪音混雜在一起，使得頻率域訊號混雜，難以有效觀察異音特徵，故天窗異音之物理意義仍須深入分析。

1-2 文獻回顧

在科技發展與算能的提升下，推動了產業與人工智慧的結合，傳統產業正面臨由人工處理轉向自動、數位化檢測的挑戰。在此背景下，異常聲音檢測(Anomaly Sound Detection)，成為一個備受關注的研究方向。多數研究選擇將音訊轉換為梅爾頻譜圖、梅爾倒頻譜係數，並使用各式各樣的模型提取異常音訊特徵，經訓練後對輸入之音訊做出辨識。

Chen [1]，以台達風扇為研究對象，並將風扇音訊轉換為頻譜圖、梅爾頻譜圖、梅爾倒頻譜係數、色度圖，再分別以上下堆疊之方式，合併為 Mel-MFCC、Mel-Chroma、Mel-MFCC-Chroma 三種複合特徵影像，與四種基礎特徵影像，透過非監督式學習，以卷積自動編碼器搭配複小波結構相異性作為損失函數，以評估輸入影像與重構影像之差異，進而對輸入影像做出辨識。

Wang [2]，以剎車來令片為研究對象，旨在對來令片做出 OK、NG 之分類，並使用加速規(Accelerometer)量測敲擊後之振動訊號，以隨機森林與梯度提升法做為模型，並比較兩種模型對於訓練之結果。

Lo [3]，為了檢測風力發電機葉片之損傷，首先以短時傅立葉轉換觀察 OK、NG 之風力發電機葉片之錄製訊號，再以 FFT 觀察正常、異常風機之頻帶差異後，並以正常風機之 FFT 檢測訊號。

Tsai [4]，為了檢測風力發電機葉片之損傷，以短時傅立葉先行觀察 OK、NG 之訊號特徵，並提出透過卷積神經網路(Convolutional Neural Network)對 STFT 影像進行訓練，唯輸入影像包含 OK、NG 兩種 STFT 二維矩陣。

Huang [5]，為了檢測風機齒輪箱振動之齒輪，軸承故障，使用單軸加速規擷取振動訊號，採用 Daubechies Wavelet 為母小波對進行小波分解，並將時頻圖作為 CNN 的輸入特徵，以此將故障作分類。

Ciou [6]，為了檢測風機傳動鏈狀態監測，以 11 個時間域，13 個頻率域之統計特徵，如平均值，標準差，最大值，峰度偏度等作為描述訊號的特徵，並以此

訓練 ANN 模型，進而檢測轉軸，齒輪，軸承等異常。

由上述的文獻回顧可知，雖然我們得以將各式各樣的特徵，不論是一維特徵，或是二維時頻特徵輸入至模型中使其自行學習，都能獲得不錯的檢測效果。此外，車用電動天窗製造廠也曾與 AI 公司合作，雖然在訓練資料中達到極高的準確率，但由於目前錄製音訊之樣件過少，模型訓練上易產生過擬合(overfitting)，導致測試資料之檢測效果不佳，難以滿足工廠品保作業之需求。

綜上所述，本研究除將訊號分類外，分析天窗各異音之物理意義則是另一項議題，故本研究不選擇將未經分析之訊號作為 AI 模型之輸入，而是由訊號處理方式擷取異音特徵，進而對天窗異音做出 OK、NG 之判讀與檢測。

1-3 研究動機與目的

透過上述參考文獻可知，雖然機器學習與深度學習已經能對不同種類之產品進行異常檢測。然而，若貿然投入機器學習與深度學習之中，可能會導致黑盒子 (Black Box) 問題。人工智慧領域中，黑盒子問題主要描述模型訓練的內部機制難以被人理解，可能是在自動特徵提取階段之時，模型學習到的特徵缺乏明確物理意義，亦或是當訓練資料與領域知識不足時，模型之訓練結果可能會誤導操作人員。為避免上述問題，本研究首先以最原始之訊號處理方法分析天窗異音。

本研究除了探尋天窗異音之物理意義外，更為實務上為在品保作業上能夠降低人員之誤判、漏檢之情況發生，因此我們期望提出一項客觀的異音指標，使其具備判定天窗樣件為 OK、NG 之能力，並輔助品保師進行品保作業。

1-4 本文簡介

我們期望選擇合適的訊號處理方式，除了能夠觀察天窗異音之物理意義，並提供一項客觀的準則做為檢測天窗異音時的參考指標。在應用 AI 等技術前，若能找出異音之物理特徵，並對訊號做出 OK、NG 之區隔，也未必需要採用機器學習與深度學習之模型進行分類作業。所以在章節安排上，我們先對天窗訊號做一些直觀的討論，比較不同的音訊時頻方法用以分析天窗異音，挑選合適的時頻方法提取天窗異音特徵作為異音指標，以合理之閾值對訊號進行 OK、NG 之分類作業，最後則是優化異音之特徵頻帶與穩健性討論，並測試本研究異音指標之泛用性，與未來的應用層面。

本文總共分為六個章節，其內容簡述如下：

- 第一章 介紹異音檢測之常見方法，並藉由文獻回顧說明本研究之方向與期望的研究成果。
- 第二章 介紹本研究之訊號量測系統，並對天窗機構與異音進行初步探討與觀察。
- 第三章 對訊號進行前處理，並比較不同的音訊時頻分析方法。
- 第四章 以適當的時頻分析方法萃取異音特徵，並提出一項客觀之異音指標，用於分類 OK、NG 訊號。
- 第五章 優化異音之特徵頻帶與穩健性討論，並以本研究之異音指標對不同款之天窗樣件進行泛用性測試與自動化檢測。
- 第六章 總結本研究分析天窗異音特徵與自動化檢測之結論，並對未來應用層面提供建議。

第二章 訊號量測與初探



2-1 無響室簡介

在進行產品聲學響應測試、聲壓量測、資料收集，以及頻率分析等研究時，為了避免氣候、地形反射條件、濕度等外在因素的干擾，通常需要建置一座環境條件可精準控制且高度封閉的場域，以有效隔絕外界影響，此即為無響室 (Anechoic Chamber)。

無響室又分為全無響(Anechoic Chamber)與半無響(Semi-Anechoic Chamber)，無響室之內部皆鋪設吸音材料，本研究產品測量之無響室為半無響室，如圖 2-1，即地板表面僅鋪設 Epoxy 沒有鋪設吸音材料，其餘壁面加裝吸音楔，如圖 2-2，並配有環境除濕機以維持最佳錄音場域。



圖 2-1 半無響室門口

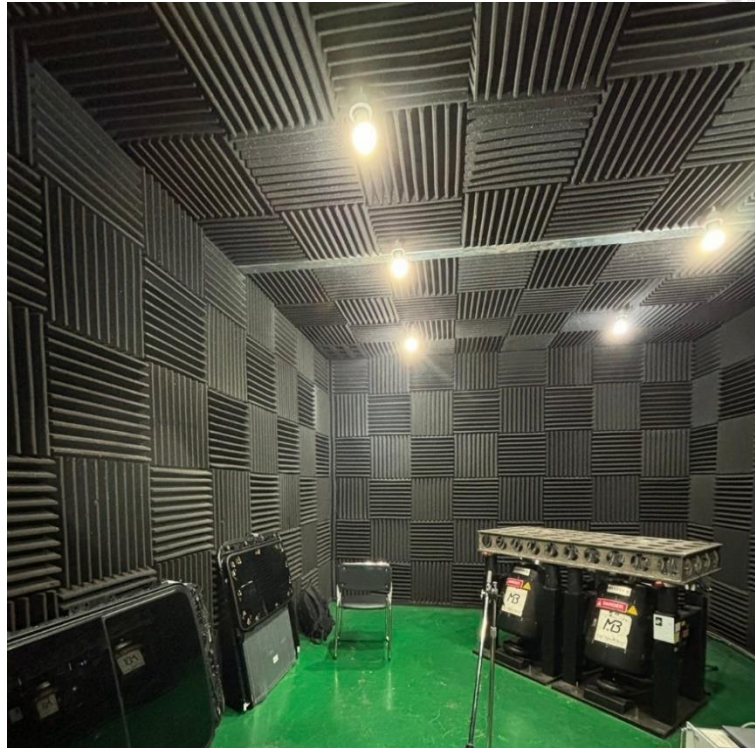


圖 2-2 半無響室內部(地板之綠色材料為 Epoxy)

2-2 量測系統

2-2-1 實驗架構

量測訊號流程如圖 2-3，首先將天窗鎖於金屬治具上，並將電容式麥克風與前置放大器架設於天窗中柱正上方 15 公分處，如圖 2-4，由線材連接至資料擷取器後，由 LabVIEW 將數位訊號儲存為 WAVE 音訊檔(Waveform Audio File Format)，完整架設如圖 2-5。

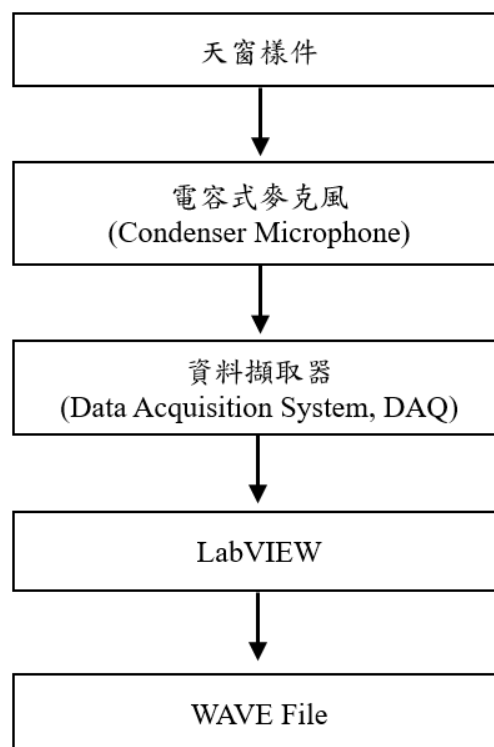


圖 2-3 量測訊號流程圖

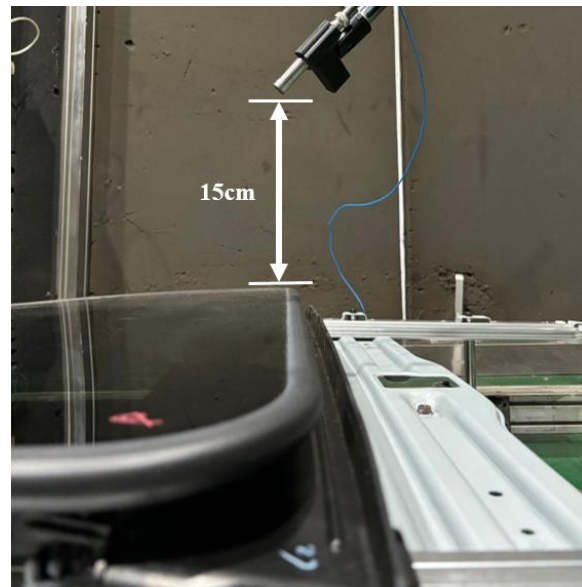


圖 2-4 麥克風與天窗放置位置



圖 2-5 完整架設圖

2-2-2 傳聲器

當空氣中的壓力波產生時，麥克風內部的敏感元件受到壓力波而產生機械作用，如變形或位移等，進而依據不同的換能原理(Transducer Principle)將機械能轉換為電子訊號，以供後續之訊號處理與分析。

本研究使用電容式麥克風(Condenser Microphone)，係由振膜與固定背極板組成之電容。當振膜受到壓力波之擾動而產生微小變形時，其與固定背極板之間的距離(d)改變，導致電容值(C)產生變化，如式(2-1)。因其頻率響應寬廣、靈敏度高、失真低，故適合高精度的聲學量測。

$$C = \frac{\epsilon A}{d} \propto \frac{A}{d} \quad (2-1)$$

式中， C 為電容(單位:法拉)， ϵ 為介電常數(單位:法拉/米)， A 為固定背極版之面積(單位:平方公尺)， d 為振膜與固定背極板之距離(單位:公尺)。

根據電容的基本定義式，當電壓固定時，電荷 Q 與電容 C 呈線性關係，電容公式定義如式(2-2)：

$$Q = C \times V \quad (2-2)$$

式中， C 為電容(單位:法拉)， Q 為電荷(單位:庫倫)， V 為電壓(單位:伏特)。

由此可知，若電荷為電壓之線性函數且電壓為定值時，則任何機械作用所引起的電容變化會反映在電荷的變化上，如式(2-3)：

$$\Delta Q = \Delta C \times V \quad (2-3)$$

式中， ΔC 為振膜與固定背極板距離變化導致之電容變化(單位:法拉)， ΔQ 為固定電壓下，電容變化導致之電荷變化(單位:庫倫)。

G.R.A.S. Sound & Vibration 為丹麥專業聲學測量設備廠商，本研究所採用該公司所生產之預極化自由音場電容式麥克風，型號 G.R.A.S. Sound & Vibration

1/2" Prepolarized Free Field Microphone 40AE，Serial NO. 82179。預極化 (Prepolarized)表示麥克風振膜本身已帶有極化電荷，不須外部極化設備，僅需藉由放大器提供之電源即可運作。自由音場(Free Field)麥克風是用於量測單一來源之聲壓，因此量測時麥克風前端會直接對準帶測音源，以達最佳錄音效果。其頻率響應為 $\pm 2\text{dB}$ 對應頻率範圍 3.15 Hz-20 kHz、 $\pm 1\text{dB}$ 對應對應頻率範圍 12.5 Hz-10kHz，如圖 2-6。麥克風尺寸如圖 2-6，直徑為 1/2"。

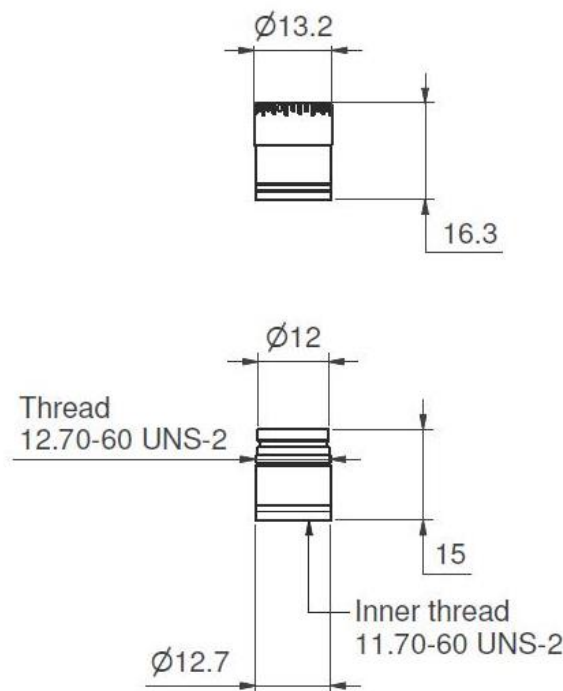


圖 2-6 麥克風尺寸(mm)

(資料來源：產品說明書)

麥克風建議之環境參數(Reference conditions)為，攝氏(Celcius)23 度，相對濕度(Relative humidity)50%，氣壓(Barometric pressure)為 101.3kPa，即一大氣壓。麥克風響應曲線如圖 2-7，麥克風規格如表 2-1。

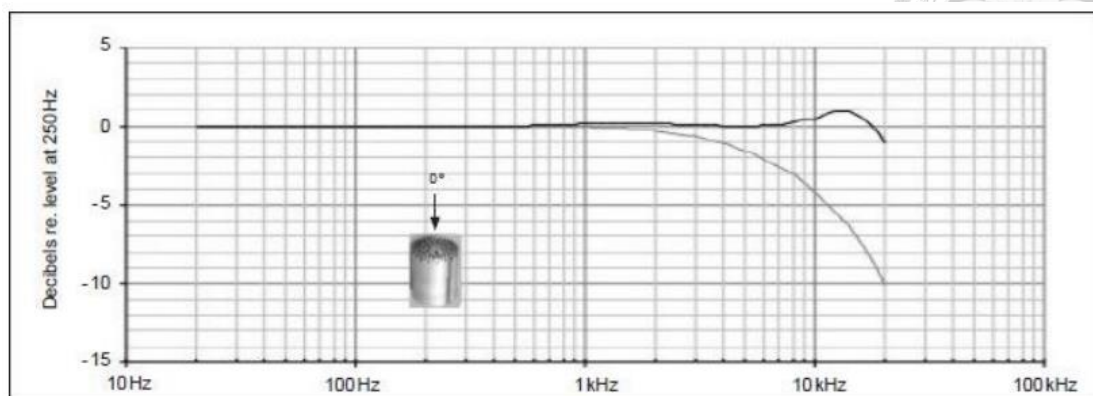


圖 2-7 上曲線為自由音場曲線，下曲線為壓力場曲線之頻率響應。
(資料來源：產品說明書)

表 2-1 電容式麥克風規格

Specifications	數值/說明
Nominal Open Circuit Sensitivity	50 mV/Pa (at 250 Hz)
Frequency Response	$\pm 2\text{dB}$, 3.15 Hz ~ 20 kHz $\pm 1\text{dB}$, 12.5 Hz ~ 10 kHz
Polarization Voltage	Preamplifier Type , 0 V
Upper Limit of Dynamic Range 3% Distortion	146 dB re. 20 μPa
Lower Limit of Dynamic Range Thermal noise	14 dB re. 20 μPa
Nominal Cartridge Capacitance Polarized	20pF
Resonance Frequency 90° Phase shift	14 kHz
Effective Front Volume	50 mm ³

Nominal at 250Hz	
Static Pressure Coefficient 250Hz at 25°C	-0.008 dB/kPa
Sensitivity vs Variation in Temperature Equiv. SPL for 1 m/s^2 perpendicular to diaphragm	62 dB re. 20 μPa
Temperature Range	-40 to +150°C
Mean Temperature Coefficient -10 to +50°C	-0.01 dB/°C
IEC 1094-12 Type Designation	WS2F
Length with Protection Grid	16.2 mm
Diameter with Protection Grid	13.2 mm
Diameter without Protection Grid	12.7 mm
Thread Protection Grid	12.7 mm 60 UNS
Thread Preamplifier	11.7 mm 60 UNS
Weight	9 g

(資料來源：產品說明書)

2-2-3 前置放大器

電容式麥克風所收集到的訊號相當微弱，屬於電容訊號。為了將其轉換為較強的電壓訊號以便後續分析，本研究在麥克風後端配置前置放大器，使其輸出更為穩定且具有更高振幅的電子訊號。前置放大器採用與麥克風相同製造公司所生產的 CCP (Constant Current Powered) Preamplifier Type 26CA, Serial No. 74864，如圖 2-8。此款前置放大器與本研究使用的麥克風具有良好的匹配度，可確保量測結果的精準與穩定，其規格如表 2-2。



圖 2-8 前置放大器

表 2-2 前置放大器規格

Specifications	數值/說明
Frequency Range (± 0.2 dB) (cable load 4.7nF)	2 Hz - 100 kHz
Input Impedance	20 G Ω , 0.4 pF
Output Impedance (Cs = 20 pF, f=1000 Hz)	< 50 Ω
Output Voltage Swing (peak)	± 8 V
Noise (measured with 20 pF ½" dummy mic.)	A- weighted: < 2.5 μ V Linear (20 Hz - 20 kHz): < 6 μ V
Attenuation	Typical -0.25 dB
Power Supply	2 - 20 mA (4 mA typical)
Temperature	Operation: -30°C to + 60°C Storage: -30°C to + 80°C
Relative Humidity	Operation: 0 - 90% Storage: 0 - 95%
Dimensions	Diameter: 12.7 mm (0.5") Length: 73 mm (2.85") Weight: 27 g (0.95 oz)

(資料來源：產品說明書)

2-2-4 音源校正器

麥克風靈敏度(Sensitivity)定義為單位聲壓作用下，麥克風輸出電壓之比值，麥克風靈敏度會隨使用時間而需要修正，因此在每次量測作業前，應使用音源校正器校準以確保輸出電壓無失真。

本研究之音源校正器為 Larson Davis Model CAL200，如圖 2-9。本設備可以產生固定 1000Hz 頻率，輸出 94dB(1 Pa)、114dB(10Pa)之固定聲壓，由式 2-4 可知，94dB 等於 1 帕斯卡(Pa)的壓力，114dB 約等於 10 帕斯卡(Pa)的壓力。

$$L_{dB} = 20 \log_{10} \left(\frac{P}{P_{ref}} \right) \quad (2-4)$$

式中， P 為固定聲壓(單位:Pa)， P_{ref} 為參考聲壓，常用值為 20×10^{-6} Pa， ΔL_{dB} 為聲壓級(單位:dB)。

校準步驟如下，首先由標準麥克風靈敏度 S_{ref} 、音源校正器之固定聲壓計算出麥克風標準輸出電壓值 V_{ref} ，如式 2-5。而後將待修正麥克風插入音源校正器中，同時量測待修正麥克風量測之輸出電壓，反推得待修正靈敏度 S_{fix} 。並對待修正麥克風之輸出電壓做修正，如式 2-6。

$$P = \frac{V_{ref}}{S_{ref}} = \frac{V_{fix}}{S_{fix}} \quad (2-5)$$

$$S_{fix} = V_{fix} \frac{S_{ref}}{V_{ref}} \quad (2-6)$$

式中， P 為音源校正器產生的聲壓(單位:Pa)， S_{ref} 為標準靈敏度(單位:V/Pa)， S_{fix} 為待修正麥克風之靈敏度(單位:V/Pa)， V_{ref} 為標準輸出電壓(單位:V)， V_{fix} 待修正麥克風之輸出電壓(單位:V)。



圖 2-9 音源校正器

2-2-5 資料擷取器

當聲壓經過傳聲器與前置放大器後，雖已將聲壓轉換為連續電壓變化，也就是類比訊號。但為了將訊號應用於如 National Instruments 所發表之 LabVIEW、廣泛運用於數位訊號處理之 Matlab 等軟體，本研究採用 National Instruments 之產品 NI-9239，如圖 2-10。

在類比電壓訊號進入類比數位轉換器前(Analog-to-Digital Converter, ADC)，為避免輸入電壓過高導致電路損壞，該模組提供過壓保護裝置(Overvoltage Protection)，並於後端接上放大器(Amplifier)，濾波器(Prefilter)後，將訊號轉換為數位訊號，方便電腦讀取，其模組的元件流程，如圖 2-11。



圖 2-10 NI-9239

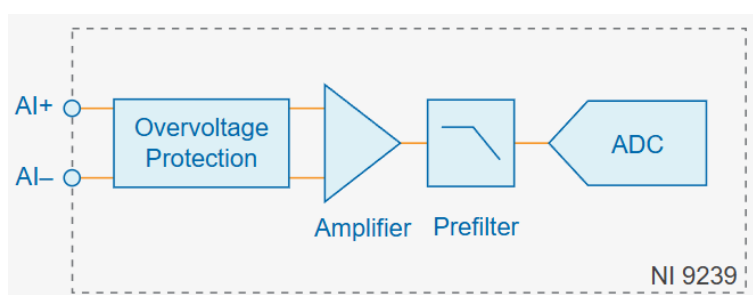


圖 2-11 NI-9239 模組

(資料來源：使用說明書)

2-2-6 線材

用於連接前置放大器與資料擷取器，本研究採用 PCB Pizotronics 所生產之 Model 070A02，如圖 2-12 中紅框標示。

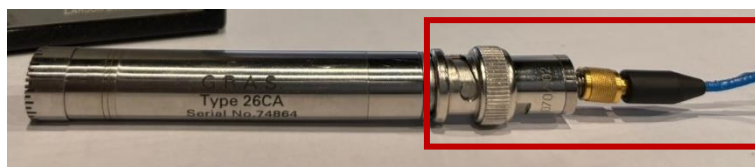


圖 2-12 Model 070A02

2-3 天窗資料介紹

本研究採車用電動天窗製造廠於台灣廠區生產之某款天窗產品為分析對象，如圖 2-13。

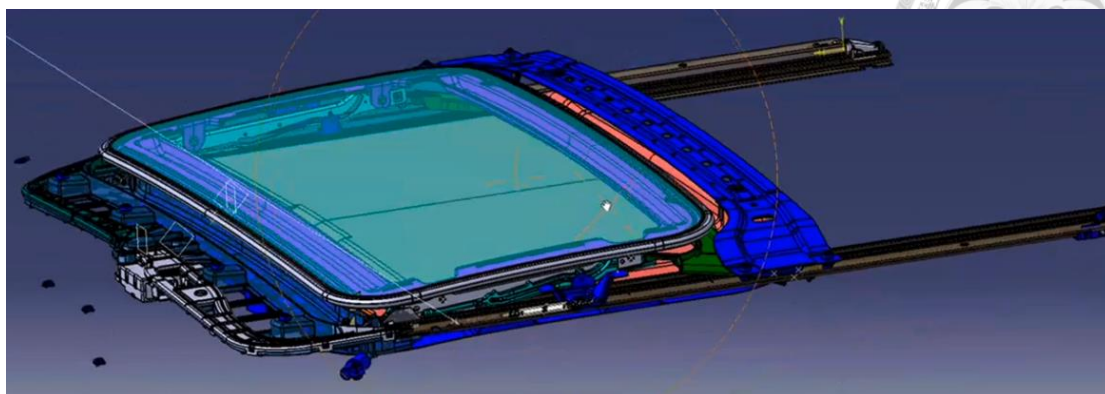


2-3-1 天窗機構動作

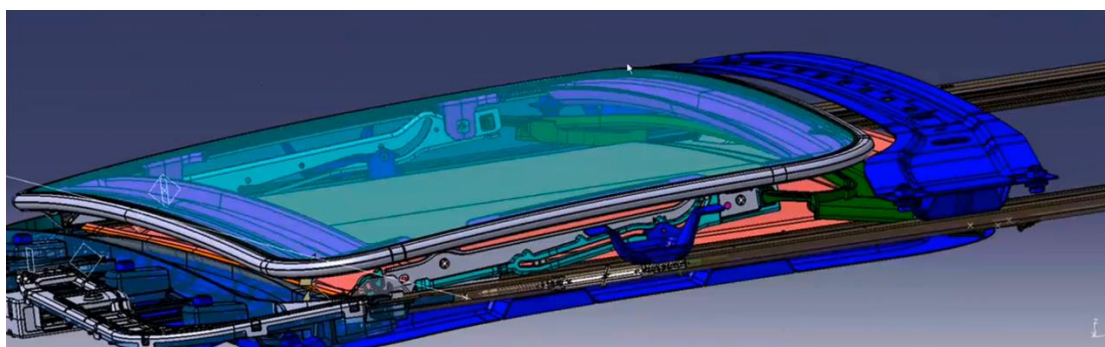
天窗之四個動作，分別為 Close-Open、Open-Close、Close-Tilt、Tilt-Close，依序代表天窗玻璃由關閉以水平移動至開啟位置、由開啟位置以水平移動回到關閉、關閉狀態上抬至通風位置，由通風位置回到關閉狀態。當天窗機構啟動後，馬達帶動拉繩，拉繩帶動滑塊，滑塊沿滑槽運行。Close 狀態之所以位於滑軌中段，係因車頂結構並非水平，因此 Close 狀態存在一傾角，如圖 2-14 中 Close 之位置。

Close-Open 動作運行時間約 6 秒，可分解為 Close 與 Open，如圖 2-15。第一步為 Close，本動作必須將天窗降至水平角度才能向後延伸，否則天窗將會以傾斜角度向上開啟，因此滑塊從滑軌中段開始向上運行，直到滑軌頂端。第二步為 Open，此時僅由馬達帶動拉繩使天窗進行水平移動，在天窗向後水平移動的瞬間，擋風板會藉由彈簧機制同步向上作動，而後開啟天窗。Open-Close 為 Close-Open 之相反動作，如圖 2-16。

Close-Tilt 動作運行時間約 1 秒，如圖 2-17。本動作僅將關閉狀態之天窗向上抬升，即滑塊由 Close 動作向下運行至滑軌底端。Tilt-Close 為 Close-Tilt 之相反動作，如圖 2-18。



