An Interactive System for Real-time Music Visualization

實時音樂視覺化多媒體互動系統

組別：A56　組員姓名：周郁欣、許暐彤

指導教授：劉奕汶　教授

Abstract

“Chromesthesia” is one of the common types of “Synesthesia” which specifically indicates the neural condition where one involuntarily associates hearing sense with colors. This only occurs in a little percentage of the population, while we have observed that many others hold their own interpretation on how music should look like due to their own life experiences. As for music therapy, music is widely used in treatments among various mental disorder patients with practices such as improvisation, singing, songwriting, and discussing music. We long to purpose a new alternative way for not just patients, but everyone to express themselves, and unique perspective of seeing music and one’s own impact.

In this paper we propose a real time music spatial visualizing interactive system designed to process audio input such as pure tones, instrumental music, and singing voice into sequential colored spatial presentation.

Two different methods are compared within pitch detecting, one utilizing the neural network based Madmom audio processing library, while the other algorithm is rather simple with noise reduction and spectrum transform procedures. Both methods generate MIDI information within specific frames, where is further moderated into 12 note names and presented in certain colors due to the design.

The proposed method yields an relatively more accurate result on real time pitch detecting. As for RNN itself, accuracy increases among longer audio inputs.

Test results show that the Peak Picking System is superior to utilizing RNN in real

time pitch detection in terms of accuracy and numerical stability.

Keywords **–** music visualization, audio processing, pitch detection

摘要

「連帶色覺 (Chromesthesia)」為較常見的一種「聯覺通感(Synesthesia)」，特指感官認知途徑上將聽覺與顏色連結的感知狀態。連帶色覺發生的機率在人群中佔比很低，但同時我們觀察到有很多人也因由自身生活經驗而常對音樂該看起來的樣子有自己獨特的詮釋。另外，音樂治療領域中，音樂被廣泛運用於各種心理疾病的治療手法，常藉由即興創作、哼唱、譜曲、討論音樂等過程來達到治療的目標。我們期望提出一種新的與音樂互動的模式，讓不只是心理疾病患者，甚至所有人都能重新表達自己、得以以獨特的角度欣賞音樂及自身對於環境事物的影響力。

此文提出一實時將音樂空間視覺化的互動系統，旨在設計一能同步將單一聲調波形、樂器演奏、人聲哼唱等多種類型之音訊轉化成空間上的連續顏色變化之系統。

音準辨識方面，此文比較兩種識別方法。其一使用神經網路的基礎上建製的音訊處理資料庫 Madmom，另則自行建構一相對單純的降噪頻域能量峰值轉換系統。兩者皆輸出時序上之音樂數位介面(MIDI)音調及強度等資料，再依據系統設計汲取音名映射相對連續色彩變化。

文中提出的系統(Simple Peak Picking System)在實時的音高辨識上與循環神經網路鋼琴音高處理器(RNN Piano Note Processor)相比呈現到相對準確的結果。對於循環神經網路鋼琴音高處理器來說，一次性輸入較長的音訊可以使辨識結果較為準確。

實驗結果顯示，此文提出之音樂數位介面能量峰值轉換系統(Simple Peak Picking System)因其在準確度及數值穩定度上表現較佳，故較循環神經網路鋼琴音高處理器更適用於的實時音高辨識。

關鍵字 **—** 音樂視覺化、數位音訊處理、音高辨識

INTRODUCTION

In the recent years, music therapy have been increasingly recognized as a new health profession, and has been widely used in the treatment of a variety of mental disorders, such as autism, dementia, anxiety, depression, and so on. Music therapy is regarded as an expressive therapy, where improvisation, singing, songwriting, and discussing music are believed to achieve treatment goals and objectives. And instead of the usual ways of interacting with music, we long to purpose a new alternative way for patients to express themselves, and to see their own impact.

