國立臺灣師範大學圖文傳播學系

碩士論文

獨立成份分析法於
可觸式聲波操控介面訊號分析之研究

An Analysis of Signal Processing by Using Independent Component Analysis for Tangible Acoustic Interfaces

研 究 生 ： 羅仁傑

指導教授： 周遵儒

中 華 民 國 102 年 8 月

誌謝

研究所四年完完整整過去，好不容易終於熬到寫誌謝文的這一刻。坦白說，這感覺真的是好極了！這段期間裡，我最要感謝我家人，特別是爸媽，雖然在學業上幫不上什麼忙但仍默默的支持著我，並忍受我竟然能念這麼久的書。我也很感謝我的姊姊與弟弟，每次當我生活陷入了財務困境時，總是二話不說金援我度過難關，讓我省去了很多困擾。感謝極電資訊蘋果師大校園店裡的各位朋友，不僅提供我一個躲老師的好地方，也給了我許多電腦專業知識及快樂回憶。我也要特別謝謝我的女朋友 Trabbit 小姐，在我論文生不出來時能忍受我的臭脾氣，在我論文插圖怎麼畫都很醜的時候能快速幫我重畫，在我分身乏術無力陪Google的時候能把它照顧得好好的，謝謝你給了我這麼多的協助並陪我度過這些煩悶的論文日子，我非常感激！

我很開心當初能選擇DCC研究室並成為DCC的一份子，雖然在這四年來周老師不知打了我幾次槍，讓我在寫論文時幾度失去方向感，但卻又如摯友般無私地提供建議與分享經驗，終於讓我完成了這艱鉅的任務。Deep、Fandy、小雞、律民、祖毅、續杯、穩容、瑋如、明信、 雅筑、 琮傑、有為、毓謙...等DCC的各位，還有常來串門子的維烝，因為你們在DCC的笑聲與陪伴，我的研究所生活才能過的如此有趣如此精彩。

最後，我還是要仰天怒吼一下！雖然自己打著圖文傳播學系的名號，做的研究卻是人機互動相關領域，寫出來的論文又像是電機工程的論文，每次被人問到“你研究所是在學些什麼東西？”都不知道該怎麼回答才好。但我不認為自己當初選擇了DCC研究室並學習如何寫程式是在這研究所中走一條難走的路，因為當自己完成這本論文時，終於能讓我拍胸脯地告訴大家，我可是科技學院畢業，拿的可是“徹徹底底”的工學院碩士學位啊！

羅仁傑（Roger Lo）謹誌於

國立臺灣師範大學

圖文傳播學系碩士班

2013.Aug

目錄

摘要 3

Abstract 4

圖目錄 5

表目錄 7

第一章　緒論 8

　第一節　研究背景與動機 8

　第二節　研究目的與問題 9

　第三節　研究範圍 10

第二章　文獻探討 12

　第一節　可觸式聲波操控介面技術 12

　第二節　獨立成份分析法原理 17

　第三節　文獻探討小結 33

第三章　研究方法 34

　第一節　實驗樣本與工具介紹 34

　第二節　符合獨立成份分析法的混合聲源假設 36

　第三節　依據獨立成份分析法原則的獨立聲源分析程序 42

第四章　實驗結果與討論 48

　第一節　多音源輸入訊號與ICA處理 48

　第二節　單音源訊號經頻率區隔與ICA處理 52

　第三節　單音源訊號經頻率區隔及高次混合與ICA處理 56

第五章　結論與建議 62

　第一節　結論 62

　第二節　建議 62

參考文獻 64

摘要

本研究利用獨立成份分析法（ Independent Component Analysis，ICA）於可觸式聲波操控介面上針對同時間不同觸碰事件所產生的混合訊號進行訊號分離。依據獨立成份分析法的處理步驟，本研究提出三種分析操作程序來對可觸式聲波操控介面下麥克風截取到的混合訊號實際進行測試，分別為 (一) 多音源輸入訊號直接進入ICA處理 (二) 單音源輸入訊號經頻率區隔分析後進行ICA處理 (三) 單音源輸入訊號經頻率區隔及高次混合後進入ICA處理。研究結果顯示，透過頻率區隔在可觸式聲波操控介面下對於非即時性的訊號分析是具有顯著的訊號分離效果。此結果也驗證了獨立成份分析法可運用在可觸式聲波操控介面上作為混合訊號分離使用。

關鍵字：獨立成份分析法、可觸式聲波操控介面、聲源訊號分離

Abstract

In this study, we tried to use independent component analysis (ICA) to separate the mixed signal caused by simultaneous touch events happened on a tangible acoustic interface. Being in accord with the processing steps of ICA, the procedures we proposed can be divided into (1) multiple audio inputs for ICA processing, (2) single audio input with frequency segmentation for ICA processing, (3) single audio input with frequency segmentation and higher-order mixed for ICA processing. Experimental results show that, using frequency segmentation for non-real time signal analysis on tangible acoustic interface has remarkable signal-separating effect. It also proves an idea that independent component analysis can be used on tangible acoustic interface for mixed-signal separation.

Keywords : independent component analysis, tangible acoustic interface, audio signal separation

圖目錄

圖 2-1　到達時間延遲法示意圖　15
圖 2-2　位置模式比對法示意圖　15
圖 2-3　雞尾酒派對問題示意圖　17
圖 2-4　中央極限定理說明圖　21
圖 2-5　獨立成份分析法步驟流程圖　23
圖 3-1　木棒敲擊與鐵棒刮擊波型圖　35
圖 3-2　本實驗所採用之可觸式聲波操控介面示意圖　35
圖 3-3　基於獨立成份分析法之混合訊號組成成份示意圖　36
圖 3-4　演唱會live音樂片段左右聲道波形圖　38
圖 3-5　演唱會live音樂片段透過ICA處理後所得之分離訊號波形圖　38
圖 3-6　演唱會live音樂片段聲音訊號頻率分佈圖　39
圖 3-7　演唱會live音樂片段透過ICA處理後所得之分離訊號頻率分佈圖　39
圖 3-8　木棒敲擊和與鐵棒刮擊原始訊號波形圖　40
圖 3-9　混合後訊號波形圖　40
圖 3-10 FastICA 分離後訊號波形圖　40
圖 3-11 木棒敲擊和與鐵棒刮擊原始訊號頻率分布圖　41
圖 3-12 混合訊號頻率分布圖　41
圖 3-13 分離後訊號頻率分布圖　41
圖 3-14 本研究所採用之獨立聲源分析三操作程序流程說明圖　42
圖 3-15 木棒敲擊與鐵棒刮擊波型圖　43
圖 3-16 木棒敲擊與鐵棒刮擊訊號頻率分布圖　43
圖 3-17 程序（一）之混合訊號組成示意圖　44
圖 3-18 程序（二）之混合訊號組成示意圖　45
圖 4-1　程序（一）：五個麥克風所截取的混合訊號聲波圖　48
圖 4-2　程序（一）：混合訊號原始波形圖　48
圖 4-3　程序（一）： ICA後所得訊號波形圖-輸出 2 個訊號　49
圖 4-4　程序（一）： ICA後所得訊號波形圖-輸出 3 個訊號　49
圖 4-5　程序（一）： ICA後所得訊號波形圖-輸出 4 個訊號　50
圖 4-6　程序（一）： ICA後所得訊號波形圖-輸出 5 個訊號　50
圖 4-7　程序（二）： 單一麥克風所錄製的聲波訊號圖　52

圖 4-8　程序（二）： 單一麥克風所錄製的聲波頻率分佈圖 52
圖 4-9　程序（二）： 利用頻率區隔所製作出的所有訊號波型圖 53
圖 4-10　程序（二）： ICA所得之訊號波型圖 - 輸出 2 個訊號 54
圖 4-11 程序（二）： ICA所得之訊號波型圖 - 輸出 3 個訊號 54
圖 4-12 程序（二）： ICA所得之訊號波型圖 - 輸出 4 個訊號 55
圖 4-13 程序（二）： ICA所得之訊號波型圖 - 輸出 5 個訊號 55
圖 4-14 程序（三）：高次混合訊號組 X_{high} 之訊號波形圖之一 56
圖 4-15 程序（三）：高次混合訊號組 X_{high} 之訊號波形圖之二 57
圖 4-16 程序（三）：高次混合訊號組 X_{high} 之訊號波形圖之三 57
圖 4-17 程序（三）：高次混合訊號組 X_{high} 之訊號波形圖之四 58
圖 4-18 程序（三）：高次混合訊號組 X_{high} 之訊號波形圖之五 58
圖 4-19 程序（三）： ICA所得之訊號波型圖 - 輸出 2 個訊號 59
圖 4-20 程序（三）： ICA所得之訊號波型圖 - 輸出 3 個訊號 59
圖 4-21 程序（三）： ICA所得之訊號波型圖 - 輸出 4 個訊號 60
圖 4-22 程序（三）： ICA所得之訊號波型圖 - 輸出 5 個訊號 60

表目錄

表 1 可觸式聲波操控介面技術優缺點比較 15
表 2 可觸式聲波操控介面技術相關應用研究 16
表 3 *FastICA* 演算法流 31

第一章　緒論

科技日新月異，人與機器的互動方式也隨著科學技術的創新與成熟，不斷地在改變著，然而要如何讓人與機器以著最自然的方式互相溝通，這不僅是所有科學家不斷在尋找的解答，也是激發本研究動機的開端。本研究是針對2004-2006年歐洲的Tai-Chi計畫（Tangible Acoustic Interfaces for Computer-Human Interfaces）中所提出的可觸式聲波操控介面（Tangible Acoustic Interface）概念 (Polotti, Sampietro, Sarti, Tubaro, & Crevoisier, 2005)，嘗試利用訊號分析的技術來擴展介面運用的可能性，因此在本章節中的內容中，將分別說明本研究的研究背景與動機、研究目的、研究問題及研究範圍。

第一節　研究背景與動機

過去20年來，在人機互動（Human-Computer interaction）的領域上產生了許多的改變。傳統的人機互動系統主要以單一使用者且任務導向為主，如鍵盤、滑鼠等輸入裝置 ; 新興的人機互動系統逐漸走向嵌入式且以多重模式來呈現，經過設計後讓使用者感覺不到裝置的存在，進而完成更自然的人機互動，如常見的webcam攝影機及微軟的 KINECT 互動控制器。普及運算（ubiquitous computing）之父 Mark Weiser 於 1991 年提出「智慧型環境」（smart environment）這個概念，他認為生活的環境可視為機器與人溝通的大介面，透過運算技術在背後支持著，能將日常生活中所用的物品一一轉變可互動式，如此一來人們就能智慧地且無障礙地處理著日常生活中的各種任務 (Elena, Omar, Stéphane, Stefano, & Houda, 2009)。

可觸式聲波操控介面（Tangible Acoustic Interface，TAI）是近幾年來被提出作為上述這種無縫式人機互動的方法之一，其作法在於將麥克風等感測元件裝設在桌子、牆壁等日常生活常見的物體表面上，使得物體表面轉變為可觸式介面，並藉由分析擷取來的聲波震動訊號，來提供系統完整描述出觸控過程，如該物體表面是何處被接

觸，以及如何被接觸等訊息 (Crevoisier & Bornand, 2008)。這種可觸式聲波操控介面方法與過往的觸控技術，如電阻/電容式觸控、紅外線矩陣式觸控等，其最主要的差別在於可觸式聲波操控介面是基於聲音在固體表面的傳播特性來進行互動，不需要複雜機械或電子裝置來配合，即可將物體本身作為訊號傳播的載體和人機互動介面(Fabiani, 2006)。如此一來不但能減少製作上的耗費，也提供了更加自然的人機互動方式，讓應用層面擴展地更全面，如在互動藝術創作、情境感知設計等等多元領域上進行應用。

大多數在可觸式聲波操控介面的相關研究聚焦於如何判定單一觸碰事件發生的位置，例如手指在木板的表面上輕敲，透過感測器去讀取傳遞於物質間的震動訊號，再計算出手指敲擊於木板的位置。相關使用的技術如『到達時間延遲法（Time Delay of Arrival）』是基於各感測器所擷取到的訊號時間差異，以及『位置模式比對法（Location Pattern Matching）』是比對擷取到的訊號與事前錄製並儲存的訊號資料庫，進而推算出訊號源產生的幾何位置(Pham et al., 2005)。然而聲音的速度隨頻率而變化所造成的頻散現象（dispersion），以及互動空間中所產生環境雜訊也會混雜於擷取到的訊號中，容易影響定位的精準度(Fabiani, 2006)。此外，上述兩項技術目前僅能同時間內處理單一觸碰事件上，若有兩者或兩者以上的觸碰事件同時產生，感測器所擷取到的訊號則為混合訊號，若無法先將各事件所產生的訊號從混合訊號中先行抽離，則系統難以進行判別。

為解決上述問題，在本文章中提出以獨立成份分析（Independent Component Analysis, ICA）的技術，嘗試將可觸式聲波操控介面中感測器所擷取的混合訊號進行訊號分析。此項技術是將混合訊號經由線性或非線性轉換，利用統計學上獨立的原理找出呈現最非高斯分佈的獨立成份，最常被應用於語音訊號分離上，可以成功地從混合的聲音訊號中分離出獨立的聲源。若能將獨立成份分析技術運用在可觸式聲波操控介面上，將感測器所截取的混合訊號依組成聲源成份進行分離，便可利用分離出的獨立聲源來擴大互動偵測的運用廣度。因此本研究依據這個原理，將獨立成份分析法推廣

運用於可觸式聲波操控介面中的混合訊號分離上，藉以提供可觸式聲波操控介面裝置開發技術上的另一選擇。

第二節　研究目的與問題

本研究希望提出能運用在可觸式聲波操控介面上達到混合訊號分離效果的系統化方式，因此欲以獨立成份分析法在可觸式聲波操控介面上，將擷取到的混合訊號進行分析，並嘗試分離出因不同互動方式所產生的不同聲源訊號。

因此本研究可歸納為以下三項目的：

一、利用獨立成份分析法對可觸式聲波操控介面發展出系統化的訊號分析方式。

二、提出利用獨立成份分析法在可觸式聲波操控介面上的訊號分離操作程序。

三、檢視獨立成份分析法在可觸式聲波操控介面中針對混合訊號的分離效果。

根據以上研究目的，本研究的研究問題可歸納如下述三項：

一、如何運用獨立成份分析法於可觸式聲波操控介面上進行訊號分析？

二、利用獨立成份分析法在可觸式聲波操控介面進行混合訊號的操作程序為何？

三、運用獨立成份分析法在可觸式聲波操控介面上將混合訊號分離的效果如何？

第三節　研究範圍

本研究之範圍界定可分為以下兩點說明：

一、訊號處理區分為即時性（real time）與非即時性（ non-real time）等兩種處理模式，本研究僅對可觸式聲波操控介面擷取得來的混合訊號進行非即時性的分析處理，至於即時性的訊號處理方式並非在本研究的研究範圍內。

二、本研究僅聚焦於如何提高獨立成份分析法在可觸式聲波操控介面中的混合訊號分離效果，至於如何利用分離出來的獨立聲源訊號進一步作為介面應用，則不在本研究的研究範圍內。

第二章　文獻探討

於本章節內容中將探討可觸式聲波操控介面與獨立成份分析法的技術原理。第一節先從可觸式聲波操控介面技術層面進行介紹，分別解說可觸式聲波操控介面的結構原理、常被用於該介面上作為觸碰事件偵測的技術比較及其相關應用研究 ; 第二節則針對獨立成份分析法之理論基礎進行說明，詳細說明演算方法及處理過程及其相關應用研究 ; 於第三小節中則將文獻探討部分進行總結性的討論。

第一節　可觸式聲波操控介面技術

現今的觸控技術發展成熟，舉凡手機螢幕、電視電腦螢幕、廣告看板、電子白板等等，觸控介面的應用已經滲透到人們的生活周遭中，種類包括了電阻式、電容式、光學式、及紅外線式等等多項觸控技術。然而絕大多數的觸控技術需要複雜機械或電子裝置來配合，這些耗費成本高且精密的觸控裝置介面，不僅難以推廣到大尺寸的應用上，也常因維護不易而不適合於所有環境情況下使用(Fabiani, 2006; Sanctis, Rovetta, Sarti, Scarparo, & Tubaro, 2006)。

可觸式聲波操控介面是利用聲音在固體中的傳導特性來進行互動。當物體與物體的表面作用時，表面所產生的聲波模型會因為物體之間作用的不同而改變，加上聲音震動在絕大多數的物質中傳導性佳，無論該物體表面是否被其他物體覆蓋，兩物體碰觸所產生的聲音訊息均能透過物體本身結構來作為傳輸管道傳遞到遠處去。我們在物體表面上設置接觸式麥克風或是加速度感測器來擷取物體內傳遞的聲音訊息，再將這些聲波模型視覺化及特徵化，經過分析後便可描述出兩物體作用的過程。

由以上敘述可知，可觸式聲波操控介面與其他觸控技術相較起來，具有以下三種特色：

一、作用表面不需侷限在任何特定大小，能有效地擴展到任何尺寸上。

二、能夠減少互動介面的製作成本。

三、可以選擇適合互動情況與環境的物體作為介面使用。

在可觸式聲波操控介面上針對觸碰事件所產生的聲音震動訊號進行分析的技術主要代表有「到達時間延遲法（Time Delay of Arrival，TDOA）」與「位置模式比對法（Location Pattern Matching，LPM）」等兩種技術(Fabiani, 2006; Pham et al., 2005)。以下將分別細部介紹這兩種技術原理：

一、 到達時間延遲法（Time Delay of Arrival，TDOA）

時間延遲估計法（time delay estimation）已經大量被運用在如雷達、聲吶等領域上作為偵測、辨別或定位物體的用途，而這種估計方法也可以被運用於開發可觸式聲波操控介面上。到達時間延遲法是利用時間延遲估計法的原理，先利用麥克風或是加速度感測器等感測裝置，將其設置在物體表面的不同位置上，每當物體表面被碰觸時，震動訊號將會在物體表面上產生並藉由物體本身將訊號傳遞出去。每個感測器所接收到訊號的時間將會被記錄下來，若已知該物質的傳導速度，則可以進一步推算出震動訊號傳遞的距離，並比對出作用點的位置。換句話說，使用到達時間延遲法來作為可觸式聲波操控介面之訊號分析的好處是，只要是在感測器所圈圍出的範圍內，任何作用位置點都可以藉由分析感測器接收訊號的時間差異推算出來。此外，到達時間延遲法的優點在於反應速度快且運算複雜量低，因此在訊號判斷上能表現出不錯的性能。

然而，到達時間延遲法主要被應用在單一觸碰事件的定位使用，無法更進一步用於同時間且多重觸碰事件的情況。到達時間延遲法的表現能力好壞依賴著所擷取的時間延遲資料之精準度與穩定性，當聲波震動訊號傳導的物質本身同質性較低時，波在同質性低的物質中傳導速度非一致，因此接收到訊號的時間將會因為波在不同物質間

的傳導速度差異而造成誤差，使得到達時間延遲法無法有效地在可觸式聲波操控介面上運作(Yang, Pham, Al-Kutubi, Ji, & Wang, 2009)。另外一個缺點則是，該方法必需設置較多的感測器來接收訊號（如圖2-1所示），缺乏足夠的資料則無法進行評估，而這樣的要求也提高了硬體設備的複雜度。

二、位置模式比對法（Location Pattern Matching，LPM）

位置模式比對法的技術原理在於先將數筆事先錄製好的訊號作為資料庫，而後將擷取到的聲波訊號圖形與資料庫進行比對，進而達到訊號識別的目的。換句話說，感測器所接收到的每個聲波訊號圖形都可以視為是某特定位置所產生的訊號內容綜合體，包含頻率、聲波的相位等等，藉由感測器所擷取到的訊號和那些已知且事先儲存的訊號資料庫進行訊號間的相關係數計算，依據係數最大值來推測從觸碰作用點發出的聲源資料為何，如位置、或接觸的方式等等。

相較於到達時間延遲法，位置模式比對法通常只需要一個感測器就可以進行（如圖2-2所示），在硬體的設置上有較大的優勢。由於位置模式比對法是利用相關係數的計算方式來找出配對者，因此也具有減少雜訊影響的效果。然而，使用位置模式比對法必須事先儲存多筆訊號資料，新資料與舊資料的圖形比對需要較大的運算量，導致在訊號判斷上的速度與性能遠低於時間延遲法。此外，雖然不同的觸碰方式會讓表面所產生的聲波模型不同，但感測器在同一時間下所擷取到的聲波訊號仍為多聲源的混合訊號，若無法事先將不同的聲源訊號從混合訊號中分離出來，則位置模式比對法仍無法在同時間且多重觸碰事件的情況下使用。以下表1將上述兩種方法依其優缺點列表比較。