(a) 俯瞰圖



(b) 側視圖

圖 2-13 天窗總成示意圖：(a)俯瞰圖；(b)側視圖

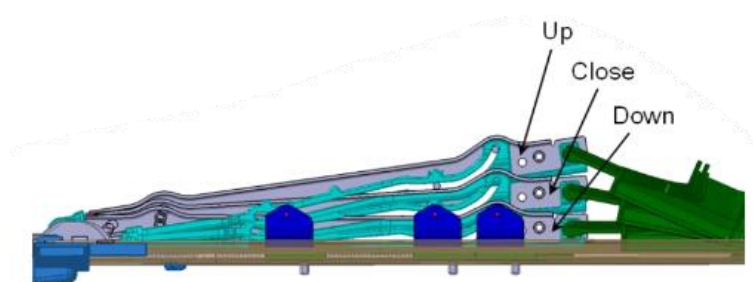


圖 2-14 天窗動作示意圖



Close 動作



Open 動作



圖 2-15 Close-Open 動作

Open 動作



Close 動作

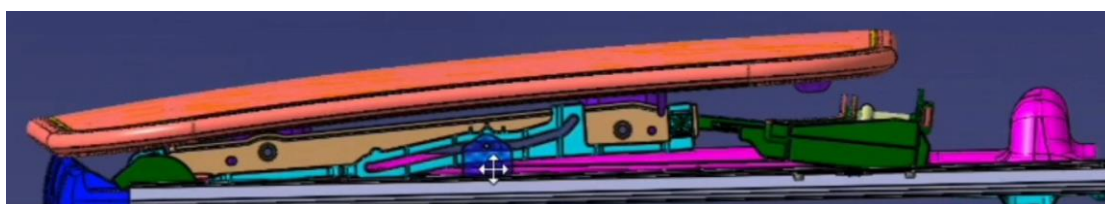


圖 2-16 Open-Close 動作

Close 動作



Tilt 動作

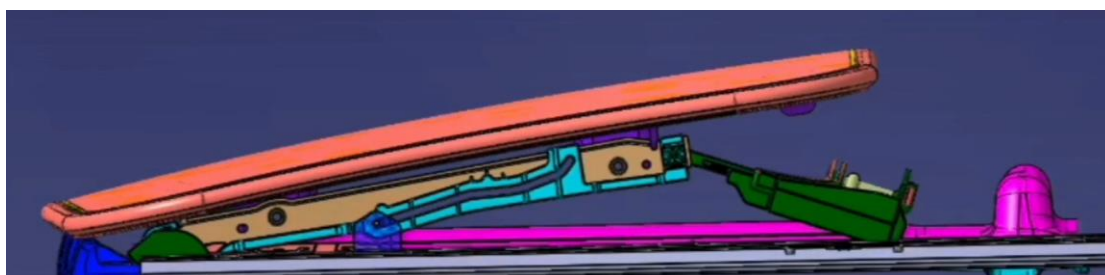
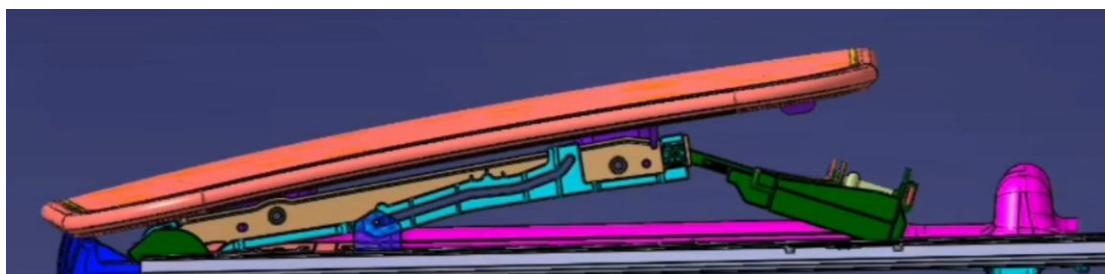


圖 2-17 Close-Tilt 動作

Tilt 動作



Close 動作



圖 2-18 Tilt-Close 動作

2-3-2 資料集

天窗樣件使用阿拉伯數字編號做紀錄，每筆樣件會錄製 7 次，舉例來說，若是第一筆天窗樣件，且錄製第一次，則檔名為 NO1-1。機構動作分為四個動作，分別為 Close-Open(簡稱 C-O)、Open-Close(簡稱 O-C)、Close-Tilt(簡稱 C-T)及 Tilt-Close(簡稱 T-C)。為便於後續對各動作中正常(OK)與異常(NG)的天窗樣件進行比較，每個動作均需設置專用代號。根據天窗異音肇因紀錄表顯示，在 C-O 動作中，共有五種機構異常引發異音，故將各異常樣件命名為 NG1，而相應的對照組則標記為 OK1。例如，第五筆天窗樣件之錄製第五次，在 C-O 動作出現 NG3 的異常，則以 C-O_NG3_NO5-5 為代號。

天窗樣件共有 58 筆，理論上可用樣件為 406 筆音訊，但因架設錄音設備時有誤，因此總訊號筆數為 255，而各異音之異常樣件如下表，正常樣件為總訊號筆數扣除異常樣件總筆數，以 NG1 舉例，NG1 異常樣件總筆數為 18，正常樣件總筆數就為 237 筆，如表 2-3。

表 2-3 Close-Open 動作之 NG1~NG5 發生異常樣件編號與異常樣本數

	發生異常之天窗樣件編號	異常總筆數
NG1	1, 5-1, 11, 56-1, 56-2, 57-1	18
NG2	1, 4-2~7, 5, 6, 7, 10, 11, 13, 23, 24, 26, 27, 29-2~7, 31-2~7, 50, 51, 53, 54, 56, 57	144
NG3	16-1, 17-1, 24-1, 25-1	4
NG4	5, 13	14
NG5	6, 9	14

2-3-3 天窗異音肇因紀錄

零組件之搭配與組裝過程可能會受生產梯次而略有不同，故每批生產天窗之異音不盡然相同，由車用電動天窗製造廠所提供之錄製訊號中，同樣也有此現象。換句話說，在錄製訊號之過程所蒐集之天窗樣件，僅在 Close-Open 動作出現異音，其他天窗動作並無出現異音，因此本研究僅針對 Close-Open 動作進行分析。根據品保師提供之 Close-Open 異音肇因紀錄表如表 2-4，又 NG6~10 為外部噪音，本研究不分析。

表 2-4 異音肇因紀錄表

	異音肇因	異音描述	異音聲響描述
NG1	滑塊與滑槽搭配不佳	撞擊音	咚
NG2	滑塊與滑槽搭配不佳	撞擊音	咚
NG3	彈簧未定位好	敲擊音	喀
NG4	擋風板回彈	撞擊音	搭
NG5	拉繩摩擦音	摩擦音	白噪音
NG6	鋁擠型治具振動音		
NG7	外部汽車移動		
NG8	0.1sec按鈕回彈音		
NG9	椅子的聲音		
NG10	鞋子的聲音		

1. 異常編號 NG1、NG2

當滑軌潤滑不足，導致滑塊開始滑動瞬間磨擦力過大出現撞擊聲。或零組件之幾何大小超過公差設定範圍，導致機構零件間搭配不良，因此發生異音。

異常發生區域如圖 2-19 中紅框。

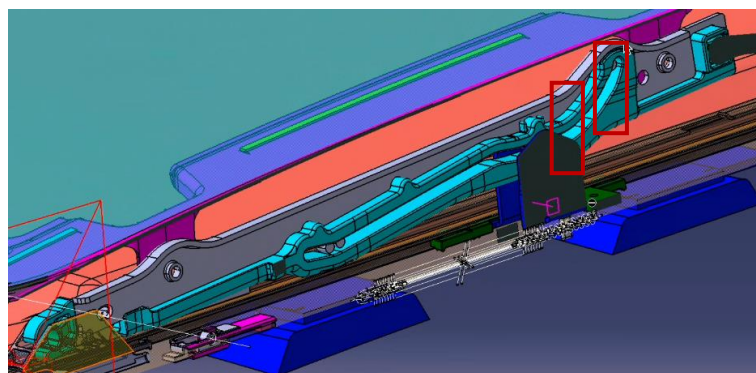


圖 2-19 天窗滑軌(淺藍色區塊)

2. 異常編號 NG3

當 Close 階段結束時，擋風板將同步升起，而此作動係賴於彈簧機構。當彈簧安裝出現過大偏差，將在擋風板回彈前出現清脆的敲擊音。異常發生區域

如圖 2-20 中紅框。

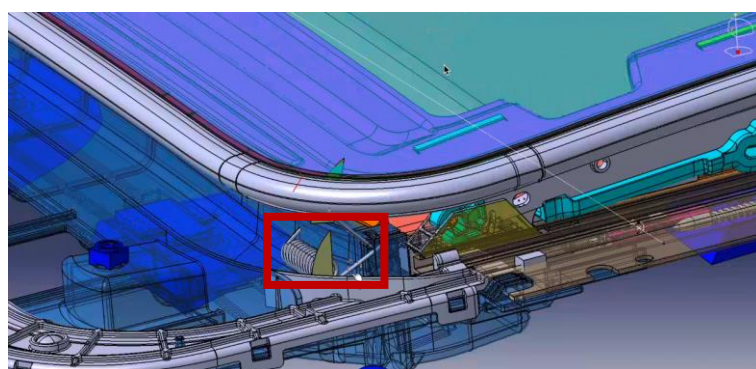


圖 2-20 彈簧機構

3. 異常編號 NG4

擋風板升起之設計採用彈簧機構來驅動擋風板的上升，並未使用電控系統調整其姿態。若機構安裝出現過大偏差，則會在彈簧釋放後，產生較大之上升能量導致撞擊聲，如圖 2-21，實心藍色為擋風板升起後，淺藍色為擋風板未升起之狀態。

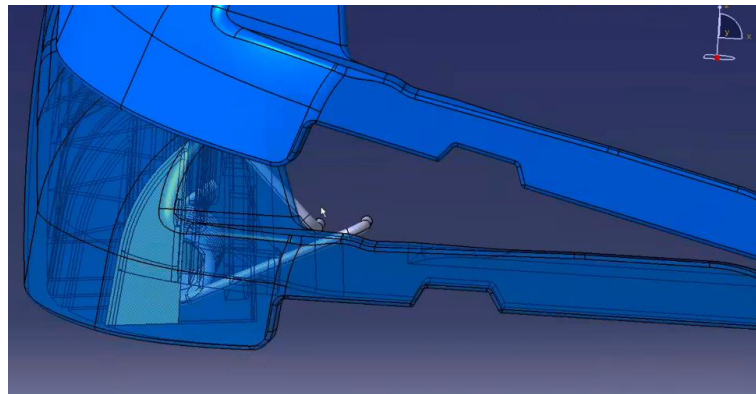


圖 2-21 擋風板機構

4. 異常編號 NG5

天窗在 Close 之位置時僅有天窗上下位移，但 Open 動作開始後，天窗玻璃將會向後水平移動直到完全開啟，類似於座椅旁的電動車窗隨著開啟而使對外開口越來越大。而拉繩摩擦音正是源於此產生，車用天窗之 Open 是仰賴於馬達對拉繩的牽引進而控制其開啟程度。而天窗開啟也並非瞬間就開啟完成，因此 NG5 被標註拉繩摩擦音為長時間的異常，是出現在天窗 Close 動作開始後，也就是 NG2 之後會出現，並持續出現到完全開啟。

2-3-4 天窗異音初探

觀察天窗音訊前，我們首先對聲學量測物理量做討論，電容式麥克風量測得到之原始訊號為電壓(V)，經靈敏度換算後，可以獲得聲壓(Pa)。進一步對聲壓取對數則可獲得聲壓級(Sound Pressure Level, SPL)，然而 SPL 為對數尺度，與人耳實際聽覺仍存在落差；若需評估人耳主觀響度(Loudness)，可將 SPL 轉換為響度級(phon)或音量感(sone)。為了取得天窗聲音的最原始物理量，本研究以聲壓(Pa)為分析單位。

2-3-4-1 時間域訊號

車用電動天窗製造廠在錄製音訊時，也同時對原始錄製音訊進行 A 加權濾波。A 加權濾波旨在將聲音頻譜依據人耳對不同頻率的敏感度進行加權，以衰減人耳對低頻與高頻不敏感的成分，並使測量結果更符合人類主觀聽感。原始錄製訊號與 A 加權訊號如圖 2-22。

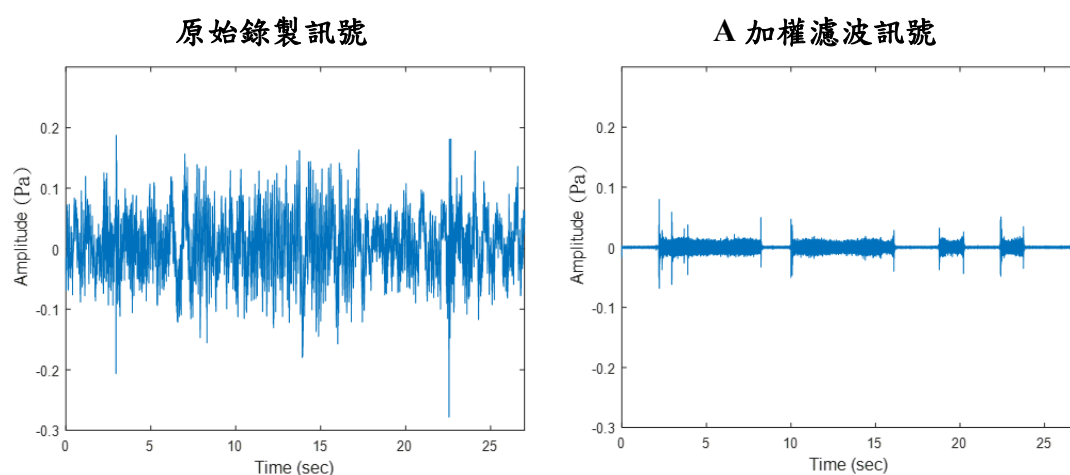


圖 2-22 原始錄製、A 加權濾波之訊號

當訊號經過 A 加權濾波後，便可直接觀察到天窗運行之四個動作 Close-Open、Open-Close、Close-Tilt、Tilt-Close，如圖 2-23。



A 加權濾波訊號

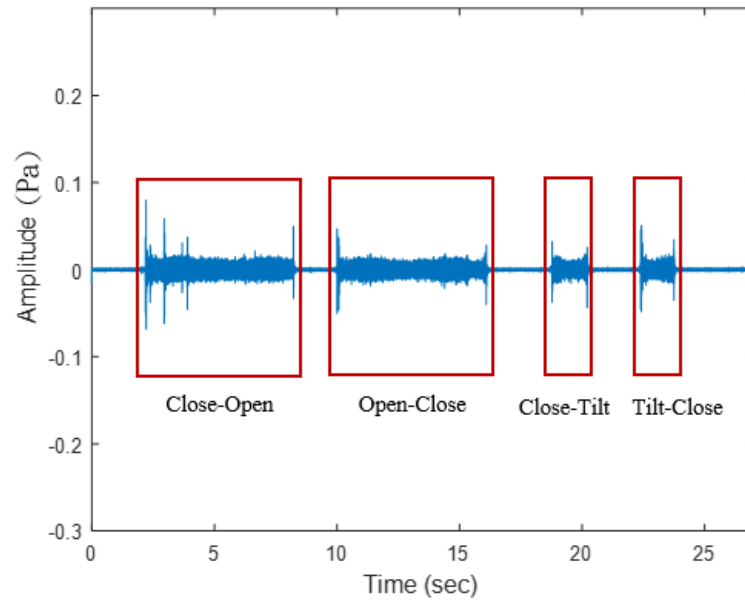


圖 2-23 天窗運行的四個動作

2-3-4-2 時間-頻率域訊號

將 Close-Open 動作之訊號提取出來後，為了快速獲得天窗異音的時間發生點與其頻率分布，首先對訊號進行短時傅立葉轉換(Short-Time Fourier Transform, STFT)，並以能量顯示其頻率貢獻。我們以樣件 NO.1-1 觀察 NG1、NG2，以樣件 NO.16 觀察 NG3，以樣件 NO.5 觀察 NG4，以樣件 NO.6 觀察 NG5，並以 NO.2 作為沒有被標註 NG 之樣件作為對照組。

經比較得知，異音的發生不僅是在特定時間出現亮帶，在品保師的判定中存在一閾值，當某一 NG 之聲響超過容許值時才會被判定為 NG 樣件，也因此 OK 樣件同樣會出現亮帶。NG5 之異常肇因為拉繩摩擦音，係因天窗向後拖移並開啟時是藉由馬達帶動拉繩並拖移天窗進而開啟，此機制是在 NG2 後作動。換句話說，發生異音之源頭為機構不匹配所造成，而固定的機構僅會造成相對應的異音，統整後可以將各 NG 之發生異常時間區段做歸納，且利於後續之特徵觀察，如表 2-5。

表 2-5 異音發生時段

	發生異常時間區段, sec
NG1	0~0.5
NG2	0.5~0.9
NG3	0.9~1.1
NG4	1.5~2
NG5	Continuously after NG2

此外我們也發現，不論 OK、NG 樣件都出現持續性亮帶，分別為錄製時的背景噪音，與天窗馬達運轉聲。背景噪音之頻帶落在 0~80Hz 區間，且其能量值非常大。天窗馬達運轉聲則約莫在 400~500Hz，900~1000Hz，1400~1500Hz 出現水平亮帶，並且忽大忽小，不規律的出現，不過這兩種持續性亮帶皆沒有出現在異音肇因紀錄表中。

1. NG1、NG2

圖 2-24 中，異常樣件之時頻圖約在 0.2 秒處、0~300Hz 出現亮帶為 NG1。約在 0.8 秒處、0~450Hz 出現亮帶則為 NG2。而正常樣件之時頻圖在同一時間附近出現亮帶，只是能量沒有 NG 大。

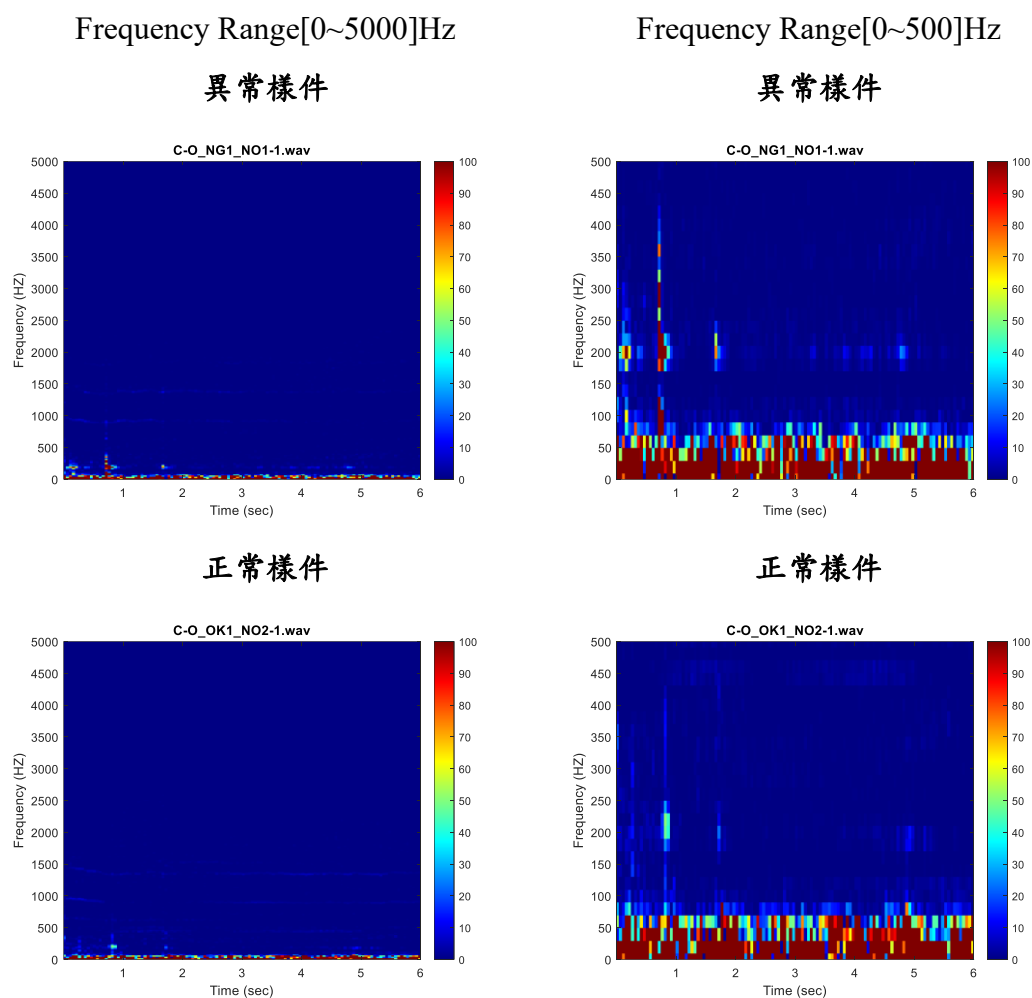


圖 2-24 NG1、NG2 之 STFT

2. NG3

圖 2-25 中，異常樣件之時頻圖中，約在 1 秒處、0~5000Hz 出現寬頻的亮帶為 NG3。正常樣件在 1 秒處則無此亮帶。

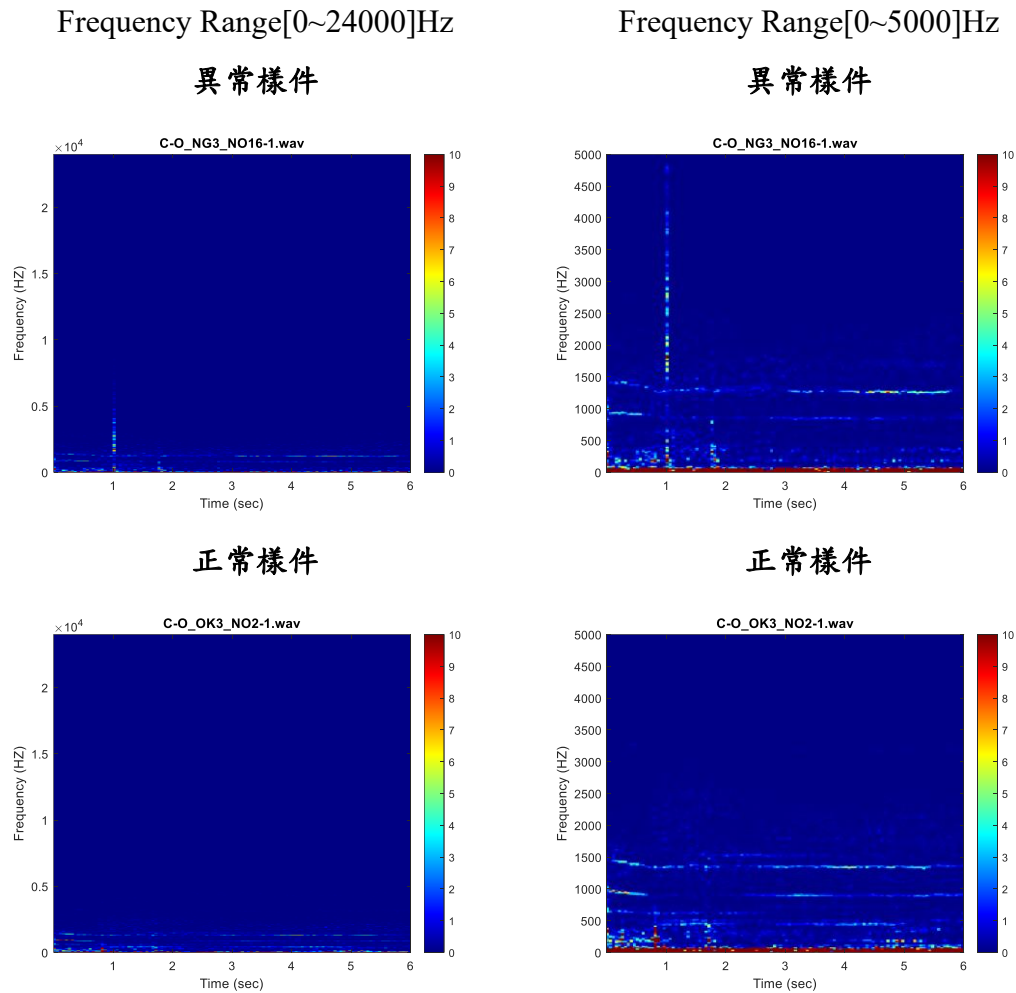


圖 2-25 NG3 之 STFT

3. NG4

圖 2-26 中，異常樣件之時頻圖中，約在 1.7 秒處、80~2900Hz 出現亮帶為 NG4。而正常樣件之時頻圖在同一時間附近出現亮帶，只是能量沒有 NG 大。

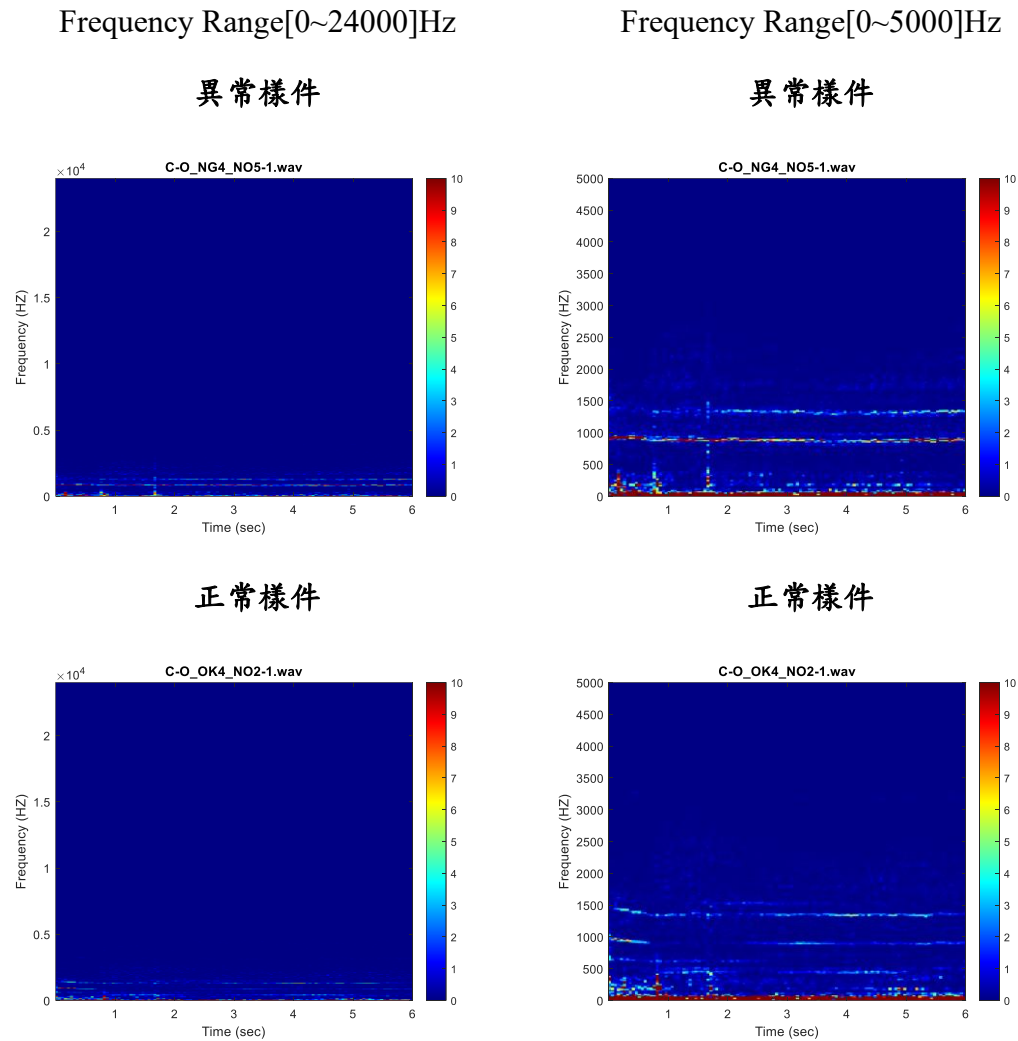


圖 2-26 NG4 之 STFT

4. NG5

圖 2-27 中，異常樣件之時頻圖中，約在 0.9 秒後、約 1500~3500Hz 出現持續性亮帶為 NG5。而正常樣件之時頻圖中，同樣在 0.9 秒後、約 1500~3500Hz 出現持續性亮帶，僅其能量不如標註 NG 大。

