

# 臺灣蟋蟀聲音資料庫為基礎的生物聲音辨識系統開發

## Developing BioAcoustic Discrimination System Based on the Cricket Sound Database of Taiwan

◎蔡尚惠<sup>1</sup> 楊正澤<sup>2</sup> 凌維<sup>3</sup> 吳世卿<sup>1</sup>

1 環球技術學院環境資源管理系助理教授

2 中興大學昆蟲學系教授

3 環球技術學院環境資源管理系研究生

### 摘要

蟋蟀聲音為種內溝通訊息，亦為種間警戒或保持生殖隔離的重要機制，近來應用於生物分類、生態調查以及休閒生態旅遊等各層面。為便於蟋蟀聲音的自動比對辨識，本研究基於臺灣蟋蟀聲音資料庫以開發生物聲音辨識系統。臺灣蟋蟀聲音資料庫係中興大學昆蟲分類研究室於 1987 年起所陸續建置，其中包括蟋蟀科 9 種、地蟋科 12 種、以及草蟋科 3 種等 497 筆類比資料，針對 24 種蟋蟀分別選取 87 筆蟋蟀呼喚聲音轉為數位資料，以為生物聲音辨識系統中之蟋蟀聲音比對之資料庫。生物聲音辨識系統的開發包含聲音訊號前處理、聲音樣本訓練以及聲音辨識三部分，並於臺灣蟋蟀聲音資料庫中同種蟋蟀各選取不同 3 筆聲音資料為測試樣本，共計 24 種蟋蟀 72 筆聲音資料，以此進行蟋蟀生物聲音辨識系統的測試，所得之辨識正確率為 81.9%；本研究開發完成之生物聲音辨識系統，未來亦可應用於不同生物聲音資料的辨識。

關鍵字：蟋蟀聲音資料庫、生物聲音辨識系統、呼喚聲、聲音行為、昆蟲聲學

### Abstract

Cricket sound could be inter- and intra-specific communication signals and being as important mechanism in alarm and reproductive behavior. It could be applied in the biological taxonomy, ecological inventory and leisure/eco-tourism. The study developed BioAcoustic Discrimination System (BDS) based on the cricket sound database from Taiwan for the convenience of identifying cricket sound data automatically. Cricket sound database of Taiwan established since 1987 continually by Entomologic Taxonomy of Laboratory in National Chung-Hsing University. There are 497 analogical for 9 species of Gryllidae, 12 species of Nemobiidae and 3 species of Trigonidiinae included in the data base. Choosing 87 data of cricket calling sound from 24 species of crickets respectively and transferring to digital data for the cricket sound comparative database in the BDS. The

BDS contained pretreatment of sound signal, training of sound sample and discrimination of sound. Furthermore, choosing 3 different sound data of each species from the 24 species of crickets respectively in the Cricket sound database of Taiwan be test samples. The result of accuracy was 81.9% via 72 data for testing BioAcoustic Discrimination System. The BioAcoustic Discrimination System what we built up could be applied to discriminate in the different kinds of biological sound data.

Keywords: Cricket sound database, BioAcoustic Discrimination System (BDS), calling sound, Acoustic behavior, Acoustics