表 **1**
可觸式聲波操控介面技術優缺點比較

方法名稱	優點	缺點
到達時間延遲法 **(TDOA)**	1. 運算量低。 2. 反應速度快。	1. 只適合用於同質性的物體表面上。 2. 無法用於同時間且多重觸碰事件的情況。
位置模式比對法 **(LPM)**	1. 只需要一個或極少數的感測器。 2. 可用在複雜形狀的物體上。	1. 反應速度較慢。 2. 在圖形比對的過程，必須要有學習階段才能提高訊號偵測的精準度與反應速度。 3. 無法用於同時間且多重觸碰事件的情況。

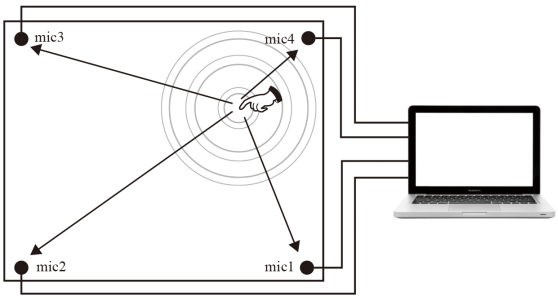

圖 **2-1** 到達時間延遲法示意圖

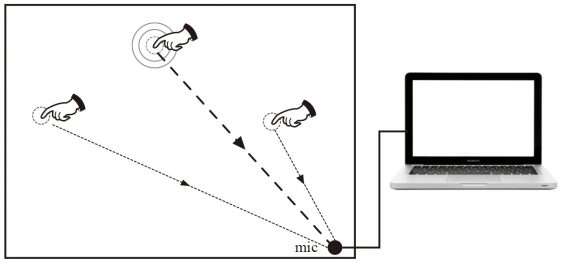

圖 **2-2** 位置模式比對法示意圖

於上述內容中，簡單的介紹了可觸式聲波操控介面中的原理及兩項訊號分析技術，接下來將於表 2 列出目前已運用可觸式聲波操控介面技術的相關應用研究。

表 **2**
可觸式聲波操控介面技術相關應用研究

年份	裝置名稱 / 引用文獻	TAI技術	互動方式	辨識依據	應用領域
2006	Sound Rose (Crevoisier, Bornand, Guichard, Matsumura, & Arakawa, 2006)	TDOA	Finger Touch (Tapping & Dragging)	Position	Touch Table (interactive installation)
2008	Scratch Input (Harrison & Hudson, 2008)	LPM	Drawing (Line, Circle, Triangle and Square)	Frequency Amplitude	Touch Table/Wall (gesture recognition)
2008	Music Wall (Hu, Tung, & Lau, 2008)	TDOA	Tapping	Position	Digital Music Player
2011	Augmenting Touch (Lopes, Jota, & Jorge, 2011)	LPM	Gesture (Finger Tap, Knock, Slap, Punch)	Amplitude Timbre	Touch Table (gesture recognition, gesture intention)
2012	Bathcratch (Shigeyuki, Yoshinobu, & Seiho, 2012)	LPM	Rubbing & Touch	Frequency	Interactive Bathtub
2012	Acoustic Barcodes (Harrison, Xiao, & Hudson, 2012)	LPM	Swiping with objects	Peak	Encoding/Decoding

第二節　獨立成份分析法原理

獨立成份分析法（Independent Component Analysis，ICA）是一種統計及計算的技術，主要是用來找出隨機變數或訊號中的隱藏因子（hidden factors）。在 ICA 的模型中，假設觀察的多變量資料是由某些未知的潛在變數（latent variable）經由線性或非線性的方法混合而成(Comon, 1994)。這些潛在的變數被稱之為觀察資料中的（Independent Component），並且具有非高斯分佈且互相獨立的特性。ICA 就是利用分析的方法來找出這些獨立成份。

一、ICA 模型的基本概念：

ICA 應用在與語音訊號處理方面，最典型的例子就是「雞尾酒派對問題」(cocktail party problem)(Ainhoren, Engelberg, & Friedman, 2008)。如圖2-3所示，假設有3個人在交談，每個說話者的聲音訊號分別為 $s_1(t)$、$s_2(t)$、$s_3(t)$，在房間中放置了3支麥克風來記錄聲音，每支麥克風收到的混合訊號分別為 $x_1(t)$、$x_2(t)$、$x_3(t)$。

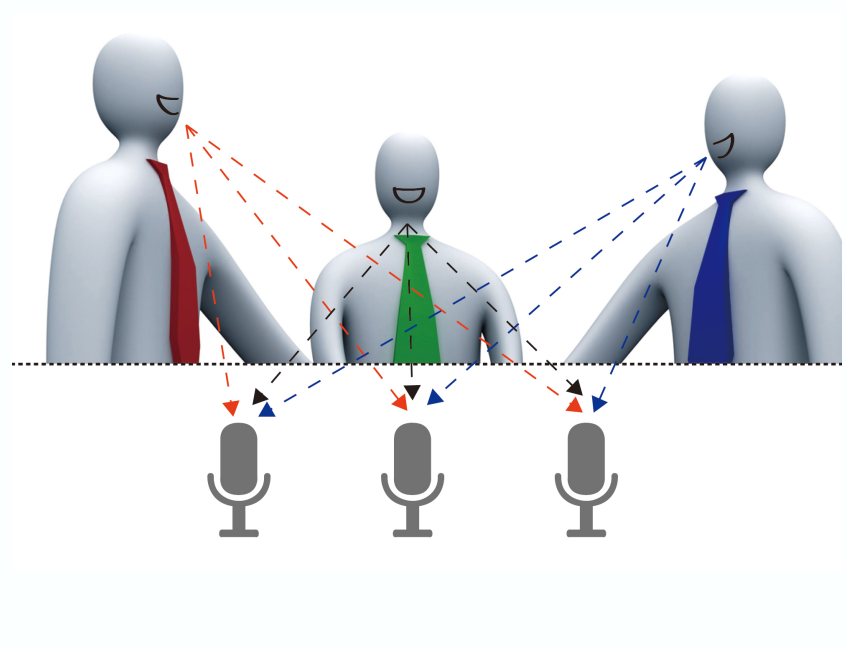

圖 **2-3** 雞尾酒派對問題示意圖

在此過程中，我們對聲源與混合過程皆無法事先知道，可擁有的資訊只有麥克風所量到的訊號。由於語音訊號從說話者到收音的麥克風中間除了直接抵達外，還會有經由不同路徑反射而來的訊號，為了簡化問題，在此先不考慮反射路徑，只考慮線性組合，將聲源與所量測到的訊號之間的關係以線性方程式表示成：

$$
\begin{aligned}
x_1(t) &= a_{11}s_1(t) + a_{12}s_2(t) + a_{13}s_3(t) \\
x_2(t) &= a_{21}s_1(t) + a_{22}s_2(t) + a_{23}s_3(t) \\
x_3(t) &= a_{31}s_1(t) + a_{32}s_2(t) + a_{33}s_3(t)
\end{aligned}
\tag{2-1}
$$

其中 a_{ij} 代表第 j 個聲源到第 i 個麥克風的放大倍率，取決於聲源與麥克風的距離。在一般的情況下，通常只能取得混合訊號 $x_1(t)$、$x_2(t)$、$x_3(t)$ 的數值，若是在 a_{11}、a_{12}、a_{13}、a_{21}、a_{22}、a_{23}、a_{31}、a_{32}、a_{33} 皆未知的情況下，要取得最原始的聲源訊號 $s_1(t)$、$s_2(t)$ 與 $s_3(t)$ 將變得相當困難。我們先將 ICA 的基本模型改寫成矩陣型式：

$$
x = As \tag{2-2}
$$

A 為 x 與 s 之間的混合矩陣（mixing matrix）。x 是由 s 線性組合而成，根據中央極限定理（central limited theorem），將一群非高斯且彼此獨立的隨機變數經過線性混合，其混合後的機率分佈會趨近於高斯分布。換句話說，x 會比 s 更近似於高斯分佈。在雞尾酒派對問題中，唯一得到訊息的只有麥克風所收錄的混合訊號 x，而說話者語音訊號 s 與實際混合情況 A 皆未知。在假設說話者的語音訊號彼此獨立的情況下，利用 ICA 來找出一個合適的解混合矩陣 W（unmixing matrix），使得接收到的訊號經過 W 的轉換能夠得到原來的聲源訊號。

x 由 W 重建後得到 y ：

$$y = Wx \tag{2-3}$$

結合上述兩公式，因此可得：

$$y = WAs \tag{2-4}$$

如何估計 W 主要是利用聲源訊號與訊號之間呈現獨立的特性，換句話說，就是經由判斷 y 中每個訊號是否獨立來決定 W。當 y 具有最大的非高斯特性時，相對的獨立性也最大，可得到 $W = A^{-1}$，則 ICA 估測出來的 y 就可當做是原來的獨立訊號源 s 。根據上述 ICA 的基本觀念，可進一步將問題擴大為有 m 個麥克風來接收 n 個聲源，如下所示：

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{21} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{bmatrix} = As \tag{2-5}$$

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1m} \\ w_{21} & w_{21} & \cdots & w_{2m} \\ \vdots & \vdots & & \vdots \\ w_{n1} & w_{n1} & \cdots & w_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} \approx s \tag{2-6}$$

二、ICA 的假設條件：

ICA 演算法本身有四項基本假設條件，在使用之前，必須先確定是否滿足這些條件(Stone, 2002)：

（一）聲源本身在統計上獨立（independent）

由於 ICA 的演算法就是用聲源彼此獨立的特性來發展出來的，因此所有的聲源必須滿足彼此獨立的特性。定義獨立的概念，在此先考慮兩個純量的隨機變數 v_1 與 v_2 ，若是 v_1 與 v_2，或是 v_2 與 v_1 沒有任何關係，就稱 v_1 與 v_2 相互獨立。更精確來說，獨立可利用 probability density 來定義，當隨機變數 $v_1, v_2, \ldots, v_n$ 其 joint probability density function 可表達為 marginal probability density function 的乘積，則可稱隨機變數 $v_1, v_2, \ldots, v_n$ 之間是互相獨立的。

$$p(v_1, v_2, v_3, \ldots, v_n) = p_1(v_1)\, p_2(v_2)\, p_3(v_3) \cdots p_n(v_n) \tag{2-7}$$

（二）各獨立聲源必須為非高斯分佈（non-Gaussian distribution）

由於 ICA 尋找獨立元素的方法來自中央極限定理（Central Limit Theorem），此定理中說明了：若將多個非高斯分佈（non-Gaussian distribution）且相互獨立的隨機變數相加之後，其整體結果會更趨近高斯分佈（Gaussian distribution），如圖2-4說明。換句話說，任意兩個隨機的訊號越偏向於非高斯分佈，則這兩個訊號之間的關聯性就越小，也就是說這兩個訊號之間越可能獨立。

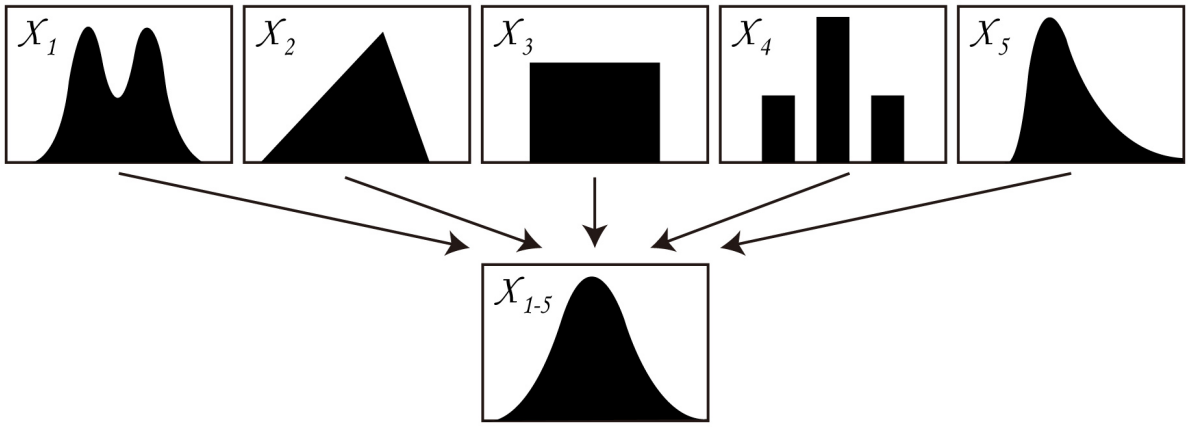

圖 **2-4** 中央極限定理說明圖

此外，高斯分佈之變數具有疊加性，當兩個高斯分佈的訊號相加之後仍然為高斯分佈。因此若一個群集訊號是由高斯訊號線性混合而成，則 ICA 將無法分離出真正的原始訊號，因此在使用 ICA 來分離訊號時，群集訊號中最多僅能允許有一個訊號為高斯分佈。

（三）聲源的數目必須小於或等於感測器的數目

假設以 m 個感測器獲得訊號 x，而原始聲源訊號 s 數量為 n 個。在公式 $x = As$ 的情況下，以線性方程式的角度來看：若 $m = n$，則 $s = A^{-1}x$，有解 ; 若 $m > n$ ，則方程式比未知數多，在這種情形下可先將維度降為 n 再執行 ICA ; 若是 $m < n$，則方程式比未知數少，在這種情況下，則缺乏足夠的條件來找出未知的獨立元素。因此聲源的數目必須小於或等於感測器的數目。

（四）感測器本身的雜訊很小，可以忽略不計

若感測器本身的雜訊很大，ICA 會將感測器雜訊視為另一個獨立聲源，進而造成聲源數目大於感測器數目的情況。

三、ICA 方法的不確定性

從 ICA 的模型中可以看出以下兩項不確定性(Hyvärinen & Oja, 2000)：

（一）無法確定各獨立元素的變異數

根據 ICA 模型 $x = As$， 第 i 個感測器所接收到的混合訊號可表示成公式(2-8)：

$$x_i = \sum_{j=1}^{n} a_{ij} s_j \tag{2-8}$$

由於混合矩陣 A 與 聲源 S 皆未知，當 S_j 被放大 k 倍時，a_{ij} 只要除以 k 就可以互相抵銷，如公式(2-9)：

$$x_i = \sum_{j=1}^{n} \left(a_{ij} / k\right)\left(s_j k\right) \tag{2-9}$$

因此假設每一個獨立聲源的變異數（Variance）均等於 1，如公式(2-10)：

$$E\left\{s_i^2\right\} = 1 \tag{2-10}$$

（二）無法確定獨立元素的順序

此項不確定性也跟 ICA 模型 $x = As$ 中未知的混合矩陣 A 與聲源 s 有關。根據公式(2-8)，每個混合訊號的次序是可以任意的被調換的，因此無法定義某個獨立訊號是第幾個。換句話說，利用 ICA 所估計出來的解混合矩陣 W 所求得的獨立成份獨立成份，與原始獨立訊號的順序是不相同的。

四、ICA 演算法的架構：

ICA 對於處理資料方面大致上可分為兩個階段。其流程如圖2-5所示：

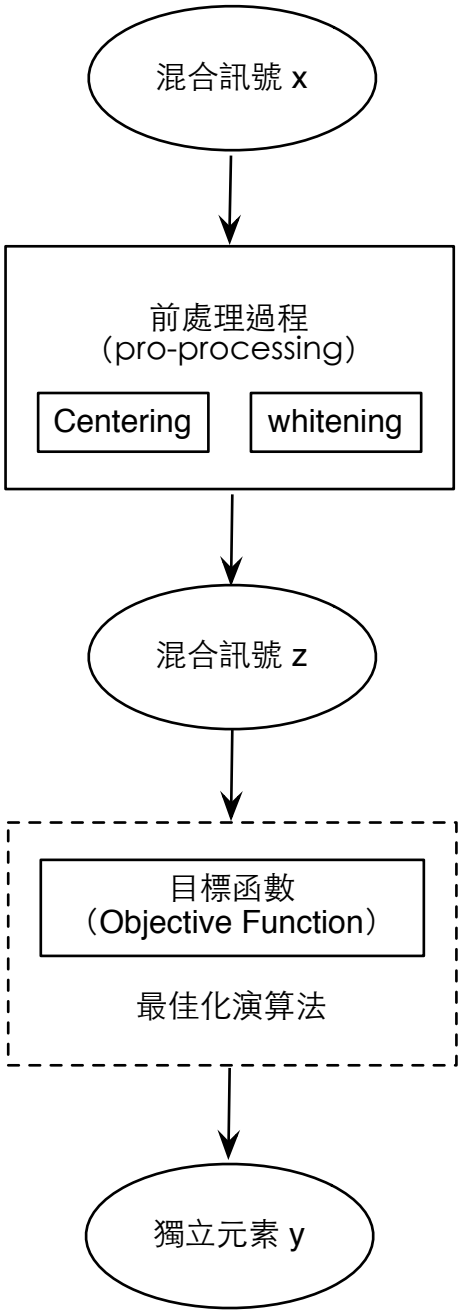

圖 **2-5** 獨立成份分析法步驟流程圖

（一）第一階段：資料前處理（Pre-processing）

ICA 是從觀察訊號中找出獨立成份獨立成份來解決訊號分離的問題，整個過程相當複雜。如果可以先將觀察訊號轉換成有意義的形式，將能簡化運算的複雜性。基於

這個理由，在進行 ICA 演算法之前，通常會先將取得的混合訊號 x 執行去均質（centering）與白化（whitening）等兩項步驟，以減輕後續動作的運算量。

1. 去均質（centering）

先將所量測到的混合訊號 x 減去其平均值 $E(x)$，使獲得的新訊號 x' 為零平均，其主要目的是要加快 ICA 的演算速度。

$$x' = x - E\{x\} \tag{2-11}$$

2. 白化（whitening）

白化的目的是利用線性轉換將原本相關變數轉換成不相關變數，使得原本變異度較大的變數都集中落在少數的新變數中。變異度較大的變數為主要成份，變異度較小的變數則可視為雜訊來源。透過這種方法，可以丟掉變異度小的訊號，進而達到抑制雜訊，加強來源訊號的效果。

假設有一零平均的隨機向量 $z = (z_1, \ldots, z_M)^T$ 滿足 $E\{zz^T\} = I$，I 為單位矩陣，則我們稱這個向量為白化向量。在 ICA 中，對於零平均的獨立訊號源 $s = [s_1, \ldots, s_N]^T$ 而言，可以得到：

$$E\{s_i s_j\} = E\{s_i\}E\{s_j\} = 0 \tag{2-12}$$

其共變異矩陣（*covariance matrix*）是單位矩陣，即 $cov(s) = I$，因此訊號源 s 是彼此不相關的。對於觀察訊號 x，我們可以找到一個線性轉換，使 x 投影到新的子空間後變成白化向量，即

$$z = Vx \tag{2-13}$$

其中 V 稱為白化矩陣（whitening matrix），z 為白化向量。常見的方法是對觀察訊號的共變異矩陣做特徵值分解（eigenvalue decomposition）。

$$E\left\{xx^T\right\} = EDE^T \tag{2-14}$$

E 是由 $E\{xx^T\}$ 的特徵向量（eigenvectors）所構成的正交矩陣（orthogonal matrix），D是由其相對應的特徵值（eigenvalues）所構成的對角矩陣（diagonal matrix）於是可得白化矩陣為：

$$V = ED^{-1/2}E^T \tag{2-15}$$

將 V 帶回 $z = Vx$ 式得 $z = ED^{-1/2}E^Tx$ ，因此可藉由檢查其共變異矩陣是否為單位矩陣，以確定經過 V 轉換後的 z 是否已為一白化向量。

$$\begin{aligned}
&E\left\{zz^T\right\} \\
&= E\left\{ED^{-1/2}E^T xx^T ED^{-1/2}E^T\right\} \\
&= ED^{-1/2}E^T E\left\{xx^T\right\}ED^{-1/2}E^T \\
&= ED^{-1/2}E^T EDE^T ED^{-1/2}E^T \\
&= ED^{-1/2}E^{-1}EDE^{-1}ED^{-1/2}E^T \\
&= ED^{-1/2}DD^{-1/2}E^T \\
&= EE \\
&= I
\end{aligned} \tag{2-16}$$

我們再將 $x = As$ 帶入 $z = Vx$ 中 ，並令 $VA = \tilde{A}$，則

$$z = VAs = \tilde{A}s \tag{2-17}$$

如果把上式中的 z 視為新的觀察訊號，我們可以說白化的步驟將原來的混合矩陣 A 進行線性轉換成一正交矩陣。

$$E\left\{zz^T\right\} = E\left\{\tilde{A}ss^T\tilde{A}^T\right\} = \tilde{A}^T E\left\{ss^T\right\}\tilde{A}^T = \tilde{A}\tilde{A}^T = I \tag{2-18}$$

經過以上證明可以發現，透過白化處理步驟能去除訊號與訊號之間的相關性，並使其變異數等於1 ，進而抑制雜訊造成的影響。

（二）第二階段：演算法

此階段主要分為兩大部分：

1. 以找出互為獨立之元素為目標，定義一個目標函數（Objective Function），用來量測隨機變數的獨立性。
2. 利用最佳化演算法（Optimization Algorithm），套用於目標函式中，以完成獨立成份之分離。