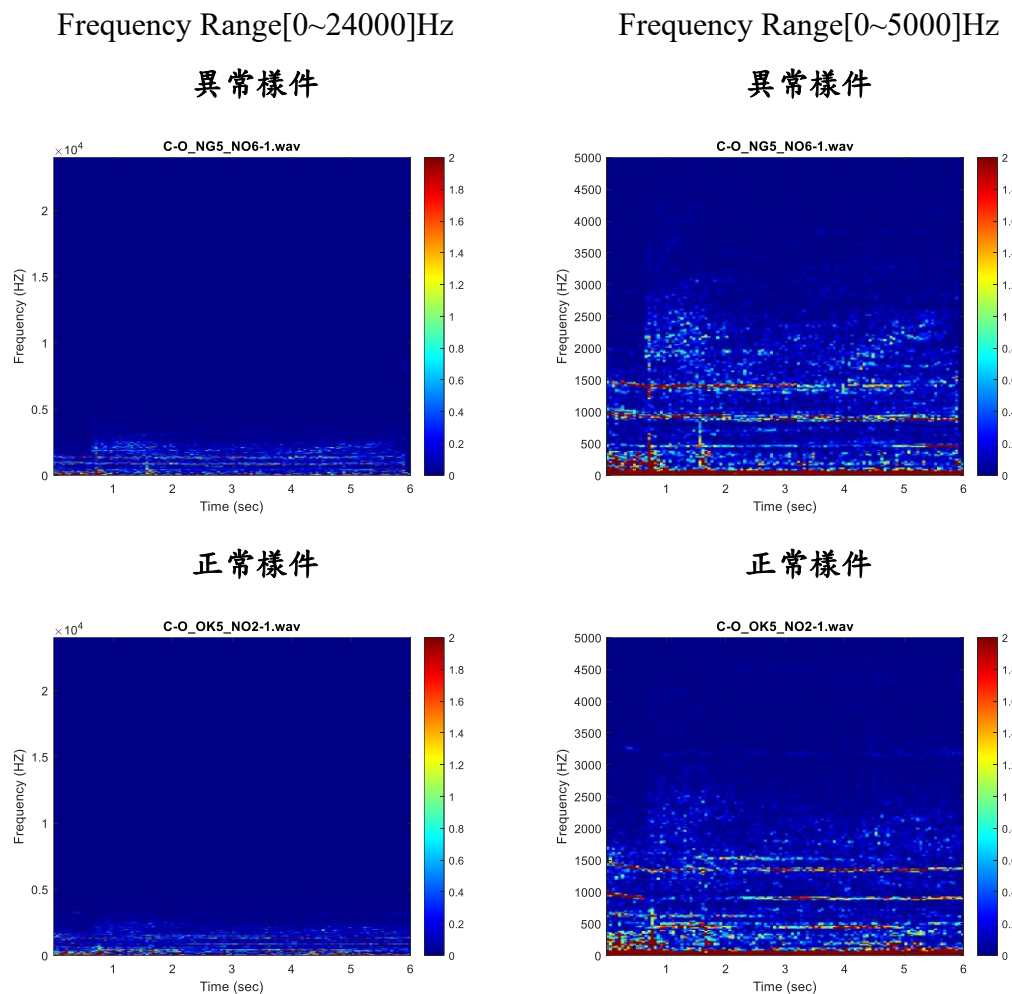


圖 2-27 NG5 之 STFT

第三章 訊號處理方法



訊號處理領域中，研究人員可依據分析目的，對蒐集到的離散訊號進行適切的表示與轉換，以此描述輸入訊號的物理現象。為觀察訊號在時間與頻率上的變化，常採用短時傅立葉轉換(Short-Time Fourier Transform, STFT)。然而，STFT 之線性頻率尺度與人耳的非線性感知特性並不完全一致。為貼近語音與一般聽覺任務的感知特徵擷取需求，近年廣泛採用感知導向的表示法。梅爾頻譜(Mel spectrogram)透過梅爾尺度重新取樣頻率軸，增加低頻的解析度，使頻率分辨率較符合人耳感受；進一步地，梅爾倒頻譜係數(Mel-Frequency Cepstral Coefficients, MFCC)則在梅爾頻譜上施以對數壓縮，再以離散餘弦轉換(Discrete Cosine Transform, DCT)達成降維與資訊濃縮，成為語音辨識與機器學習中常用的聲學特徵。

本研究旨在區分天窗正常(OK)與異常(NG)樣件。為了讓後續的分類模型能夠擷取並解釋具物理意義的訊號特徵，本章首先對原始的一維天窗音訊進行前處理，以簡化分析流程。隨後比較不同的訊號處理方法，作為後續特徵擷取的依據。

3-1 訊號前處理

為進一步探討異音特徵，本研究先對原始音訊施以適切之前處理，並僅保留後續分析所需之訊號區段。

3-1-1 切割訊號

完整的天窗錄製訊號包含四個動作，從目前蒐集到的音訊中，異音只出現在 Close-Open 動作，且過長的訊號難以觀察異音特徵，故對訊號做切割。原始錄製訊號經 A 加權濾波後，能清楚觀察到天窗動作之間有一振幅極小之區間，稱作靜音段。故我們可以從 A 加權訊號找到一閾值能夠區切割天窗的四個動作，如圖 2-23。A 加權之理論將於本章後續進一步探討。

切割方法如下述：

1. 對 A 加權訊號做絕對值，以訊號之一倍平均值加兩倍標準差作為天窗是否動作的閾值(Threshold)，如圖 3-1 之紅色水平線。

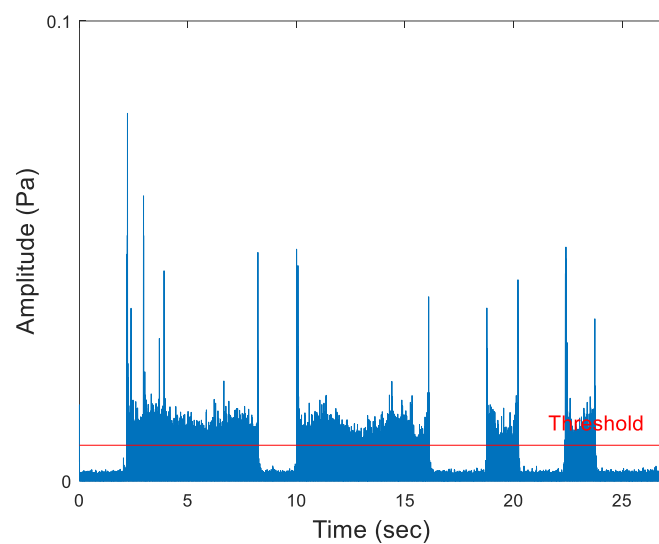


圖 3-1 A 加權振幅絕對值訊號

2. 為了將四個動作分群，若超過閾值之訊號前後時間差距大於 0.5 秒，則判定為下一動作之起點，如圖 3-2。

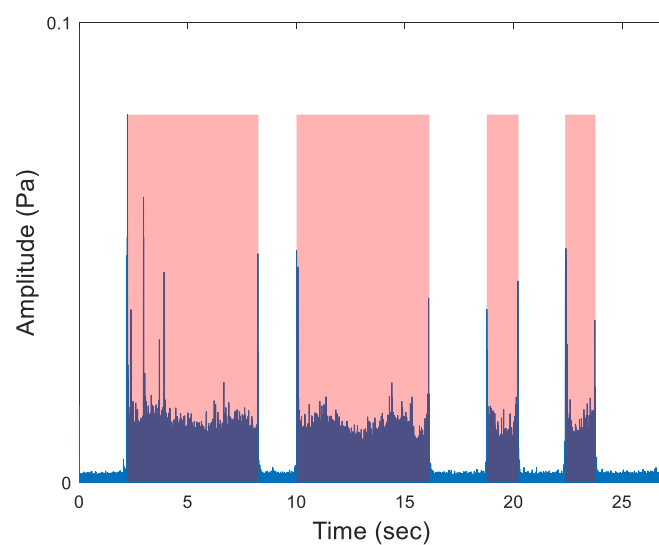


圖 3-2 A 加權振幅絕對值訊號

3. 將 A 加權訊號之切割起始點儲存，如圖 3-3，再次切割原始錄製訊號，如圖 3-4。



A 加權訊號

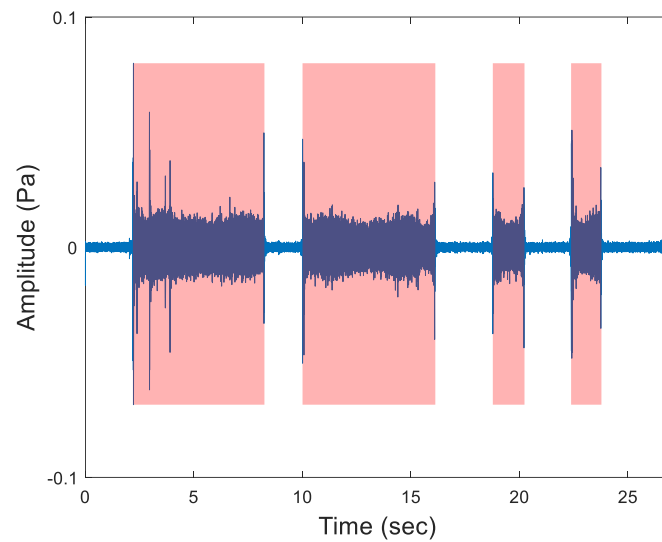


圖 3-3 A 加權訊號切割結果

原始錄製訊號

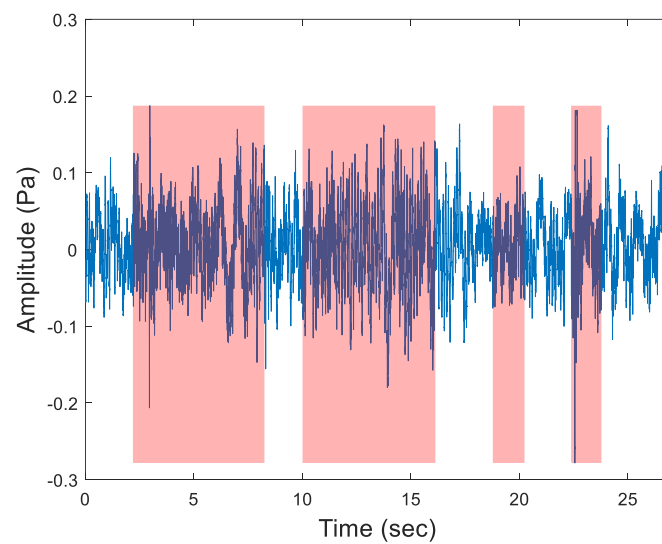


圖 3-4 原始錄製訊號切割結果

4. 將 Close-Open 動作訊號萃取出來後，最後再將頭尾之無效訊號切除，紅框所代表之訊號為天窗馬達未作動段，綠色框代表馬達開啟聲，因為聲響較大且沒有被歸類在異音肇因紀錄表中，故將其切除，如圖 3-5。

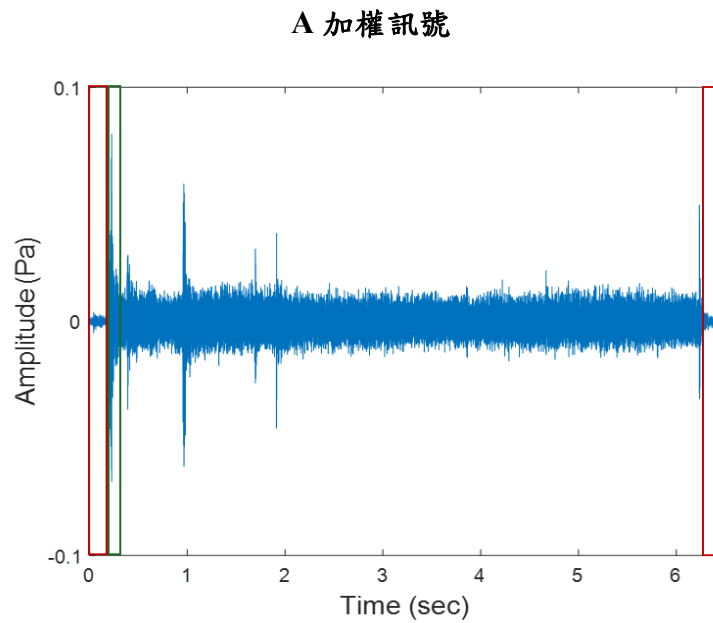


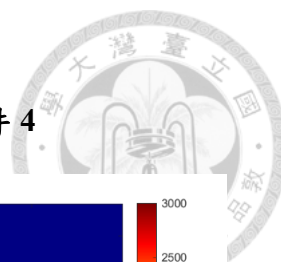
圖 3-5 切除無效訊號

3-1-2 濾波

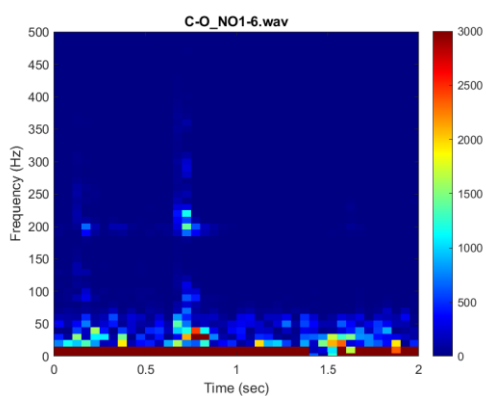
為避免遺失異音特徵，本研究採用原始錄製訊號進行分析，而非基於 A 加權濾波後的訊號。儘管 A 加權濾波能夠模擬人耳聽覺特性，但其濾波過程會削弱異音特徵所在頻段的能量，不利於後續對異音特徵的辨識。此外，雖然錄音環境部分為無響室，但錄音地點鄰近產線，錄製音訊中仍可能摻雜大型貨車振動與產線背景噪音。若這些背景噪音能量過大，將導致後續對異音之判讀產生偏差，因此有必要對原始錄製訊號進行適當濾波處理。

首先，背景噪音之特性應為持續，且其特徵頻帶應當差不多，為濾除背景噪音，本研究對所有被標註有 NG1、2、4 與無標註異常之天窗錄製訊號進行短時傅立葉轉換(STFT)，並將頻率範圍限制於 0 至 500 Hz，以利觀察低頻亮帶隨時間的變化，如圖。從 Close-Open 動作的 STFT 結果可見，不論天窗樣件是否產生異常聲響，大約在 80Hz 以上已經看不到背景噪音了，如圖 3-6 至圖 3-10。

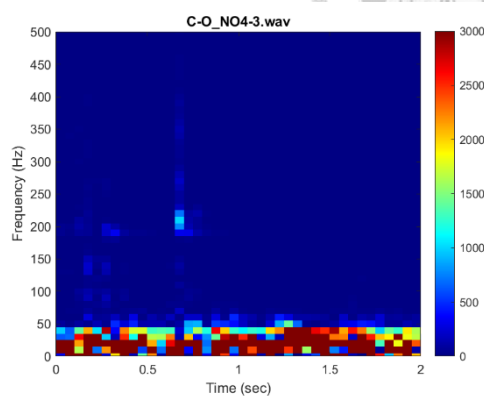
同樣的，在 2-3-4-2 節中我們也發現 NG1~NG5 在低頻約 0~80Hz 部分出現能量非常大的持續性亮帶，為了降低背景噪音的影響，本研究選擇對原始錄製訊號進行 80Hz 之高通濾波。



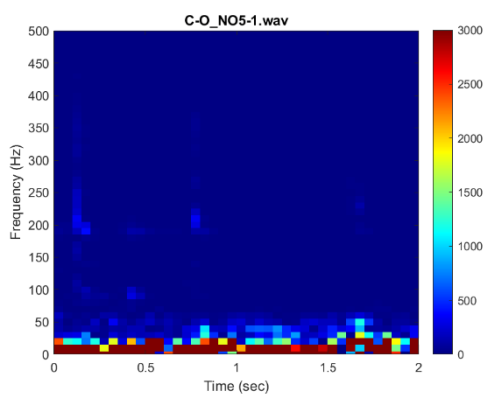
天窗樣件 1



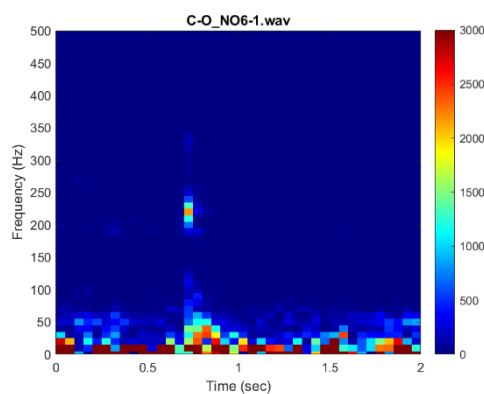
天窗樣件 4



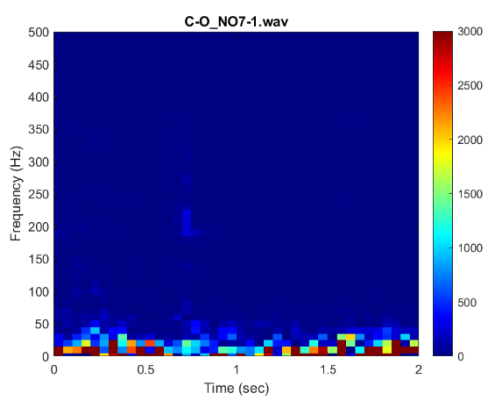
天窗樣件 5



天窗樣件 6



天窗樣件 7



天窗樣件 10

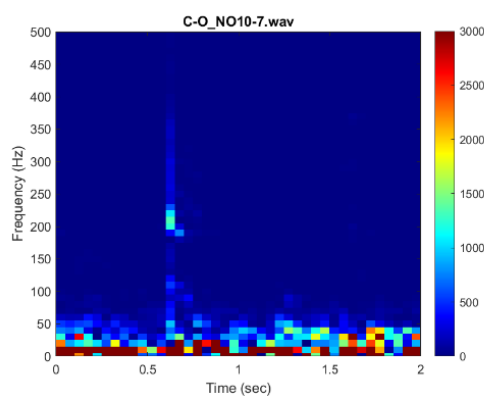
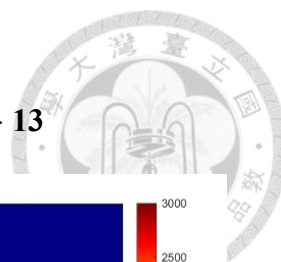
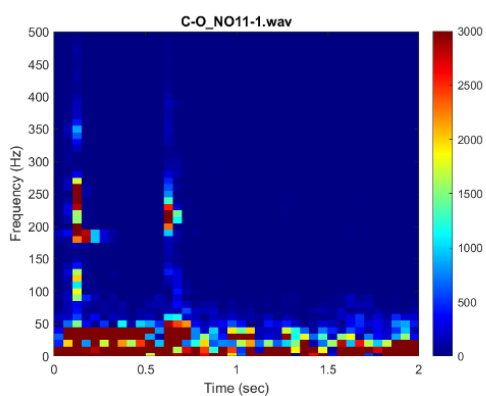


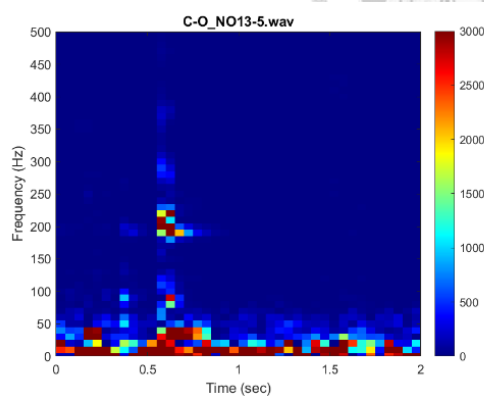
圖 3-6 被標註 NG1、2、4 之天窗樣件 STFT



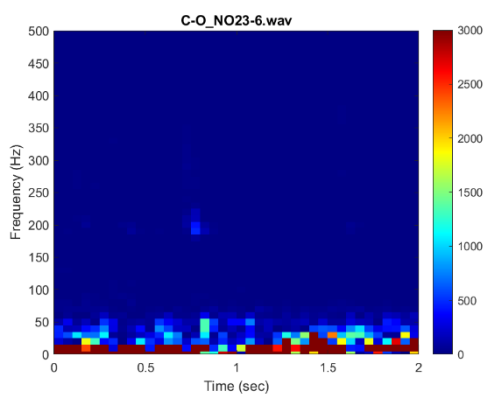
天窗樣件 11



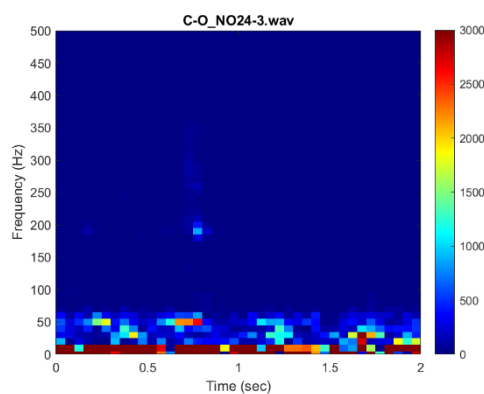
天窗樣件 13



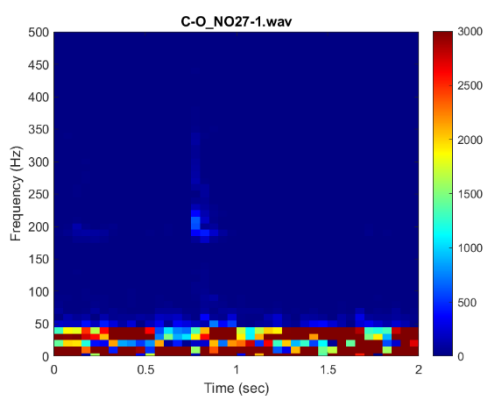
天窗樣件 23



天窗樣件 24



天窗樣件 27



天窗樣件 50

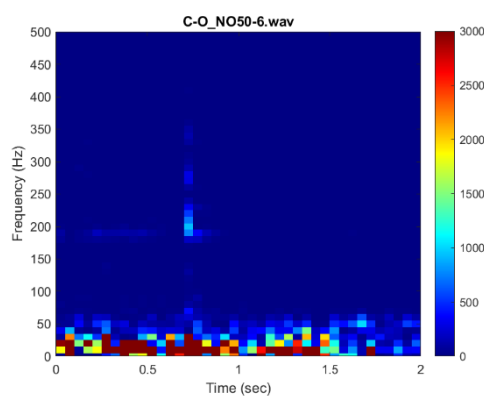
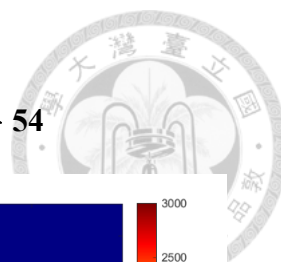
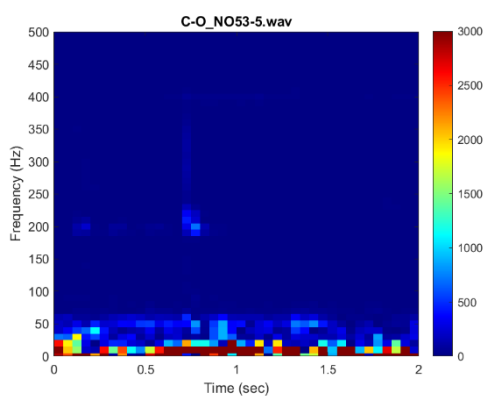


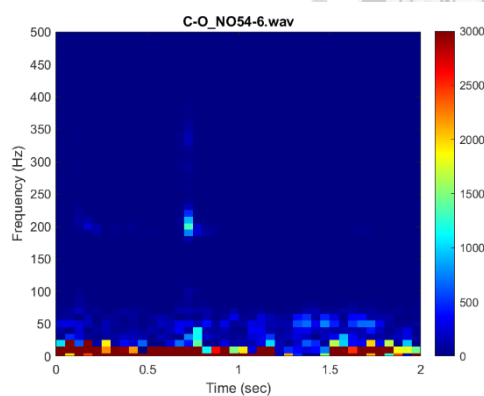
圖 3-7 被標註 NG1、2、4 之天窗樣件 STFT



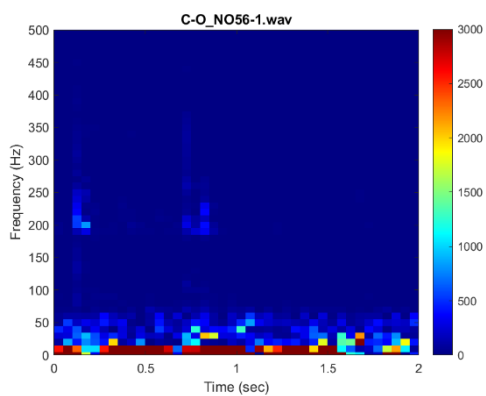
天窗樣件 53



天窗樣件 54



天窗樣件 56



天窗樣件 57

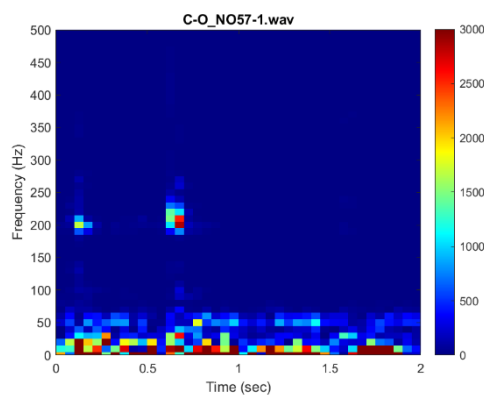
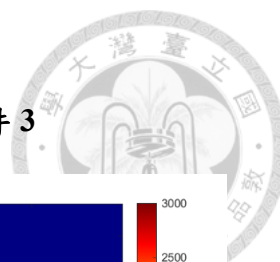
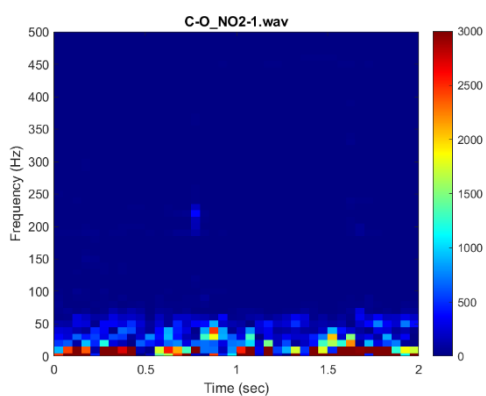


