

Ping-Pong Paddle Type Estimation using Acoustic Characteristics

藉由聲學特性進行乒乓球球拍特性情蒐

組別：A123 組員：周德灝 指導教授：黃柏鈞 教授

1. 摘要

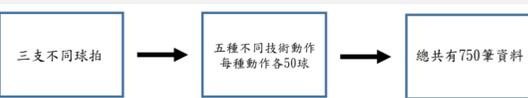
在現今的乒乓球界裡，球拍的種類百百種，不同的球拍所造成的擊球手感也不同，因此每個專業運動員都在找尋屬於自己的合適球拍，這對於分析對手情蒐方面也非常重要，面對不同對手時，若能夠提前透過音訊而不用實體目測或接觸即可判斷出對手是使用何種球拍、何種技術，進而可以賽前模擬策略，則能知己知彼而百戰百勝。

我想要結合科技的方法，利用音訊處理和機器學習來進行球拍分類，而使用此方法的優點是並不需實體目測和接觸，就能獲取球拍的特徵資料。而實作過程中，首先要收集資料，使用三種不同材質的球拍和五種不同的技術動作，每一種技術動作為50球，收集完後會有15種不同種類的集合，總共大約有750筆音訊資料。經過實作聲音事件切割、音訊特徵提取和機器學習，目前能判斷出不同球拍的準確率約為75%。

2. 實驗器材與資料收集

在這次實作專題裡需要三支不同材質的球拍和五種不同的技術動作結合，先透過手機在安靜的環境下錄取音訊資料，再利用MATLAB做音訊分析，最後將分析後的結果使用Python做機器學習的模型分類。

- 球拍材質
 - 純木材質
 - 純木材質裡加上碳纖
 - 純檜木材質
- 五種不同技術動作，每種技術動作為50球
 - 攻球
 - 扣殺
 - 切球
 - 前衝
 - 弧圈



圖[2] 資料收集過程示意圖

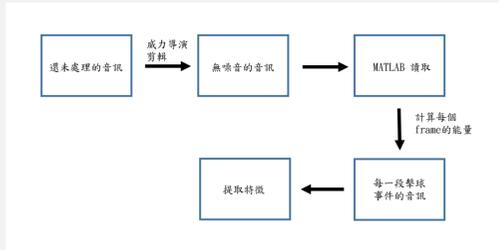


圖[1] 106年中華民國 乒乓球國手

3. 演算法設計

將收集好的音訊分為兩大階段：

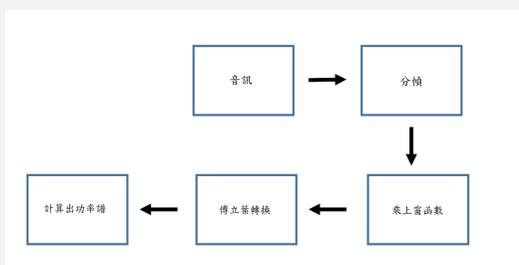
第一階段為利用MATLAB來提取特徵，而要提取特徵前，要先將一整段擊球音訊分割出每次擊球的事件，透過計算每一個frame的能量，利用每個frame的能量大小比較，可清楚得知每個擊球事件，進而再將分割完的每個事件提取出特徵。



圖[3] 音訊切割事件方法

音訊特徵提取方法：Mel-Frequency Cepstral Coefficients(MFCC)

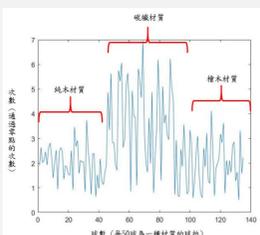
- 將音訊分幀。
- 經過分幀完的每個音訊要乘以一個窗函數，增加音訊的連續性。
- 經過窗函數的音訊，做傅立葉轉換。
- 將訊號的頻譜平方得到功率譜，進而將功率譜通過Mel濾波器取得音訊特徵。



圖[4] mfcc音訊特徵提取方法

音訊特徵提取方法：過零率

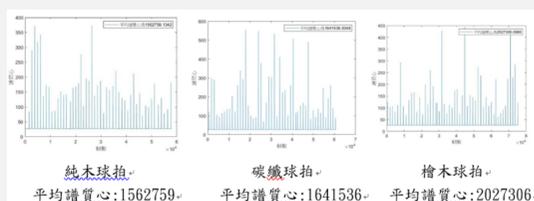
計算每個frame裡，音訊通過零點的次數，根據實驗結果能發現，雜訊的過零率均大於清晰可辨視音訊的過零率。



圖[5] 音訊特徵過零率

音訊特徵提取方法：譜質心

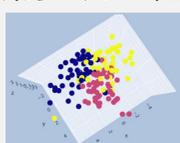
據不同的球拍，相對應也有不同的音色，而譜質心是頻率成分的重心，也是聲音信號的頻率分佈和能量分佈的重要信息。以扣殺技術動作為例：