## 壹、前言

聲音行為(acoustic behavior)係指伴隨富有生物性意義的行為過程，如昆蟲綱(Insecta)直翅目(Orthoptera)蟋蟀(cricket)聲音行為意指以聲音為溝通訊號的行為，而發出的聲音依其頻率可分為聽得到的聲音(audible sound)；波長約為 20 Hz-20 KHz，以及超音波(ultrasound)(楊正澤，1999)。昆蟲鳴聲除為種內傳遞訊息，亦為種間警戒或保持生殖隔離的重要機制，而由於不同種昆蟲之鳴聲具其獨特性，是故昆蟲聲學(insect acoustics)可有效地應用於昆蟲分類學研究的重要特徵之一；如楊正澤(1993)應用於臺灣之蟋蟀科(Gryllidae)中 11 屬 30 種的生物系統分類，Yang and Yang(1995)分析臺灣大蟋蟀(*Brachytrupes portentosus* (Licht.))之形態學(morphology)與聲學特性，賴啟芳(2000)探討東方區蟋蟀科中白緣眉紋蟋蟀(*Teleogryllus mitratus* (Burmeister))及烏頭眉紋蟋蟀(*T. occipitalis* (Serville))之二種蟋蟀的地理分布、形態學及聲學特性，Liu et al.(1998)研究地蟋科(Nemobiidae)2 屬 10 種之聲學及分類，邱碧雲(1993)結合外部形態、雄性生殖器、聲音及棲地(habitat)特性等特徵，綜合判定草蟋科(Trigonidiidae)為 6 屬 18 種。此外，蟋蟀聲音亦可應用於生物資源調查(Riede, 1993; 1998；楊正澤與蔡尚惠，2004)，以及休閒與生態旅遊(Warren et al., 2006)等層面。如由於生物資源調查具難以避免的破壞性取樣，且現今之調查方法亦有部分缺漏，而透過大尺度之生物聲音遙測調查技術，可補強此等缺失，是故 Riede(1993, 1998)曾提出藉由生物聲學(bioacoustics)監測生物多樣性，可用於瞭解熱點(hotspot)之生物多樣性(biodiversity)，以為保育工作之參考，其中特別論及直翅目昆蟲的聲音調查(acoustic survey)方法；並將之應用於熱帶雨林之直翅目昆蟲調查，而凌維(2006)亦以楊正澤與蔡尚惠(2004)所提之三角形聽音辨位法(triangular sound allocation, TSA)等鳴蟲調查技術，應用於監測雲林縣古坑鄉地區蟋蟀聲音之豐富度(abundance)。此外，生物聲音經蒐集後，需加以辨識(discrimination)分析，然除一般專家經訓練之人耳辨識，為使聲音資料辨識之分析過程標準化和自動化，本研究基於臺灣蟋蟀聲音資料庫，開發「生物聲音辨識系統」(BioAcoustic Discrimination System, BDS)

## 貳、材料與方法

### 一、研究材料

直翅目昆蟲中已被鑑定者估計約 70,000 種(Riede, 1998)。又全世界蟋蟀估計 3,000 種左右，依其棲地不同可分為地棲性、草棲性與樹棲性，而臺灣已知約計 100 多種，約占全世界的二十分之一。而蟋蟀總科(Superfamily Grylloidea)中含甲蟋科(Cachoplistidae)、叢蟋科(金銛蛉科, Eneopteridae)、蟋蟀科、鉦蟋科(Mogoplistidae)、蟻蟋科(Myrmecophilidae)、地蟋科、樹蟋科(Oecanthidae)、長蟋科(Pentacentridae)、鐘蟋科(金鐵兒科, Phalangopsidae)、毛翅蟋科(Pteroplistidae)、鐵蟋科(Scleropteridae)、草蟋科(吉蛉科、蝘蛉科)等(中華昆蟲學會, 1994)。本研究所使用之中興大學昆蟲系臺灣蟋蟀聲音資料庫，主要為 1987-2001 年之蟋蟀科、地蟋科以及草蟋科等蟋蟀聲音類比資料。

### 二、研究工具

楊正澤(1999)指出蟋蟀的聲音屬單調(monotone)，是故音色(sound quality)可略而不論，而其聲音重複特性則可透過人為方法分析；又所錄得之聲音訊號，依儀器之性能與所需的資料不同，可分別使用各種儀器與電腦軟體加以分析。職是之故，本研究利用 CoolEdit Pro 2.1 分析不同種類蟋蟀之聲音結構特徵，如以波形圖(waveform, oscillogram)展示之長唧聲長度(trill duration, TD)、長唧聲間隔(trill interval, TI)、長唧聲總長(trill length, TL)等時域特徵(time domain characteristics)，以及透過頻譜圖(spectragraphy)或聲紋圖(sonagraph)展示主頻率(principle frequency)、共振頻率(resonant frequency)等頻域特徵(frequency domain characteristics)。此外，主要參酌林增安(2004)之相關聲音辨識(speech recognition)研究，結合 GoldWave v5.08 聲音分析軟體，使用 Visual C++ 6.0 程式語言開發生物聲音辨識系統(BDS)。