（1）目標函數（Objective Function）

根據中央極限定理（Central Limit Theorem），獨立成份的分佈最不為高斯分佈，因此要以非高斯分佈當做目標函數去進行獨立成份分析估計，就必須對非高斯分佈進行量化、公式化定義，進而產生一個非高斯的量測值來當做指標去進行獨立成份分析。在統計學上，通常以峰度（kurtosis）與負熵（neo-entropy）等兩項概念來做為非高斯程度量測使用。

◆ 峰度（kurtosis）:

峰度的定義在統計學上是隨機變數 y 的四次動差（moment）減去 3 倍隨機變數 y 的二次動差之平方。

$$kurt(y) = E\left\{y^4\right\} - 3\left(E\left\{y^2\right\}\right)^2 \tag{2-19}$$

如果 y 是由平均值為 0 ，變異度為 1 的隨機變數所組成，則 $E\{y^2\}= 1$，上式可以簡化為：

$$kurt(y) = E\{y^4\} - 3 \tag{2-20}$$

透過峰度可以知道該隨機變數 y 是何種分佈。峰度為零，y 為高斯分佈（gaussian distribution ）；峰度不為零，代表 y 為非高斯分佈（non-gaussian distribution）。峰度可為正值，其代表超高斯分佈（super-gaussian distribution）；峰度為負值時，其代表次高斯分佈（sub-gaussian distribution）。當峰度的絕對值越大，其非高斯的特性越強；峰度的絕對值越小，代表越接近高斯分佈。因此峰度的絕對值或平方可被用來作為目標函數，用來量測訊號的獨立性。

雖然使用峰度來量測訊號的非高斯特性相當方便且計算簡單，但峰度是屬於四次方的量測，在如此高次方的計算下，峰度會對於偏離值（outliers）過於敏感，也就是說當有小小的偏離值時，就會嚴重地改變峰度值，因此對於非高斯程度的量測來說，峰度並不是穩定且耐用的評估方法。

◆ 負熵（neo-entropy）:

另一種用來量測非高斯程度的方法為負熵（neo-entropy）。熵（entropy）在資訊理論中是代表資訊量的量測單位，用來描述隨機變數觀測值不確定性的多寡。對於一連續的隨機變數 y ，如果它的機率密度分佈函數為 $p(y)$，熵的定義為：

$$H(y) = -\int_{-\infty}^{+\infty} p(y) \ln p(y) dy \tag{2-21}$$

根據資訊理論，在所有具有相同共變異矩陣（covariance matrix）的分佈當中，高斯分佈具有最大的熵值，因此若以高斯分佈作為參考，則可以用熵來描述該連續變數 y 與高斯分佈之間的偏離程度，也就是非高斯特性。基於這樣的概念，我們可以將負熵 J 定義為：

$$J(y) = H(y_{gauss}) - H(y) \tag{2-22}$$

其中 y_{gauss} 是與 y 有相同共變異矩陣的高斯分佈隨機變數。當隨機變數 y 也是高斯分佈時，負熵才會為零，又因為高斯分佈的變數其熵最大，所以對所有隨機變數的負熵永遠不為負值。尋找負熵的最大值，即為尋找最非高斯分佈的隨機變數。然而，負熵的計算相當複雜，為了提高估計的方便性，通常會採用較簡化的近似方式，如以下公式：

$$J(y) \approx \left[E\{G(y)\} - E\{G(v)\}\right]^2 \tag{2-23}$$

上式中 G 為對照方程式，v 為平均值為零、變異度為 1 的高斯分佈變數。一般來說，對照方程式 G 可選擇如下：

$$G_1(y) = \frac{1}{a_1}\log\cosh a_1 y \tag{2-24}$$

$$G_2(y) = -\exp\left(-\frac{y^2}{2}\right) \tag{2-25}$$

上述式子，$1 \le a_1 \le 2$，對於對照方程式並無強硬的規定，只要該方程式不要是多項式與二次函數即可，若能挑選正確的對照方式，則可以求出更好的結果。

（2）最佳化演算法

當目標函數給定後，可以採用一般常用的最佳化演算法來最佳化目標函數，而最佳化演算法的選擇，將取決於演算法的收斂速度（convergence speed）、記憶體需求（memory requirements）等特性。由於考量到收斂速度的快慢與效能高低，本研究用於目標函數上的最佳化演算法將採用 FastICA 演算法來進行。

◆ FastICA 演算法：

FastICA 演算法是一種基於類神經網路演算法所推導而來，利用快速收斂的定點（fixed-point）演算方式，將大量的樣本資料進行批次處理(Hyvärinen & Oja, 2000)。

FastICA 演算法可選擇基於峰度或是基於負熵等等的目標函數，由於採用負熵來進行獨立性的判斷較為穩定，因此本研究將以負熵最大值作為目標函數的搜尋方向，並利用 FastICA 演算法來依序擷取獨立訊號源。

假設欲量測的 $y = w^T z$ ，公式 (2-23) 則可改寫成：

$$J\left(w^T z\right) \approx \left[E\left\{G\left(w^T z\right)\right\} - E\left\{G(v)\right\}\right]^2 \tag{2-26}$$

FastICA 演算法式根據定點疊代（fixed-point iteration）的方式來找出$J\,(w^T z)$ 的最大值。根據 Kuhan-Tucker 最佳化條件（Kuhan-Tucker condition），在 $E\{G(w^T z)^2\} = \|w\|^2 = 1$的限制下， $E\{G(w^T z)^2\}$ 需在滿足公式 (2-27) 的情況下才可獲得 w 的最大值。

$$E\left\{zG\left(w^T z\right)\right\} + \beta w = 0 \tag{2-27}$$

公式 (2-27) 中，β 是一個常數。接著利用定點演算法來解公式 (2-27)，以 $F(w)$ 代表上式等號左邊的函數，可得 $F(w)$ 的 Jacobian 矩陣 $JF(w)$ 如下：

$$JF(w) = E\left\{zz^T G'\left(w^T z\right)\right\} - \beta I \tag{2-28}$$

定點演算法就是在求每次該對 w 修正多少 Δw

$$\Delta w = \frac{F(w)}{JF(w)} \tag{2-29}$$

由上式可知，要求得 Δw 必須計算 $JF(w)$ 的逆矩陣，因此為了簡化求逆矩陣的過程，加上資料已經經過白化的前處理動作，則可將公式(2-28) 修改成公式(2-30)：

$$E\left\{zz^T G'\left(w^T z\right)\right\} \approx E\left\{zz^T\right\} E\left\{G'\left(w^T z\right)\right\} = E\left\{G'\left(w^T z\right)\right\} I \tag{2-30}$$

於是 Jacobian 矩陣變成了對角矩陣（diagonal matrix），並且比較容易求得逆矩陣。根據牛頓法，可以推出如以下的疊代公式：

$$w^+ = w + \Delta w = w - \frac{\left[E\left\{zG\left(w^T z\right)\right\} - \beta w\right]}{\left[E\left\{G'\left(w^T z\right)\right\} - \beta\right]}$$

$$w = \frac{w^+}{\| w^+ \|} \tag{2-31}$$

其中w^+ 代表新的 w 值。由於 β 是個未知的常數，將等式的兩邊乘上$\beta - E\{G'(w^T z)\}$ 則可簡化成：

$$w^+ = E\left\{zG\left(w^T z\right)\right\} - E\left\{zG'\left(w^T z\right)\right\} w \tag{2-32}$$

若是 w^+ 與 w 同方向即代表收斂，此時便可找到獨立成份。由於獨立成份不只一個，為了避免每次疊代會收斂至相同的 w 值，因此在找出第 n+1 個 w 時，需將前 n 個找出 w 的方向減去，以確保每次估測出來的獨立成份都不一樣，如下式：

$$w_{n+1} = w_n - \sum_{i=1}^{n+1} (w_{n+1}^T w_i) w_i \tag{2-33}$$

綜合以上所述，*FastICA* 演算法的整體步驟流程如下表3所示：

表 **3**
FastICA 演算法流程

Step 1	輸入混合訊號 x
Step 2	將混合訊號 x，透過centering過程，得到平均為零的訊號 x_c
Step 3	將訊號 x_c 作 whitening，得到 variance 為 1的訊號 z
Step 4	選擇需要估計的獨立成份個數m 設定疊代次數 $p=1$
Step 5	隨機選擇一個初始的 w_p 與適當的對照方程式 G
Step 6	令 $w_{p+1}=E\left\{zG\left(w_p^T z\right)\right\}-E\left\{zG'\left(w_p^T z\right)\right\}w_p$
Step 7	計算 $w_{p+1}=w_p-\sum_{i=1}^{p-1}(w_{n+1}^T w_i)w_i$
Step 8	令 $w_p=\frac{w_{p+1}}{\|\| w_{p+1} \|\|}$
Step 9	假如 w_p 不收斂的話，返回 Step 6
Step 10	令 $p=p+1$，如果 $p<m$，則返回 Step 5

獨立成份分析的技術主要被當作是盲訊號分離（Blind Signal Separation）的一種工具，大部份被應用於語音訊號處理上，例如環境下語音事件的偵測與減低聲音雜訊(Rebordao, Islam Molla, Hirose, & Minematsu, 2008)、雲端快速聲音訊號分離服務(Liang, Wang, Chou, & Chen, 2011)，以及結合頻率遮蔽（spectral masking）方法來提高吵雜環境下的語音辨識度(de Souza Siqueira Versiani, Rodrigues, de Souza, de Matos Moreira, & Yehia, 2012) 等等相關研究。除了語音訊號分析之外，獨立成份分析方法也被運用於其

他類型的訊號分析上，例如腦電波訊號的處理(Carvalhaes, Perreau-Guimaraes, Grosenick, & Suppes, 2009)、心肺聲音訊號分離與自動去除雜音(Ayari, Ksouri, & Alouani, 2013)，以及針對銀行文件上被油墨透印過的影像進行辨識(Chu & Chen, 2012)等等應用。在此特別要針對謝佩琳所提出的獨立成份分析法運用於萃取影片字幕上之研究(謝珮琳, 民94)提出討論。由於傳統在影片上的文字萃取，往往是利用廢時且沒效率的人工審查方式來進行處理。於該研究中，將影片片段視為是由一連串的影片畫面所組成，而每一張影片畫面是經由文字成份與非文字成份以線性關係所組成的混合訊號，利用獨立成份分析法能找出每張影片畫面中的特徵值，並計算出文字成份與非文字成份。於該研究中提出三種方式來進行影片畫面中的特徵分析，首先是針對連續影像的序列特徵表示法，透過輸入影片中具有相同字幕的連續畫面並分別轉為灰階影像作為輸入訊號，以利進行獨立成份分析 ; 其次為RGB影像特徵表示法，是以單張影片畫面中每個畫素所存在的r、g、b三個色頻分別轉為灰階影像作為多張影像的輸出，以利進行獨立成份分析 ; 第三種方式則為高次特徵表示法，該方法不僅將單張影像畫面中r、g、b三個色頻視為三張灰階影像，更考慮影像畫素與它周圍畫素之間的關係，以上、下、左、右的差異再製造出4張混合影像，並利用影像多次混合的法則，加以創造出二次、三次的混合訊號作為多張影像的輸出，以利進行獨立成份分析。實驗效果也發現，利用研究中所提出三種方法，在文字的萃取上有顯著的效果，特別是高次特徵表示方式，對於背景有更好的濾除效果。這不僅是對影片文字萃取提出一個完善且自動化處理的方式，也證實了獨立成份分析法能被應用的範圍相當廣泛，對於訊號分離的效果也相當令人期待。

第三節　文獻探討小結

在可觸式聲波操控介面的相關討論中，可以發現使用在可觸式聲波操控介面上的技術均各有其優點與缺點。反應速度較快的方法，如到達時間延遲法，往往需要嚴格控制環境因素，減低雜訊的影響，才能滿足其對精準度的要求。相反地，能提供更多聲源訊號資訊的方法，如位置模式比對法，卻具有運算量高、反應速度較慢等缺點。因此，在開發可觸式聲波操控介面前，必須先考量其裝置目的再謹慎選擇方法。此操控介面外，此兩種技術常被用來作為聲源的位置判斷使用，這是因為聲音訊號處理技術運用在定位上性能表現較為穩定，不易受到環境雜訊的影響。若要使用聲音訊號處理技術用在識別聲源類型時，最常用的方法是針對聲音的頻率來進行處理，然而環境雜訊影響因素會影響系統分析的穩定性，這樣的狀況也侷限了可觸式聲波操控介面的應用層面。

根據獨立成份分析法的演算法則，可以發現獨立成份分析法能藉由統計計算方式來將混合訊號中的獨立元素分析出來。有鑑於此方法在聲音訊號分離的顯著效果，我們可以對於「獨立成份分析法可運用在可觸式聲波操控介面裝置上，針對感測器所擷取之混合訊號進行聲源分離使用」這樣的假設保持樂觀看法。從上述獨立成份分析法之相關應用中也發現到，當利用獨立成份分析法在訊號時域上的分離效果無法突顯時，可進一步結合了頻域上的分析方式來進行測試。特別是當已知混合訊號有限而欲求得的成份數量未知時，可以透過頻域切割及高次混合等方式製造更多的輸入訊號來符合獨立成份分析法的處理程序。本研究將藉由這樣的概念，進一步運用於可觸式聲波操控介面中，分別針對多麥克風（多音源輸入）與單一麥克風（單音源輸入）所截取到的混合訊號來進行訊號分離的測試。當不同的聲源能被有效的從混合訊號中分離出來，不僅雜訊會藉此被過濾，可觸式聲波操控介面也能利用如位置模式比對法等技術做到同時間下多觸碰事件的辨識，可運用的互動領域也將因此拓寬。

第三章 研究方法

由於本研究目的在於利用獨立成份分析法針對可觸式聲波操控介面發展出系統化的訊號分析架構，因此本研究將利用可觸式聲波操控介面所實際錄製到的聲源訊號進行實驗與探討。於本章節中，將依序介紹實驗工具及本研究如何把獨立成份分析演算法運用於可觸式聲波操控介面的混合訊號分析上，嘗試萃取出因不同觸碰事件所產生的獨立聲源訊號。

第一節 實驗樣本與工具介紹

可觸式聲波操控介面最初的構想是希望將生活中常見事物，如書桌、餐桌、窗戶等任何物體表面都能轉變成可互動的空間，但由於在高度異向性（anisotopic）物質內，如大型木製桌，聲波不會以直線方向傳遞，且傳遞速度也不一致(Fabiani, 2006)。為減少環境影響因素並維持可觸式聲波操控介面的概念初衷，在本研究中選擇常用來建造傢俱的密迪板（Medium Density Fiberboard）作為可觸式聲波操控介面的作用物質。密迪板是由木削或是木材纖維混合膠合劑後熱壓製成，其密度及加工性質與一般的木材相似，板材的邊緣光滑易於實施型削處理，不會有粒片板或邊緣粗糙的現象(Fabiani, 2006)。此外由於密迪板是由非常小尺寸的微粒所組成，可被假設為等向性（isotropic）物質，其衰減係數非常高，可避免因邊緣反射造成的波擾動現象（disturbance）(Fabiani, 2006)。在密迪板內傳遞的波也以較慢的速度傳遞著，測量上較為容易(Christensen & Godsill, 2011)。

為了要測試獨立成份分析法在混合聲源訊號下的分離效果，於實驗中將分別採用鐵棒與木棒同時地在長120cm寬60cm厚5mm的密迪板表面上進行作用，鐵棒以刮擊的方式作用於介面上，木棒則以敲擊的方式作用於介面上，進而產生兩種不同聲源訊號（圖3-1）。密迪板表面上貼覆著接觸式麥克風來擷取表面震動聲音訊號，利用 MOTU

UltraLite-mk3 多軌道錄音介面依取樣頻率 44100 Hz、位元深度16bit ，配合Sound Forge Pro 軟體錄製成單(多)軌的 wav 格式音訊檔案，整體裝置示意圖請詳見圖3-2。

本實驗將採用MATLAB R2012a版本進行訊號分析的編程，並配合 FastICA Toolbox 進行獨立成份分析。

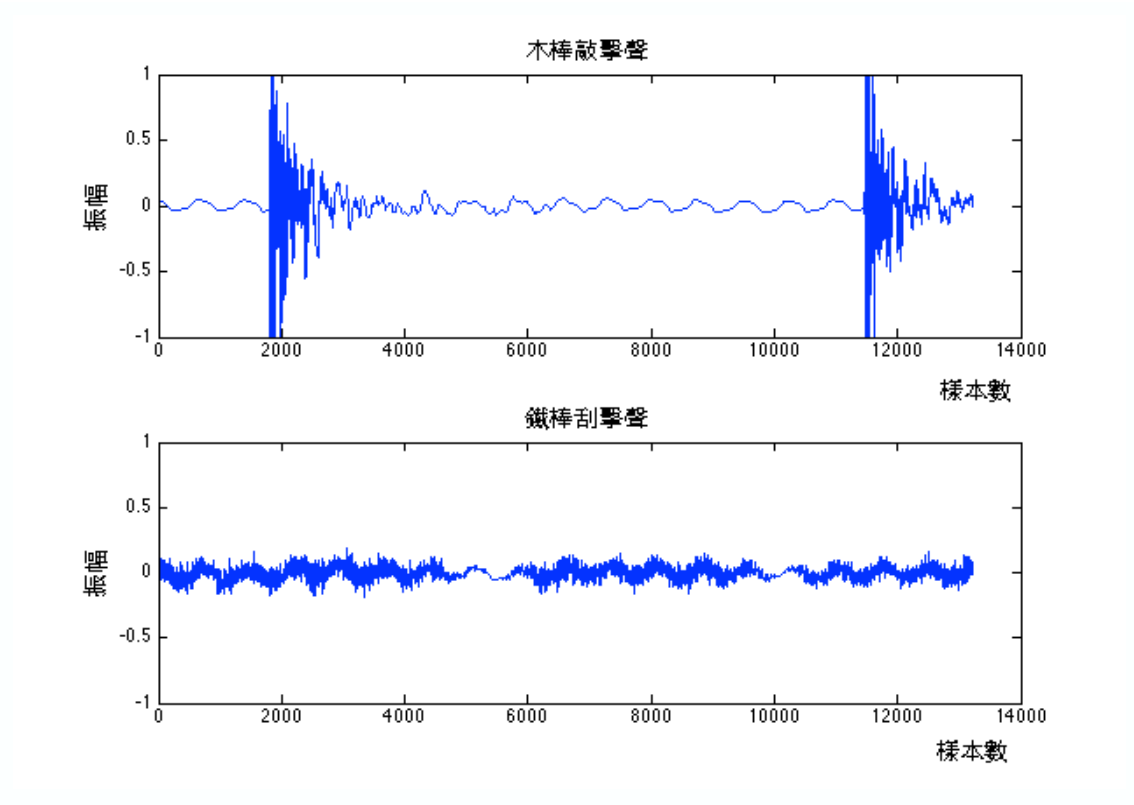

圖 **3-1** 木棒敲擊與鐵棒刮擊波型圖

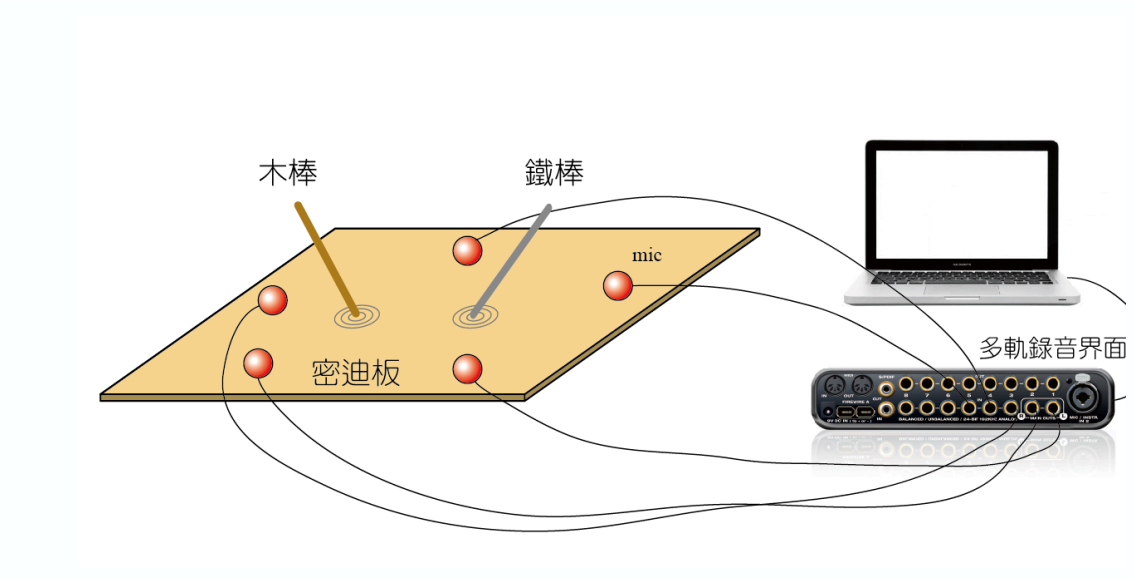

圖 **3-2** 本實驗所採用之可觸式聲波操控介面示意圖

第二節　符合獨立成份分析法的混合聲源假設

在本研究中，假設每個麥克風截取的聲音訊號就是一組混合訊號 X ，此訊號是由木棒敲擊聲源訊號 S_k 與鐵棒刮擊聲源訊號 S_S 經由線性關係混合而成。（見圖3-3）