圖[6] 音訊特徵譜質心

第二階段為將以上得到的音訊特徵，利用Python機器學習，獲得最終結果。我使用的方法為Support Vector Machine (SVM)。

以攻球為例：



藍色:純木球拍
粉色:碳纖球拍
黃色:檜木球拍

圖[7] 特徵從16維降為3維所形成的分布圖

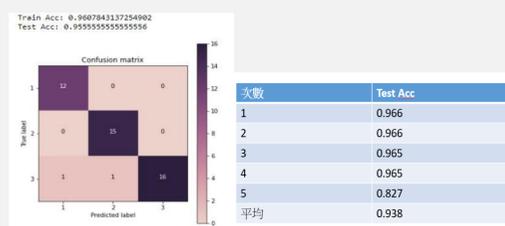
4. 實驗結果與分析

在這次專題裡，已完成兩大階段實驗目標

第一階段目標為音訊的特徵提取，先將音訊的每個frame的能量大小分割為每個擊球事件，再從中挑取特徵，這次實作專題總共實驗了三支不同的球拍，五種不同的技術動作，而每一支球拍做五種技術動作，而每一種技術動作為五十球，因此一段音訊裡會有五十個擊到球的事件。再將這五十個事件一一通過MFCC、過零率、譜值心等音訊特徵的演算法提取出特徵。

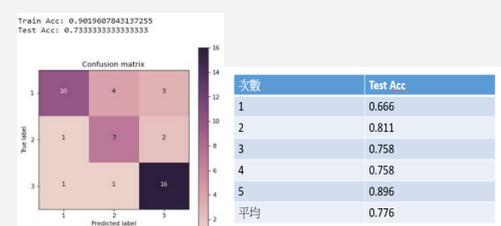
第二階段目標為利用python機器學習，使用第一階段所提取出的目標進而標註訓練，將全部的特徵資料分為訓練資料和驗證資料，而訓練資料佔全部資料的70%，剩下的30%為驗證資料，進而再使用交叉驗證法，我將資料依據4:1比例分割為訓練資料與驗證資料，迭代5次且每一次都會有一個答案，再將五次答案取平均得到更為精準的準確率。以下將兩大階段的結果呈現。

攻球 平均準確率:93.8%



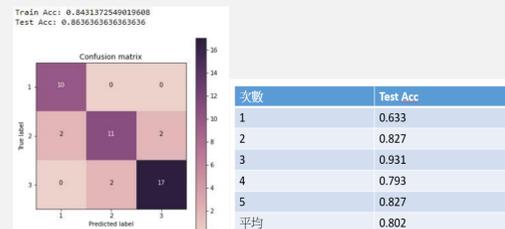
圖[8] 攻球準確率

扣殺 平均準確率:77.6%



圖[9] 扣殺準確率

前衝 平均準確率:80.2%



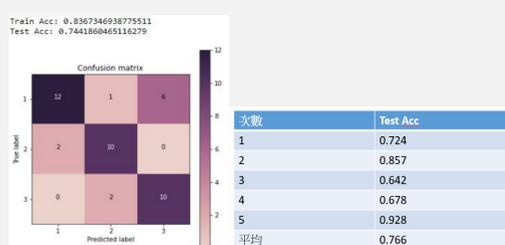
圖[10] 前衝準確率

切球 平均準確率:87.7%



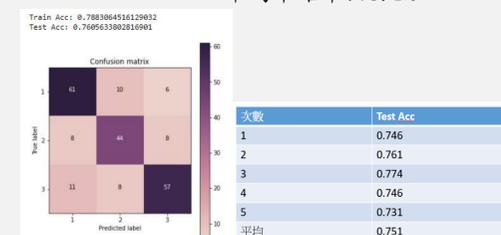
圖[11] 切球準確率

弧圈 平均準確率:76.6%



圖[12] 弧圈準確率

三支球拍 五種不同技術 (全部750筆資料)
平均準確率:75.1%



圖[13] 全部資料訓練完的準確率

根據以上結果能得知，當撞擊力量較小時，球拍對於球所造成的音訊影響較明顯，因此分析結果的準確率都有八成以上，但當撞擊力量較大或無撞擊時，分析結果的準確率會較遜色；弧圈這個技術動作可能是因為無撞擊只有球拍上的膠皮與球造成聲響，而非球拍材質所影響，因此訓練出的模型分類準確率只有76.6%；而扣殺是一種大力撞擊的技術動作，而當力量足夠大時所造成的音訊會較複雜且不穩定，因此訓練出的模型分類準確率只有77.6%。

5. 結論

經過這次專題實作，我所實作的演算法能有效地將不同的球拍利用各個音訊特徵的不同，進而將球拍分類。在音訊特徵的提取實做MFCC、過零率、譜值心等諸多音訊特徵後，透過這些音訊來得出有效的結果，在將來的工作裡可以在優化演算法進而不只分類球拍，而是能分出不同球拍擊到不同球種和使用不同的技術動作，能有效地分析出結果，且加上球的飛行速度、球的轉速和球的落點，能即時顯示在外接的螢幕上，對於乒乓球的觀賞性或情蒐都是一大進步！