Enjoying music is an experience that most people share, but expressions arising from them varies within all individuals, and can be presented in different forms, words, emotions, body moves, or even visual effects. “Synesthesia” is the neurological condition when stimulating one of the senses simultaneously stimulates one’s other senses, an involuntary response through different cognitive pathways. Since this may involve any combination of senses, Synesthesia may have 60 – 80 subtypes, and

“Chromesthesia” is one of the common types which specifically indicates ones who associate sounds with colors. Seeing through a synesthete’s perspective can give the world an additional dimension, while studies proposed that this only occurs within a very little percentage of the population.

This paper introduces a real-time system that bridges not just those who suffer with mental disorders but all of us to the chance of experiencing seeing music, or even defining the corresponding colors based on one’s own life experiences. This system translates real-time music and singing into colors, spatially presented in colored lights, implemented on R-pi3. There are a lot of methods when it comes to pitch detection, some include neural network ensemble and others require heavy pre-processing. In this paper, comparing the results to a neural network based audio signal processing library written in Python called Madmom, we tried to build a system which is rather simple and fast, not using any datasets or beforehand modeling. Requiring just the board and a mic, our system is designed to process audio input such as pure tones, instrumental music, and singing voice into sequential colored spatial presentation.

Result Pictures

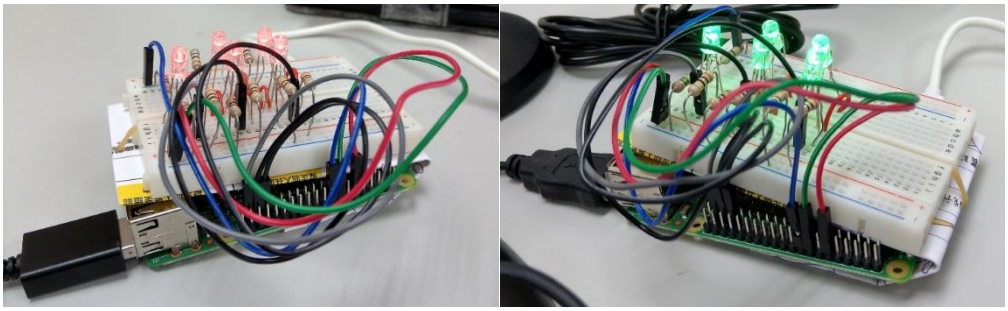

Figure 1 LEDs Light Up Colors According to Detected Notes

Figure 2 System's Appearance

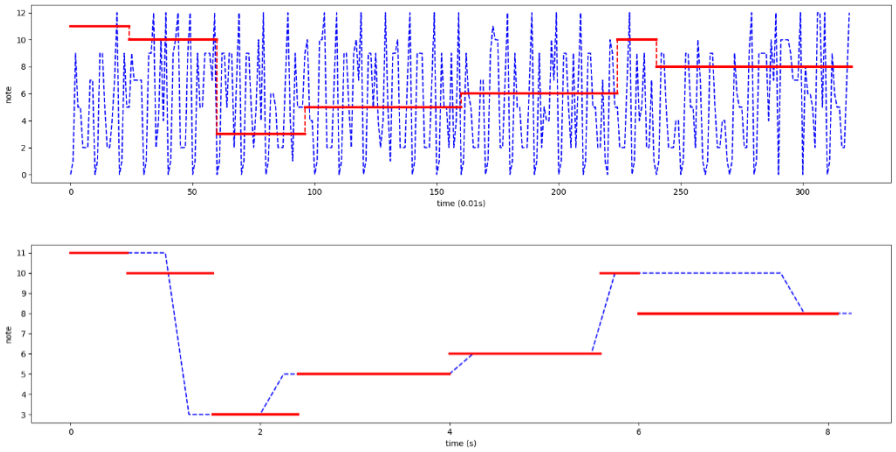

Figure 3 Results of Pure tones Detection with RNN and SPP

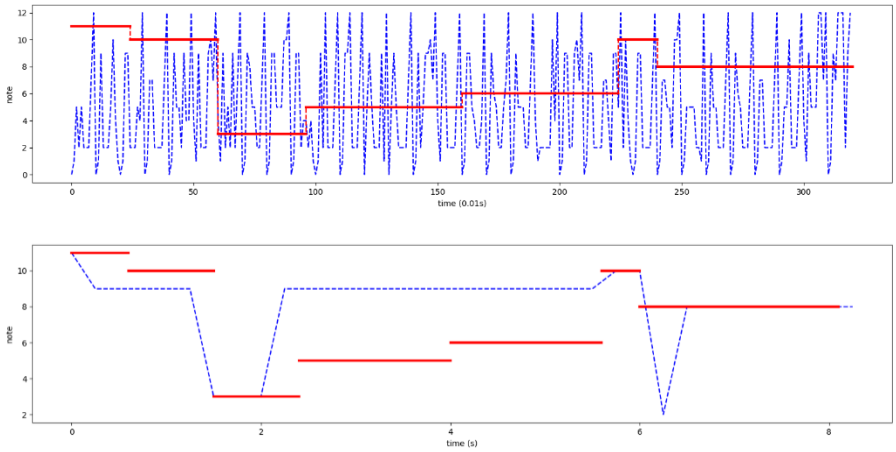

Figure 4　Results of Piano Notes Detection with RNN and SPP

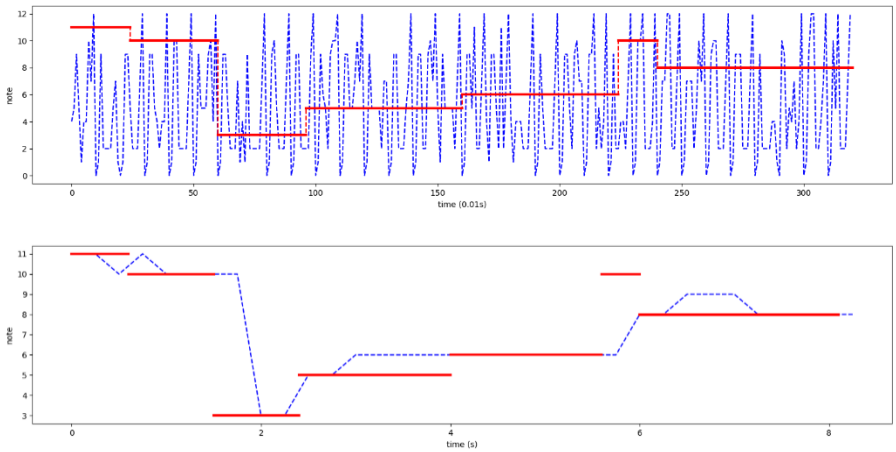

Figure 5　Results of Singing Voice Detection with RNN and SPP

Epilogue

周郁欣：

Professor Liu has always been emphasizing the importance of “creating want ourselves long to use”, which is also the mindset where I came up with several ideas surrounding music. I truly appreciate that I get to group up with my partner whom our thoughts meet, and together give it a try about a subject we are both into. Professor Liu was as supportive as ever, providing suggestions and encouragements every now and then. Though the topic isn’t as profound as many other groups’ are, but it definitely benefits myself and inspired me a lot within the process. Not only putting theorems in textbooks into practice, but also got myself embraced within a whole new environment of graduates and various study subjects. It was a pleasant journey.

許暐彤：

剛開始對專題要做什麼一點概念都沒有。看著台達館內的海報，覺得學長姐做的每個專題都好厲害。總覺得這些東西離我好遠。直到在找指導教授的途中遇到郁欣學姊，聽著她理想的 final result 我彷彿也能看到一台聲光效果機出現在我眼前。而在和劉奕汶教授討論後，我才知道專題不光光只有那些看起來好深奧的理論，還能有這麼生活化、這麼有趣的題目。或者應該說，那些看起來很深奧、離我很遠的理論，其實都像這樣被其他人運用著、藏在我們的生活之中。在這將近一年的相處中，教授也一直扮演著支持的角色。雖然這次專題成果與一開始的想像仍有些許差距，但教授的那句「做自己想要用的東西」已經成為我這次最重大的收穫了。