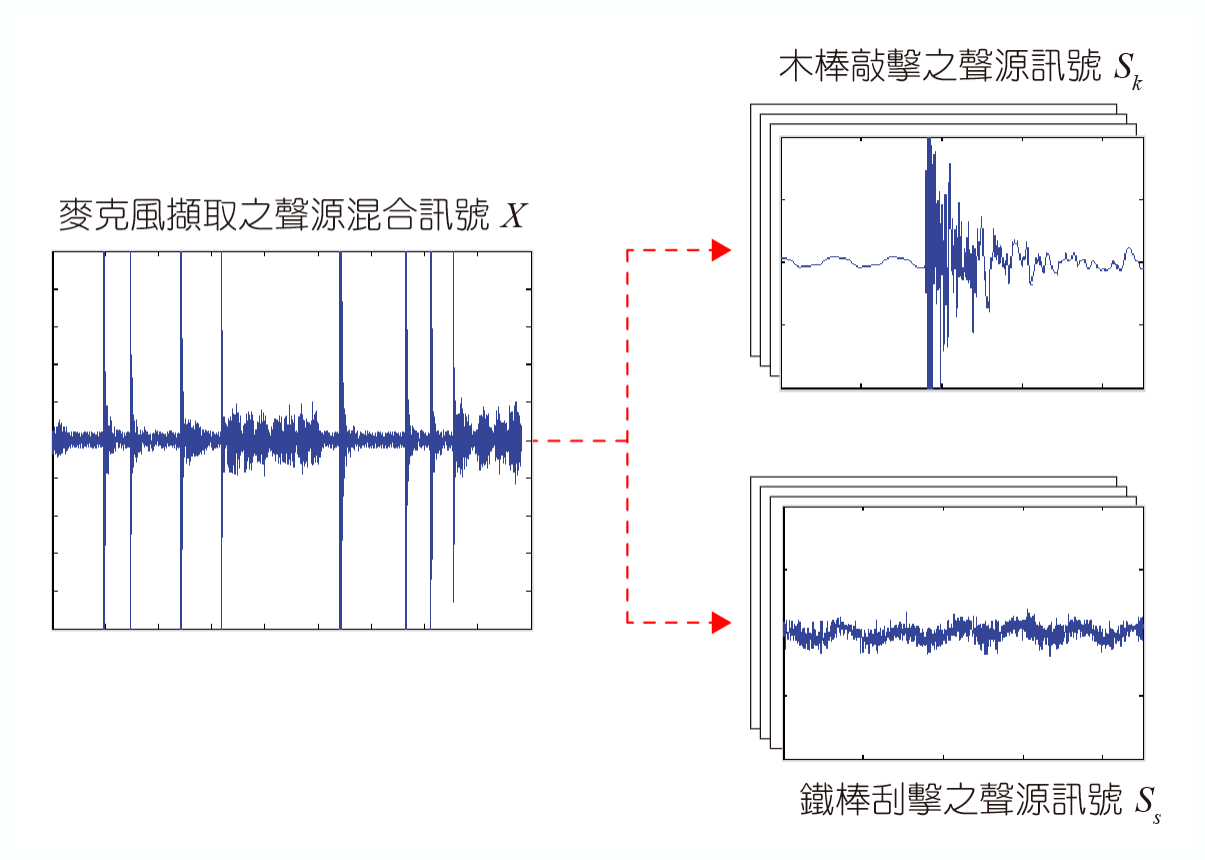

圖 **3-3** 基於獨立成份分析法之混合訊號組成成份示意圖

以上述的假設，可建構出於符合本研究之聲源訊號的獨立成份分析模型，如公式(3-1)。在此，木棒敲擊聲源訊號 S_k 與鐵棒刮擊聲源訊號 S_S 被視為是兩獨立特徵值，而麥克風所截取的混合訊號 X_i 是由木棒敲擊聲源訊號 S_k 與鐵棒刮擊聲源訊號 S_S 透過混合權重 a_{ik} 及 a_{is} 所混合而成。

$$X_i(t) = a_{ik}S_k + a_{is}S_s \tag{3-1}$$

若有m個麥克風同時間作為截取具有n項特徵值的混合訊號使用時，公式(3-1)可進一步擴大成為公式(3-2)：

$$\begin{bmatrix} X_1 \\ X_2 \\ X_3 \\ \vdots \\ X_m \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} S_1 \\ S_2 \\ \vdots \\ S_n \end{bmatrix} \tag{3-2}$$

從文獻當中可以得知，獨立成份分析法主要是在只有混合訊號 X 已知的情況下，去找一個解混合矩陣 W，經過轉換後可以得到原本獨立的訊號。換句話說，只要能把混合訊號依據特徵值表示如公式(3-2)的形式，便可透過獨立成份分析處理程序計算出藏在混合聲源訊號中的獨立聲源成份。

我們先依據公式(3-2)概念，利用一段演唱會live片段音樂進行獨立成分分析測試。此片段音樂為取樣頻率44100Hz的wav格式音樂，長度約10秒鐘，內容包含鼓聲與吉他演奏聲。將此段音樂依左右聲道拆成兩段音軌分別來聽，發現兩音軌聽得出均含有共同的聲音組成，但由於組成的比例不同混雜在一起，因此聽覺上有明顯的差別。此外，將兩段音軌繪製波形圖來看（如圖3-4），兩段波形雖在振幅上有所差異，但整體波形走向大致相同，因此可以將此兩段音樂片段視為是符合獨立成份分析法假設下的混合聲源訊號。

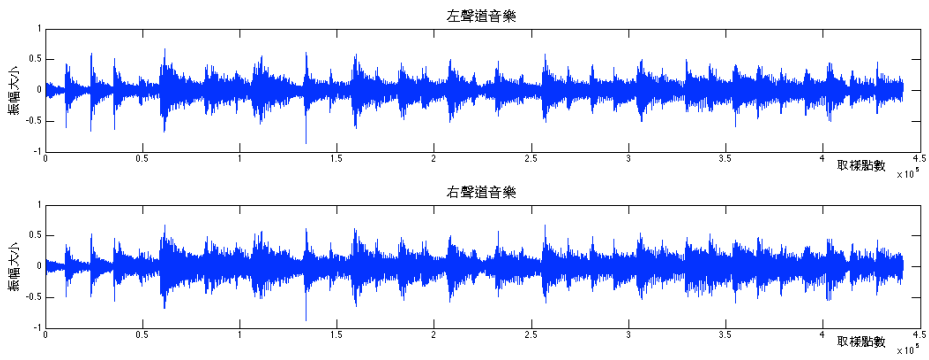

圖 **3-4** 演唱會**live**音樂片段左右聲道波形圖

透過獨立成份分析法進行分析後，分離出兩段聲音訊號，其波形圖如圖3-5所示。從圖3-4與圖3-5中可以發現，獨立成份分析法確實對混合的聲音訊號產生了作用，使得原本看似相同的聲音分離過後產生波形上的差異。除了波形上的不同，分離過後的訊號在聽覺上也加大了差異。在訊號2中，吉他的聲音被凸顯了出來，而鼓聲與其它背景音樂聲則變為混雜，且音量降低許多；然而訊號1與原始兩音軌的聽覺效果較為相似，沒有太多變化。

轉為頻率角度來觀察，原始兩音軌的頻率分佈如圖3-6所示，而透過獨立成份分析法處理過後的分離訊號之頻率分佈則見圖3-7。兩組訊號在頻率的分佈上並沒有改變太多，表示聲源訊號大致上都存在，只是所佔比例多寡有所差異，與聽覺上所顯示出來的效果大致相同。

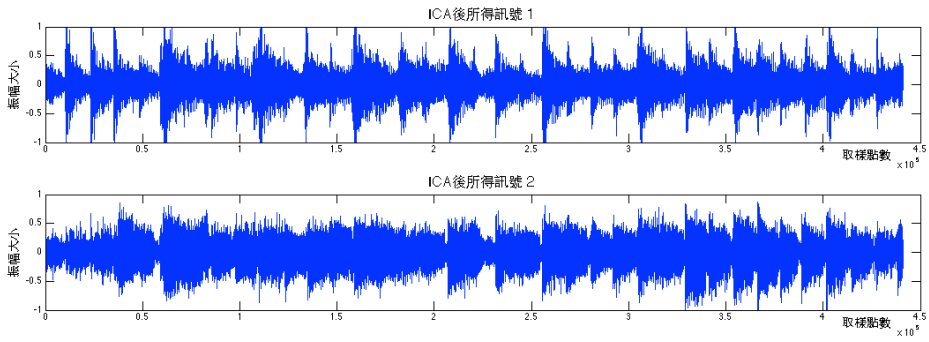

圖 **3-5** 演唱會**live**音樂片段透過**ICA**處理後所得之分離訊號波形圖

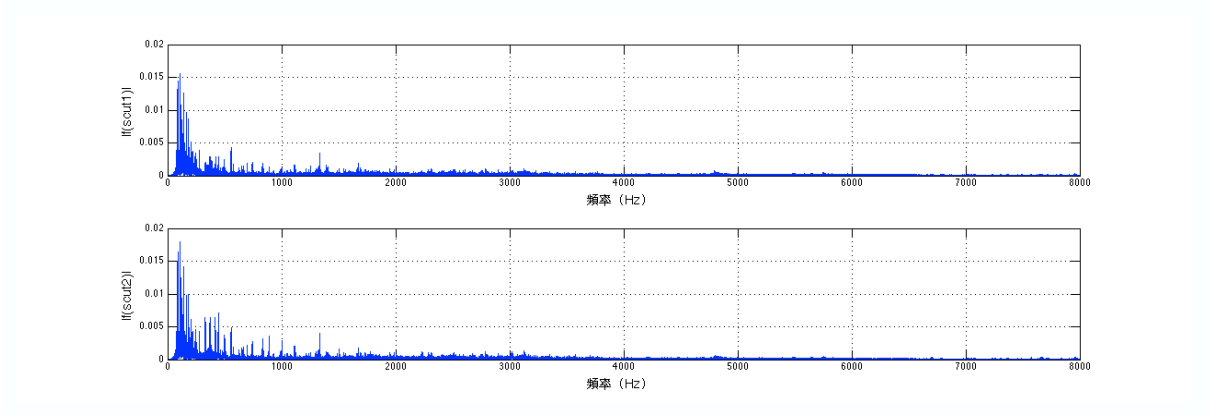

圖 **3-6** 演唱會**live**音樂片段聲音訊號頻率分佈圖

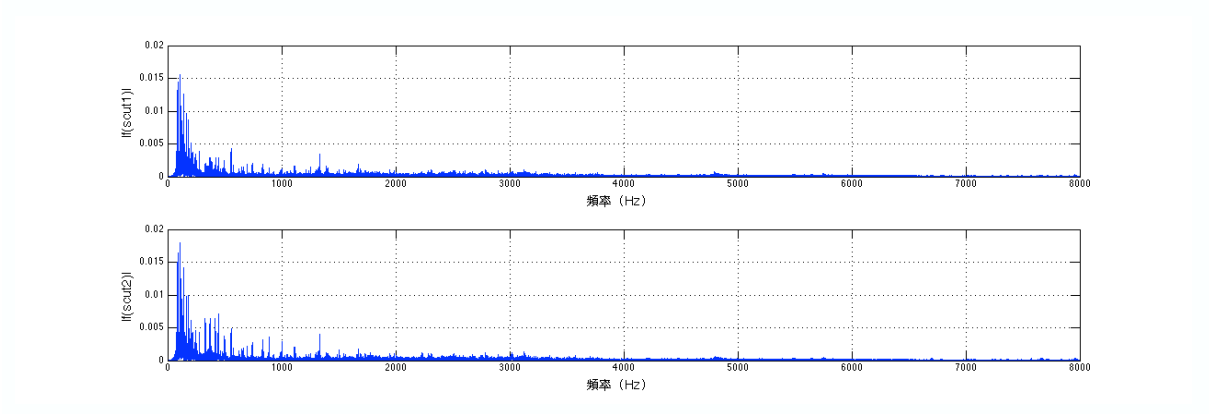

圖 **3-7** 演唱會**live**音樂片段透過**ICA**處理後所得之分離訊號頻率分佈圖

我們再依據公式(3-2)概念，嘗試以模擬混合的方式來驗證獨立成份分析法是否能依據此架構來達到獨立成份分離的效果。首先先針對木棒敲擊和與鐵棒刮擊的聲音訊號分別單獨利用麥克風實際錄製（如圖3-8），並給予兩段訊號不同的權值後以MATLAB進行人工模擬混合，再透過 *FastICA* 演算法進行訊號分離的步驟。

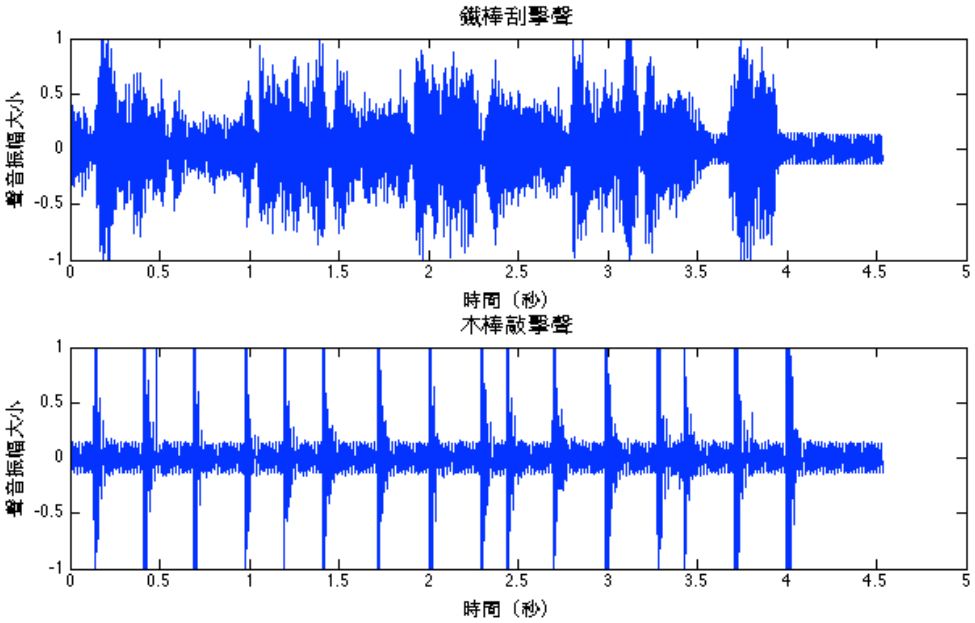

圖 **3-8** 木棒敲擊和與鐵棒刮擊原始訊號波形圖

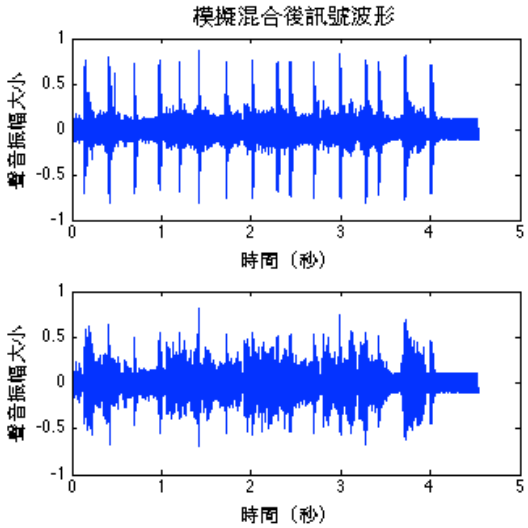

圖 **3-9** 混合後訊號波形圖

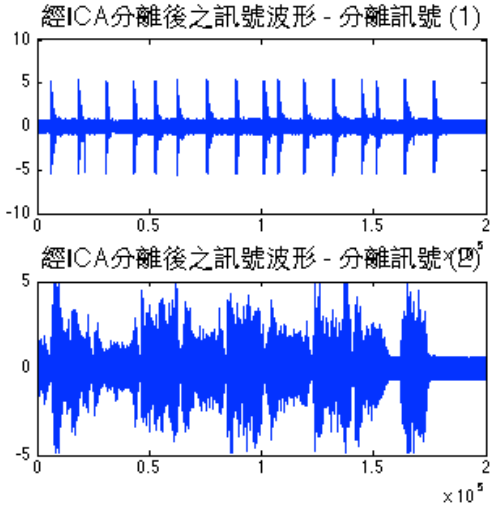

圖 **3-10 FastICA**分離後訊號波形圖

透過人工混合出的兩段混合訊號在波形上雖能看出有包含兩原訊號的輪廓（如圖3-9所示），但聽覺效果上卻已與原訊號大不相同。經過*FastICA*進行分析後，得出兩個獨立元素，其波形與原信號之波形較為相似，如圖3-10所示，藉此可推斷，不同權重值的混合，並不會對於獨立元素分析法在分離訊號上造成太大影響。

將原始訊號、混合訊號以及分離出來的訊號均轉換成頻域的角度來觀察，可以發現訊號的混合改變了原始訊號的頻率分佈，使得兩段混合訊號的頻率分佈趨近相似。然而透過獨立成份分析之後所得的分離訊號，其頻率分布則轉為近似於原始訊號所呈現出的頻率分佈，如圖3-11至圖3-13所示。由上述現象可推論，在符合獨立成份分析法的混合聲源假設情況下，利用獨立成份分析法在混合訊號上進行分析，由時域及頻域分佈應可觀察出獨立成份的分離效果。因此，於下一小節內，本研究將提出三種依據獨立成份分析法原則的操作程序，直接針對可觸式聲波操控介面截取到的混合聲音片段進行分析處理。

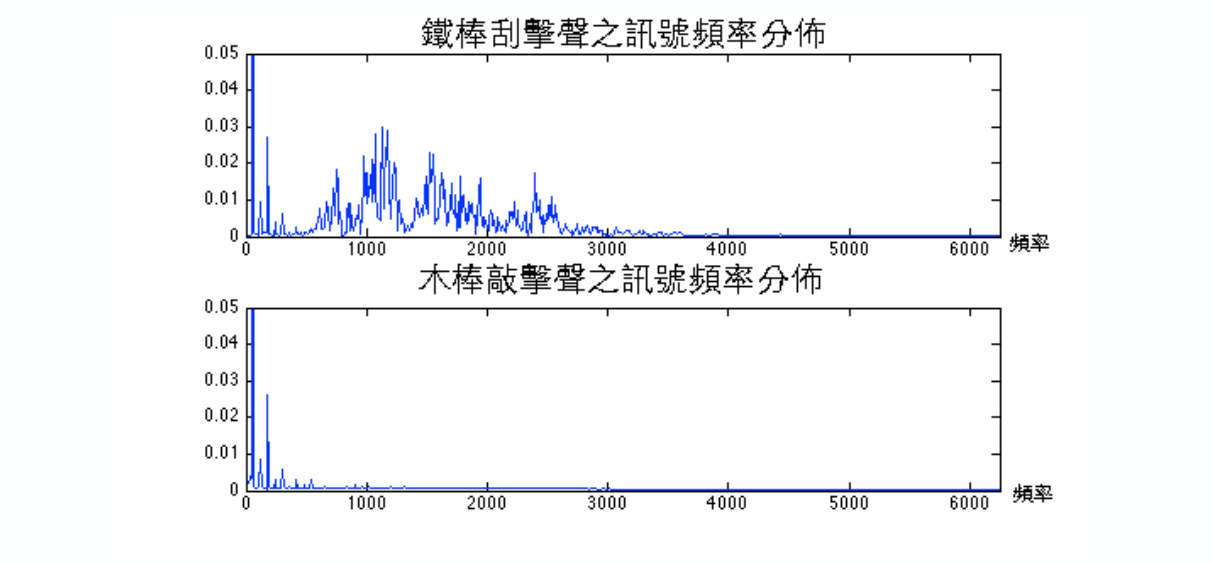

圖 **3-11** 木棒敲擊和與鐵棒刮擊原始訊號頻率分布圖

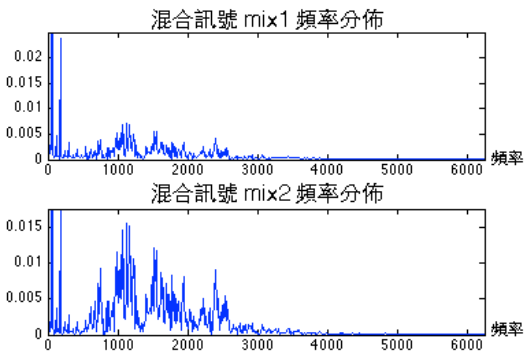

圖 **3-12** 混合訊號頻率分布圖

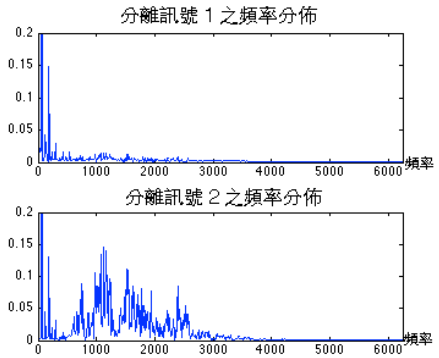

圖 **3-13** 分離後訊號頻率分布圖

第三節　依據獨立成份分析法原則的獨立聲源分析程序

在本小節中將介紹本研究所採用的三種依據獨立成份分析法原則的獨立聲源分析操作程序來針對可觸式聲波操控介面截取到的混合聲音片段進行分析處理，分別為(一)　多音源輸入訊號直接進入ICA處理。(二)　單音源輸入訊號經頻率區隔分析後進行ICA處理。(三)　單音源輸入訊號經頻率區隔及高次混合後進入ICA處理。操作流程如圖3-14所示，並於下內容進行細部介紹。