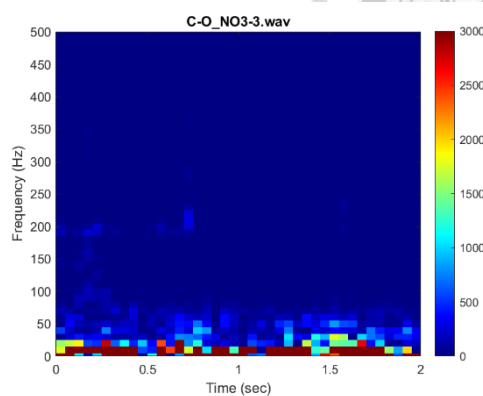
圖 3-8 被標註 NG1、2、4 之天窗樣件 STFT



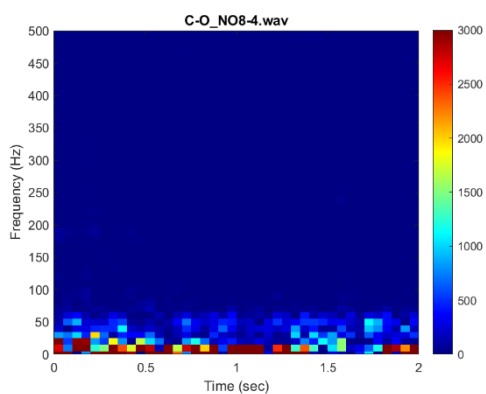
天窗樣件 2



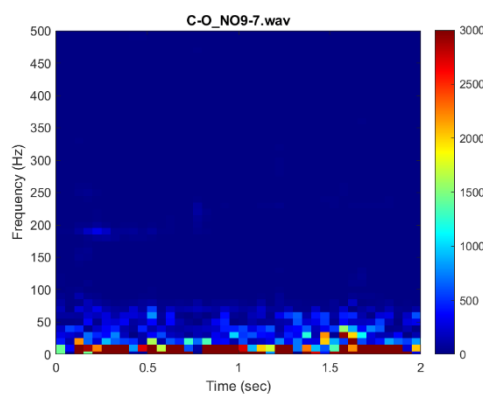
天窗樣件 3



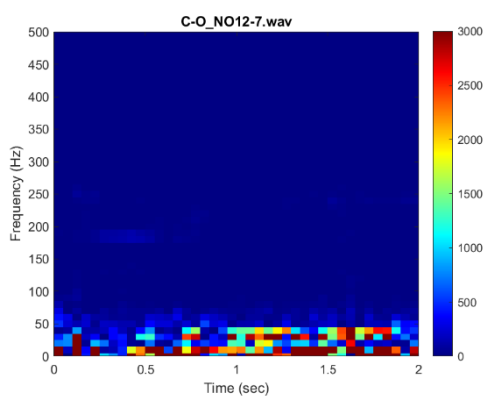
天窗樣件 8



天窗樣件 9



天窗樣件 12



天窗樣件 14

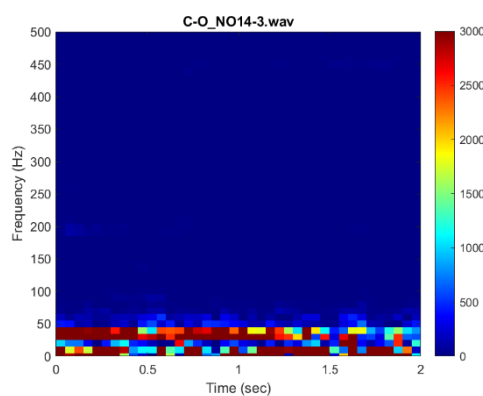
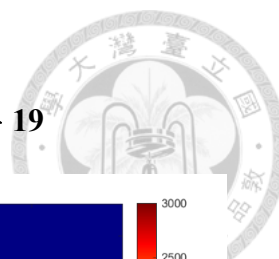
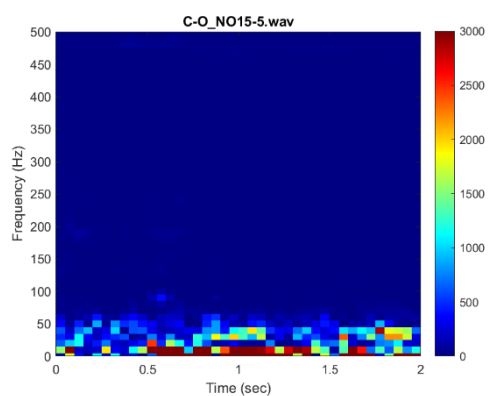


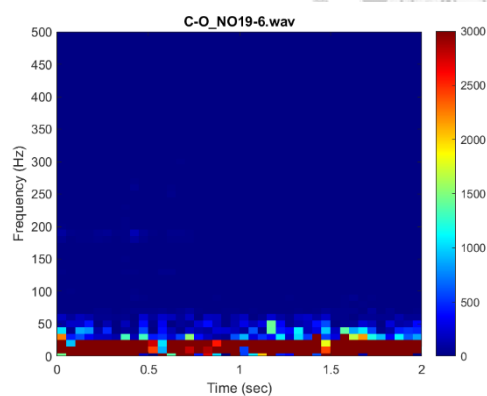
圖 3-9 無標註異常之天窗樣件 STFT



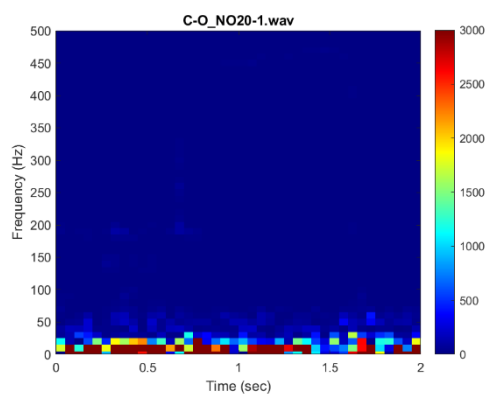
天窗樣件 15



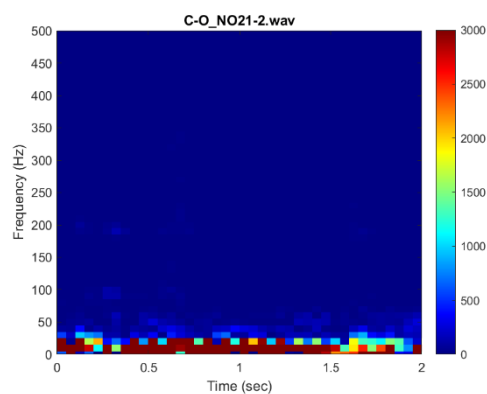
天窗樣件 19



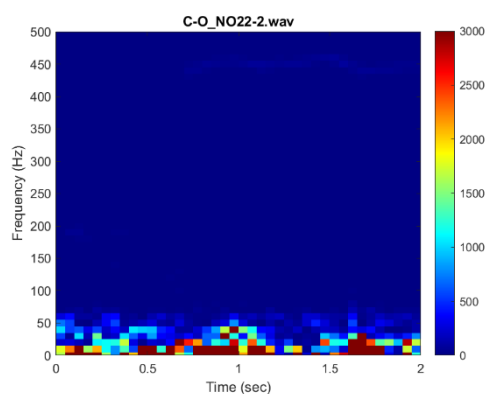
天窗樣件 20



天窗樣件 21



天窗樣件 22



天窗樣件 26

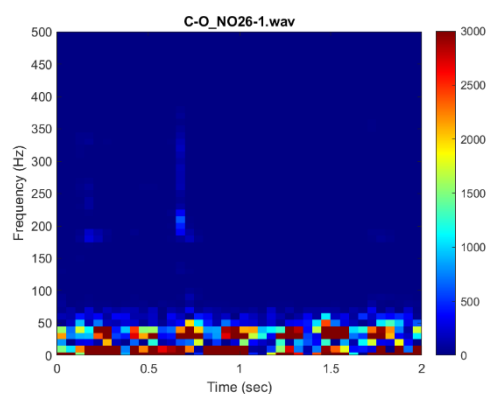


圖 3-10 無標註異常之天窗樣件 STFT

3-2 A 加權濾波器

A 加權濾波器(A-weighting)，如式 3-1，A 加權濾波曲線如圖 3-11，圖中 x 軸為頻率軸，y 軸為原始訊號之加權因子。是一種用於噪音測量的頻率加權方法，其主要目的是模擬人耳對不同頻率聲音的聽覺靈敏度。人耳對中頻(大約 1000 ~5000 Hz)最為敏感，而對低頻和極高頻的響應則較弱，因此 A 加權濾波器在這些頻段分別施加了較大的衰減。這樣一來，經過 A 加權處理後的聲音訊號，將更符合噪音對人耳影響的實際情況。

$$A(f) = 20 \log_{10} \left(\frac{12200^2 \cdot f^4}{(f^2 + 20.6^2)(f^2 + 12200^2) \sqrt{(f^2 + 107.7^2)(f^2 + 737.9^2)}} \right) \quad (3-1)$$

式中， f 為原始訊號之頻率， $A(f)$ 為加權值。

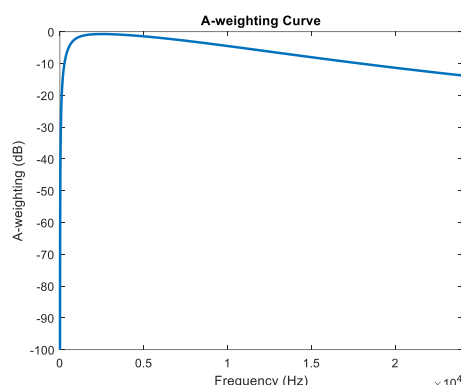


圖 3-11 A-weighting Curve

車用電動天窗製造廠在分析天窗錄製音訊前，為了對訊號做出更符合人耳在行車時對異音的敏感度，因此除了原始錄製訊號，也同步輸出 A 加權濾波訊號，如圖 3-12，並以紅色框標示出天窗四個動作。

A 加權濾波雖能將原始錄音轉換為符合人耳聽覺靈敏度之加權頻譜，然其濾波曲線對全頻帶進行調整，且在低頻與高頻成分尤為抑制。若異音之特徵頻帶落於衰減較大之區域，則異音之特徵頻帶將被削弱，並可能導致異音判讀失真。

根據第 2-3-4-2 節的 STFT 觀察結果，NG1 與 NG2 的異音在 100~500 Hz 之間出現明顯能量亮帶。根據 A 加權理論，當異音之特徵頻帶位於 0~1 kHz 區段時，A 加權將使其能量顯著下降；換言之，對 NG1 與 NG2 而言，若採用 A 加權，則 100~400 Hz 的異音頻帶將幾乎消失。為了佐證此一差異，本研究對 A 加權濾波訊號與原始錄製訊號之 NG1 時段訊號進行快速傅立葉轉換(FFT)，如圖 3-13。結果可見，雖 A 加權有效抑制背景噪音，卻同時削弱了異音的特徵頻帶。因此，本研究僅由 A 加權訊號作為切割天窗四個動作之依據。

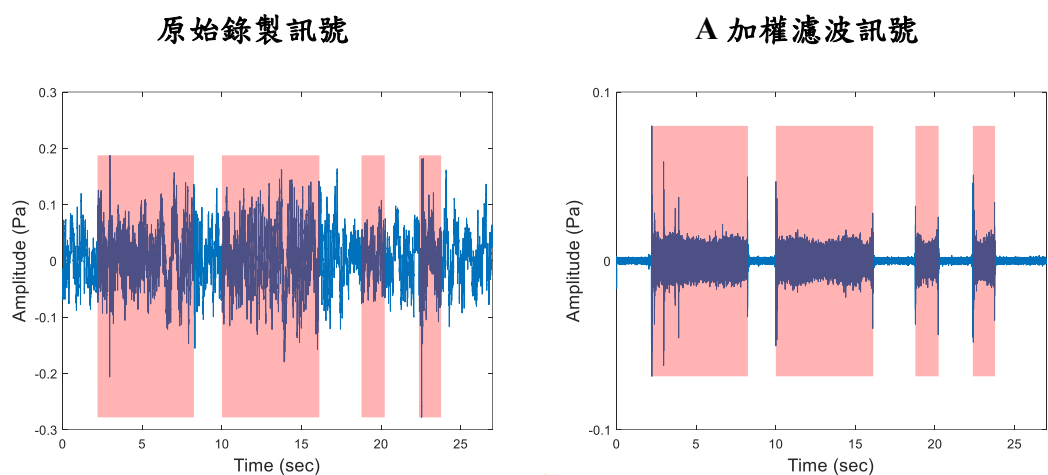


圖 3-12 原始錄製、A 加權之時間域訊號

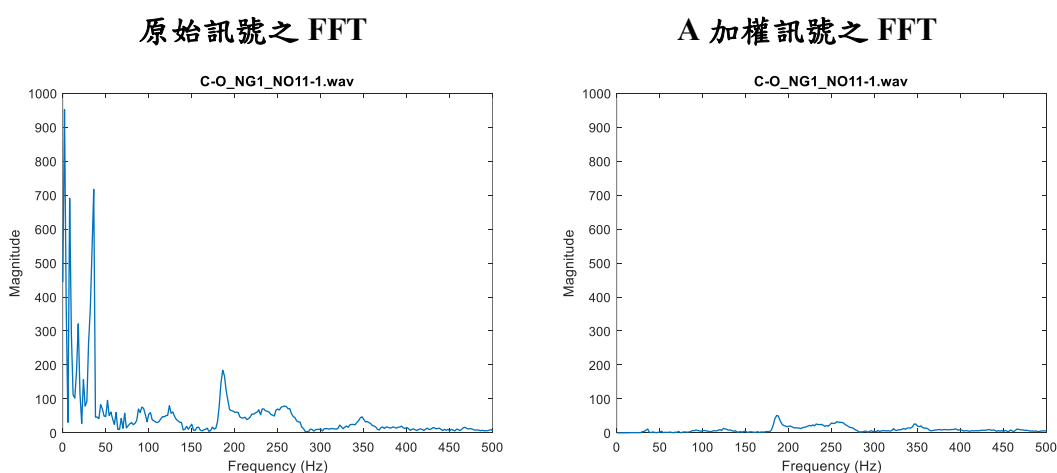


圖 3-13 原始訊號、A 加權訊號之 FFT

3-3 傅立葉轉換

十八世紀末到十九世紀初，法國數學家約瑟夫·傅立葉在研究熱傳導現象時，首次提出將任意週期函數以正弦和餘弦波形線性組合的思路，並在其著作《熱的解析理論》中詳細闡述了傅立葉級數的原理。隨後，他又將此方法推廣至非週期性訊號的處理，形成傅立葉積分，最終演變為今日所稱的傅立葉轉換，其關鍵在於將時間域的訊號映射到頻率域，以揭示訊號中各頻率成分的貢獻。

在訊號分析中，積分變換是一種通過將原始函數與核函數進行積分運算，而將該函數從一個函數空間映射到另一個函數空間的方式，如式 3-2，而傅立葉轉換即是利用複數指數函數作為核函數。

$$X(s) = \int_{-\infty}^{+\infty} K(s,t)x(t)dt \quad (3-2)$$

式中， $X(s)$ 為變換後的函數， $K(s,t)$ 為核函數， $x(t)$ 為原始函數。

3-3-1 傅立葉轉換原理

考慮一連續有限時間域訊號 $x(t)$ ，若其絕對值可積分(Absolute Integrable)且積分值收斂(converges)，如式 3-3，則傅立葉正轉換(Fourier transform)，如式 3-4，傅立葉逆轉換(Inverse-Fourier transform)存在，如式 3-5。

$$\int_{-\infty}^{+\infty} |x(t)| dt < \infty \quad (3-3)$$

傅立葉正轉換(Fourier transform)如下：

$$\mathfrak{F}\{x(t)\} = X(\omega) = \int_{-\infty}^{+\infty} x(t)e^{-i\omega t} dt \quad (3-4)$$

傅立葉逆轉換(Inverse-Fourier transform)則為：

$$\mathfrak{F}^{-1}\{X(\omega)\} = x(t) = \int_{-\infty}^{+\infty} X(\omega)e^{i\omega t} dt \quad (3-5)$$

式中， t 為時間(sec)， ω 為角頻率，由式(3-4)可將原始訊號解讀為藉由傅立葉核函數為基底展開。

又因

$$e^{i\omega t} = \cos(\omega t) + i \cdot \sin(\omega t) \quad (3-6)$$

可將核函數分解為奇函數(Odd Function)與偶函數(Even Function)，如式(3-6)。

換句話說，傅立葉轉換之物理意義就是利用複數指數函數將原始訊號展開正、餘弦之週期函數，如式 3-7：

$$\begin{aligned} X(\omega) &= \int_{-\infty}^{+\infty} x(t)e^{-i\omega t} dt = \int_{-\infty}^{+\infty} x(t)(\cos(\omega t) - i \cdot \sin(\omega t))dt \\ &= \int_{-\infty}^{+\infty} x(t)\cos(\omega t)dt + i \cdot \left(-\int_{-\infty}^{+\infty} x(t)\sin(\omega t)dt\right) \end{aligned} \quad (3-7)$$

因

$$\begin{aligned} X_{RE}(-\omega) &= X_{RE}(\omega) \\ X_{Im}(-\omega) &= -X_{Im}(\omega) \end{aligned} \quad (3-8)$$

得

$$X(\omega) = X_{RE}(\omega) + i \cdot X_{Im}(\omega) \quad (3-9)$$

式 3-8 表明若原始訊號為偶函數， $X(\omega) = X_{RE}(\omega)$ ；原始訊號為奇函數，則 $X(\omega) = i \cdot X_{Im}(\omega)$ ；式 3-9 表明若原始訊號為非奇非偶函數。

傅立葉轉換藉由傅立葉核函數為基底展開，其對訊號之整段時間做積分後，失去時間變化之成分，僅留下該時間內之頻率大小分布。而當一訊號經傅立葉轉換從時間域完全轉換至頻率域後，將失去頻率隨時間變化資訊。本研究之天窗異音出現在不同時間，僅由傅立葉展開不足以對各異音進行討論，因其無法描述聲音頻率隨時間之變化。

3-4 短時傅立葉轉換

短時傅立葉轉換(Short-Time Fourier Transform, STFT)的概念可追溯至 1946 年由丹尼斯·加博爾(Dennis Gabor)所提出的加博爾變換(Gabor transform)。具體來說，他在對較長之訊號做傅立葉轉換前，首先引入窗函數的概念，並使用窗函數將較長之訊號切割成許多一小段訊號，再對每一小段訊號單獨進行傅立葉轉換，以同時獲取時間和頻率資訊。。

後來的研究者便在這一思想上加以擴展，完善了我們如今所稱的 STFT 分析方法。相較於一次性處理整段訊號僅能獲得一維頻率資訊，STFT 以窗函數對訊號分段進行傅立葉轉換後，不僅保留了頻域特性，同時也將時間維度納入分析，將原本單一的頻率域表示，延伸為兼具時間與頻率的二維時頻圖譜。

3-4-1 短時傅立葉轉換原理

短時傅立葉轉換為將原始訊號 $x(\tau)$ 乘上一選定的窗函數(Window Function)，並將此窗函數隨時間間隔平移，如式 3-10：

$$STFT(t, f) = \int_{-\infty}^{+\infty} x(\tau)w(t-\tau)e^{-i2\pi f\tau} d\tau \quad (3-10)$$

式中， $x(\tau)$ 為原始訊號， $w(t-\tau)$ 為窗函數。

為確保異音成分不受窗函數的旁瓣特性而遭到衰減，本研究在短時傅立葉轉換中採用了如圖 3-14 所示的矩形窗(Rectangular window)。而由式 3-10 之計算值為複數譜之絕對值，若再對其取平方即可得到能量頻譜(Power Spectrogram)，其物理意義即為訊號之各頻率分量隨時間之能量變化。如式 3-11。

$$|STFT(t, f)|^2 = \left| \int_{-\infty}^{+\infty} x(\tau)w(t-\tau)e^{-i2\pi f\tau} d\tau \right|^2 \quad (3-11)$$

式中， $|STFT(t, f)|^2$ 代表 STFT 能量譜。

短時傅立葉轉換係將時間長度為 T 之訊號分割為多個時間長度較短的 ΔT ，再分別計算長度 ΔT 之傅立葉轉換，最後以色階圖表示傅立葉頻譜，而時頻圖中之長度等於 ΔT 之傅立葉頻譜稱為一個幀(Frame)。

當我們使用短時傅立葉轉換進行訊號分析時，因為是對每一個幀進行傅立葉轉換，故頻率解析度為 $\frac{1}{\Delta T}$ 。因此，當我們需要較高的頻率解析度，將犧牲時間解析度，反之亦然，當需要較高的時間解析度，頻率解析度就會下降。然而，當我們需要較高之頻率解析度，導致時間解析度不足時，可以藉由調整重疊率使各幀的時間域訊號重疊。

因此，計算 STFT 時有兩個參數可以做調整，其一為上述之窗函數長度(以 ΔT 表示)，這項參數將決定時頻圖之時間解析度與頻率解析度，其二為重疊率(以 p 表示)，當較高之頻率解析度導致時間解析度不足時，可以此參數進一步調整時間解析度，如式 3-12。

$$\Delta t = \Delta T(1 - p) \quad (3-12)$$

式中， Δt 為 STFT 之時間解析度(sec)， ΔT 為窗函數大小(sec)， p 為重疊率(%)。

計算 STFT 時，若採用 0% overlap， $\Delta t = \Delta T$ ，此時每一個 Frame 所擷取之訊號沒有重疊，如圖 3-15。若採用 50% overlap， $\Delta t = \frac{1}{2} \Delta T$ ，則相鄰之 Frame 擁有一半的重疊訊號，如圖 3-16。若採用 90% overlap， $\Delta t = \frac{1}{10} \Delta T$ ，則相鄰之 Frame 擁有 90% 的重疊訊號，如圖 3-17。因此，雖然我們得以透過調整 Overlap 進而提升時頻圖之時間解析度，然而 Overlap 卻使相鄰 frame 之訊號重複計算，造成冗餘。

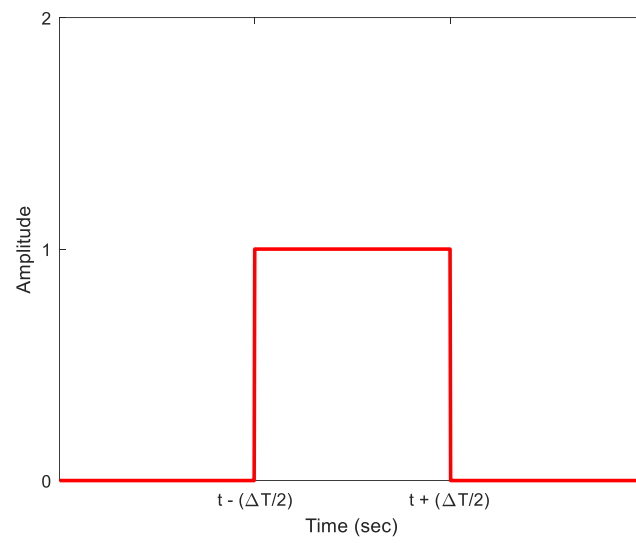


圖 3-14 矩形窗(Rectangular window)

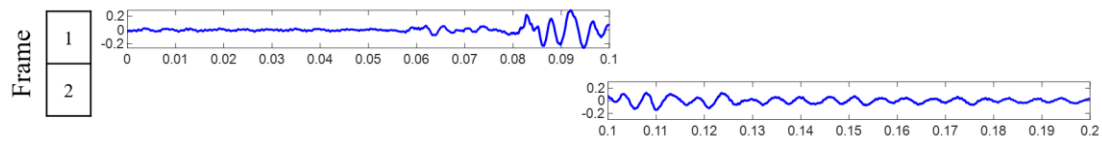
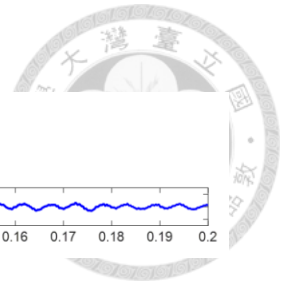


圖 3-15 0% overlap 之 Frame 取用訊號

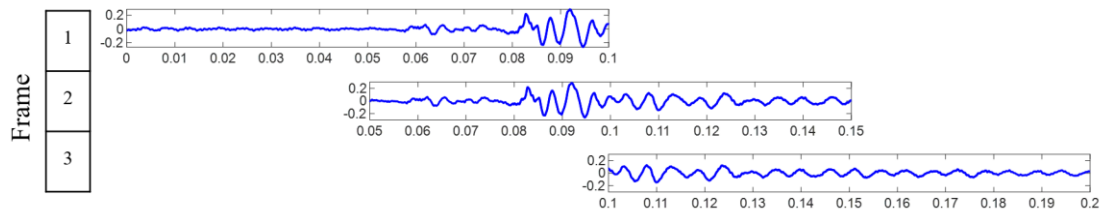


圖 3-16 50% overlap 之 Frame 取用訊號

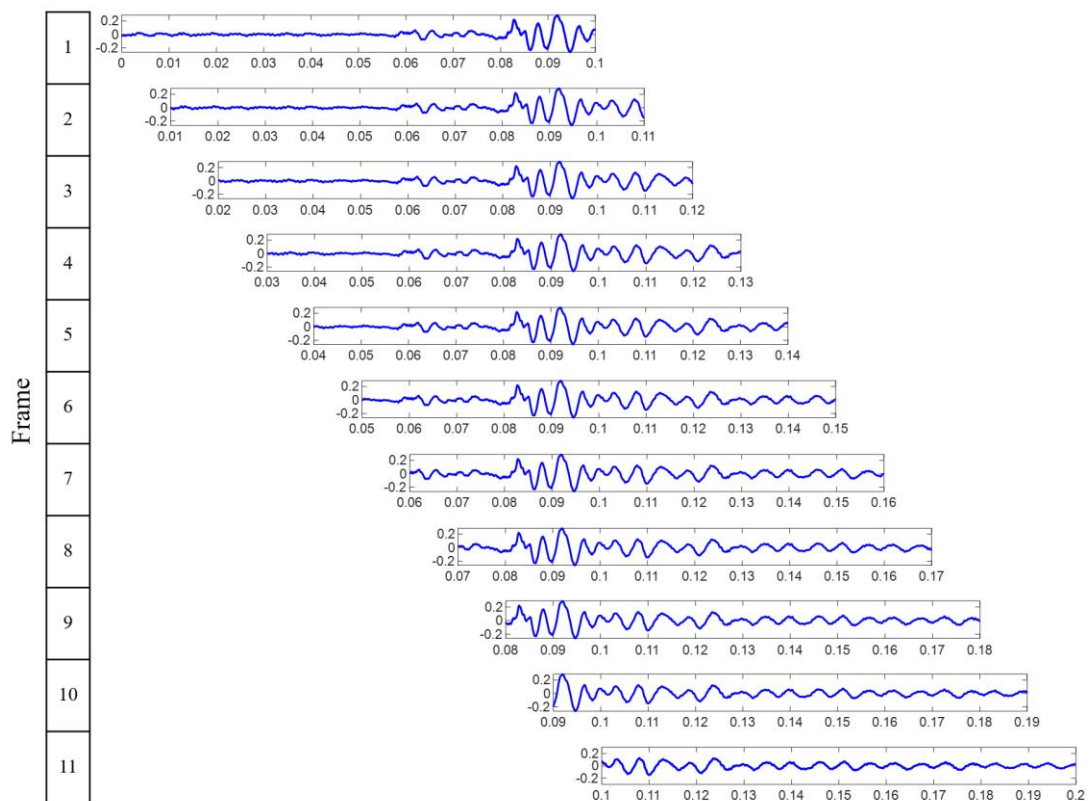


圖 3-17 90% overlap 之 Frame 取用訊號

3-5 梅爾頻譜

在人類語音與聲訊分析中，時間-頻率表示(Time-Frequency representation)扮演關鍵角色。短時傅立葉轉換(Short-Time Fourier Transform, STFT)雖然能夠揭示訊號在不同時間點的頻譜分佈，但其採用線性頻率尺度(linear frequency scale)並不契合人耳對頻率的非線性色覺。具體而言，人耳對低頻段(約 0-1 kHz)的頻率變化相當敏感，而對高頻段(約 2-8 kHz 以上)則較不敏感；若直接以等寬線性頻率處理語音或音樂，便無法強調人耳最易察覺的頻段資訊，亦會導致特徵維度偏高、運算冗餘。為了解決此一問題，研究者在 1930 年代至 1940 年代便開始探討「Mel 音高」概念，並陸續將其應用於語音分析；其中，最具代表性的做法便是「Mel-spectrogram」(梅爾頻譜)的引入。

Mel 尺度最早由 Stevens 等人於 1937 年提出，用以量化人類聽覺對音高(pitch)的非線性感知。當時 Stevens 等人透過實驗發現，人耳判斷「相同聽感間隔」所對應的頻率並非等差，而是近似指對數關係。基於此，他們定義了梅爾尺度(Mel scale)，可由式 3-13 將頻率轉換為梅爾尺度，反轉換如式 3-14，圖 3-18 即為梅爾刻度與頻率轉換之對應圖。

$$m = 2595 \times \log\left(1 + \frac{f}{700}\right) \quad (3-13)$$

$$f = 700 \times (10^{\frac{m}{2595}} - 1) \quad (3-14)$$

式中， m 為梅爾尺度。

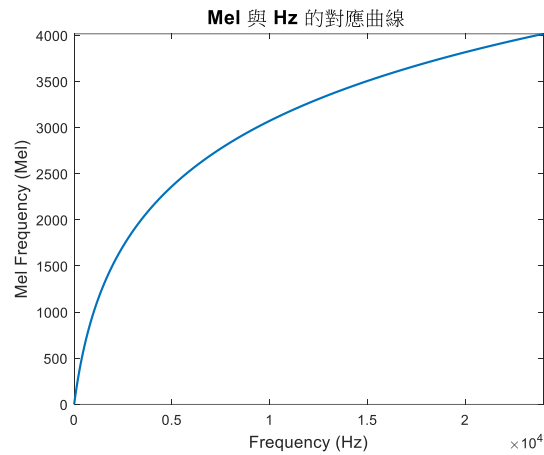


圖 3-18 Mel scale 與 Hertz 之對應曲線

在進行梅爾頻譜(Mel spectrogram)計算時，僅將線性頻率軸映射到 Mel 標度還不足以模擬人耳的頻譜響應，必須在 STFT 所計算的能量譜上，以梅爾三角濾波器組(Mel triangular filter bank)對做加權，如圖 3-19(以 80~1000Hz 之頻帶區間、32 條三角濾波器為例)，最終才會獲得梅爾頻譜。