## 參、結果與討論

### 一、臺灣蟋蟀聲音資料庫之建立

聲音是蟋蟀總科(Grylloidea)昆蟲的主要通訊方式，對物種之求偶、交尾等生命活動具有重要作用(蘆榮勝等, 2002)。而廉振民與李愷(2002)認為蟋蟀鳴聲與其生活習性相適應，而將其區分為三種類型，即呼喚聲音(calling sound, CAS)，又稱常鳴聲；求偶聲音(courtship sound, COS)，又稱調情聲；攻擊聲音(aggressive sound, AGS)，又稱打鬥聲。又 Alexander(1962)指出依行為背景不同，雄蟋蟀一般具有呼喚聲音(CAS)、求偶聲音(COS)、求偶中斷聲音(interruption sound, IRS)、交配後聲音(postcopulate sound, PCS)、攻擊聲音

(AGS) 以及巢洞辨認聲音 (nest recognition sound, NRS) 等六種不同聲音，此等聲音曲目 (acoustic repertoire) 各具不同的功能，欲達成其物種生存延續之目的，則各種蟋蟀必須巧妙地安排其聲音所表達的曲目。而本研究所建立之臺灣蟋蟀聲音資料庫中係以呼喚聲 (CAS) 為主，且將原類比聲音資料轉換為數位資料檔。

臺灣蟋蟀聲音資料庫中含蟋蟀科之 *Brachytrupes portentosus* Lichtenstein (BraPor)、*Gryllus bimaculatus* De Geer (GryBim)、*Loxoblemmus doenitzi* Stein (LoxDoe)、*L. abortus* Yang et Yang (LoxAbo)、*Mitius splendens* Shiraki (MitShi)、*Teleogryllus occipitalis* Serville (TelOcc)、*T. mitratus* Burmeister (TelMit)、*Loxoblemmus sylvestris* Matsuura (LoxSyl)、*Velarifictorus sukhadae* Bhowmik (VelSuk) 等 9 種 265 筆，地蟋蟀科之 *Dianemobius mikado* Shiraki (DiaMik)、*Dianemobius kimurae* Shiraki (DiaKim)、*D. fascipes* Walker (DiaFas)、*D. chibae* Shiraki (DiaChi)、*D. furumagiensis* Ohmachi et Furukawa (DiaFur)、*D. protransversus* Liu et Yang (DiaPro)、*Pteronemobius jucundus* Liu et al. (PteJuc)、*P. fascipes* Walker (PteFas)、*P. wulaina* Liu et al. (PteWul)、*P. kimurae* Shiraki (PteKim)、*P. caudatus* Shiraki (PteCau)、*P. ohmachi* Shiraki (PteOhm) 等 12 種 162 筆，以及草蟋科之 *Trigonidium cicideloide* Rambur (TriCic)、*T. humbertianus* Saussure (TriHum)、*Metioche kotoshoensis* Shiraki (MetKot) 等 3 種 70 筆，共計 497 筆蟋蟀聲音資料。

## 二、生物聲音辨識系統之開發

生物聲音辨識系統 (BDS) 的開發亦參考如 Lee et al. (2006) 以平均梅爾倒頻譜係數 (averaged Mel frequency cepstral coefficient, AMFCC) 萃取聲音特徵，再透過線性區別分析演算法 (linear discriminant analysis) 降低特徵維度開發生物聲紋自動辨識 (Automatic recognition of animal vocalizations)，並以蛙、蟋蟀、鳥聲等檢測其辨識率。又張勇富 (2002) 以 ActiveX 技術，整合 Matlab 程式語言和 Delphi 發展環境，開發鳥音辨識系統 (bird sound recognition system)。另 Ingrisich et al. (2004) 所提之德國直翅目典藏 (German Orthoptera Collections = Deutsche Orthopteren-Sammlungen, DORSA, <http://www.dorsa.de>) 中，透過前處理 (pre-processing) 和濾波 (filtering)、單一脈衝 (pulse) 分段 (segmentation)、萃取典型特徵 (extraction of typical features) 以及自動分類 (automated classification) 等步驟所開發之自動曲目辨識 (automatic song recognition) 系統，主要應用於蟋蟀總科與螞蟻總科 (Tettigonioidae) 之聲音比對。此系統包含聲音訊號前處理、聲音樣本訓練以及聲音辨識三部分；以聲音訊號前處理而言，先自臺灣蟋蟀聲音資料庫之 497 筆資料中，選取 87 筆蟋蟀呼喚聲，每筆資料萃取 9 秒之聲音資料，再利用 GoldWave v5.08，將原 22,050 Hz 之蟋蟀聲音轉換為單聲道、16 bit、8 KHz 取樣頻率，檔案格式為 \*.wav 檔，其主要目的為減少雜訊，且可