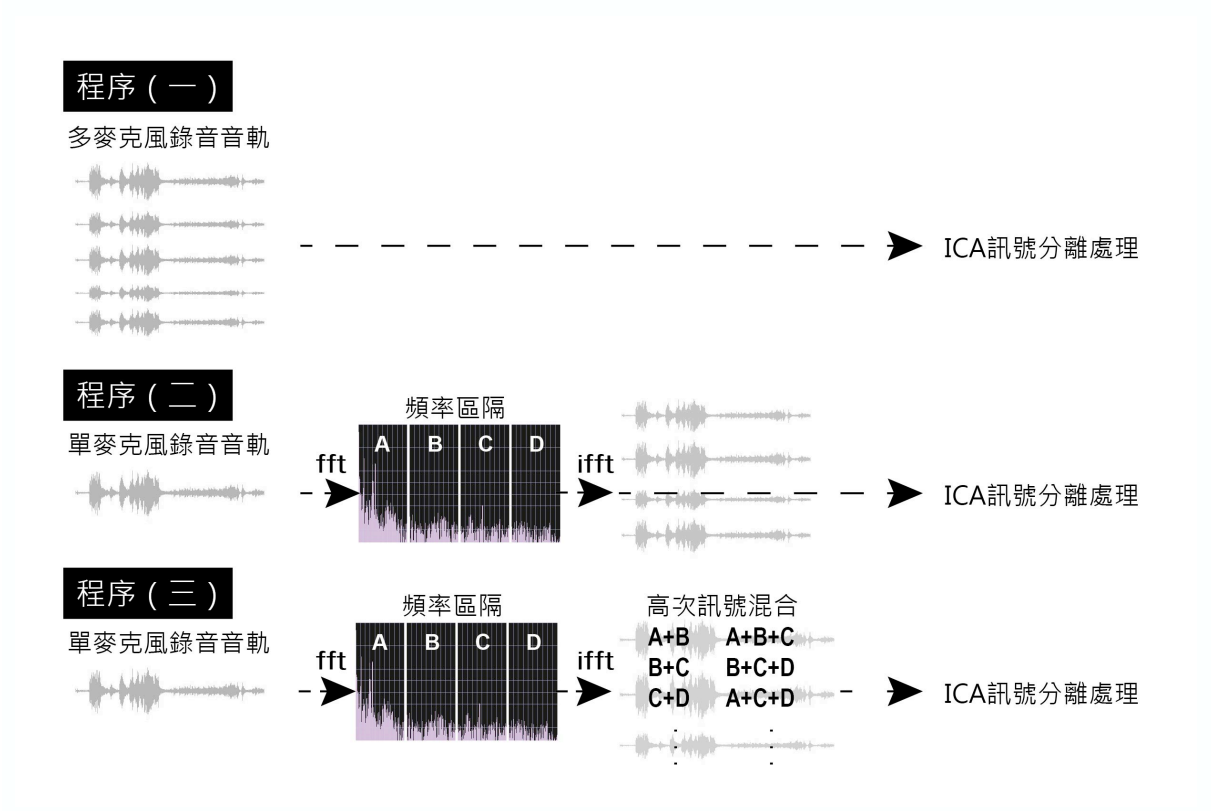

圖 **3-14** 本研究所採用之獨立聲源分析三操作程序流程說明圖

程序（一）： 多音源輸入訊號直接進入ICA處理

由於木棒敲擊聲音與鐵棒刮擊聲音，在聽覺上、波型顯示上及頻率分佈上均有明顯的不同（見圖3-15，3-16），因此依據公式(3-1)的架構，將一段混合訊號中所包含的所有木棒敲擊聲音視為一獨立訊號源，而所有鐵棒刮擊聲音則視為另一獨立訊號源，兩獨立訊號源因作用點距離麥克風的遠近而有不同混合係數來對應其組成的比例。

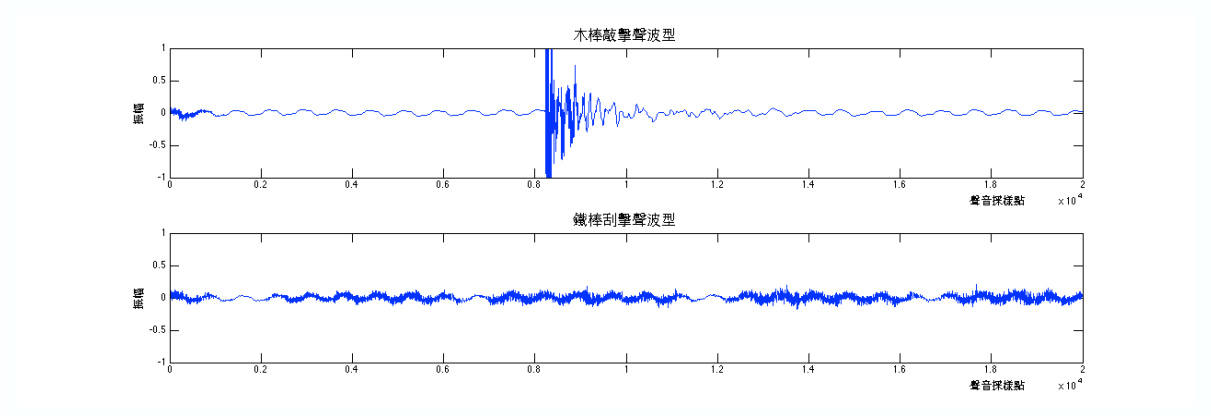

圖 **3-15** 木棒敲擊與鐵棒刮擊波型圖

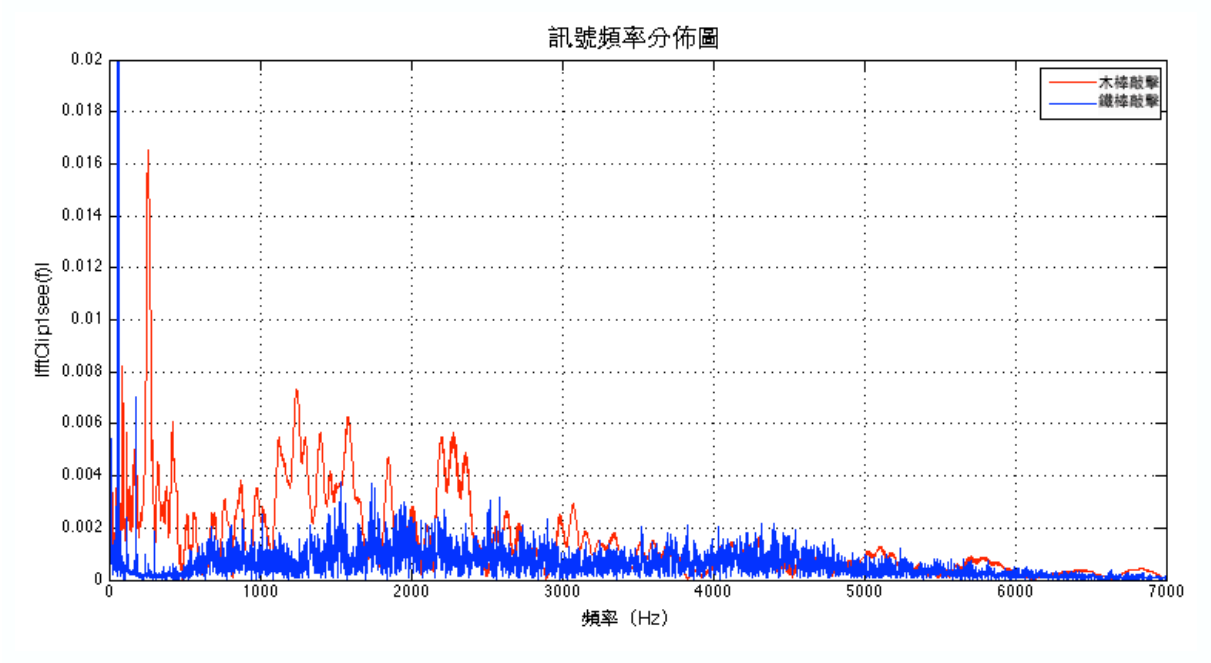

圖 **3-16** 木棒敲擊與鐵棒刮擊訊號頻率分布圖

我們將密迪板的表面任意設置5個接觸式麥克風，利用多軌道錄音界面同時間進行錄製木棒敲擊與鐵棒刮擊的聲音，得到混合訊號 X_{cont} ，如公式(3-3)，再將 X_{cont} 透過 *FastICA* 演算法進行獨立成份分析測試（見圖*3*-17 說明）。

$$X_{cont} = \begin{bmatrix} x_{mic1} \\ x_{mic2} \\ x_{mic3} \\ x_{mic4} \\ x_{mic5} \end{bmatrix} = \begin{bmatrix} a_{mic1k} & a_{mic1s} \\ a_{mic2k} & a_{mic2s} \\ a_{mic3k} & a_{mic3s} \\ a_{mic4k} & a_{mic4s} \\ a_{mic5k} & a_{mic5s} \end{bmatrix} \begin{bmatrix} S_{knock} \\ S_{saratch} \end{bmatrix} \tag{3-3}$$

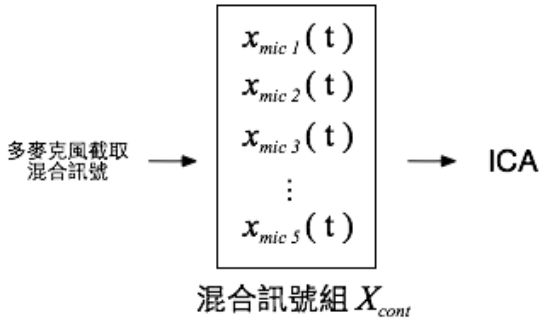

圖 **3-17** 程序（一）之混合訊號組成示意圖

程序（二）： 單音源輸入訊號經頻率區隔分析後進行ICA處理

訊號分析方式大致有時間域與頻率域之兩種分析方式。在時間域之分析，最常用的是時間函數（time function），其表示式為 $x(t)$。將訊號的時間函數經過傅立葉轉換（Fourier Transform）後可得頻率域之函數表示式 $x(f)$；相對的，將訊號的頻率函數 $x(f)$ 透過反傅立葉轉換（Inverse Fourier Transform）後也可轉換回時間域下的時間函數 $x(t)$。

此操作程序（二）中所採用的方法是利用單一麥克風所錄下的聲音片段先透過頻率區隔處理，再來進行獨立成份分析。由於獨立成份分析法有一非常重要的原則，就是混合的成分其數目不得小於欲求得的獨立成份數量。因此，針對單一麥克風所錄製得來的聲音片段，無法直接利用獨立成份分析法來處理，必須先依據該混合訊號本身的特性來製作更多訊號輸出，才能進行分析。為解決這個問題，先將截取來的時間域下混合訊號轉換成頻率域來觀察，了解該段訊號在頻譜上主要分布在哪些頻率範圍

內，並分別切割出該段頻率下的頻域訊號，透過反傳立葉轉換公式輸出成時域訊號，藉此可製造出一組混合訊號組 X_{sect} 以利*FastICA*演算法執行獨立成份分析測試（如圖3-18）。

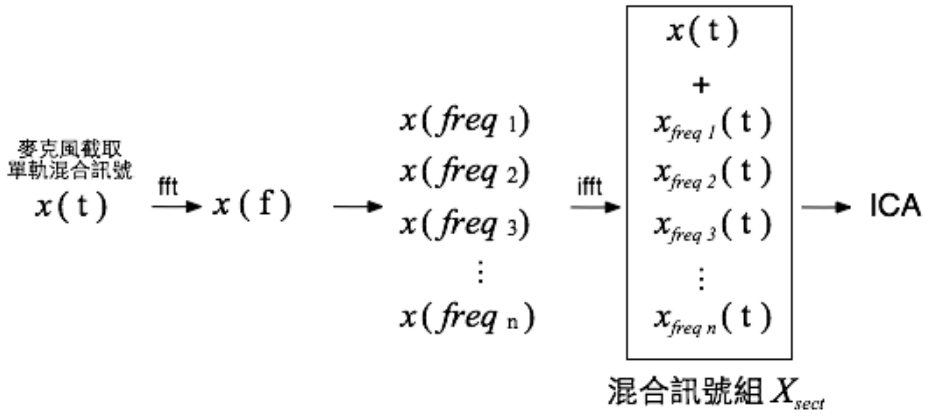

圖 **3-18** 程序（二）之混合訊號組成示意圖

程序（三）： 單音源輸入訊號經頻率區隔及高次混合後進入ICA處理

同操作程序（二）所採用的步驟，在程序（三）中也是採用單一麥克風所錄下的聲音片段，並儘可能地從該聲音片段中製造出更多聲音訊號來提供獨立成份分析法進行分析處理。於此操作程序中，除了頻率區隔所製造出來的 $X_{freq(1\sim n)}$ 等聲音訊號之外，更進階的考慮聲音訊號的多次混合法則。公式(3-4) 與 公式(3-5) 分別表示2次與3次的聲音訊號。將得到高次的混合聲音訊號利用公式(3-6)正規化，得到混合訊號組 X_{high} ，如公式(3-7)所示，即可透過*FastICA*演算法進行 X_{high} 的獨立成份分析。

$$x^{(2)}_{(a,b,c)} = x_a * x_b$$

$$a \ , \ b \in \{x(t), x_{freq1}(t), x_{freq2}(t), \cdots x_{freqn}(t)\} \quad and \ a \neq b \tag{3-4}$$

$$x^{(3)}_{(a,b,c)} = x_a * x_b * x_c$$

$$a \ , \ b \ , \ c \in \{x(t), x_{freq1}(t), x_{freq2}(t), \cdots x_{freqn}(t)\} \quad and \ a \neq b \neq c \tag{3-5}$$

$$x_{new} = \frac{(x-\mu)}{3\sigma} \tag{3-6}$$

$$x_{high} = \begin{bmatrix} x_{rec.freq1} \\ x_{rec.freq2} \\ x_{rec.freq3} \\ \vdots \\ x_{freq(n-1).freq(n)} \\ x_{rec.freq1.freq2} \\ x_{rec.freq1.freq3} \\ x_{rec.freq1.freq4} \\ \vdots \\ x_{freq(n-2).freq(n-1).freq(n)} \end{bmatrix} \tag{3-7}$$

第四章　實驗結果與討論

在上述內容提到，本研究針對獨立成份分析法概念提出三種利用在可觸式聲波操控介面上的混合訊號分析操作程序，實驗樣本為運用可觸式聲波操控介面時，同時由木棒敲擊表面以及鐵棒刮擊表面而產生的混合聲音，經由貼覆在作用表面的接觸式麥克風錄製得來。採取的三種操作程序分別為 (一) 多音源輸入訊號直接進入ICA處理。(二) 單音源輸入訊號經頻率區隔分析後進行ICA處理。(三) 單音源輸入訊號經頻率區隔及高次混合後進入ICA處理。其分析目的主要以求得可區別的獨立元素訊號為主，因此在效果評估上，將針對所求得的訊號波型圖與原始混合訊號的波型圖是否俱有明顯差異作為主要區分原則。以下將針對三種分析程序所得到的實驗結果分別進行探討。

第一節　多音源輸入訊號與ICA處理

於程序（一）多音源輸入訊號直接進入ICA處理分析操作程序中，密迪板的表面被任意設置5個接觸式單聲道麥克風以截取並錄製木棒與鐵棒同時作用於板面的聲音，透過MATLAB分別讀取五段聲音訊號資料，得到的混合訊號聲波圖，如圖4-1所示。仔細觀察原始波形圖(見圖4-2)，可以看出木棒敲擊表面所產生的聲音是短暫具有較高振幅的波型，而鐵棒刮擊表面所產生的聲音是連續、段時間內震動次數多且振幅較低的波型。而當鐵棒與木棒同時作用時，其產生的波型則同時具有上述兩種特性，雖然可大致從振幅大小來推斷出哪些時間區段下有發生動作，但卻無法明確地指出兩動作各自發生或結束的時間點為何。

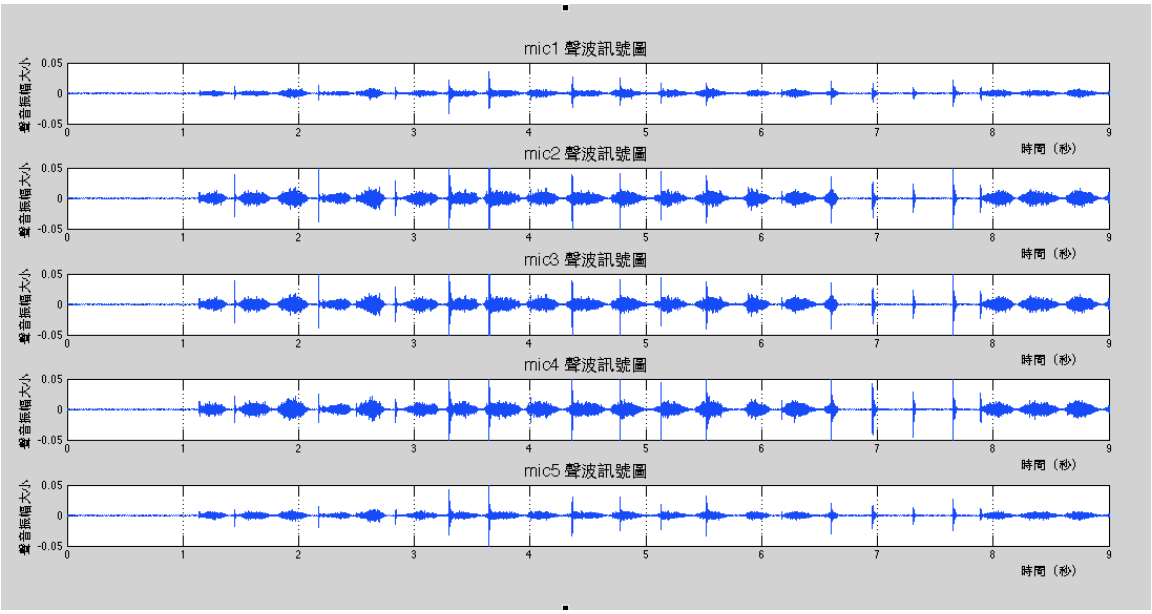

圖 **4-1** 程序（一）：五個麥克風所截取的混合訊號聲波圖

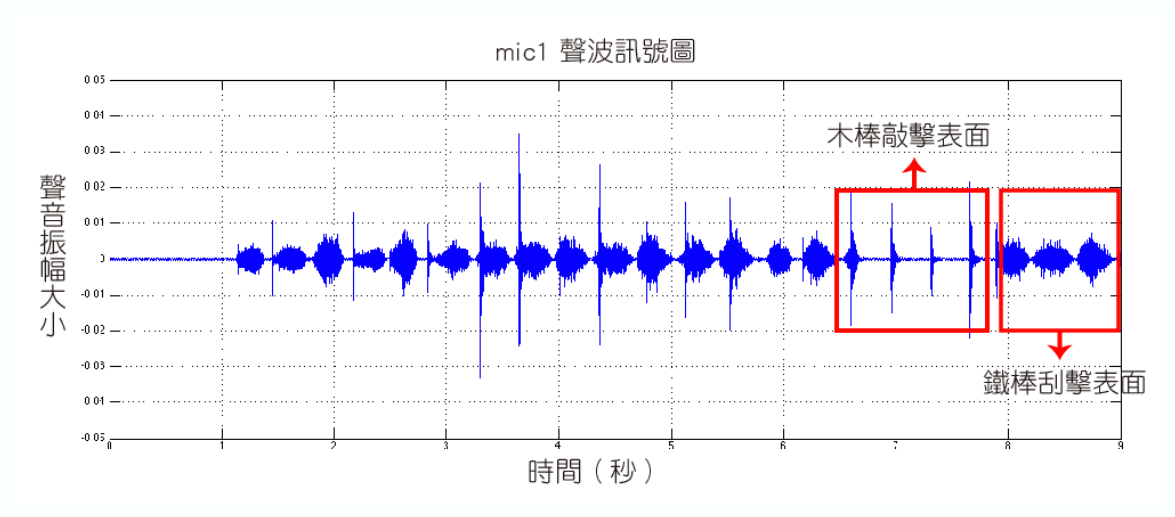

圖 **4-2** 程序（一）：混合訊號原始波形圖

透過FastICA演算法進行分析，嘗試輸出2~5個分析後訊號以進一步觀察。所得結果發現，無論透過該方法去嘗試輸出幾個獨立訊號，雖有一些區段的振幅大小略為變動，但所得的訊號波型與原始訊號均大同小異(見圖4-3至圖4-6)，並無法辨識有出明顯的訊號分離現象發生。