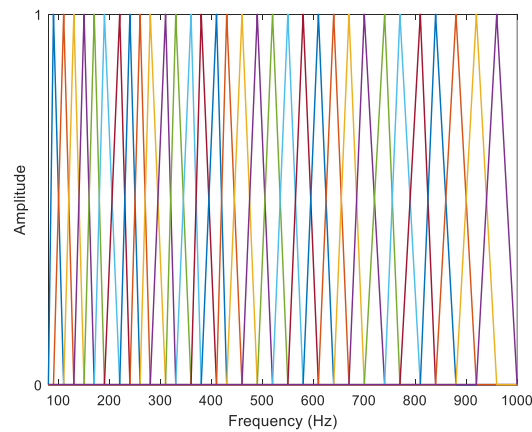


圖 3-19 梅爾三角濾波器組

3-6 梅爾倒頻譜係數

上述介紹了傅立葉轉換，並由短時傅立葉轉換進一步獲得訊號之時間-頻率特徵，並引進梅爾尺度並對功率譜乘上三角濾波器獲得梅爾頻譜以模擬人耳對頻率之非線性感受。而梅爾倒頻譜係數(Mel-Frequency Cepstral Coefficients, MFCC)，是將梅爾頻譜取對數，再計算離散餘弦轉換(Discrete Cosine Transform, DCT)，最終獲得倒頻譜(Cepstrum)，如式 3-15，通常取前 13 個係數。

算法如下：

$$MFCC_p(t) = \sum_{m=0}^{M-1} \log(Mel(t, m)) \cos\left[\frac{\pi p}{M} \left(m + \frac{1}{2}\right)\right] \quad (3-15)$$

式中， $MFCC_p$ 為倒頻譜係數($p=0, 1, \dots, 12$)， M 為梅爾濾波器個數， $Mel(t, m)$ 為梅爾頻譜在時間 t 上，第 m 個梅爾濾波器的能量。

3-7 比較 STFT、Mel spectrogram、MFCC

在候選的訊號處理技術中，雖各有優缺，然首先可予排除者為將訊號直接做傅立葉轉換(FT)，由於傅立葉轉換在數學上僅提供整段訊號的頻率分佈，缺乏時間解析度，故無法判定異音的確切發生時刻。其次，A 加權濾波雖能將原始錄音轉換為符合人耳聽覺靈敏度之加權頻譜，然其濾波曲線對全頻帶進行調整，對低頻與高頻成分尤為抑制。若異音之主要特徵落於低頻區域，經 A 加權後異音能量將大幅衰減，導致異音判讀失真。

因此，本節內容將聚焦於比較三種時頻方法，分別為短時傅立葉轉換(STFT)、梅爾頻譜(Mel spectrogram)、梅爾倒頻譜係數(MFCC)，如圖 3-20 至圖 3-24。

首先，梅爾倒頻譜係數經過 DCT 轉換後的係數不再對應於具體的頻率分量，因此不易觀察，通常僅取前 13 個係數作為特徵向量。經比較，MFCC 對於異音的判讀並不是很清楚，因此不採用。剩下的兩種時頻方法為 STFT、Mel spectrogram。同樣的我們分別觀察兩種時頻方法，繪製梅爾頻譜時則設定為常用的 32 個三角濾波器。

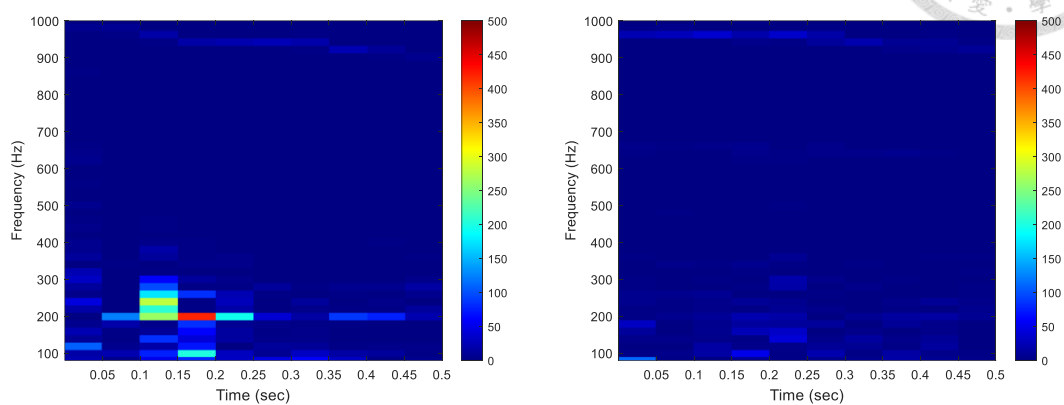
比較 STFT、Mel spectrogram 後，雖然各類異常在特定時間皆出現亮帶，但對於梅爾頻譜而言，經過三角濾波器的加權後，異音的特徵頻帶與原始頻率成分已不相同。對本研究而言，這可能不利於異音特徵探尋，因為若直接套用梅爾頻譜或其他濾波方法，將導致特徵頻帶分布失真。

1. NG1 發生時段(0~0.5 sec)

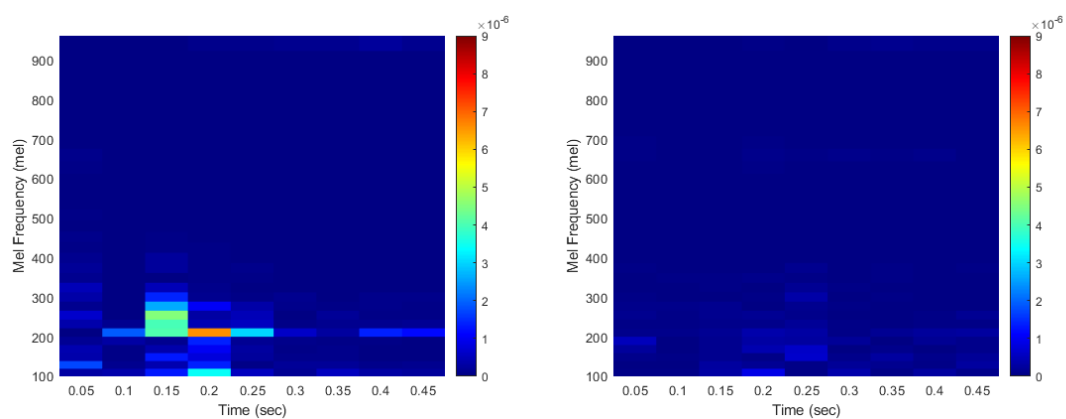
異常樣件

正常樣件

STFT



Mel spectrogram



MFCC

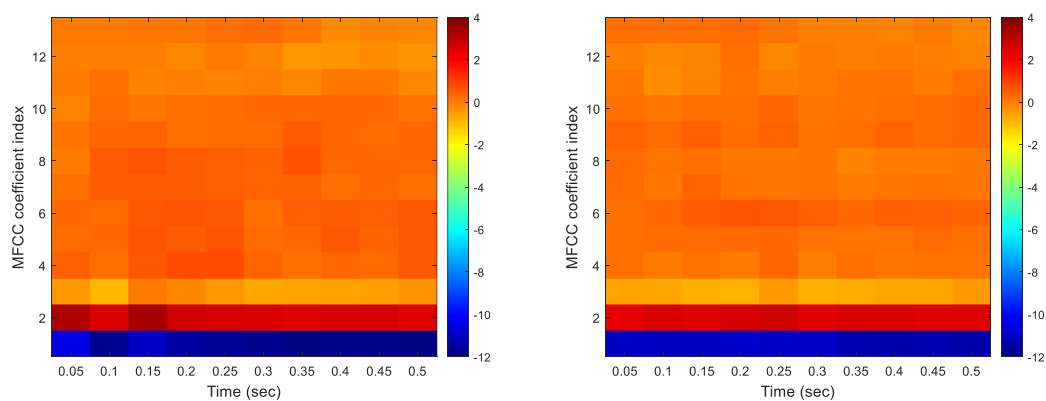


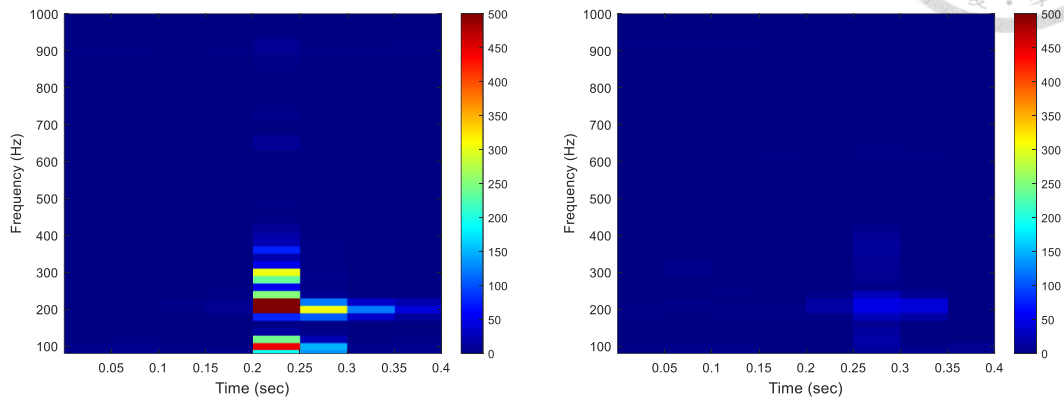
圖 3-20 NG1 之三種時頻分析比較

2. NG2 發生時段(0.5~0.9 sec)

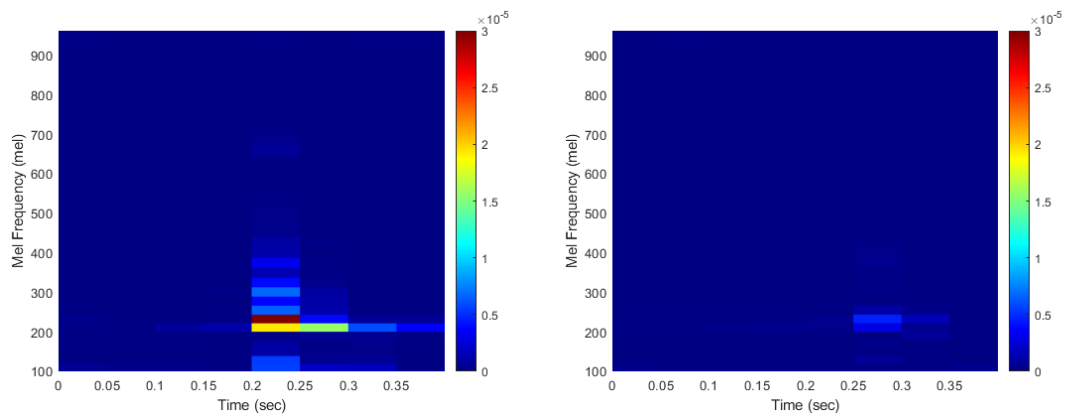
異常樣件

正常樣件

STFT



Mel spectrogram



MFCC

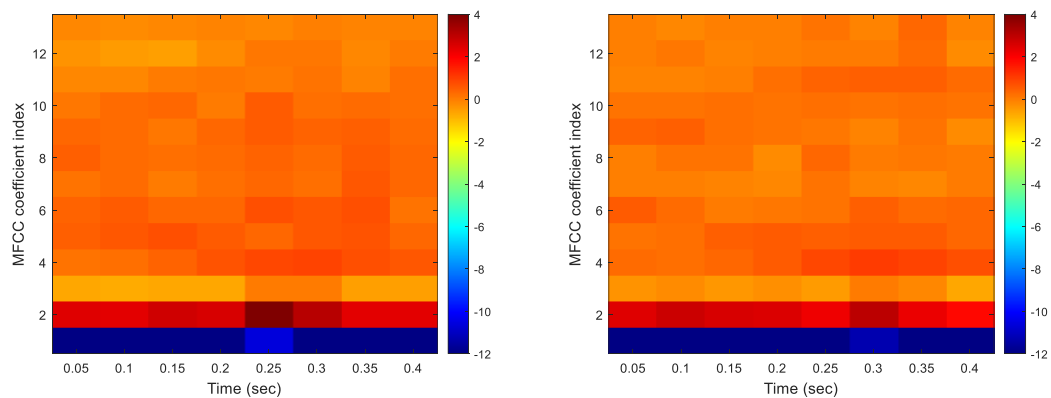


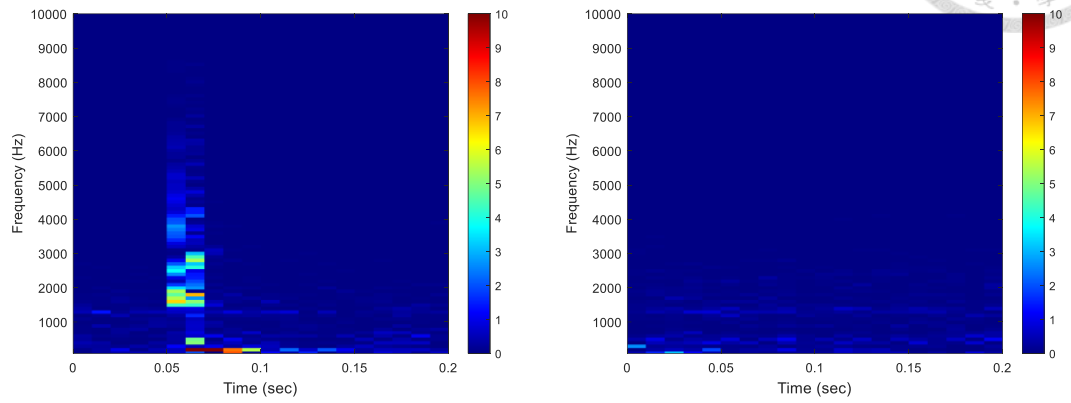
圖 3-21 NG2 之三種時頻分析比較

3. NG3 發生時段(0.9~1.1 sec)

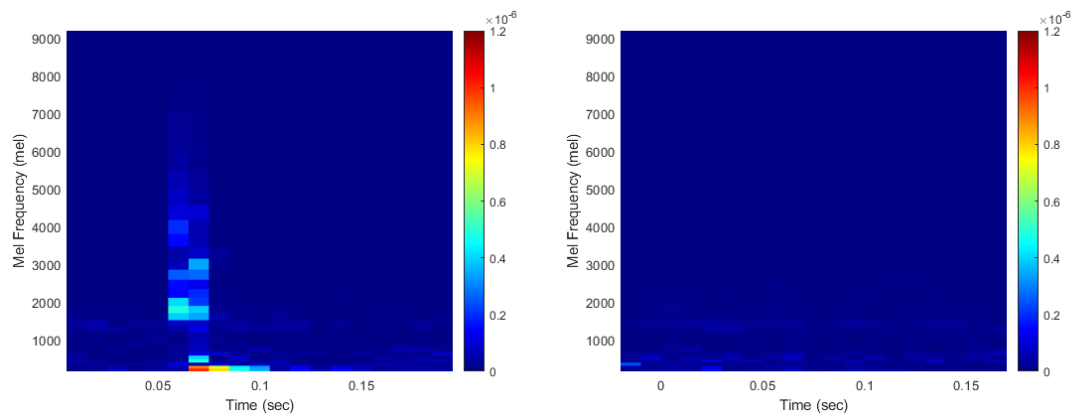
異常樣件

正常樣件

STFT



Mel spectrogram



MFCC

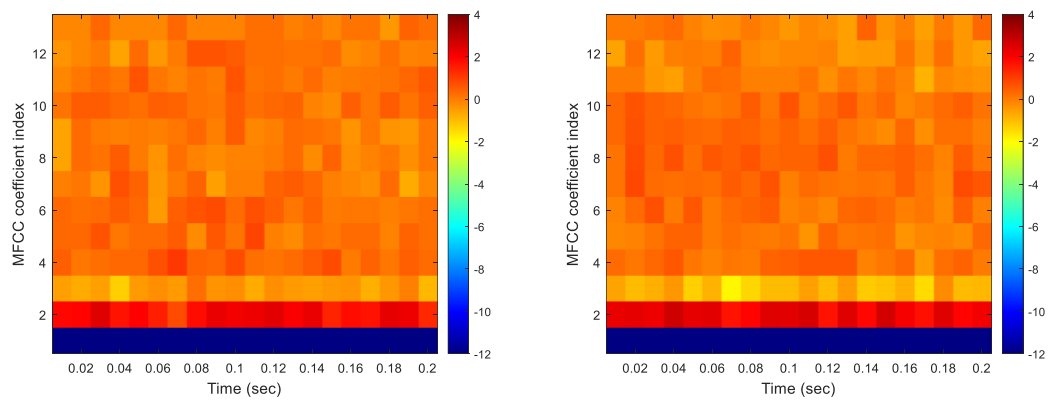


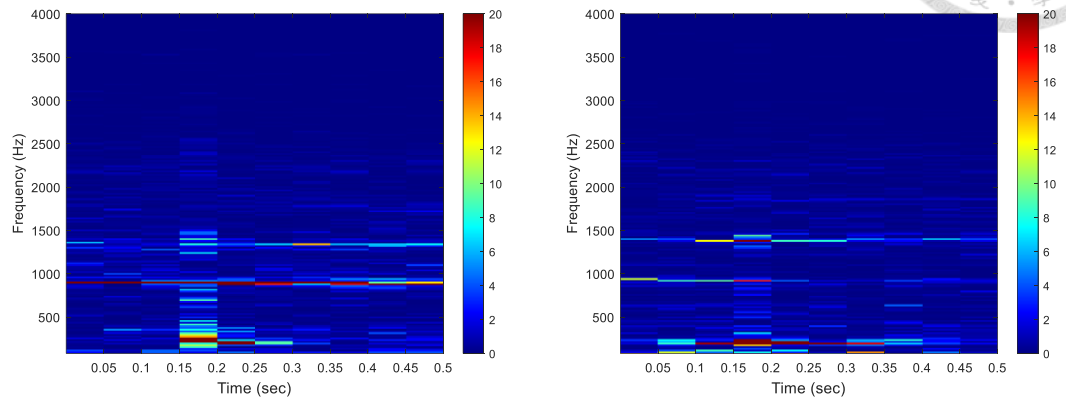
圖 3-22 NG3 之三種時頻分析比較

4. NG4 發生時段(1.5~2.0 sec)

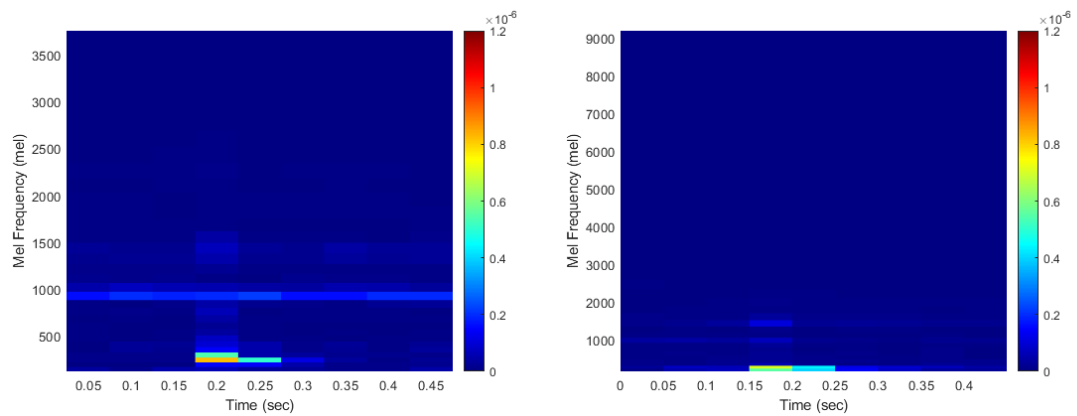
異常樣件

正常樣件

STFT



Mel spectrogram



MFCC

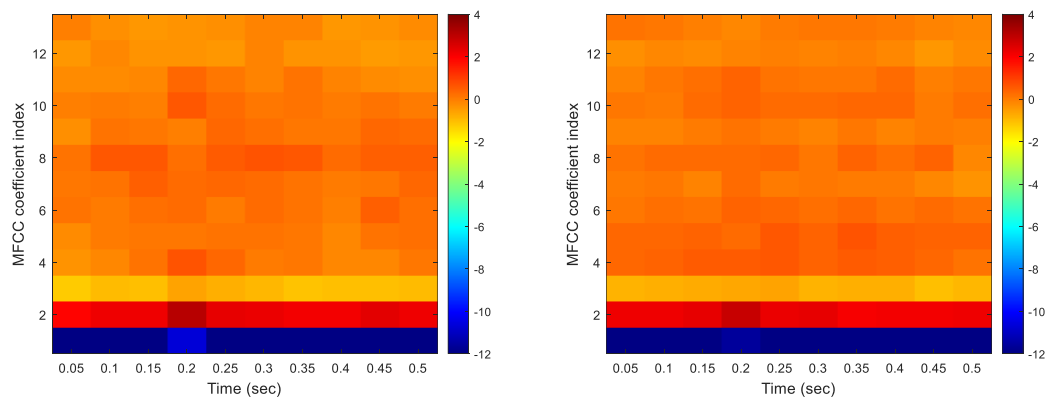


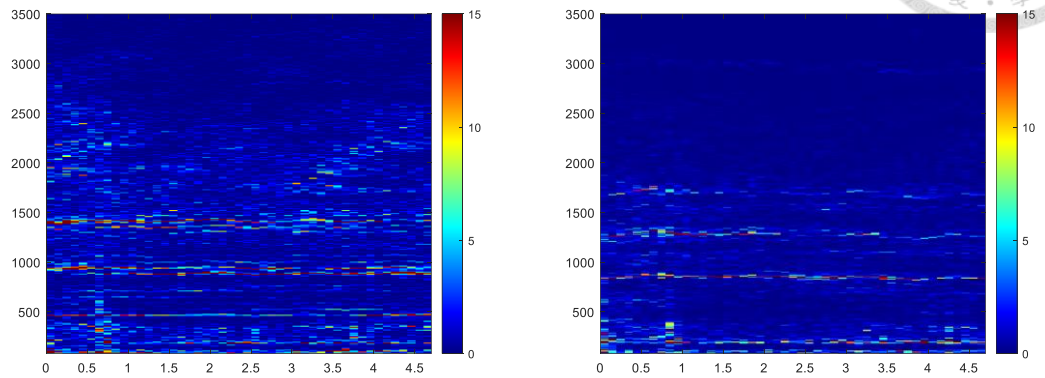
圖 3-23 NG4 之三種時頻分析比較

5. NG5 發生時段(0.9~5.6 sec)

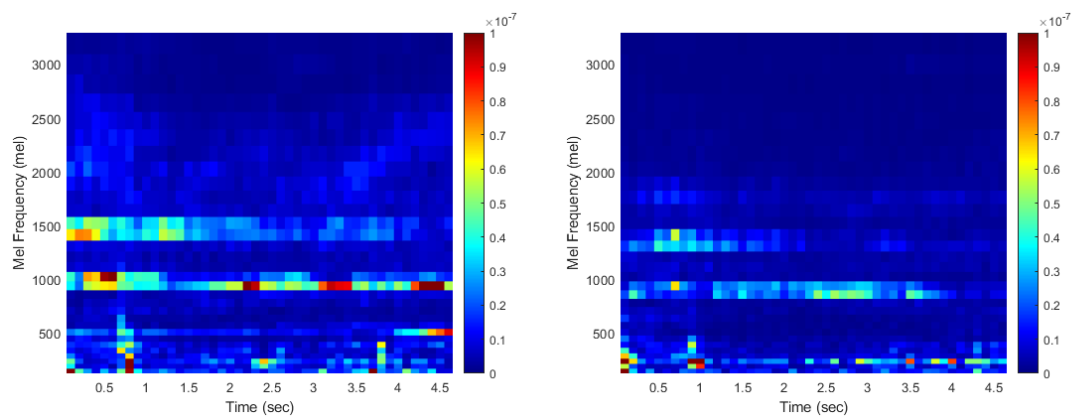
異常樣件

正常樣件

STFT



Mel spectrogram



MFCC

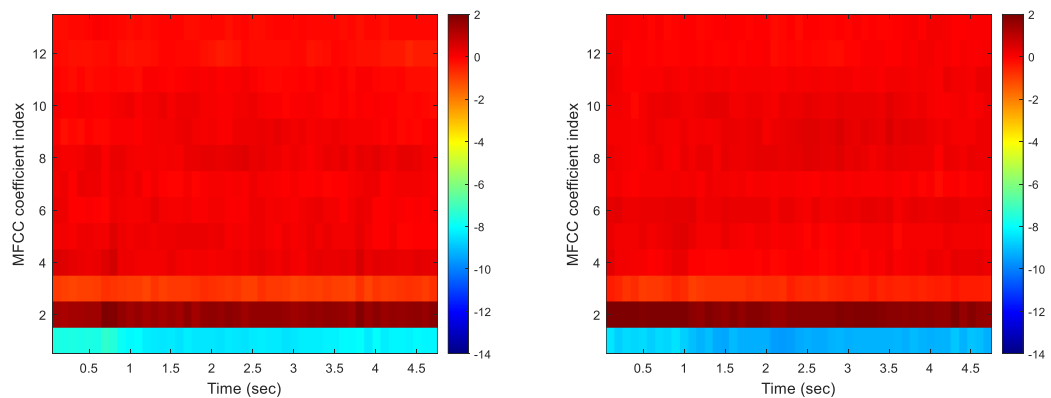


圖 3-24 NG5 之三種時頻分析比較

第四章 異音特徵擷取與判讀



本研究之目的在提出一個可以偵測異音的方法，因此必須先建立一客觀的異音指標，以作為判定天窗品質的依據，不必仰賴品管師的主觀經驗。

4-1 異音的特徵頻帶與時間範圍

4-1-1 異音之特徵頻帶， $\mathcal{F}_{anomaly}$

本研究以 STFT 所觀察到之亮帶來找出異音可能的特徵頻帶，如表 4-1。並以此對全頻帶之訊號做帶通濾波，並經品保師聆聽確認此頻帶之聲響即為我們所關注的異音。

為了獲取大略的異音特徵頻帶，我們觀察最大聲響異音樣件之 STFT，並以紅色框加以標示，如圖 4-1 至圖 4-5，異音亮帶如表 4-1。值得注意的是，各異常肇因之正常樣件 STFT 也有出現微弱的亮帶，只是能量值沒有異常樣件大。

表 4-1 由 STFT 觀察之異音亮帶(Hz)

$\mathcal{F}_{anomaly}$ 異常	$f_{low} \sim f_{high}$
NG1	80~500
NG2	80~500
NG3	100~10000
NG4	500~3300
NG5	1700~3500