提高爾後辨識比對的速度，又聲音樣本訓練與聲音辨識之使用方法與步驟，茲說明如下：

### (一) 聲音樣本訓練

聲音樣本訓練主要目的是萃取樣本特徵，係透過預強調(pre-emphasis)、音框化(framing, frame blocking)、短時距能量(short-time energy)測量、越零率(zero-crossing rate)、漢明視窗(Hamming window)之音框切割，並以快速傅利葉轉換(fast Fourier transform, FFT)、梅爾倒頻譜係數(Wang et al., 2002; Leung and Siu, 2004)以及線性預測係數(linear predictive cepstrum coefficient, LPCC)求得最佳聲音特徵；其中音框化(Frame blocking)是將蟋蟀聲音樣本分為  $n$  個樣點，利用 GoldWave v5.08 將  $n$  個樣點，以每音框為 170 個樣點的方式進行切割。

### (二) 聲音辨識

將經聲音特徵萃取後之蟋蟀聲音樣本，透過特徵值向量量化(vector quantization, VQ)、隱藏式馬可夫模型(Stadermann and Rigoll, 2006)進行聲音辨識比對；蟋蟀聲音訊號經過特徵值組合之後，為增加辨識效率與資料處理效能，透過向量量化的技術處理聲音特徵參數，向量量化是將  $k$  維的向量空間中，分成有限區域(通常是 8-12 維)，生物聲音辨識系統(BDS)係以 12 維為基準，利用向量量化找出一組編碼向量(code vector)代表蟋蟀聲音在空間中全部的特徵向量，編碼向量為每筆蟋蟀聲音碼之參考(speaker's codebook)，可降低計算複雜度及減少資料儲存量，此外，將蟋蟀聲音訊號進行轉換訓練過程中，訓練樣本即為每一筆蟋蟀聲音資料為一模型，而辨識蟋蟀聲音的方法，即是比對該筆蟋蟀聲音訊號最接近的模型，研究中藉由維特比(Viterbi)演算法進行轉換運算，用以找尋最佳路徑，於隱藏式馬可夫模型(HMM)狀態轉移中，必須記錄轉移至各狀態最大機率值與先前狀態計算出的結果，並找出結果中具有最大機率值的狀態，記錄找出最佳路徑(optimal path)所歷經的狀態序列。

### (三) 生物聲音辨識系統測試

自臺灣蟋蟀聲音資料庫中另外擷取同種蟋蟀之不同個體的 3 筆聲音資料為測試樣本，共計 24 種蟋蟀 72 筆聲音資料，進行生物聲音辨識系統的測試，所得之辨識正確率為 81.9%(表一)；測試結果發現蟋蟀聲音樣本數多者，則其辨識率較高，而樣本數少則反之，故樣本數較少的蟋蟀聲音較難辨識。

表一、生物聲音辨識系統之辨識率測試

數 量	科 名			合 計
	Gryllidae	Nemobiidae	Trigonidiidae	
樣本數	27	36	9	72
正確數	23	31	5	59
錯誤數	4	5	4	13
正確率(%)	85.2	86.1	55.5	81.9

## 肆、結論

本研究臺灣蟋蟀聲音資料庫中之蟋蟀科、地蟋蟀科以及草蟋科等24種蟋蟀聲音資料為基礎，利用Visual C++ 6.0程式語言，結合GoldWave v5.08所開發之生物聲音辨識系統(BDS)，其辨識率為81.9%，未來需建立並擴充臺灣蟋蟀聲音資料庫之蟋蟀聲音的比對樣本，如增加其他的物種，以及各類聲音行為之蟋蟀聲音，是故亦可應用於直翅目直翅目的蝈蝈(katydid)、螻蛄(mole cricket)，或如同翅目(Homoptera)昆蟲中的蟬(cicada)、兩生類、鳥類等；此外，再參酌聲音辨識之相關文獻，採用更有效之聲音特徵萃取和辨識技術，以改良生物聲音辨識系統(BDS)。