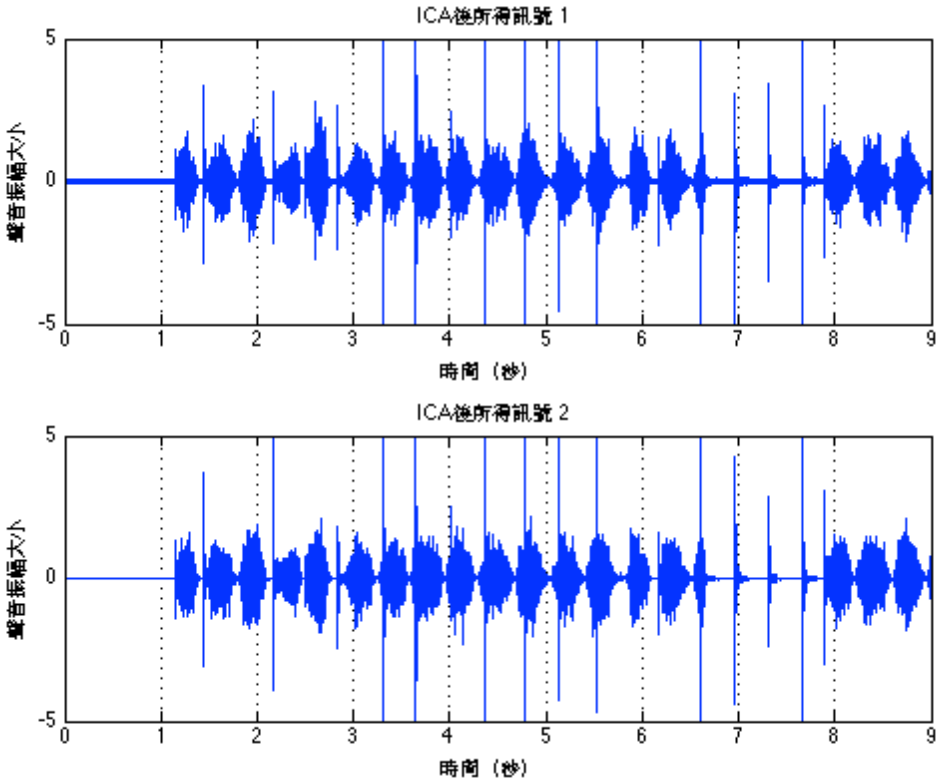

圖 **4-3** 程序（一）： **ICA**後所得訊號波形圖**-**輸出 **2** 個訊號

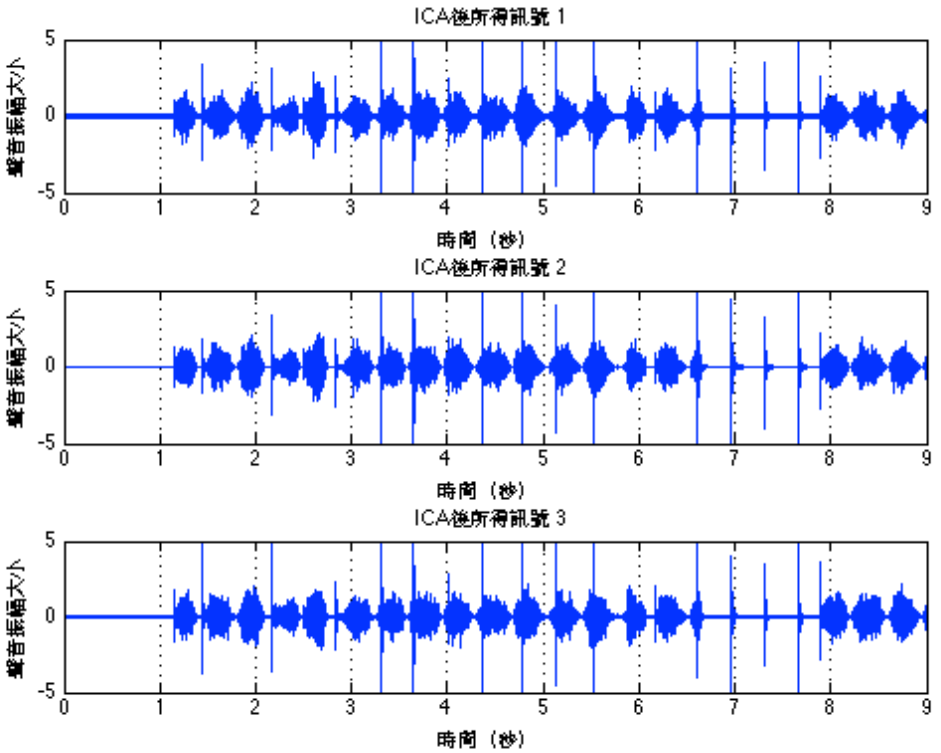

圖 **4-4** 程序（一）： **ICA**後所得訊號波形圖**-**輸出 **3** 個訊號

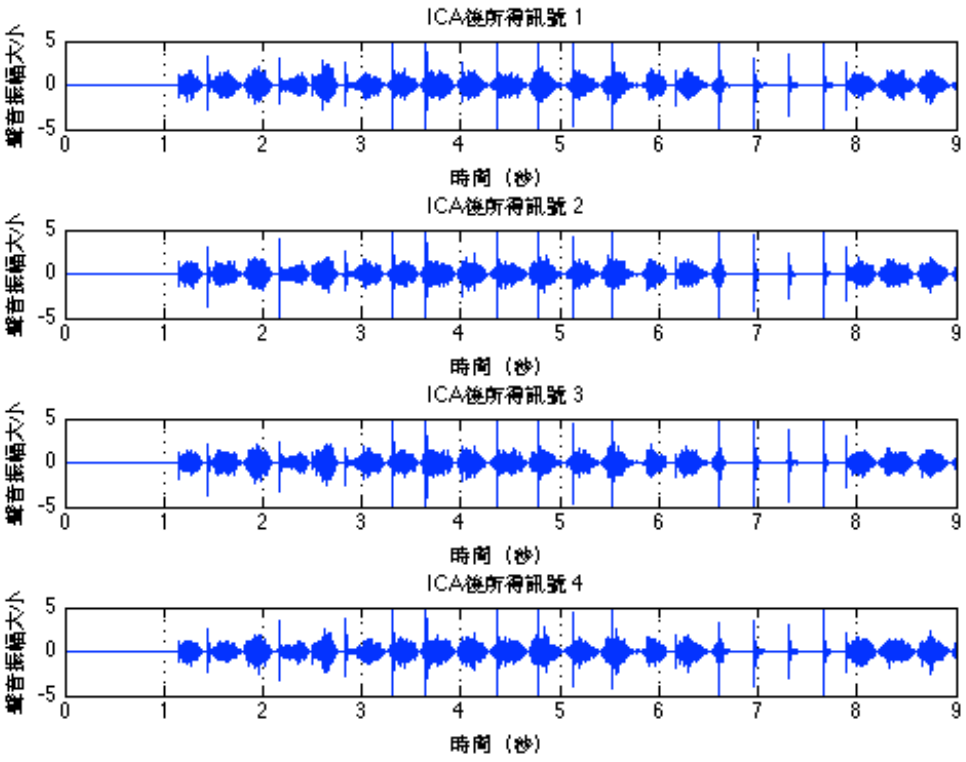

圖 **4-5** 程序（一）： **ICA**後所得訊號波形圖**-**輸出 **4** 個訊號

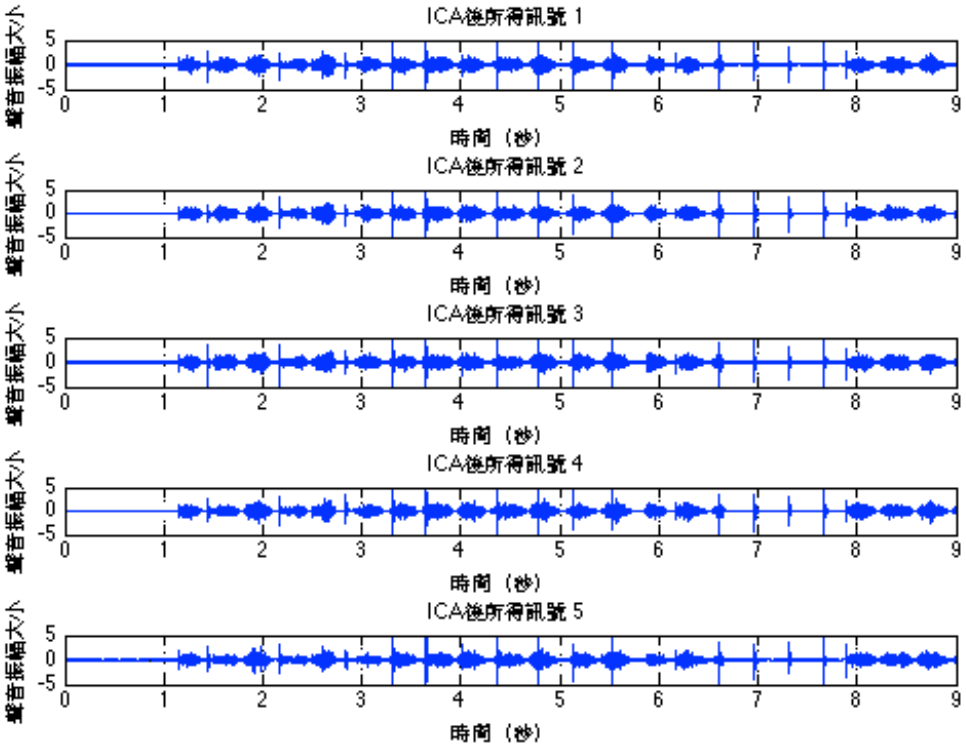

圖 **4-6** 程序（一）： **ICA**後所得訊號波形圖**-**輸出 **5** 個訊號

第二節　單音源訊號經頻率區隔與ICA處理

於程序（二）單音源輸入訊號經頻率區隔分析後再利用ICA進行聲源訊號萃取，密迪板的表面僅被設置1個接觸式單聲道麥克風以截取並錄製木棒與鐵棒同時作用於板面的聲音，所收錄到的混合訊號的波形如圖4-7所示，由振幅的大小與波出現的頻率仍然可以大致看出木棒敲擊所造成的訊號以及鐵棒刮擊所造成的訊號在該混合訊號中之發生的時間區段為何。將該訊號轉換為頻域角度來觀察，頻率分佈顯示，此段混合訊號主要分佈在0 Hz ~ 8000 Hz之間，且分別在頻率區段 0 ~ 1000Hz、1000Hz ~ 2000Hz、2000Hz ~ 3000Hz、及 3000Hz ~ 5000Hz之間有較多的訊號分布。因此，於此階段時間中將針對以上四頻域區段進行頻域時域訊號轉換，進而產生新訊號(公式4-1)。

$$X_{\sec t} = \begin{bmatrix} x_{freq0-1000} \\ x_{freq1000-2000} \\ x_{freq2000-3000} \\ x_{freq3000-5000} \\ x_{rec} \end{bmatrix} \tag{4-1}$$

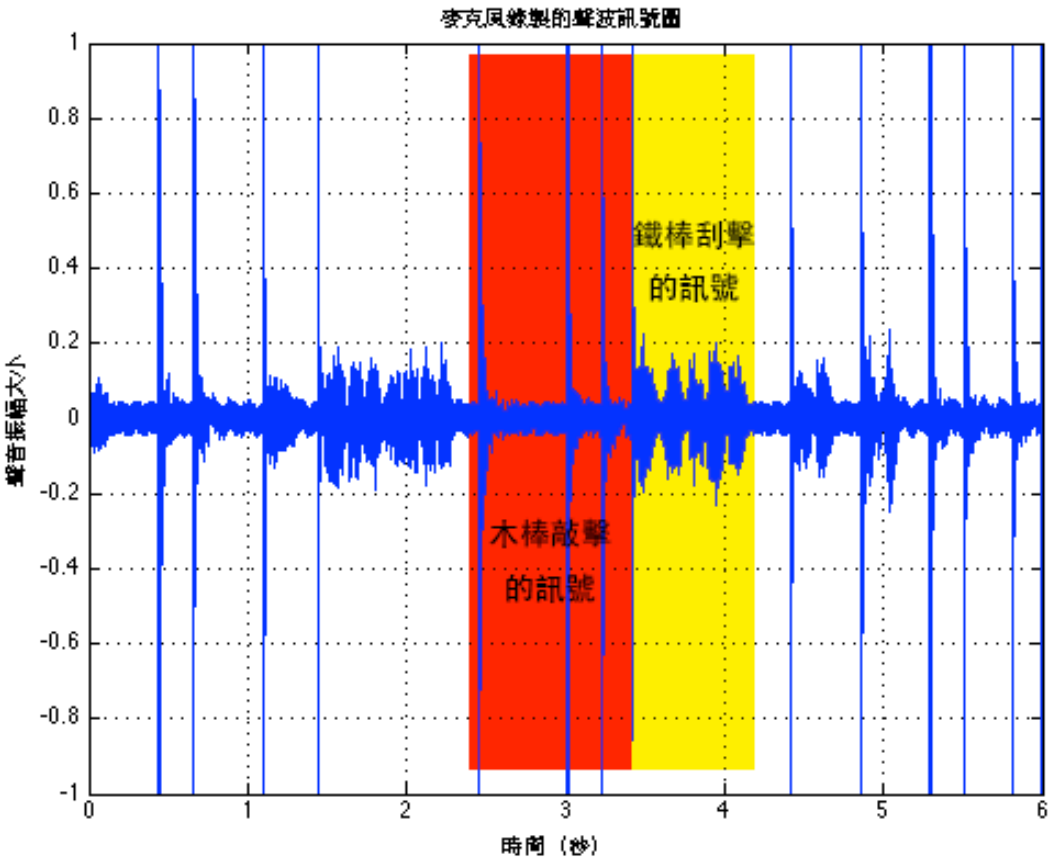

圖 **4-7** 程序（二）：單一麥克風所錄製的聲波訊號圖

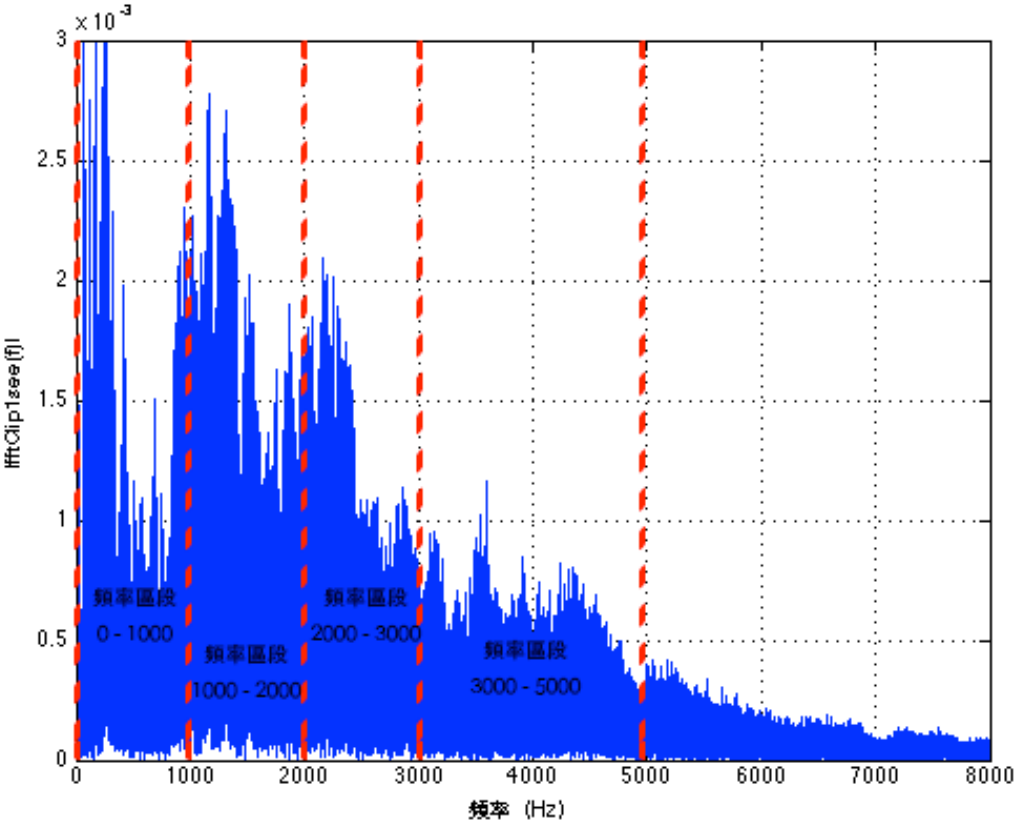

圖 **4-8** 程序（二）： 單一麥克風所錄製的聲波頻率分佈圖

將X_{sect} 繪製出波型來觀察（見圖4-8），可以發現從頻率區段來切割混合訊號，同時也具有過濾訊號的效果，除了較高頻的訊號與較低頻的訊號可被藉此分開，一些環境雜訊也透過該方法被過濾掉，因此從波型圖上已大致上能看出木棒敲擊所造成的訊號聲源的波型從混合訊號中被凸顯出來。接著透過FastICA進行獨立成份分析處理時，也試著嘗試分別輸出2~5個分析後訊號進行觀察（見圖4-9至圖4-13）。由分析結果波形圖中可以發現木棒敲擊表面所造成的訊號聲已經很明顯地可以由混合訊號中被抽離出來，透過ICA輸出的訊號越多，能藉此分離出來的訊號細節就更加的詳細。