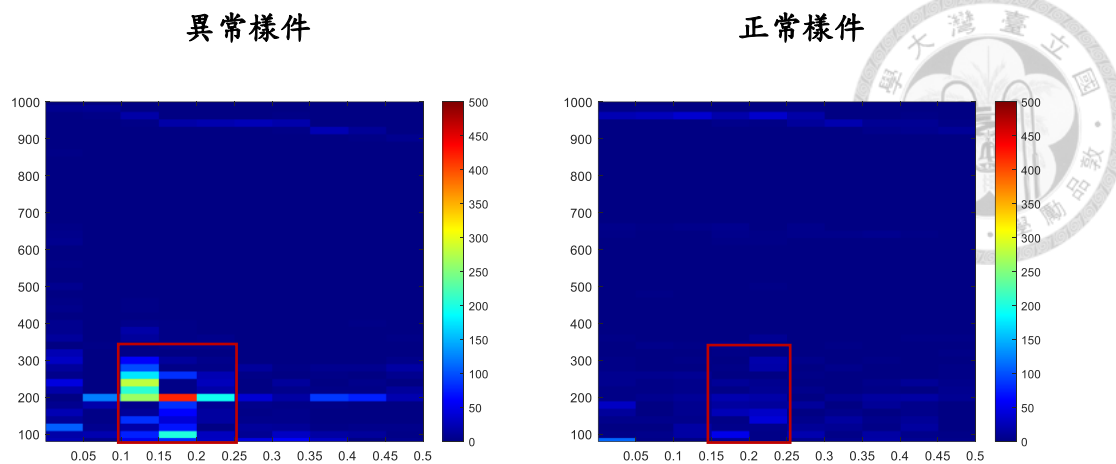


圖 4-1 NG1 發生時段(0~0.5 sec)之 STFT

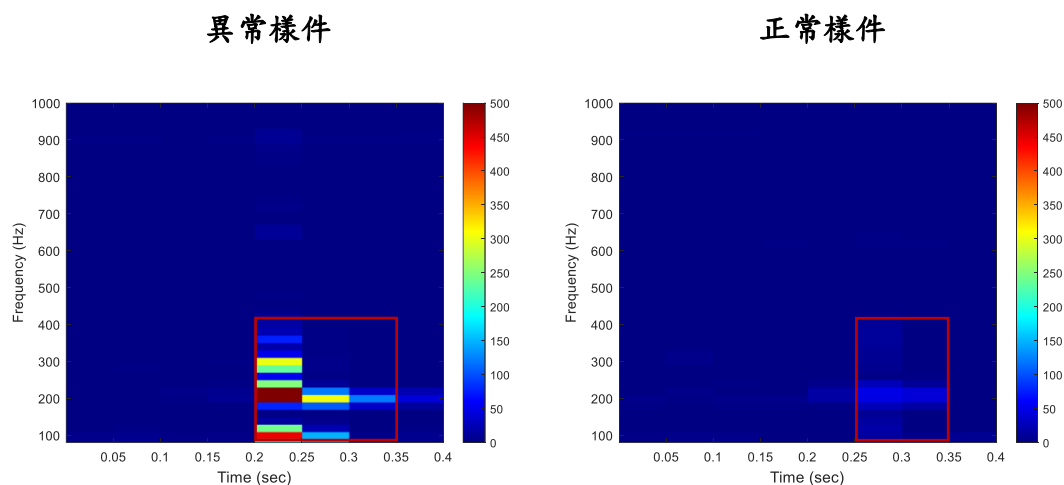


圖 4-2 NG2 發生時段(0.5~0.9 sec)之 STFT

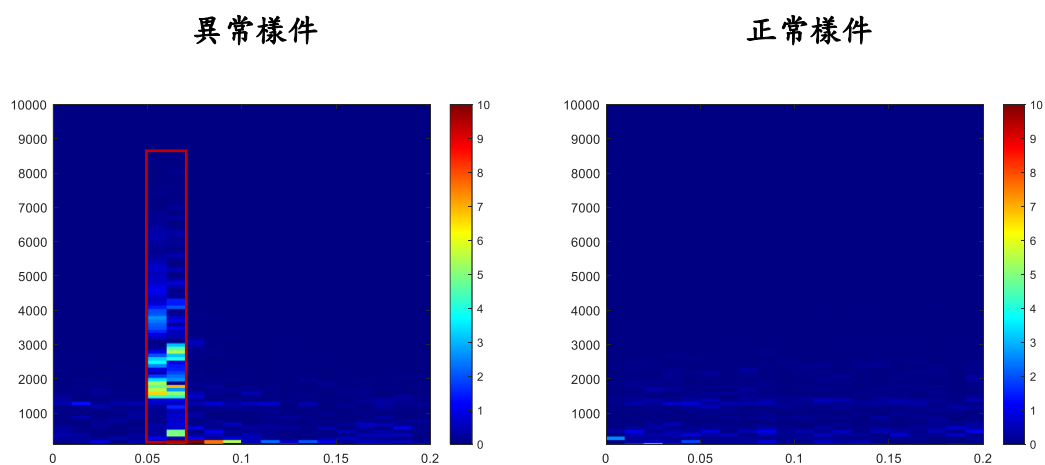


圖 4-3 NG3 發生時段(0.9~1.1 sec)之 STFT

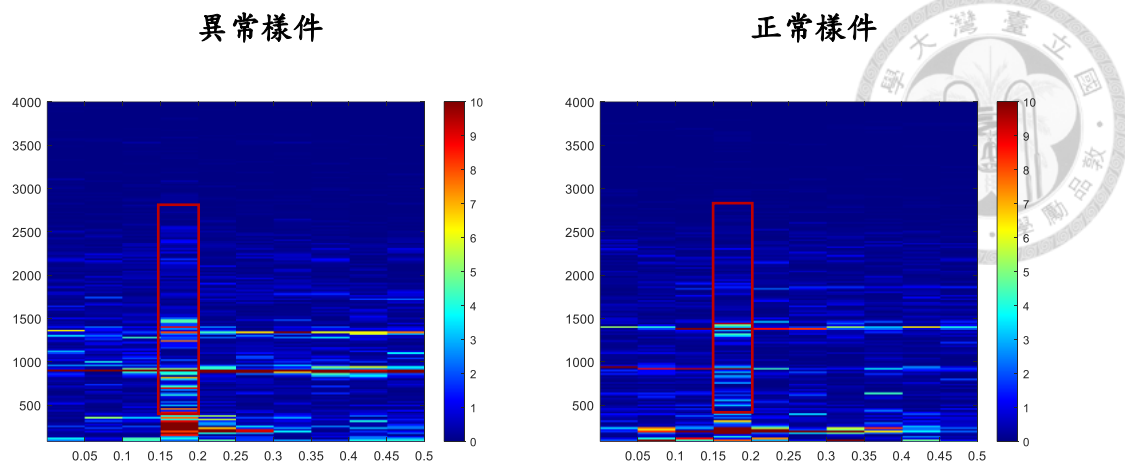


圖 4-4 NG4 發生時段(1.5~2 sec)之 STFT

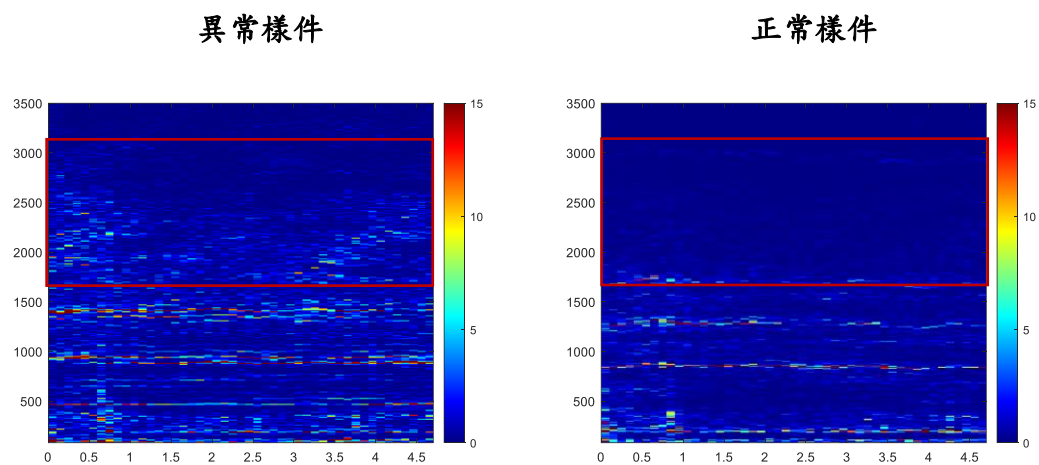


圖 4-5 NG5 發生時段(0.9~5.6 sec)之 STFT

4-1-2 異音之長度， $D_{anomaly}$

從 STFT 可知，當異音發生時，某一時間區間的能量會變大。為了確保能量增大確由異音所造成，我們先以 4-1-1 節之異音特徵頻帶對時間域訊號做濾波，並觀察濾波後之時間域訊號，如圖 4-6。然而，時間域曲線難以觀察因此我們繪製能量歷時曲線，以 0.01 秒計算一次平均值以平滑化，如圖 4-7。

經平滑化後的曲線就能夠清楚的觀察到在某一時段能量會突然變大，此即為異音發生的區間。為了標記異音發生的起始點，我們可以在 OK 樣件中取一段平坦無異音發生段之最大能量作為閾值，超出閾值即為異發生區間，如表 4-2。

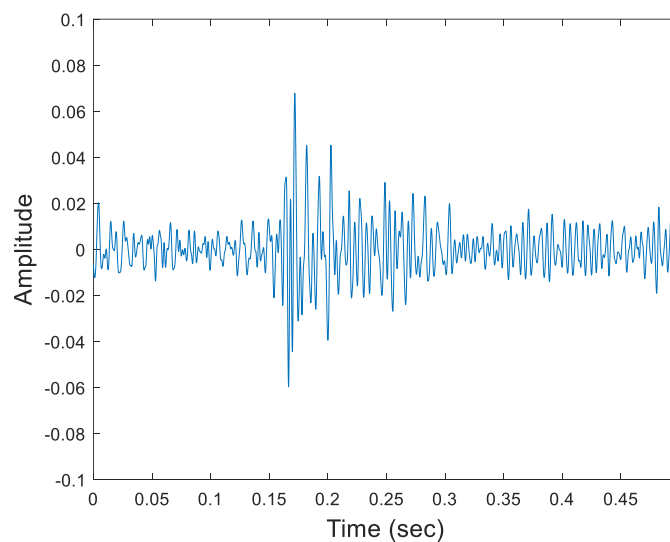


圖 4-6 時間域訊號難以觀察異音之起始~終止點

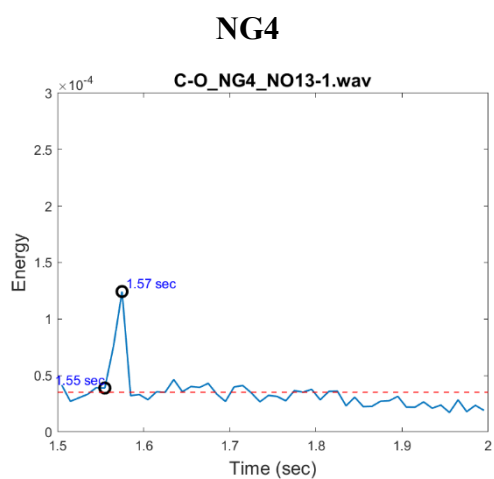
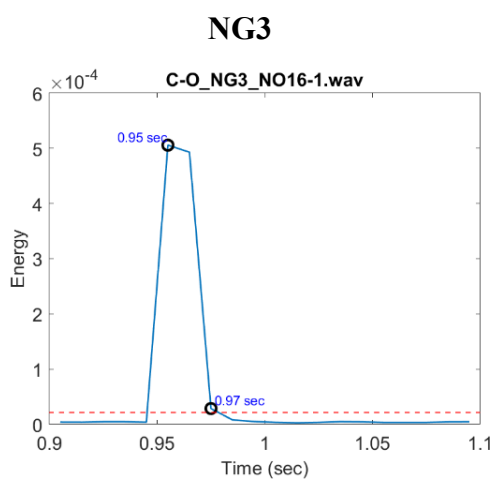
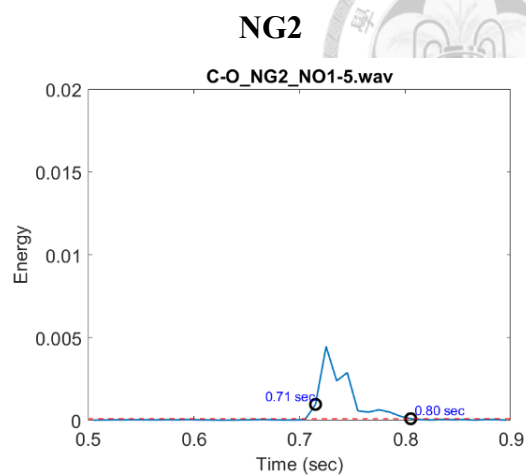
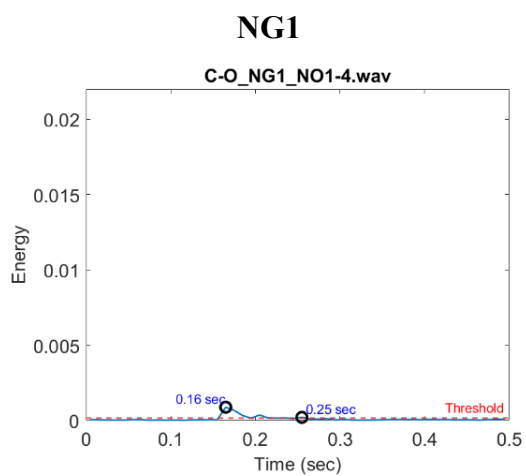


圖 4-7 各異常之能量歷時曲線

表 4-2 異音長度(sec)

$D_{anomaly}$ 異常	時間長度
NG1	0.1
NG2	0.1
NG3	0.03
NG4	0.03





4-2 異音判讀

4-1 節中，我們對異音之發生時間長度、頻帶特徵已經有初步的分析。考量本研究之目的為期望對音訊做出 OK/NG 之判斷，因此我們需要設立一指標用於分類目的。

4-2-1 異音指標

由 STFT 可觀察到當異音發生時，在時頻圖中特定時間之 $\mathcal{F}_{anomaly}$ 出現較大能量值。不過，由於 NG1~NG4 之異音發生區間不固定，因此我們提出對時間域訊號進行掃描，進而找出異音區間。具體作法如下，為了完整擷取異音區間，我們設定掃描窗之時間長度為 $D_{anomaly}$ ，掃描之移動步幅設定為 10% 的 $D_{anomaly}$ ，如圖 4-8，並對掃描窗訊號做傅立葉轉換後，計算特徵頻帶 $\mathcal{F}_{anomaly}$ 之能量積分值，因為異音的出現將使能量值變大，因此最大能量積分值即為真實異音之所在區間，我們則將該能量值定義為異音指標。由於 NG5 為持續性異音，因此不需要掃描，故以整段訊號之特徵頻帶能量積分值作為異音指標。

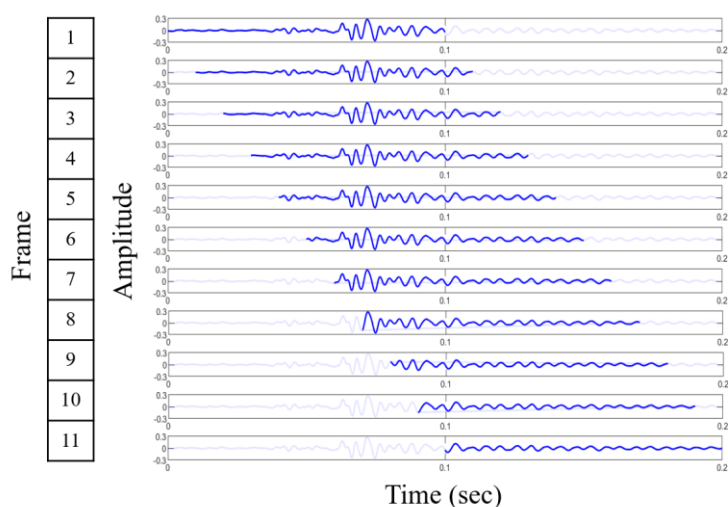


圖 4-8 掃描訊號示意圖

首先，對原始訊號 $x(t)$ 之第 n 個 Frame 表示為 $x_n(t)$ ，進行傅立葉轉換。

$$X_n(f) = \int_{t_{start}}^{t_{end}} x_n(t) e^{-2i\pi ft} dt \quad (4-1)$$

式中， $t_{start} = (n-1)h$, $n = 1, 2, 3, \dots$, $t_{end} = t_{start} + D_{anomaly}$, $h = D_{anomaly} \times 10\%$ 。

經傅立葉轉換後，我們將取其能量貢獻，並對 $\mathcal{F}_{anomaly}$ 之頻帶積分。

$$E_n = \int_{f_{low}}^{f_{high}} |X_n(f)|^2 df \quad (4-2)$$

式中， $|X_n(f)|^2$ 代表能量譜， f_{low} 代表 $\mathcal{F}_{anomaly}$ 之頻帶下限， f_{high} 代表 $\mathcal{F}_{anomaly}$ 之頻帶上限。

因此，訊號之異音特徵頻帶能量積分值向量如下。

$$\mathbf{E} = [E_1, E_2, E_3, \dots] \quad (4-3)$$

式中， $\mathbf{E}$ 代表所有掃描窗之特徵頻帶能量積分值向量。

最終，異音指標則為所有特徵頻帶能量積分之最大值。

$$E_c = \max(\mathbf{E}) \quad (4-4)$$

式中， E_c 代表異音特徵頻帶之能量，其中 c 為 *characteristic*。

計算所有訊號之異音指標後，當某一訊號之 E_c 超出設定閾值，即視為異音。

以上計算異音指標的步驟，若是令 $\Delta T = D_{anomaly}$ ， $\Delta t = h$ ，則 X_n 與 STFT 所計算之第 n 個 Frame 的頻譜是完全一致的。因此在操作時，僅需依照異音長度 $D_{anomaly}$ 與所需之掃描精度 h ，就可以算出異音指標。

4-2-2 分類閾值

欲決定一音訊是否存在異音，必須先選擇一個合理的閾值，但閾值不能隨意判定，最合理的方式是依據已標註的資料，選出一個可以分辨出 OK、NG 訊號的標準。



本研究之第一種分類閾值如下：

$$E_{Threshold} = \min_{NG} E_c \quad (4-5)$$

為了最嚴謹的對異音進行判讀，也就是較接近品保上之所需，寧可有較多的警告，也不能出現漏網之魚，因此採用最低之 NG 積分值作為區分 OK/NG 之標準，即不會有 NG 樣件誤判為 OK 樣件，如式 4-5。

本研究之第二種分類閾值如下：

$$E_{Threshold} = \max_{OK} E_c \quad (4-6)$$

同時，我們也針對 $\max_{OK} E_c$ 做討論，即不會有 OK 樣件誤判為 NG 樣件，用於希望降低 False Positive 時的另一種分類閾值，如式 4-6。

本研究設定之兩種閾值，是依據現有訊號所訂定之值。不過在實際品保應用上，是能依照需求在這兩種閾值之間做調整的。

4-3 異音頻帶優化

異音紀錄表中，品保師對天窗之 OK、NG 訊號的判斷經驗，僅停留在個人記憶中。儘管品保師能全頻帶之音訊進行判讀，但究竟哪些子頻帶最具異音特徵，尚未明確界定。儘管我們已透過 STFT 時頻圖初步描繪出異音所屬之大致頻段，仍需深入分析該頻帶內最能區分 OK、NG 訊號之特徵頻帶。

目前異音的 $\mathcal{F}_{anomaly}$ 是由 STFT 所觀察亮帶之頻帶區間，為了評估所有 (f_{low}, f_{high}) 組合，我們構建了一個三維曲面，能夠自動調整 $\mathcal{F}_{anomaly}$ 之 f_{low}, f_{high} ，並以色階顯示對應的分類準確率。相較於採用人工逐頻段探詢，不僅耗時且容易陷入局部極值（類似梯度下降法收斂到局部而非全域最優解），如圖 4-9 之 Local Max。

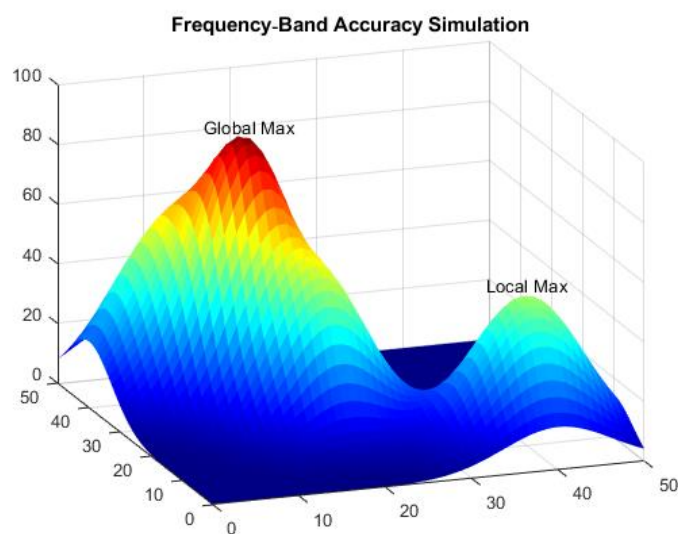


圖 4-9 局部極值示意圖

對應之 Δf 繪製，
呈現一倒三角形，

如圖 4-10。

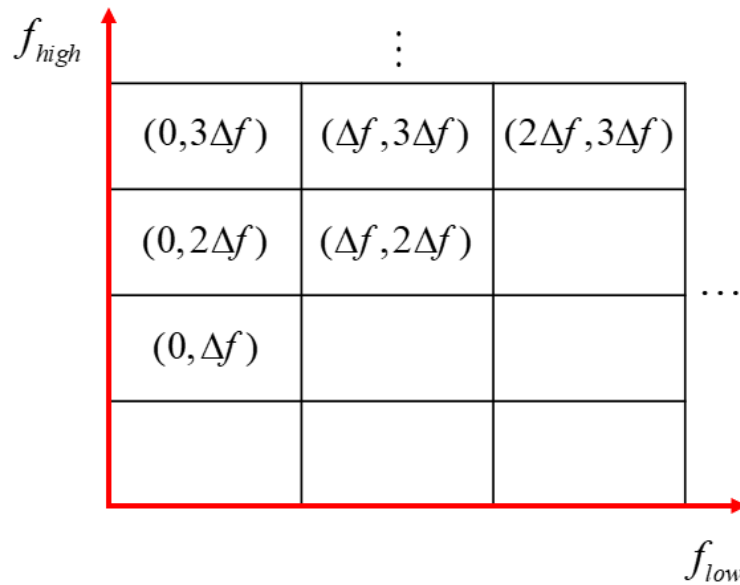


圖 4-10 自動調整 f_{low}, f_{high} 之二維矩陣

4-4 頻帶穩健性

經 $\text{Accuracy}(f_{low}, f_{high})$ 之計算後，二維矩陣中會有不只一種頻帶組合能夠獲得最好的準確率。為了有系統性的挑選較能穩定辨別 OK、NG 樣件之頻帶，本研究採用一項指標為該頻帶之 $\min_{NG} E_c - \max_{OK} E_c$ 差值為評估指標，如圖 4-11。當差值越大，則代表此頻帶組合最能有效區辨 OK、NG 之樣件。這邊以 $E_{Threshold} = \max_{OK} E_c$ 為閾值，並以 NG2 舉例，同時將最大差值之頻帶標示在 $\text{Accuracy}(f_{low}, f_{high})$ 上，觀察其所處的位置，以白色米字標示，如圖 4-12。

由圖 4-12 可發現所選頻帶周邊之準確率不佳，為了讓選擇頻帶之更為穩健，本研究採用 3*3 kernel summation 考量周邊頻帶之準確率，並繪製 $\text{Accuracy}_{3 \times 3}(f_{low}, f_{high})$ ，如圖 4-13 所示。經過穩健性考量後，我們便能在 $\text{Accuracy}_{3 \times 3}(f_{low}, f_{high})$ 中找到穩健頻帶區域，如圖 4-14 之實心白色標示所示。再由 $\text{Accuracy}_{3 \times 3}(f_{low}, f_{high})$ 穩健頻帶區域回到 $\text{Accuracy}(f_{low}, f_{high})$ 重新挑選最大 $\min_{NG} E_c - \max_{OK} E_c$ 頻帶，以藍色米字標示，如圖 4-15 所示。

我們可以對照圖 4-12 與 4-15，經過穩健性考量後，最佳頻帶將從原本的高峰，轉至高原處，也就是周邊頻帶較安全的區域。

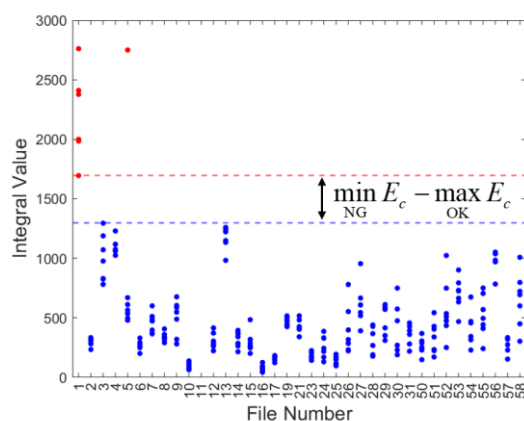


圖 4-11 頻帶之 $\min E_c - \max E_c$ 差值示意圖

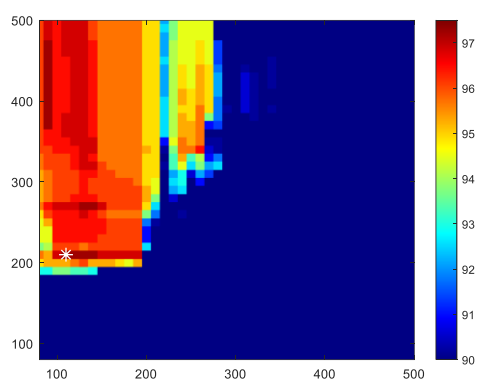


圖 4-12 $\text{Accuracy}(f_{low}, f_{high})$

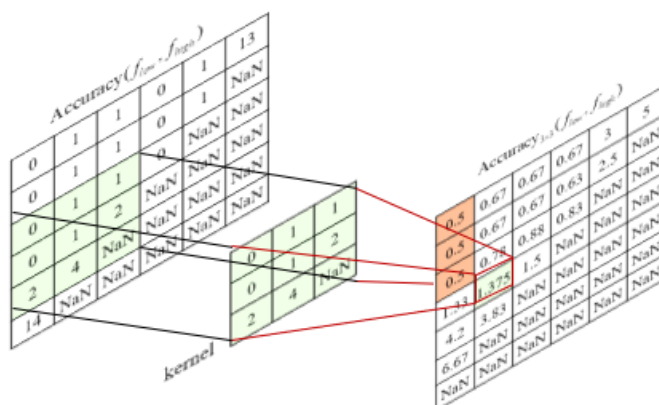


圖 4-13 3*3 kernel summation

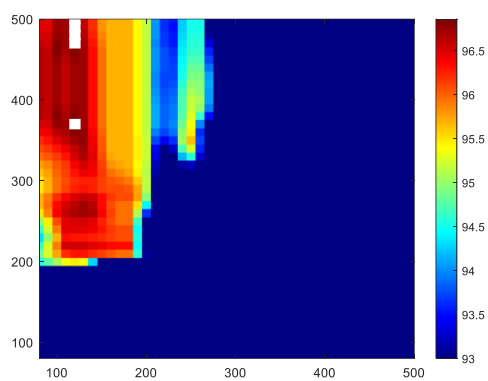


圖 4-14 NG4 之 $\text{Accuracy}_{3 \times 3}(f_{low}, f_{high})$

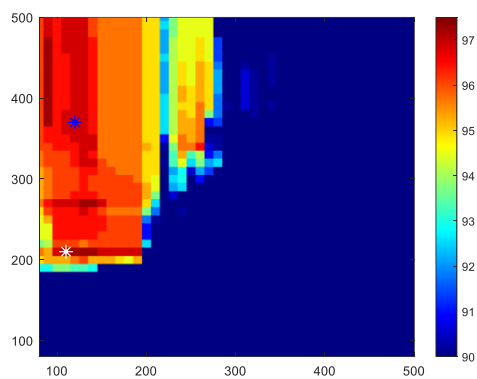


圖 4-15 重新挑選之最佳頻帶於 NG4 $\text{Accuracy}(f_{low}, f_{high})$ 之位置

第五章 結果與討論



5-1 觀察異音特徵頻帶之能量分布

第四章中，我們由 STFT 觀察異音的特徵頻帶，再由能量歷時曲線獲取異音長度，並以 $D_{anomaly}$ 對訊號做掃描，計算掃描窗之特徵頻帶能量積分值作為分類依據。經計算，各異常肇因之異音指標分布如圖 5-1，圖中紅點代表被標註 NG 訊號，藍點代表 OK 訊號。圖中能夠明顯觀察到被標註 NG 之訊號相較 OK 訊號之能量積分值來的大，換句話說，異音指標具備區分 OK、NG 訊號之能力。而為了分類 OK、NG 訊號，我們分別以兩種閾值計算準確率，如表 5-1。