## 伍、參考文獻

1. 中華昆蟲學會，昆蟲綱科以上學名中文對照表，中華昆蟲特刊，第九期，第 1-40 頁(1994)。
2. 林增安，聲音辨識與對話策略，國立雲林科技大學電子工程系碩士論文，雲林縣(2004)。
3. 邱碧雲，臺灣的草蟋蟀亞科(直翅目：蟋蟀科)，國立中興大學昆蟲學研究所碩士論文，臺中市(1993)。
4. 凌維，蟋蟀聲音調查方法及生物聲音辨識系統之開發—以臺灣蟋蟀聲音資料庫為基礎，環球技術學院環境資源管理系碩士論文，雲林縣(2006)。
5. 張勇富，以語料分析為主的鳥音辨識系統研究，國立東華大學電機工程學系碩士論文，花蓮市，第 65 頁(2002)。
6. 廉振民、李愷，蟋蟀常見鳴聲類型的比較研究(直翅目：蟋蟀總科)，昆蟲分類學報，第二十四卷，第一期，第 45-51 頁(2002)。
7. 楊正澤、蔡尚惠，鳴蟲調查技術及其應用—直翅目昆蟲，「第五次野生動物研究與調查方法」研討會論文集，野生動物保護基金會，第 67-91 頁(2004)。
8. 楊正澤，臺灣蟋蟀亞科(直翅目：蟋蟀科)生物系統分類，國立中興大學昆蟲學研究所博士

論文，臺中市(1993)。

9. 楊正澤，蟋蟀聲學特徵分析，昆蟲分類及進化研討會專刊，第 175-197 頁(1999)。
10. 劉淑惠，臺灣地蟋蟀亞科(直翅目：蟋蟀科)，國立中興大學昆蟲學研究所碩士論文，臺中市(1991)。
11. 賴啟芳，東方區 *Teleogryllus mitratus* (Burmeister) 及 *T. occipitalis* (Serville) 地理分布、形態學及聲學特性之探討，國立中興大學昆蟲學研究所碩士論文，臺中市(2000)。
12. 蘆榮勝、楊培林、石福明、李成華，歷山自然保護區四種蟋蟀鳴聲結構的比較研究(直翅目：蟋蟀總科)，動物分類學報，第二十七卷，第三期，第 491-497 頁(2002)。
13. Alexander, R. D., Evolutionary change in cricket acoustic communication. *Evolution*, Vol.16, No.4, pp. 443-467(1962).
14. Ingrisch, S., K. Klaus, K.-H. Lampe and C. Dietrich., DORSA- a “virtual museum” of German Orthoptera collections. *Memorie Soc. Entomol. Ital.*, Vol.82, No.2, pp. 349-356(2004).
15. Lee, C.-H., C.-H. Chou, C.-C. Hanand and R.-Z. Huang., Automatic recognition of animal vocalizations using averaged MFCC and linear discriminant analysis. *Patt. Recog. Lett.*, Vol.27, No.2, pp.93-101(2006).
16. Leung, K.-Y. and M. Siu., Integration of acoustic and articulatory information with application to speech recognition. *Inform. Fusion*, Vol.5, No.2, pp.141-151(2004).
17. Liu, S.-H., J.-T. Yang, H.-K. Mok and C.-T. Yang., Acoustics and taxonomy of Nemobiidae (Orthoptera) from Taiwan. *J. Taiwan Mus.*, Vol.51, No.1, pp.55-124(1998).
18. Riede, K., Monitoring biodiversity: analysis of Amazonian rainforest sounds. *Ambio.*, No.22, pp. 546-548(1993).
19. Riede, K., Acoustic monitoring of Orthoptera and its potential for conservation. *J. Insect Conserv.*, No.2, pp. 217-223(1998).
20. Stadermann, J. and G. Rigoll., Hybrid NN/HMM acoustic modeling techniques for distributed speech recognition. *Speech Commun.*, Vol.48, No.8, pp.1037-1046(2006).
21. Wang, J.-C., J.-F. Wang and Y.-S. Weng., Chip design of MFCC extraction for speech recognition. *Integration, the VLSI J.*, Vol.32, No.1-2, pp.111-131(2002).
22. Warren, P. S., M. Katti, M. Ermann and A. Brazel., Urban bioacoustics: it's not just noise. *Animal Behav.* Vol.71, No.3, pp.491-502(2006).
23. Yang, J.-T. and C.-T. Yang., Morphology and male calling sounds of *Brachytrupes portentosus* (Licht.) (Orthoptera: Gryllidae). *J. Taiwan Mus.*, Vol.48, No.2, pp.1-9(1995).