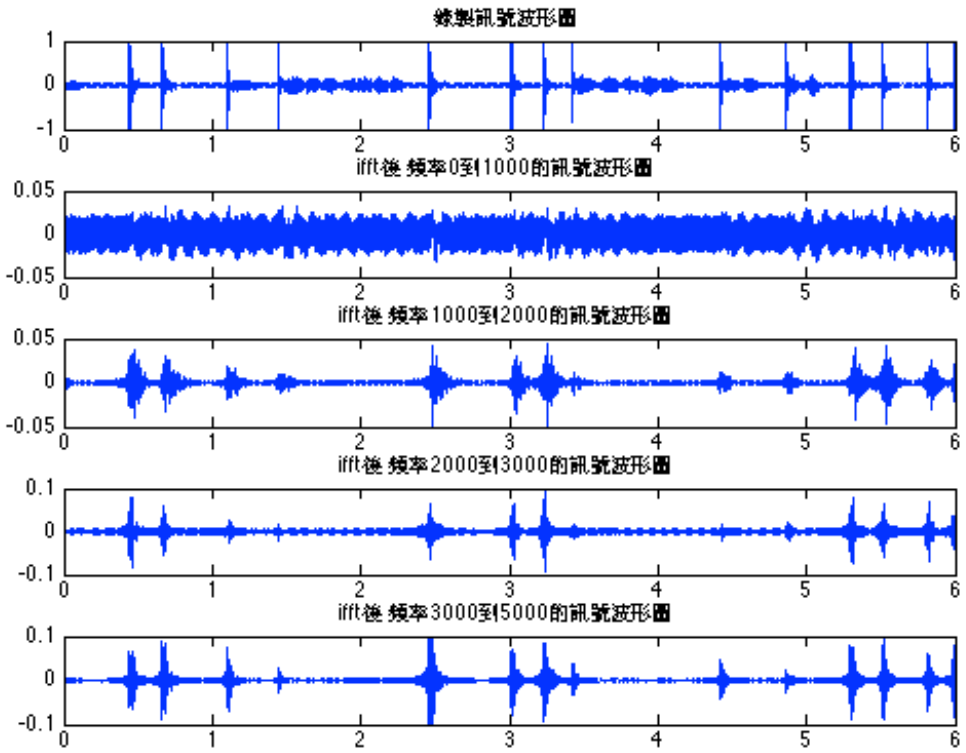

圖 **4-9** 程序（二）： 利用頻率區隔所製作出的所有訊號波型圖

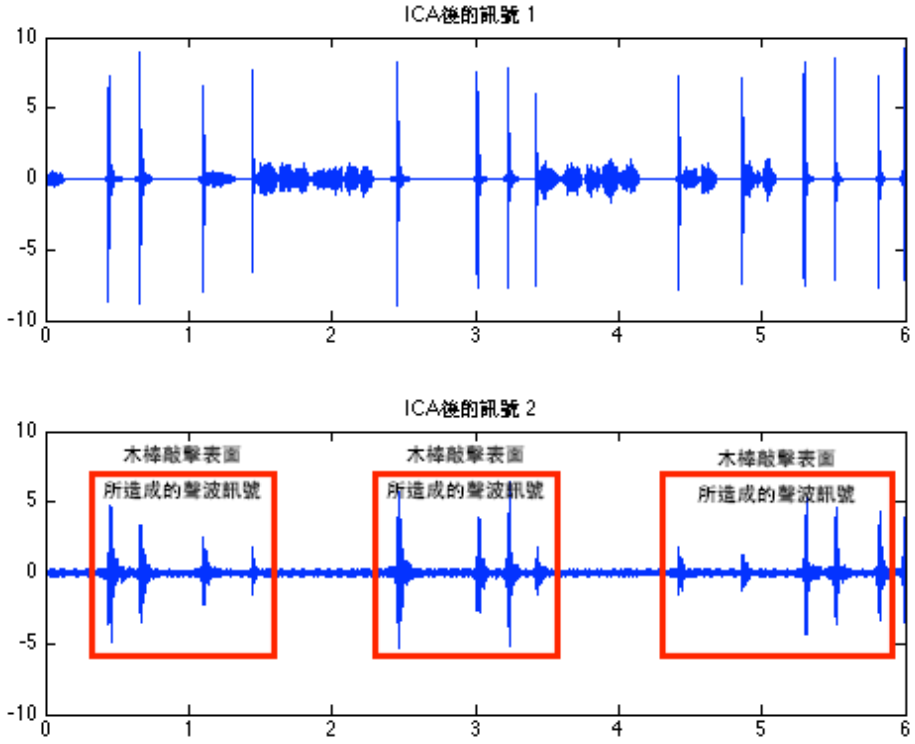

圖 **4-10** 程序（二）： **ICA**所得之訊號波型圖 **-** 輸出 **2** 個訊號

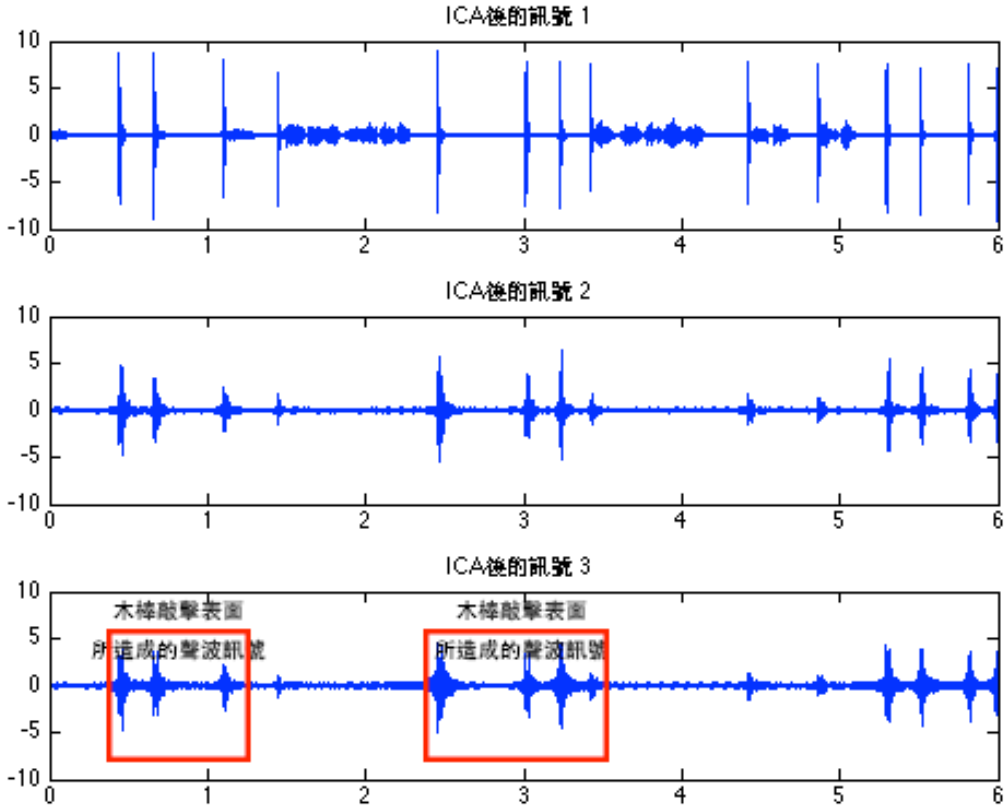

圖 **4-11** 程序（二）： **ICA**所得之訊號波型圖 **-** 輸出 **3** 個訊號

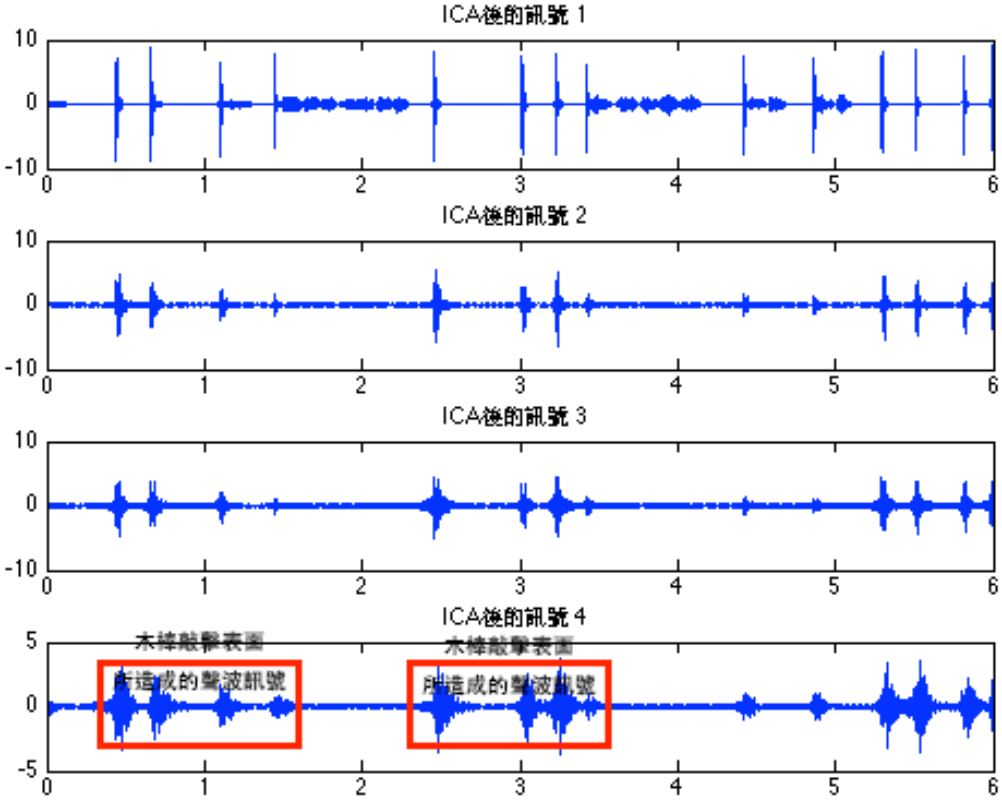

圖 **4-12** 程序（二）： **ICA**所得之訊號波型圖 **-** 輸出 **4** 個訊號

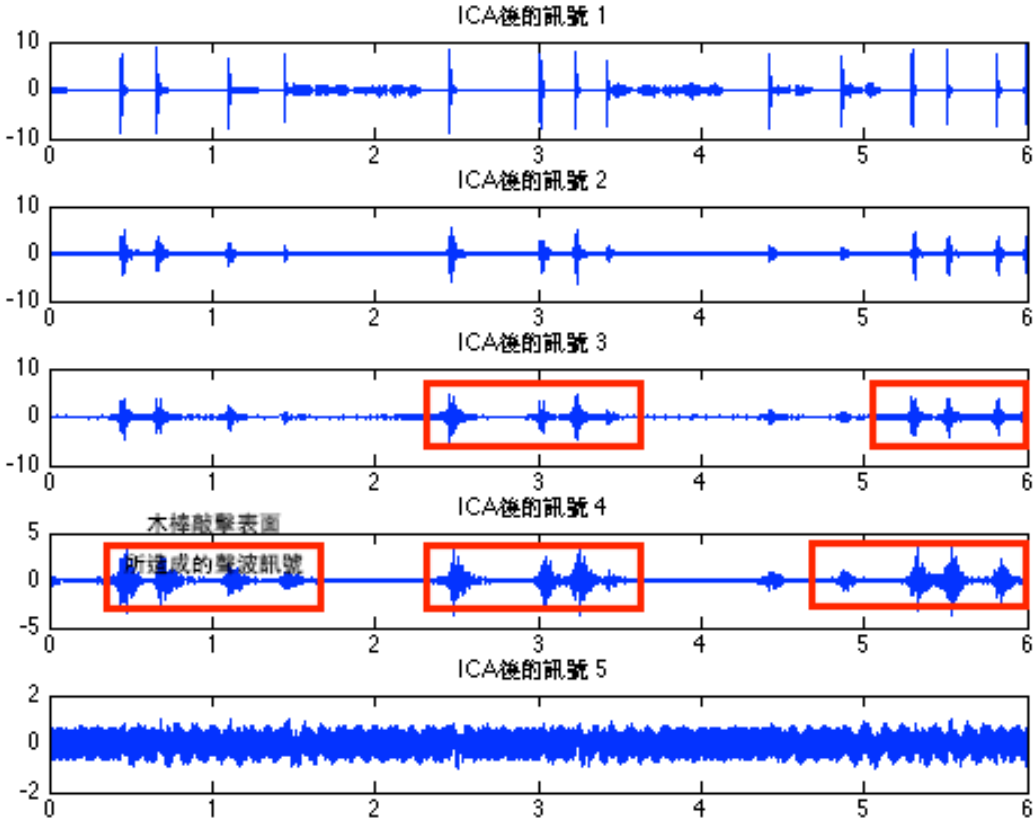

圖 **4-13** 程序（二）： **ICA**所得之訊號波型圖 **-** 輸出 **5** 個訊號

第三節　單音源訊號經頻率區隔及高次混合與ICA處理

於程序（三）單音源輸入訊號經頻率區隔及高次混合後再利用**ICA**進行聲源訊號萃取，此階段利用操作程序（二）中麥克風所截取到的混合訊號以及由頻域切割轉換而來的區段訊號，並依據公式(3-4)與公式(3-5)製作各出10段二次訊號與10段三次訊號，合計高次混合訊號組 X_{high} 共有25段的訊號來進行獨立成份分析，訊號波形圖如圖4-14至圖4-18所示。由於訊號數量眾多，加上運算需求量較大，使得程序（三）在進行獨立成份分析時的運算遠較程序（一）與程序（二）所消耗的時間更多。

藉由獨立成份分析法進行訊號分析後，嘗試由25段原始混合訊號中輸出2~5個分析後訊號進行觀察，結果顯示獨立成份分析法在程序（三）中僅針對木棒敲擊表面所造成的大振幅的訊號進行分離，且分離出來的訊號波形已與原訊號的波形大不相同，造成失真的現象(參考圖4-19至圖4-22)。