綜上所述，本研究所提出之異音指標，能夠有效對 OK、NG 訊號進行分類，並成功將品保人員之經驗轉換為客觀的物理意義。

表 5-1 不同閾值下之最佳準確率(%)

閾值 異常	$\mathcal{F}_{anomaly}$ (Hz)	$E_{Threshold} = \min_{NG} E_c$	$E_{Threshold} = \max_{OK} E_c$
NG1	80~500	97.65	99.2
NG2	80~500	97.25	94.9
NG3	100~10000	98.04	98.4
NG4	500~3500	100	100
NG5	1700~3500	100	100

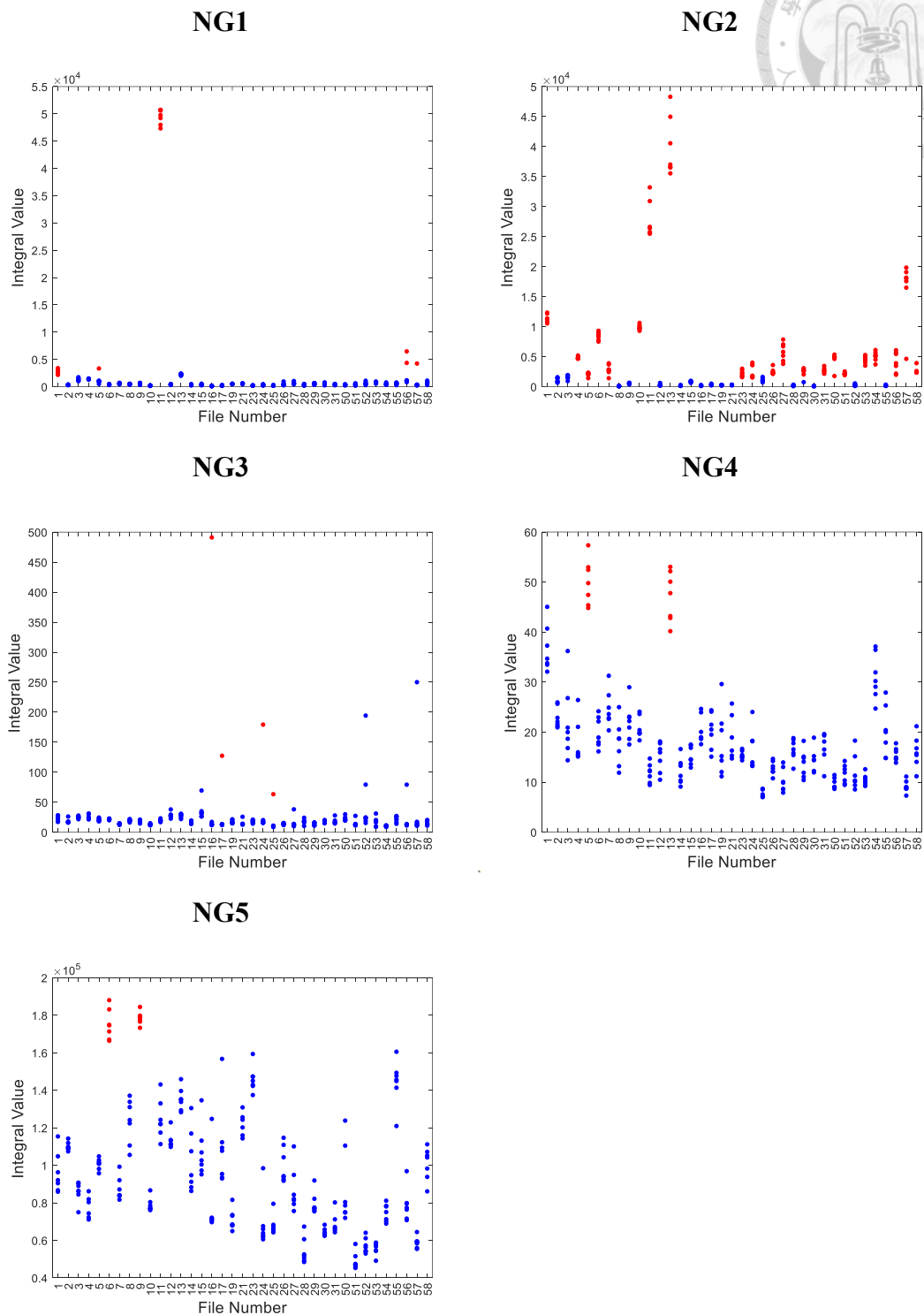


圖 5-1 各 NG 訊號之能量分布

5-2 自動調整(f_{low}, f_{high})，並計算各種組合之準確率

為避免失去異音之特徵頻帶，我們在 4-1-1 節是以較保守的方式描述特徵頻帶，然而，在此亮帶區間內是否有更精確的頻帶能夠更好區分 OK、NG 之訊號則是我們進一步要探尋的。

經計算，本研究提出之兩種閾值所對應之 $\text{Accuracy}(f_{low}, f_{high})$ 如圖 5-2，圖 5-3，而對應之最高準確率如表 5-2。 $\text{Accuracy}(f_{low}, f_{high})$ 中，當頻帶組合越接近異音之實際分佈時，分類之準確率將會越佳，反之當頻帶組合不為異音之所在，其準確率就會大幅下降，也就是說該頻帶下的能量值沒有分辨異音之能力。

經過對 STFT 亮帶區間之(f_{low}, f_{high}) 組合之準確率探討後，我們發現分類準確率仍有所提升，這也間接表示了人耳所能聆聽到之頻帶係具有一定程度的限制，且難以分辨非常細微的頻率變化。事實上，聲波的頻率和強度需達到一特定值範圍內，才能引起人耳的聽覺。人耳能感受到的振動頻率範圍約為 20-20000 Hz。隨著年齡的增長，聽覺上限會降低，這就是所謂的主觀因子，也就是當天窗產品出貨後有可能面臨的窘境，其有一定機率為出貨方認為此產品在可接受範圍內，但購買方並不認為該產品係 OK 樣件。

表 5-2 不同閾值下之最高準確率(%)

閾值 異常	$E_{Threshold} = \min_{NG} E_c$	$E_{Threshold} = \max_{OK} E_c$
NG1	100	100
NG2	97.3	99.6
NG3	100	100
NG4	100	100
NG5	100	100

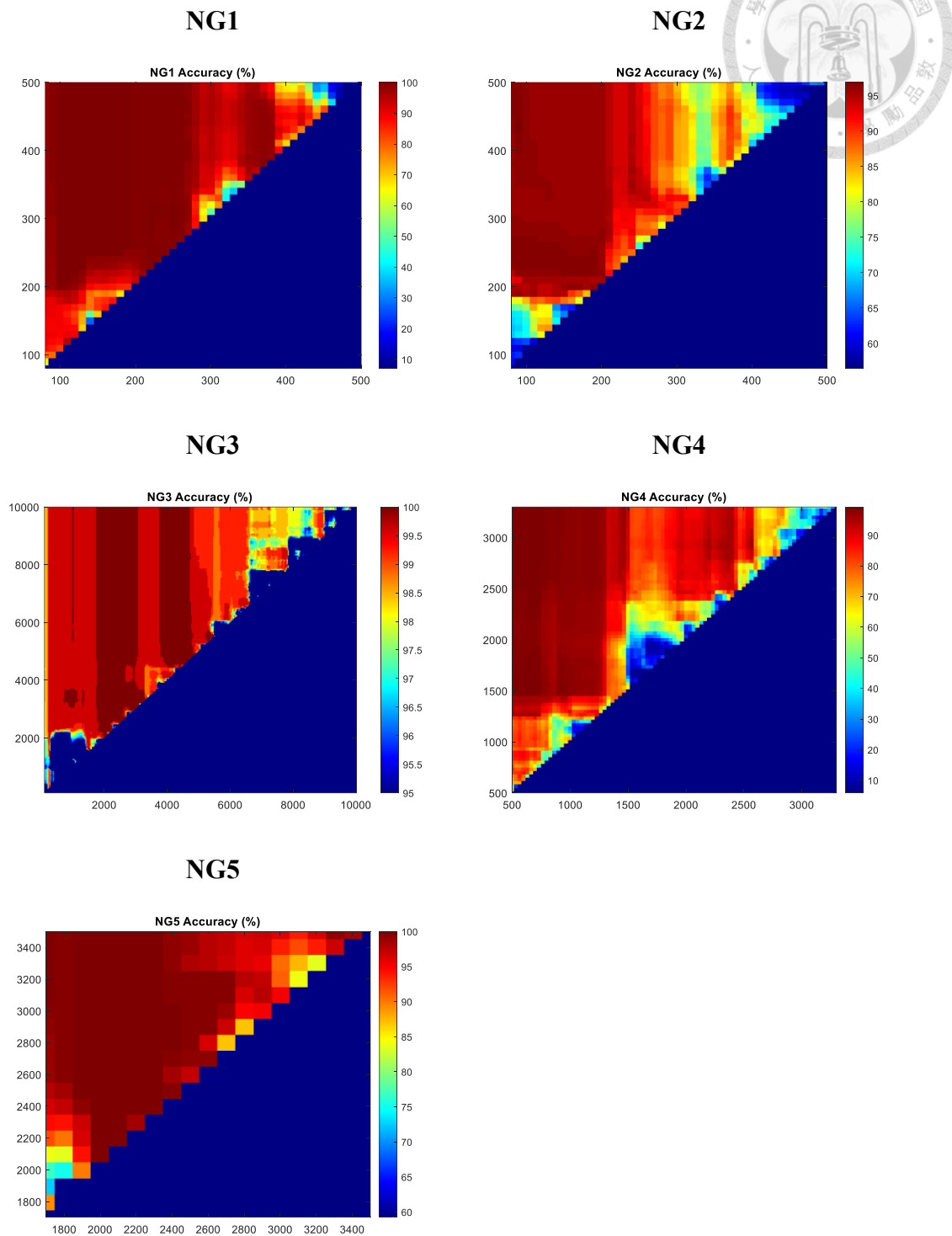


圖 5-2 以 $E_{Threshold} = \min_{NG} E_c$ 之 NG1~NG5 Accuracy(f_{low}, f_{high})

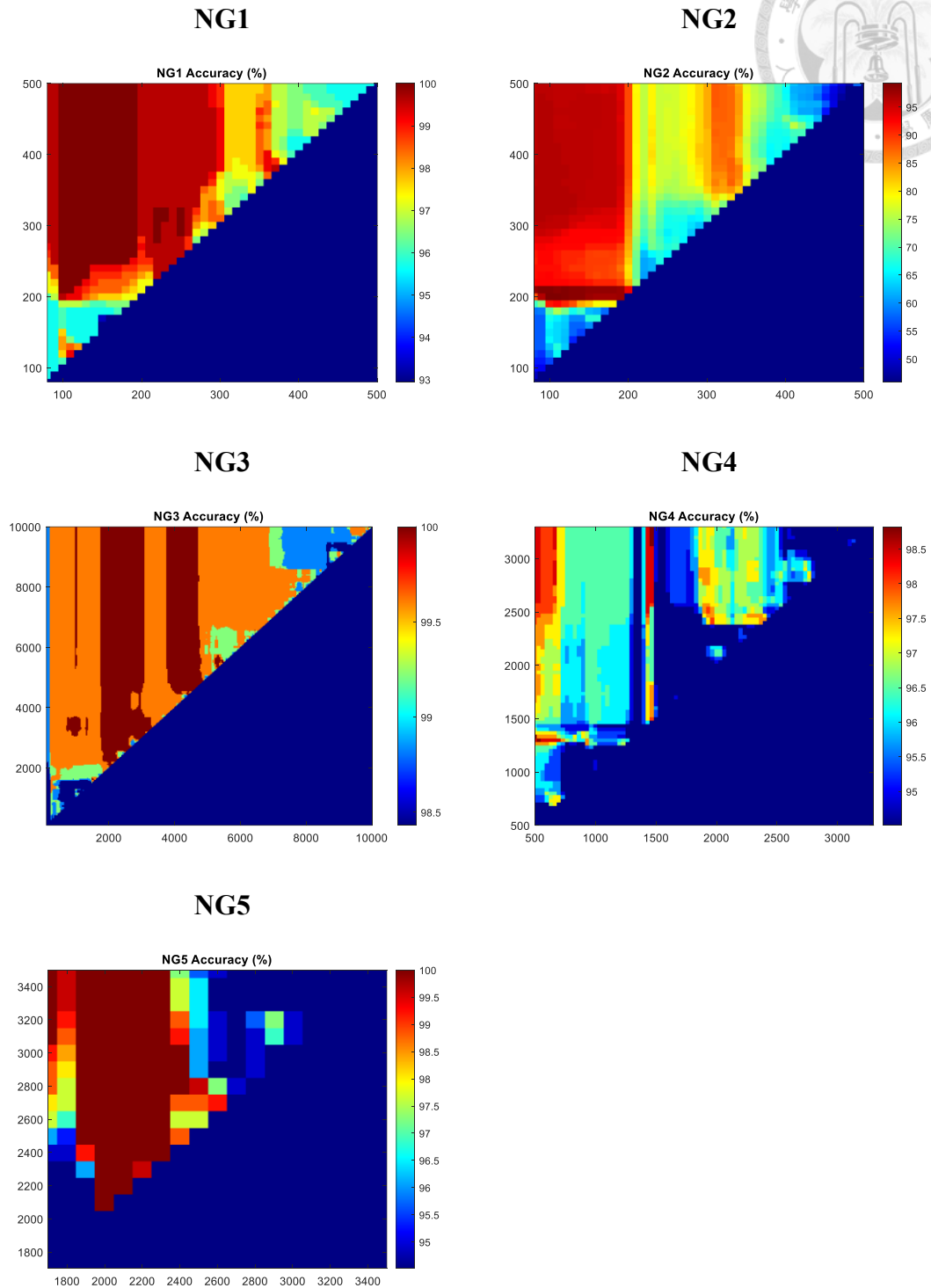


圖 5-3 以 $E_{Threshold} = \max_{OK} E_c$ 之 NG1~NG5 Accuracy(f_{low}, f_{high})

5-3 異音頻帶之穩健性

經 5-2 節所計算的 $\text{Accuracy}(f_{low}, f_{high})$ 矩陣中，將不只有單一 (f_{low}, f_{high}) 組合能夠獲得最高準確率。因此本研究以 $\min_{NG} E_c - \max_{OK} E_c$ 差值為指標，當差值越大，代表此頻帶組合最能有效區辨 OK、NG 之樣件，各異常肇因之最大差值頻帶如表 5-3。

相較於 NG1、NG3、NG4、NG5 之準確率已達 100%，如表 5-2，我們可以對 NG2 做一些探討，由表 5-3 中之 NG2 頻帶繪製能量分布圖，如圖 5-4。雖然最佳準確率已足夠高，但能量分布圖中仍有少數訊號可能受到人為因素的影響，導致 OK、NG 訊號之異音指標無法明確區隔。

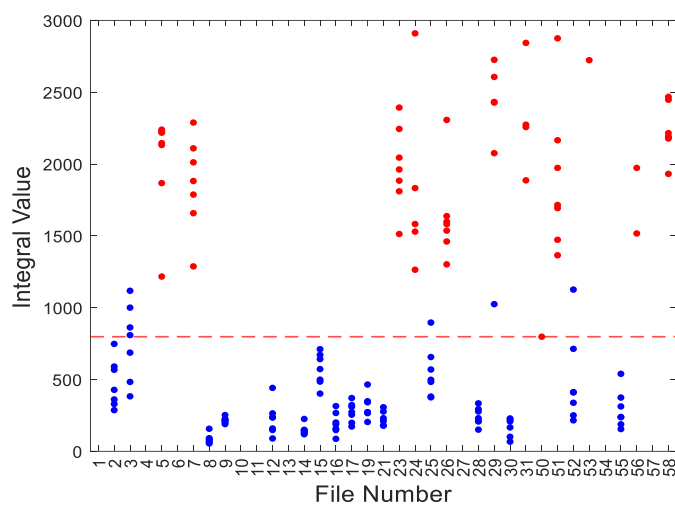
表 5-3 不同閾值下之最大 $\min_{NG} E_c - \max_{OK} E_c$ 差值頻帶(Hz)

異常 \ 閾值	$E_{Threshold} = \min_{NG} E_c$	$E_{Threshold} = \max_{OK} E_c$
NG1	220~310	220~310
NG2	110~210	100~210
NG3	1966~6800	1966~6800
NG4	533~3200	533~3200
NG5	2000~3500	2000~3500



以 $E_{Threshold} = \min_{NG} E_c$ 為分類閾值

NG2



以 $E_{Threshold} = \max_{OK} E_c$ 為分類閾值

NG2

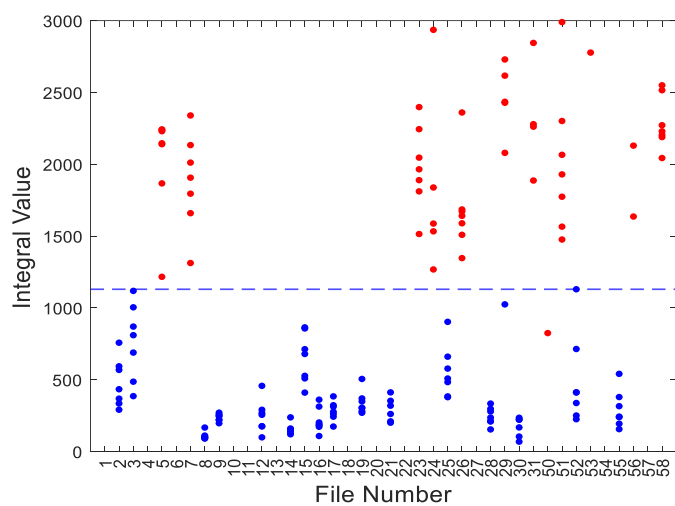


圖 5-4 不同閾值下之 NG2 能量分布圖

以下開始討論穩健性，我們先將表 5-3 所選頻帶標示在 $\text{Accuracy}(f_{low}, f_{high})$ 上，並觀察其所處的位置，並以白色米字加以標示，如圖 5-5，圖 5-6。經觀察，我們可以得知以 $E_{Threshold} = \min_{NG} E_c$ 為分類閾值下的 NG1、NG2 與 $E_{Threshold} = \max_{OK} E_c$ 為分類閾值下的 NG1、NG2 之最佳頻帶出現在不穩健的位置，因此我們以 4-4 節所介紹之 3*3 kernel summation 考量周邊頻帶準確率。

經過穩健性計算後，我們便能在 $\text{Accuracy}_{3 \times 3}(f_{low}, f_{high})$ 之穩健頻帶區域，如圖 5-7 中白色區域所示，回到 $\text{Accuracy}(f_{low}, f_{high})$ 重新挑選最大 $\min_{NG} E_c - \max_{OK} E_c$ 頻帶，如圖 5-8 中藍色米字所示。經穩健性討論後之最佳頻帶選擇與準確率、 $E_{Threshold}$ 如表 5-4、表 5-5。

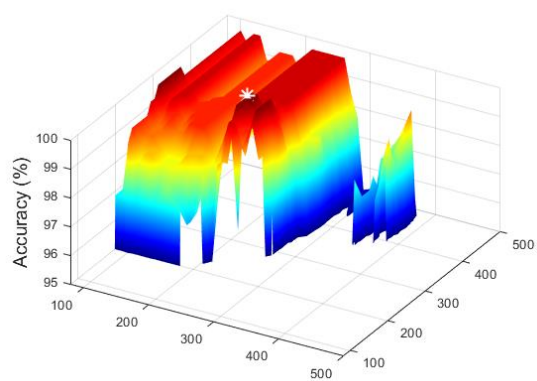
找到各異音之最佳頻帶後，我們對特徵頻帶做帶通濾波並匯出訊號，可以避免人耳對於背景噪音，馬達音等不屬於異音之頻率聲響，以此提高人耳對異音的敏感程度。最終，本研究匯出之音訊資料，一方面可供品保人員進行異音特徵辨識與聆聽訓練，提升異音熟悉度；另一方面亦可作為後續 AI 模型訓練之異音特徵來源。

本研究提出之兩種分類閾值下所分析的最佳頻帶，僅 NG2 不同，因此在音訊表 5-7 中，以 $E_{Threshold} = \max_{OK} E_c$ 為閾值僅嵌入 NG2 之最佳頻帶音訊，剩餘異音類別之最佳頻帶則與 $E_{Threshold} = \min_{NG} E_c$ 為閾值相同，音訊如表 5-6。並由此音訊可知，本研究所分析之最佳 $\mathcal{F}_{anomaly}$ 之聲音確實出現在原始訊號之異音當中。

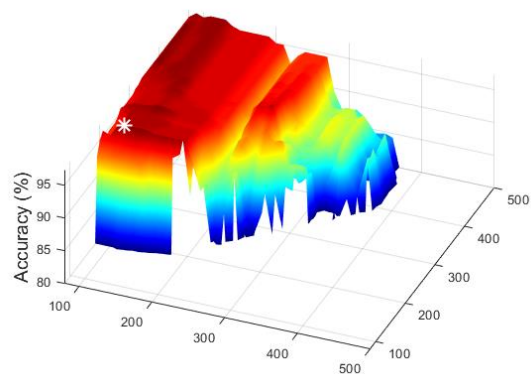
以 $E_{Threshold} = \min_{NG} E_c$ 為分類閾值



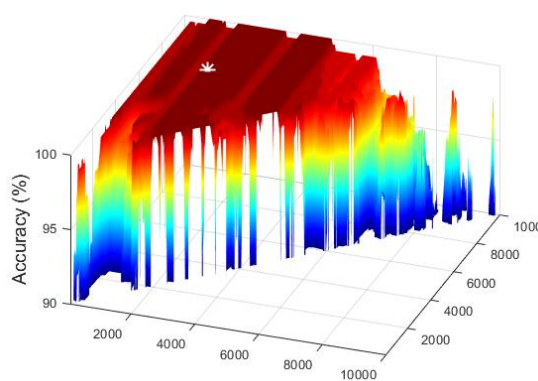
NG1



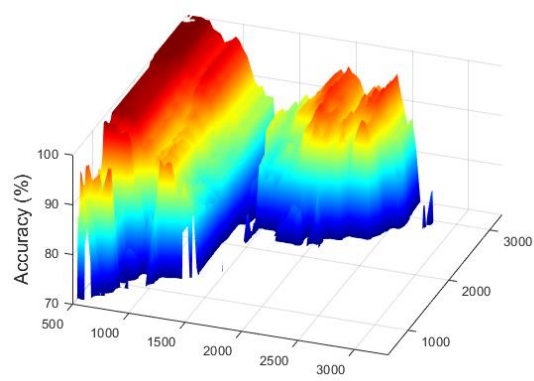
NG2



NG3



NG4



NG5

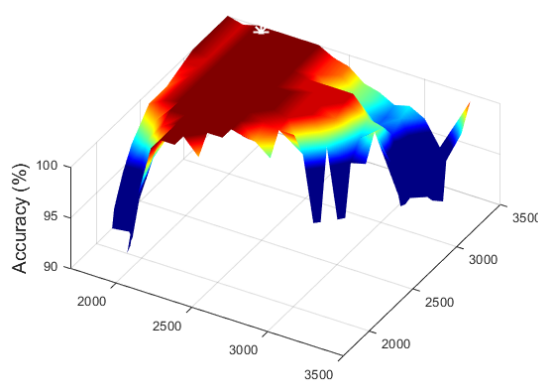
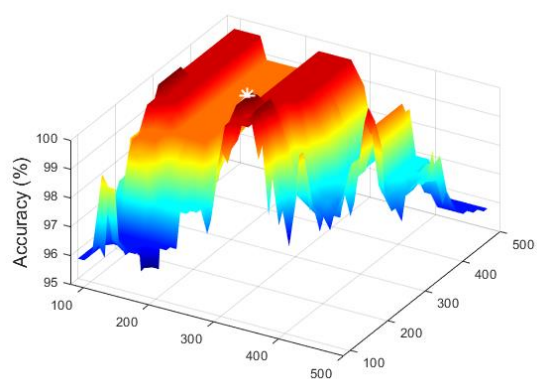


圖 5-5 NG1~NG5 最佳頻帶於 $Accuracy(f_{low}, f_{high})$ 位置

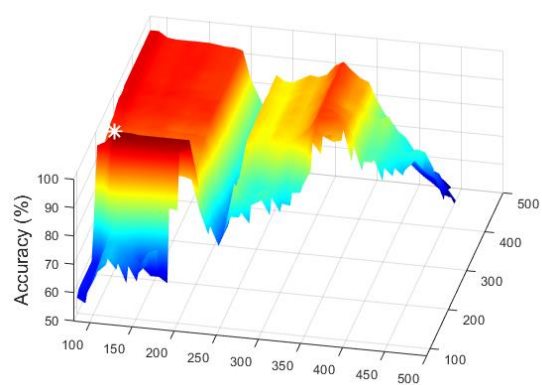
以 $E_{Threshold} = \max_{OK} E_c$ 為分類閾值



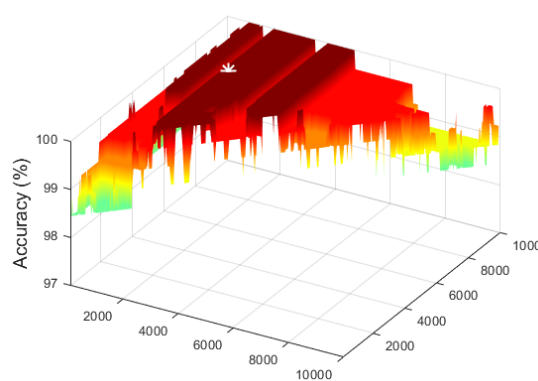
NG1



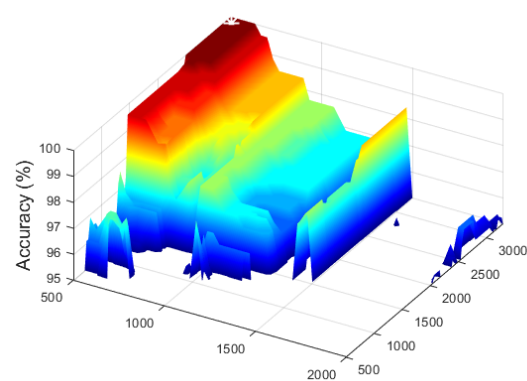
NG2



NG3



NG4



NG5

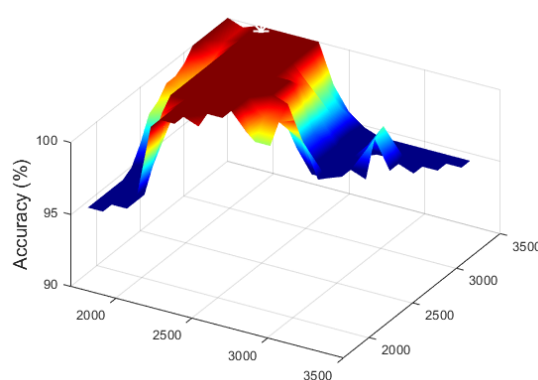
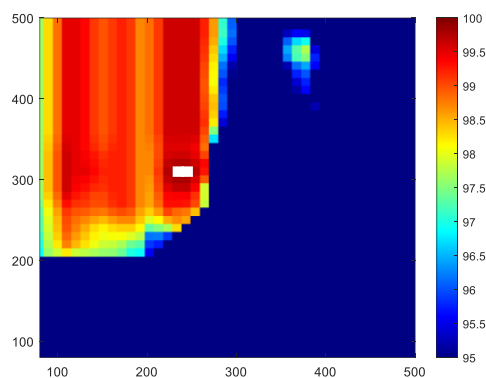


圖 5-6 NG1~NG5 最佳頻帶於 $Accuracy(f_{low}, f_{high})$ 位置

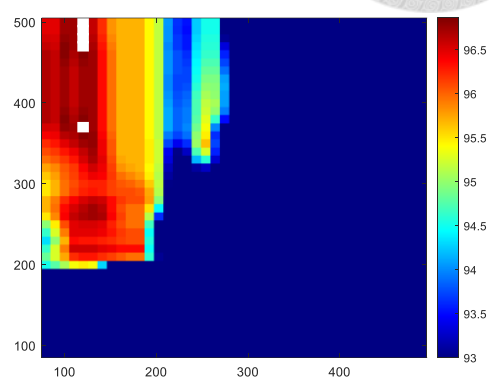


以 $E_{Threshold} = \min_{NG} E_c$ 為分類閾值

NG1

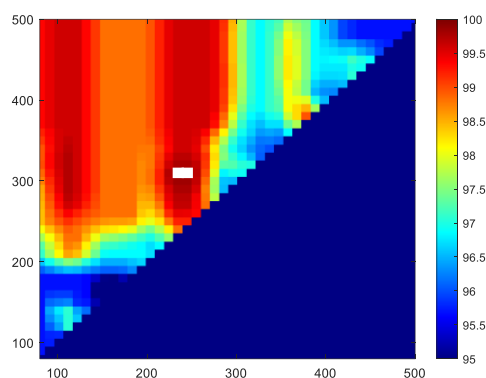


NG2



以 $E_{Threshold} = \max_{OK} E_c$ 為分類閾值

NG1



NG2

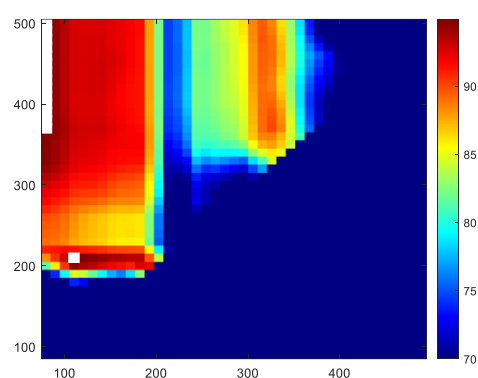
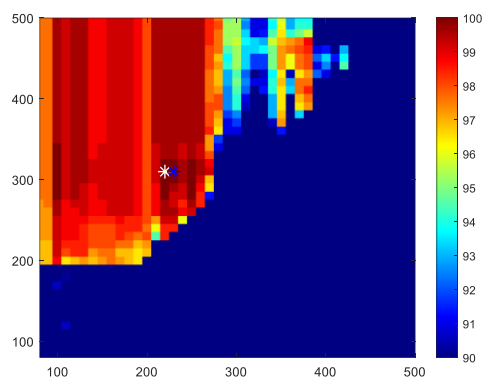


圖 5-7 $Accuracy_{3 \times 3}(f_{low}, f_{high})$

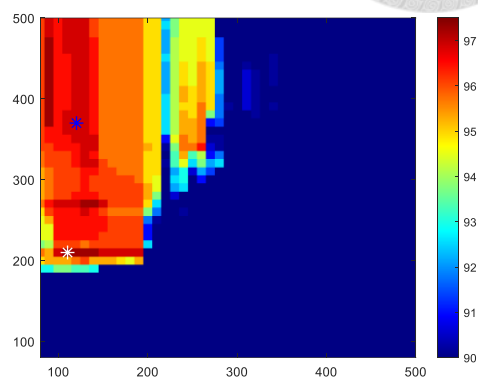


以 $E_{Threshold} = \min_{NG} E_c$ 為分類閾值

NG1

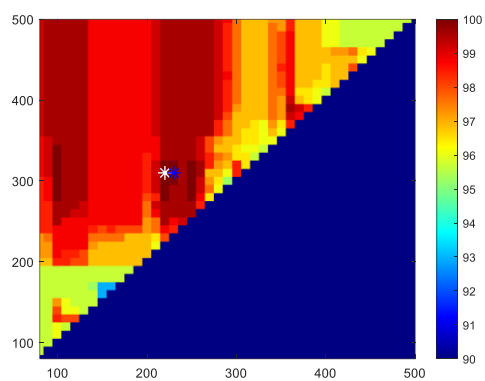


NG2



以 $E_{Threshold} = \max_{OK} E_c$ 為分類閾值

NG1



NG2

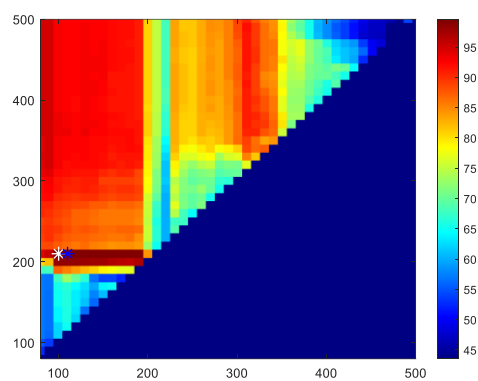


圖 5-8 重新挑選之最佳頻帶於 $\text{Accuracy}(f_{low}, f_{high})$ 之位置

表 5-4 最佳頻帶與對應之準確率、 $E_{Threshold}$



分類閥值	$E_{Threshold} = \min_{NG} E_c$		
變因 異常	最佳頻帶(Hz)	Accuracy (%)	$E_{Threshold}$
NG1	230~310	100	515.3
NG2	120~370	96.9	1821.9
NG3	1966~6800	100	36
NG4	533~3200	100	71.6
NG5	2000~3500	100	100338.6

表 5-5 最佳頻帶與對應之準確率、 $E_{Threshold}$

分類閥值	$E_{Threshold} = \max_{OK} E_c$		
變因 異常	最佳頻帶(Hz)	Accuracy (%)	$E_{Threshold}$
NG1	230~310	100	502.2
NG2	110~210	99.6	1125.9
NG3	1966~6800	100	21.6
NG4	533~3200	100	66.8
NG5	2000~3500	100	81969.3

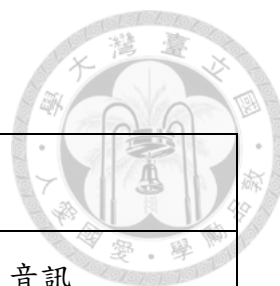


表 5-6 異音之音訊

$E_{Threshold} = \min_{NG} E_c$	
原始音訊	最佳 $\mathcal{F}_{anomaly}$ 音訊
NG1(滑塊與滑槽搭配不佳)	
	
NG2(滑塊與滑槽搭配不佳)	
	
NG3(彈簧未定位好)	
	
NG4(擋風板回彈音)	
	
NG5(拉繩摩擦音)	
	

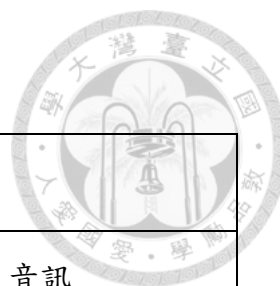


表 5-7 異音之音訊

$E_{Threshold} = \max_{OK} E_c$	
原始音訊	最佳 $\mathcal{F}_{anomaly}$ 音訊
NG2(滑塊與滑槽搭配不佳)	
	

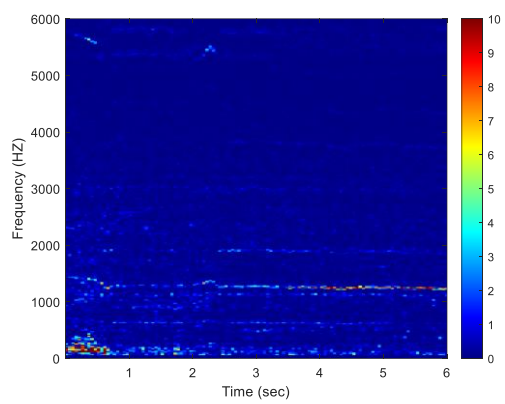
5-4 泛化性

為了測試本研究提出方法之泛化性，我們使用墨西哥廠區所生產的另一款車款的天窗，其機構配置不同，做為比較對象。首先，墨西哥天窗產品之錄製環境並不像台灣訊號在無響室內錄製，故其背景噪音沒有被抑制，所以聽起來較為混雜，又錄製時偶會摻雜產線上之噪音，或是錄製人員之走動聲等等外部噪音，為本研究之異音指標測試增加了難度與複雜性。

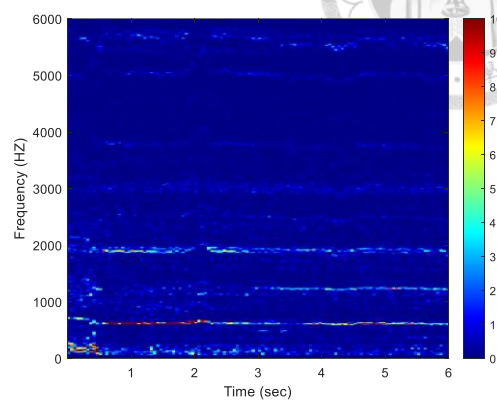
在前述介紹之訊號分析時，由於我們熟悉天窗之作動機制，因此我們能將各類異音之發生時段做區隔，並獲得極佳的檢測效果，但對於不熟悉其運作模式之天窗機構，卻難以做出時段的假設，因此我們直接對整段天窗訊號進行檢測，如圖 5-10 至圖 5-13。具體來說，我們可以計算 NG1~NG4 所有掃描窗之特徵頻帶能量積分值向量，如式 4-3，就能可獲得整段訊號之能量曲線，並於能量曲線上以紅線標示 $E_{Threshold}$ ，便可檢測整段訊號中是否出現異音。NG5 為長時間的特徵頻帶積分值為異音指標，因此僅有一個能量值，為了方便觀察，我們仍將 NG5 之能量值繪製成水平線，並以紅線標示 $E_{Threshold}$ 。

經聆聽，所有樣件(樣件一、二、三、四)之背景音都較大聲。另外，樣件三在 0~0.5 秒間出現中高頻之敲擊聲響，且在 1~2 秒之間出現低頻聲響，樣件四在 0~0.5 秒間出現中高頻之敲擊聲響。各樣件之 STFT，如圖 5-9。不過，由於本款天窗樣件未進行異音分析，故異常肇因有待分析。

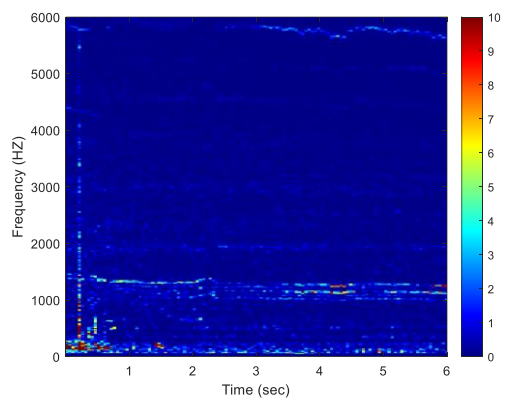
樣件一



樣件二



樣件三



樣件四

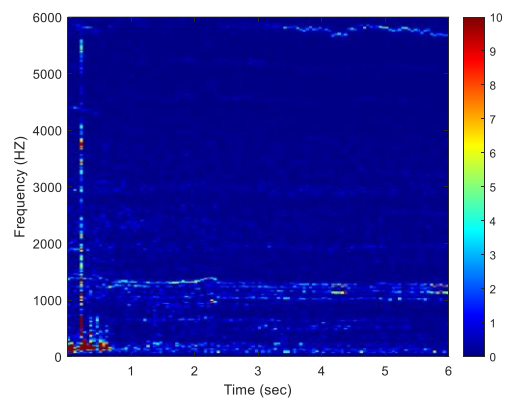


圖 5-9 墨西哥天窗樣件之 STFT

檢測結果如下：

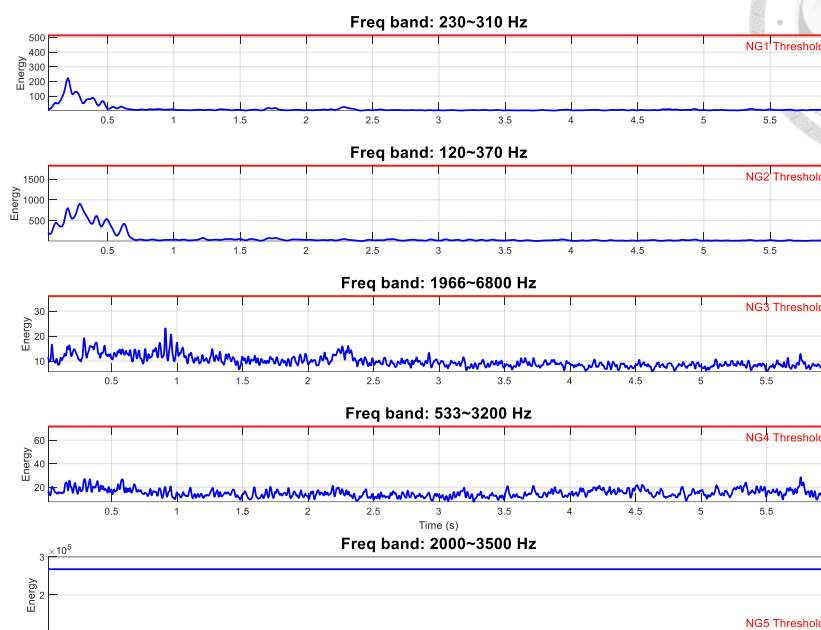
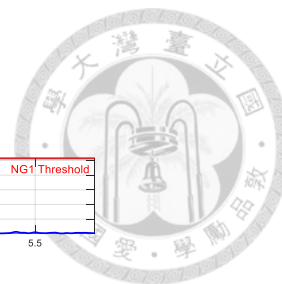


圖 5-10 墨西哥樣件一之泛用性測試

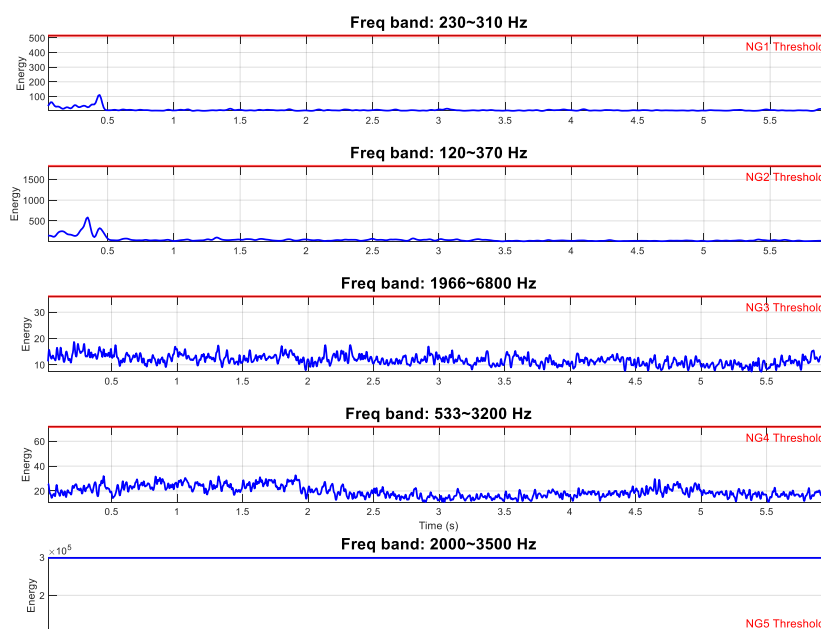


圖 5-11 墨西哥樣件二之泛用性測試

檢測結果如下：

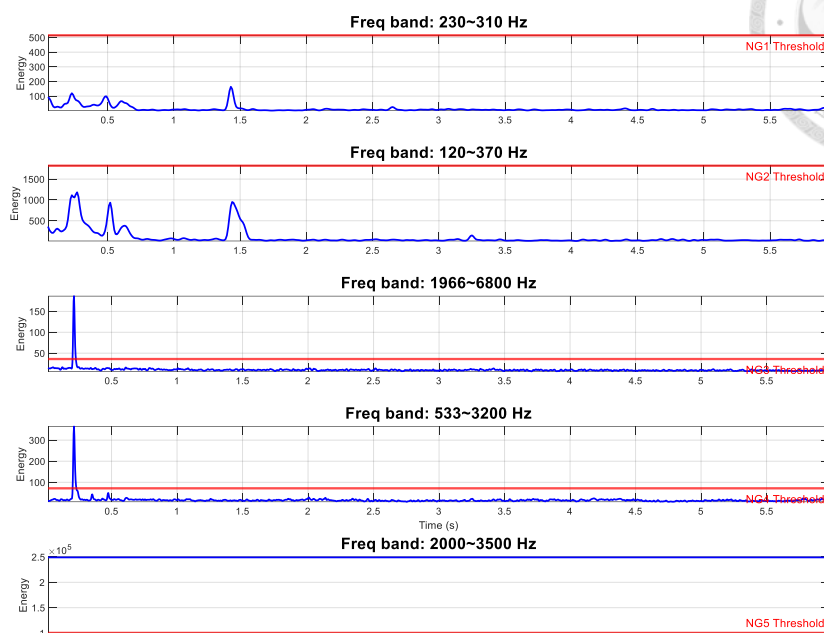
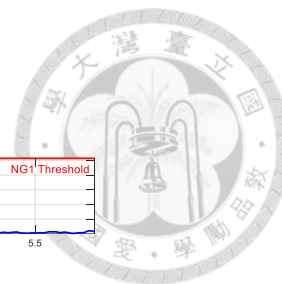


圖 5-12 墨西哥樣件三之泛用性測試

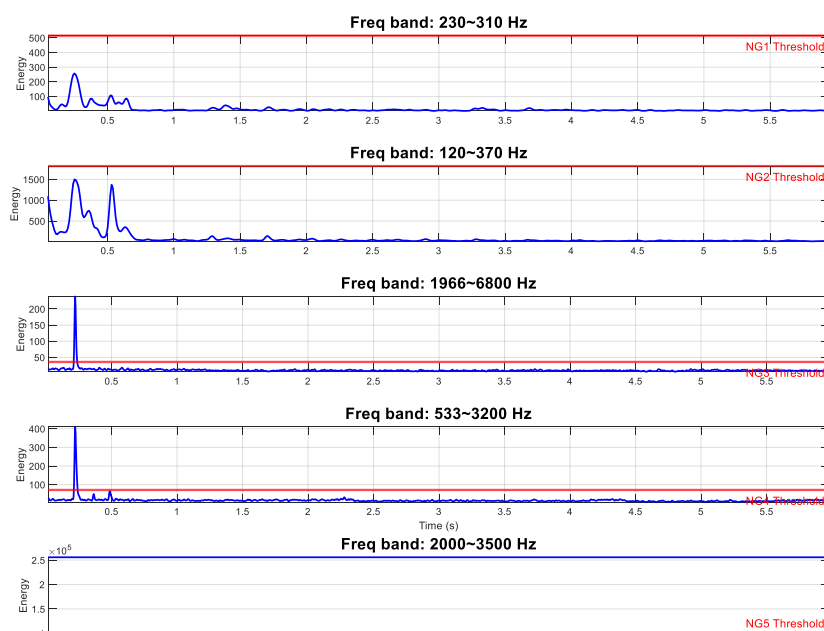


圖 5-13 墨西哥樣件四之泛用性測試

經檢測，樣件一、二之 NG1~4 之能量皆沒有超出閾值。樣件三於 0~0.5 秒之高頻聲響出現時，NG3、4 一同超出閾值，而 1.5 秒之低頻聲響則沒有超過閾值。樣件四於 0~0.5 秒之高頻聲響出現時，NG3、4 一同超出閾值。

然而，墨西哥樣件與台灣樣件之異音聽起來皆不相同，因此我們繪製台灣與墨西哥樣件之 STFT。由圖 5-14 可知，墨西哥樣件之中高頻異音與台灣樣件之 NG3、NG4 之頻帶分布不同，再由圖 5-15 發現，墨西哥樣件之 STFT 出現較多的持續性亮帶，這是因為台灣訊號是於無響室內錄製，因此在拉繩摩擦音(NG5)之特徵頻帶中沒有背景噪音，而墨西哥訊號是直接產線上錄製，故墨西哥訊號背景音較為雜亂，導致背景音與特徵頻帶重疊。

總而言之，若要檢測不同型號之天窗，需重新分析天窗異音之頻帶，並計算分類閾值，才能夠正確的分類異音，否則以不同型號天窗之異音指標將造成誤判。

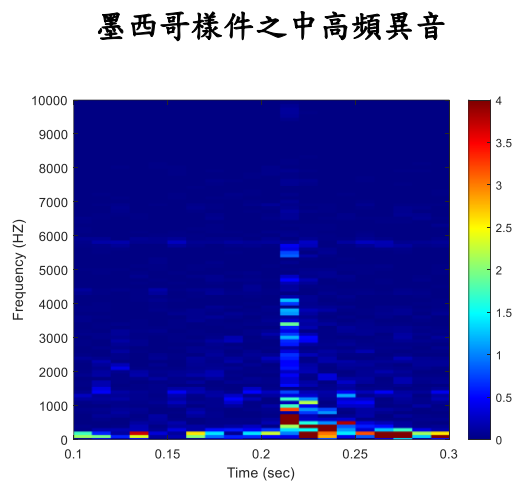
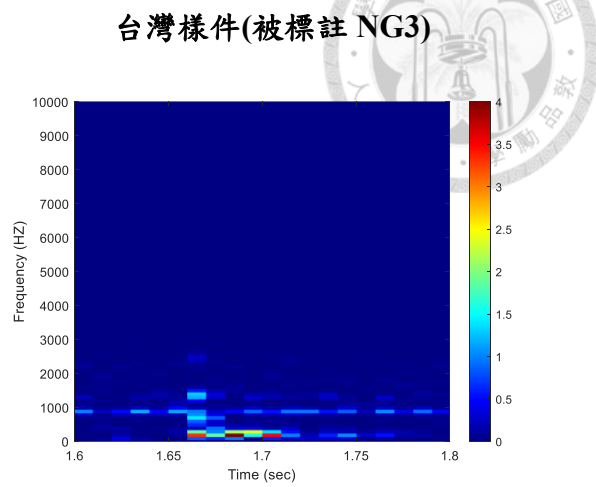
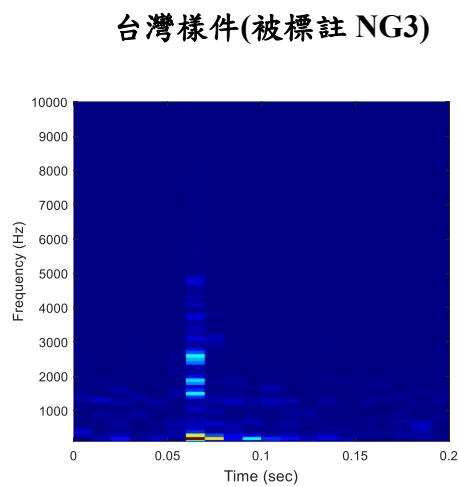


圖 5-14 台灣與墨西哥樣件之中高頻異音之 STFT 比較

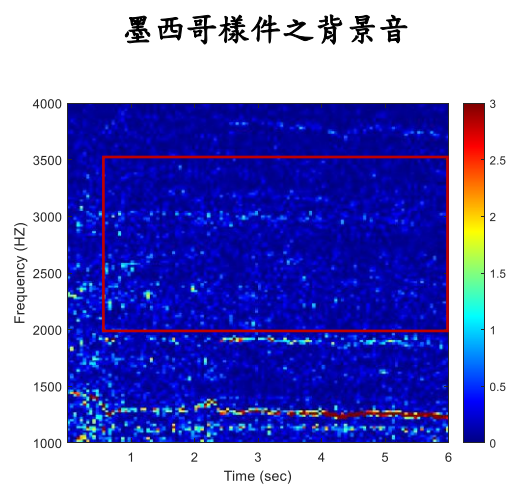
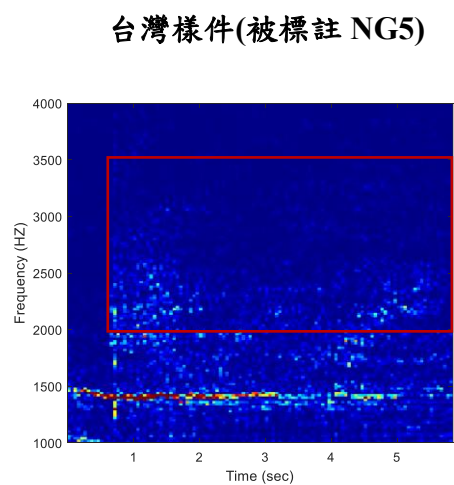


圖 5-15 台灣與墨西哥樣件之 NG5 特徵頻帶之 STFT 比較

第六章 總結與未來展望



本研究與車用電動天窗製造廠合作，提出一套以訊號處理為核心的異音偵測方法，分析之檢測音訊係由合作廠商提供與標註。

我們由 STFT 能量譜的異音亮帶找出特徵頻帶，並以特徵頻帶帶通濾波之訊號繪製能量歷時曲線以訂定異音長度。最後，以異音發生區間特徵頻帶的能量積分值作為異音指標。研究中，亦發展出一系統化方式，除能自動搜尋準確率最佳的特徵頻帶，還同時考慮了該頻帶的穩健性。

經過分析測試，可以得出以下結論：

1. 比較 STFT、Mel spectrogram 及 MFCC 等三種方法所產生的時頻圖，STFT 能量譜在異音發生時會產生明顯亮帶，異音特徵最明顯，最適合作 NG、OK 音訊判讀，故採用此方法進行所有的音訊分析。
2. 以異音發生區間之特徵頻帶能量積分作為偵測異音之指標可有效區分 OK、NG 訊號。
3. 本研究採用兩種閾值，第一種是以 $E_{Threshold} = \min_{NG} E_c$ ，採用此閾值不會將 NG 訊號被誤判為 OK，品管最嚴格；第二種是以 $E_{Threshold} = \max_{OK} E_c$ ，採用閾值將不會有 OK 訊號被誤判為 NG，若希望降低 False Positive 則採用此指標。實際應用係可依工廠需求，在兩種閾值之間自行調整。
4. 可由 $Accuracy(f_{low}, f_{high})$ 二維圖決定較佳準確率之異音指標積分頻帶，但選擇積分頻帶時應考慮穩健性。本研究所得之最佳特徵頻帶，如表 6-1、6-2。

表 6-1 以 $E_{Threshold} = \min_{NG} E_c$ 為分類閾值

變因 異常	最佳特徵頻帶(Hz)	Accuracy (%)
NG1	230~310	100
NG2	120~370	96.9
NG3	1966~6800	100
NG4	533~3200	100
NG5	2000~3500	100

表 6-2 以 $E_{Threshold} = \max_{OK} E_c$ 為分類閾值

變因 異常	最佳特徵頻帶(Hz)	Accuracy (%)
NG1	230~310	100
NG2	110~210	99.6
NG3	1966~6800	100
NG4	533~3200	100
NG5	2000~3500	100

5. 不同型號天窗的異音特徵頻帶與本研究之異音頻帶不同，檢測時可能出現誤判，因此針對不同型號之天窗進行異音檢測，需以本研究之方法重新分析訊號。

綜上所述，本研究所提出之異音偵測方法可有效分辨 OK 與 NG 訊號，但若將本方法應用至不同機構或環境之訊號，則需重新修訂異音指標以增加準確率與泛用性。未來在實務應用上，除了可發展為自動檢測系統，輔助品保師更有效率的執行品保工作，亦可利用各種異音的特徵生成大量的訊號，作為品保人員的訓練素材。

參考資料



- [1] 陳冠達(2023), “以複合音訊特徵及叢集分群實現非監督式異音檢測系統”, 國立清華大學 AI 碩士論文.
- [2] 王翊丞(2023), “機器學習應用於煞車來令片振動訊號缺陷檢測”, 國立臺灣大學工科海洋碩士論文.
- [3] 羅芳均(2017), “短時傅立葉轉換於風力發電機葉片表層損傷即時診斷之應用”, 國立臺灣大學工科海洋碩士論文.
- [4] 蔡東辰(2021), “運轉噪音與卷積神經網路於風力發電機葉片表層損傷即時診斷之應用”, 國立臺灣大學工科海洋碩士論文.
- [5] 黃柏諭(2021), “小波分析與卷積神經網路於風機齒輪箱振動故障診斷之應用”, 國立臺灣大學工科海洋碩士論文.
- [6] 邱冠穎(2021), “基於振動訊號與神經網路之風機傳動鏈狀態監測及流程設計”, 國立臺灣大學工科海洋碩士論文.