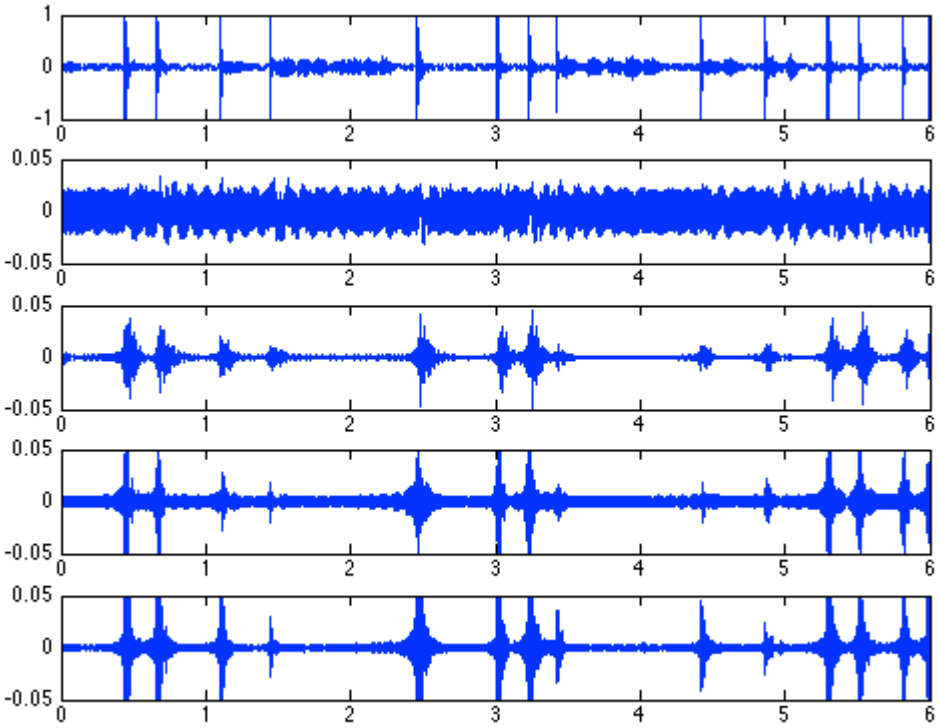

圖 **4-14** 程序（三）：高次混合訊號組 X_{high} 之訊號波形圖之一

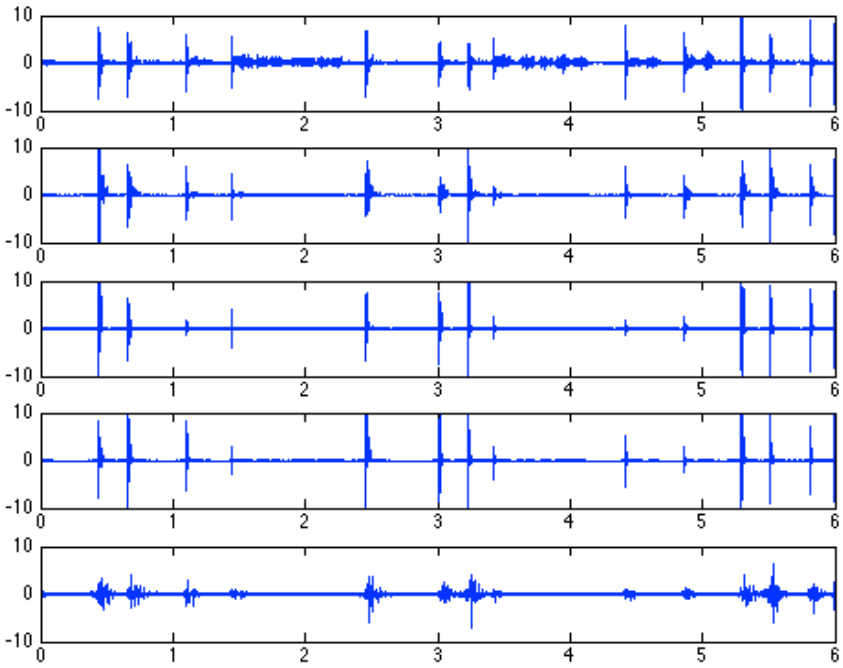

圖 **4-15** 程序（三）： 高次混合訊號組 X_{high} 之訊號波形圖之二

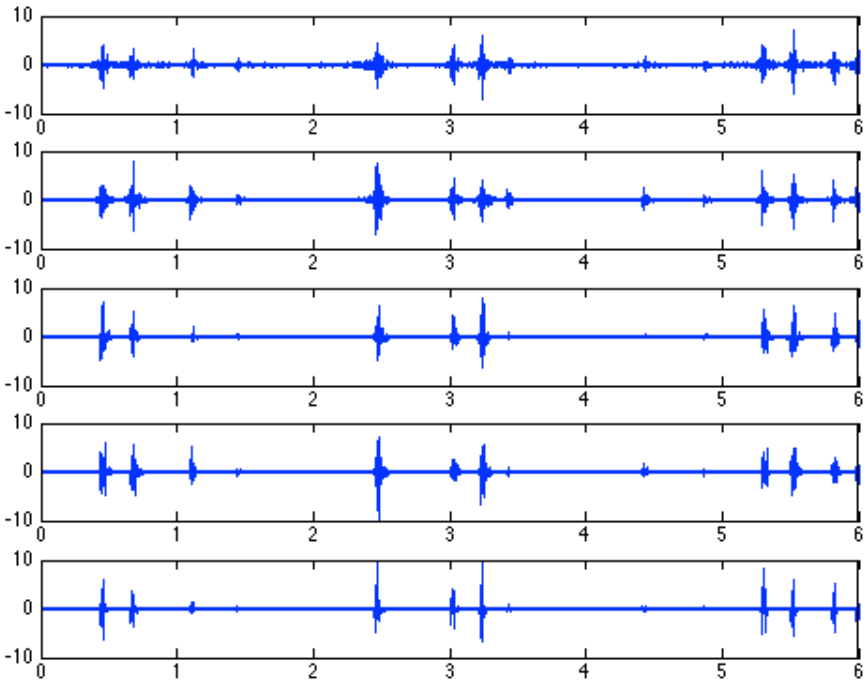

圖 **4-16** 程序（三）： 高次混合訊號組 X_{high} 之訊號波形圖之三

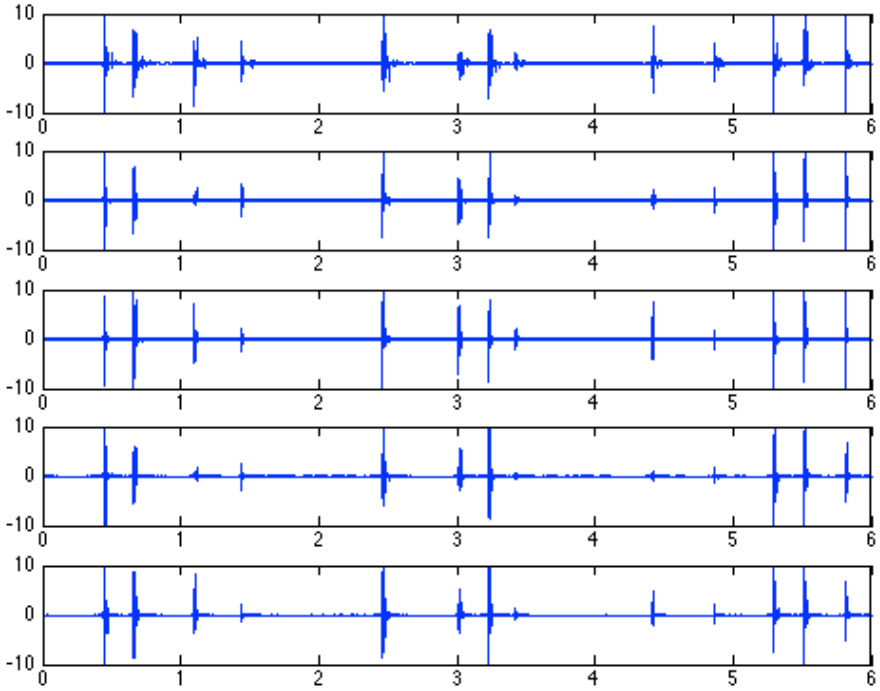

圖 **4-17** 程序（三）： 高次混合訊號組 X_{high} 之訊號波形圖之四

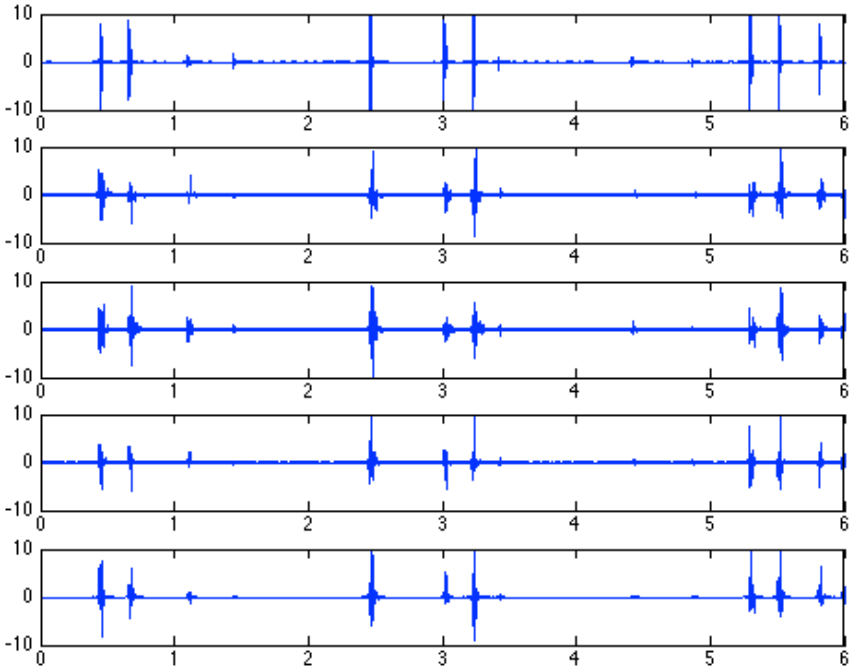

圖 **4-18** 程序（三）： 高次混合訊號組 X_{high} 之訊號波形圖之五

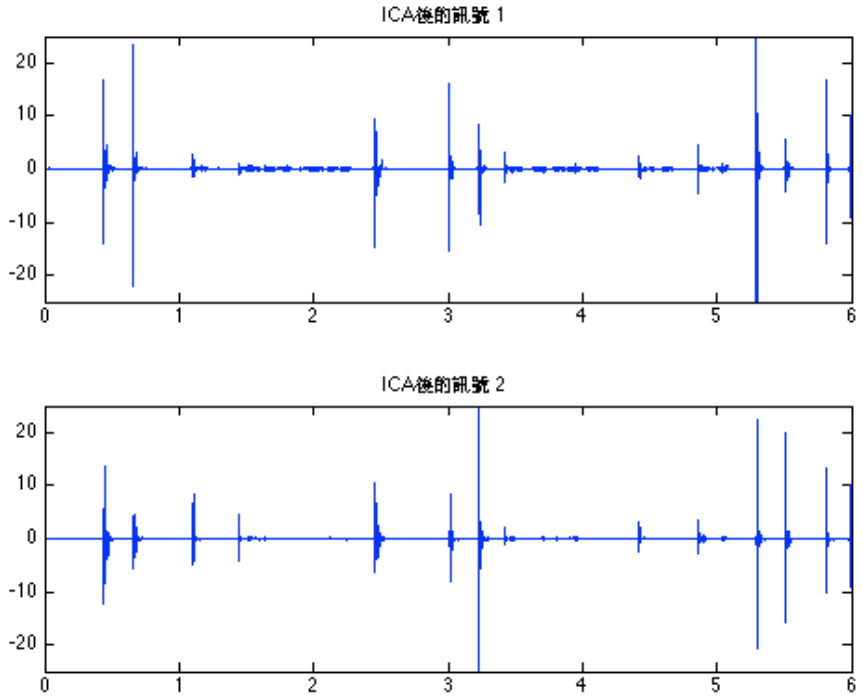

圖 **4-19** 程序（三）： **ICA**所得之訊號波型圖 **-** 輸出 **2** 個訊號

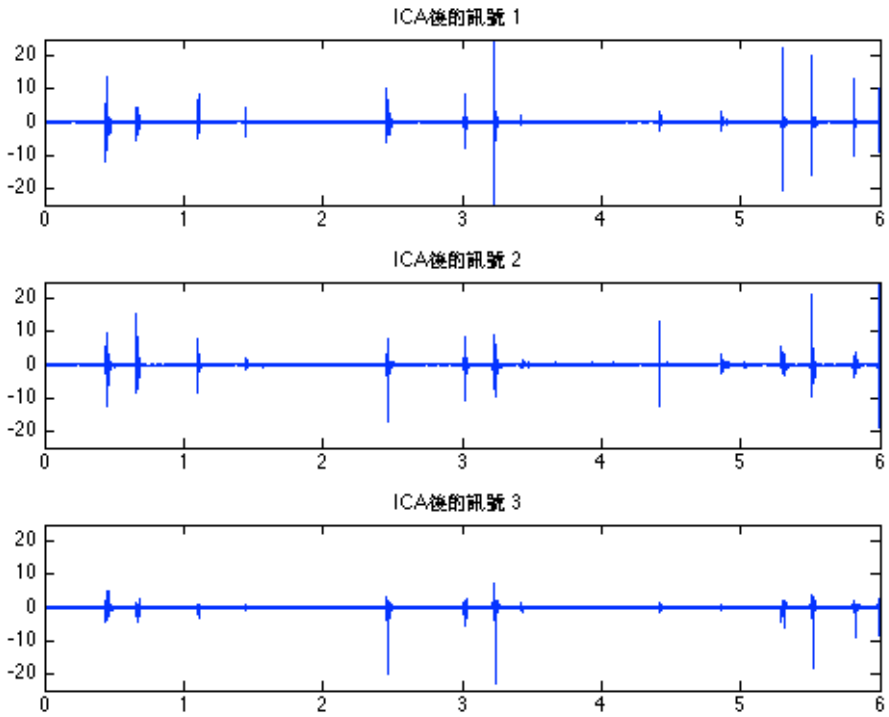

圖 **4-20** 程序（三）： **ICA**所得之訊號波型圖 **-** 輸出 **3** 個訊號

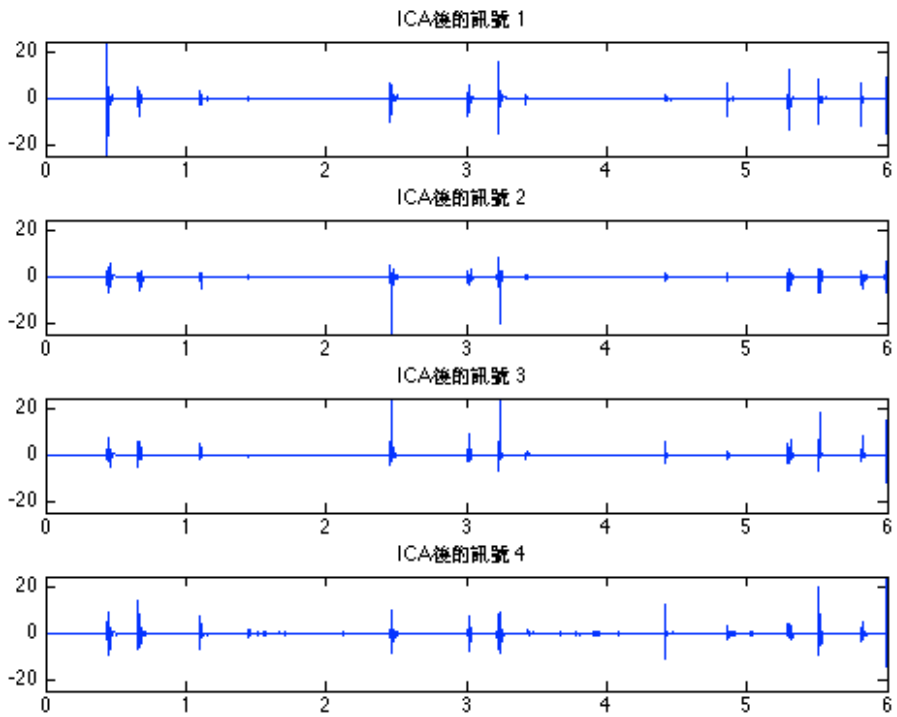

圖 **4-21** 程序（三）： **ICA**所得之訊號波型圖 **-** 輸出 **4** 個訊號

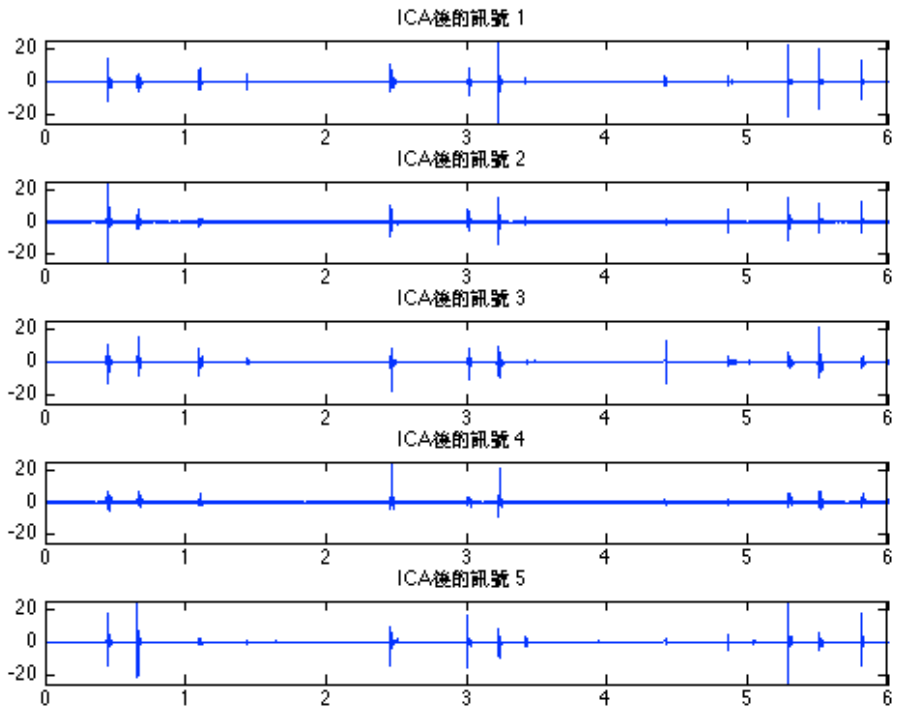

圖 **4-22** 程序（三）： **ICA**所得之訊號波型圖 **-** 輸出 **5** 個訊號

由上述三項實驗操作程序所得之實驗結果顯示，在實驗操作程序（二）中，透過分析混合訊號的頻率分佈，切割出頻率分佈的區段來進行獨立成份分析，其訊號分離效果相較於直接將多音源輸入訊號進行獨立成份分析處理，或是將單音源輸入訊號外加頻率區隔與高次混合後所得來的分析效果，明顯有效許多，且有過濾訊號雜訊的能力。然而，透過本實驗所採用的分析程序，所得到的最佳分離效果僅對木棒敲擊表面聲音的分離效果較為顯著，而鐵棒刮擊表面的聲音則仍混雜無法分離。

經研究者推測，其原因可能在於本研究中所採用的兩種不同聲音訊號源：木棒敲擊表面聲音與鐵棒刮擊表面聲音，在波形圖上已有明顯的差別；若轉為頻率來看，木棒敲擊表面聲音在頻率分佈上為低頻訊號居多，而鐵棒刮擊表面聲音則略居高頻的範圍內。兩者在頻率上分佈得明顯差異，或許是造成實驗操作程序（二）所得之木棒敲擊訊號具顯著分離效果之因素。

在實驗操作程序（一）中利用多音源的輸入來進行分析，其結果卻與預期效果差異甚大，無法將任何訊號分離出來；依據以盲訊號分離技術的角度來看，利用獨立成份分析法於盲訊號分離上，其優點在於可以在不知道訊號源數量的情況下而進行分析，然而本實驗利用多麥克風所收錄的訊號來進行獨立成份分析，其效果與假設不符，原因可能在於聲音訊號於固體物質中傳導具複雜的特性，而未完善考量聲音訊號傳導的物理因素使得訊號分離失敗。

在實驗操作程序（三）中所採用的高次混合方法所得結果也與實驗預期結果差異甚大，令實驗者感到相當訝異。依據文獻探討中所提到之影片字幕萃取實驗(謝珮琳, 民94)所得結果來看，高次混合加強了像素與像素之間的關係性，也造成獨立成份分析後的顯著分離效果。反觀本次實驗中針對聲音訊號進行高次混合處理，所得到的分離訊號卻明顯失真且造成無法辨識的結果，其原因則需後續研究進一步探討。

第五章　結論與建議

本章分為兩小節，第一節的結論部分，將針對前述第四章實驗結果進行彙整與結論。第二節的建議部分將根據研究結論提出相關建議，以供未來研究者參考。

第一節　結論

本研究利用獨立成份分析法針對可觸式聲波操控介面發展出系統化的訊號分析架構，並提出三種依據獨立成份分析法概念所發展出的混合聲音訊號分析操作程序，分別為，分別為 (一) 多音源輸入訊號直接進入ICA處理， (二) 單音源輸入訊號經頻率區隔分析後進行ICA處理 ，以及 (三) 單音源輸入訊號經頻率區隔及高次混合後進入ICA處理。分析的目的在於將可觸式聲波操控介面與物體因觸碰所產生的聲音混合訊號進行訊號分離。

實驗結果發現，透過分析單一麥克風所收錄的混合聲音訊號之頻率分佈，並依據頻率分佈的區段進行訊號切割與重建，製造出更多混合訊號來進行獨立成份分析，其訊號分離效果相較於直接將多音源輸入訊號進行獨立成份分析處理，或是將單音源輸入訊號外加頻率區隔與高次混合後所得來的分析效果明顯有效許多，能有效將木棒敲擊密迪板表面所造成的聲音分離出來。此研究結果也證實了獨立成份分析法可用於可觸式聲波操控介面上作為聲源訊號分離使用，提供可觸式聲波操控介面在裝置開發技術上的另一選擇。

第二節　建議

回歸到本研究的目的來看，導入獨立成份分析技術於可觸式聲波操控界面上使用的目的，主要在於希望利用觸控方式進行互動的人機界面系統能運用此方法來擴展更多元的互動事件辨識，並減低設備的建設成本。對於系統來說，只要能萃取到與原始

混合訊號有一定程度差異的分離訊號，無論該分離訊號是否具有任何實質意義，只要系統能對此分離訊號進行辨識並作出回應，就代表已達到有效的分離效果。因此，當利用獨立成份分析法來做為可觸式聲波操控介面的裝置開發技術時，應該先回歸到界面的設計層面及目的進行考量。若能先將欲獲得的訊號類型進行特徵分析，並規劃系統的辨識範圍，便能大大提高此操控介面的互動性與完整性。

由於本研究採用非即時性的方式來進行訊號分析，且對於系統進行訊號分析的效能層面並無納入考量，因此本實驗所採用的訊號分析方法、演算法以及研究結果不完全能應用於所有的可觸式聲波操控介面上。此外，在混合訊號的分離效果評估上也應更進一步採實驗數據驗證才能更加完善。因此，對於未來後續的研究方向有下列五點建議，以供未來研究參考：（1）將聲音的傳導特性納入考量，進一步討論多音源輸入實際在獨立成份分析法下的可行性。（2）改良在頻率區隔過程上的人工辨識方式，提高頻率區隔的精準度。（3） 針對聲音訊號的特性去發展有效的高次混合，以利單音源輸入下的獨立成份分析過程使用。（4）加入多種物體與界面互動所產生的不同聲源訊號進行測試，進一步驗證獨立成份分析法在可觸式聲波操控介面下的可行性。（5）利用獨立成份分析法來為即時性的可觸式聲波操控介面進行開發設計。

參考文獻

Ainhoren, Y., Engelberg, S., & Friedman, S. (2008). The cocktail party problem [instrumentation notes]. *Instrumentation & Measurement Magazine, IEEE, 11*(3), 44-48. doi: 10.1109/MIM.2008.4534378

Ayari, F., Ksouri, M., & Alouani, A. T. (2013, Jan 20-22). *Computer based analysis for heart and lung signals separation.* Paper presented at the 2013 International Conference on Computer Medical Applications (ICCMA), Sousse.

Carvalhaes, C.G., Perreau-Guimaraes, M., Grosenick, L., & Suppes, P. (2009, June 28 - July 1). *EEG classification by ICA source selection of Laplacian-filtered data.* Paper presented at the IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI 2009), Boston, MA.

Christensen, J. E. N., & Godsill, S. J. (2011, Oct. 16-19). *Bayesian classification of acoustical waveforms under environmental variability,.* Paper presented at the 2011 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, New Paltz, NY.

Chu, B. Y., & Chen, Y. B. (2012, Sept. 18-20). *Reduction of bleed-through effect in images of chinese bank items.* Paper presented at the 2012 International Conference on Frontiers in Handwriting Recognition (ICFHR), Bari.

Comon, P. (1994). Independent component analysis, a new concept? *Signal Processing, 36*(3), 287-314. doi: 10.1016/0165-1684(94)90029-9

Crevoisier, A., & Bornand, C. (2008). *Transforming daily life objects into tactile interfaces.* Paper presented at the 3rd European Conference on Smart Sensing and Context, Zurich, Switzerland.

Crevoisier, A., Bornand, C., Guichard, A., Matsumura, S., & Arakawa, C. (2006). *Sound rose: creating music and images with a touch table.* Paper presented at the Proceedings of the 2006 conference on New Interfaces for Musical Expression (NIME06), Paris, France.

de Souza Siqueira Versiani, T., Rodrigues, G. F., de Souza, A. C. S., de Matos Moreira, J., & Yehia, H. C. (2012, July 3-4). *Binary spectral masking for speech recognition systems.* Paper presented at the 35th International Conference on Telecommunications and Signal Processing (TSP), Prague.

Elena, M., Omar, A. K. , Stéphane, P., Stefano, C., & Houda, C. D. (2009). *Generic framework for transforming everyday objects into interactive surfaces.* Paper presented at the 13th International Conference on Human-Computer Interaction. Ambient, Ubiquitous and Intelligent Interaction, San Diego, USA.

Fabiani, M. (2006). *Development of a tangible human-machine interface exploiting in-solid vibrational signals acquired by multiple sensors.* (Master's thesis).

Harrison, C., & Hudson, S. E. . (2008). *Scratch input: creating large, inexpensive, unpowered and mobile finger input surfaces.* Paper presented at the Proceedings of the 21st annual ACM symposium on User interface software and technology, Monterey, CA, USA.

Harrison, C., Xiao, R., & Hudson, S. E. (2012). *Acoustic barcodes: passive, durable and inexpensive notched identification tags.* Paper presented at the Proceedings of the 25th annual ACM symposium on User interface software and technology, Cambridge, Massachusetts, USA.

Hu, C., Tung, K., & Lau, L. (2008). *Music wall: a tangible user interface using tapping as an interactive technique*. Paper presented at the Proceedings of the 8th Asia-Pacific conference on Computer-Human Interaction, Seoul, Korea.

Hyvärinen, A., & Oja, E. (2000). Independent component analysis: algorithms and applications. *Neural Netwworks, 13*(4-5), 411-430. doi: 10.1016/s0893-6080(00)00026-5

Liang, T. Y., Wang, T. H., Chou, M. T., & Chen, S. W. (2011, Sept. 18-21). *A cloud computing service for fast audio source signal separation.* Paper presented at the 2011 IEEE International Workshop on Machine Learning for Signal Processing (MLSP), Beijing, China.

Lopes, P., Jota, R., & Jorge, J. A. (2011). *Augmenting touch interaction through acoustic sensing*. Paper presented at the Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces, Kobe, Japan.

Pham, D. T., Al-Kutubi, M., Ji, Z., Yang, M., Wang, Z., & Catheline, S. (2005). *Tangible Acoustic Interface Approaches.* Paper presented at the Proceedings of IPROMS 2005 Virtual Conference.

Polotti, P., Sampietro, M., Sarti, A., Tubaro, S., & Crevoisier, A. (2005, September 20-22). *Acoustic localization of tactile interactions for the development of novel tangible interfaces.* Paper presented at the Proc. of the 8th Int. Conference on Digital Audio Effects (DAFX-05), Madrid, Spain.

Rebordao, A., Islam Molla, M. K., Hirose, K., & Minematsu, N. (2008, 7-9 July 2008). *ICA's suitability assisted by voice activity detection.* Paper presented at the International Conference on Audio, Language and Image Processing (ICALIP 2008). .

Sanctis, G. D., Rovetta, D., Sarti, A., Scarparo, G., & Tubaro, S. (2006, September 4-8). *Localization of tactile interactions through TDOA analysis: geometric vs. inversion-based method*. Paper presented at the 14th European Signal Processing Conference(EUSIPCO 2006), Florence, Italy.

Shigeyuki, H., Yoshinobu, S., & Seiho, H. (2012). Bathcratch: touch and sound-based DJ controller implemented on a bathtub. In A. Nijholt, T. Romão & D. Reidsma (Eds.), *Advances in Computer Entertainment* (Vol. 7624, pp. 44-56): Springer Berlin Heidelberg.

Stone, J. V. (2002). Independent component analysis: an introduction. *TRENDS in Cognitive Sciences, 6*(2), 59-64. doi: 10.1016/S1364-6613(00)01813-1

Yang, M., Pham, D. T., Al-Kutubi, M., Ji, Z., & Wang, Z. (2009, June 23-26). *A new computer interface based on in-solid acoustic source localization.* Paper presented at the 7th IEEE International Conference on Industrial Informatics (INDIN 2009).

謝珮琳. (民94). 以獨立成份分析法進行影片文字萃取之研究. (碩士論文), 國立臺灣師範大學. Retrieved from http://etds.lib.ntnu.edu.tw/cgi-bin/gs/ntnugsweb.cgi?o=dntnucdr&i=sid=%22GN0693720011%22.