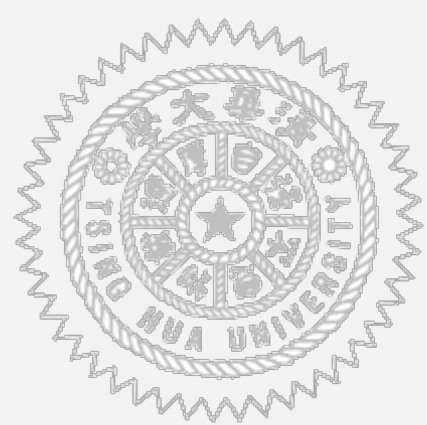


# 【自動樂譜辨識】

## Automatic Music Score Recognition



組別:A01

指導教授:劉奕汶

組員:曾子芸、黃瑋玲、江昀軒

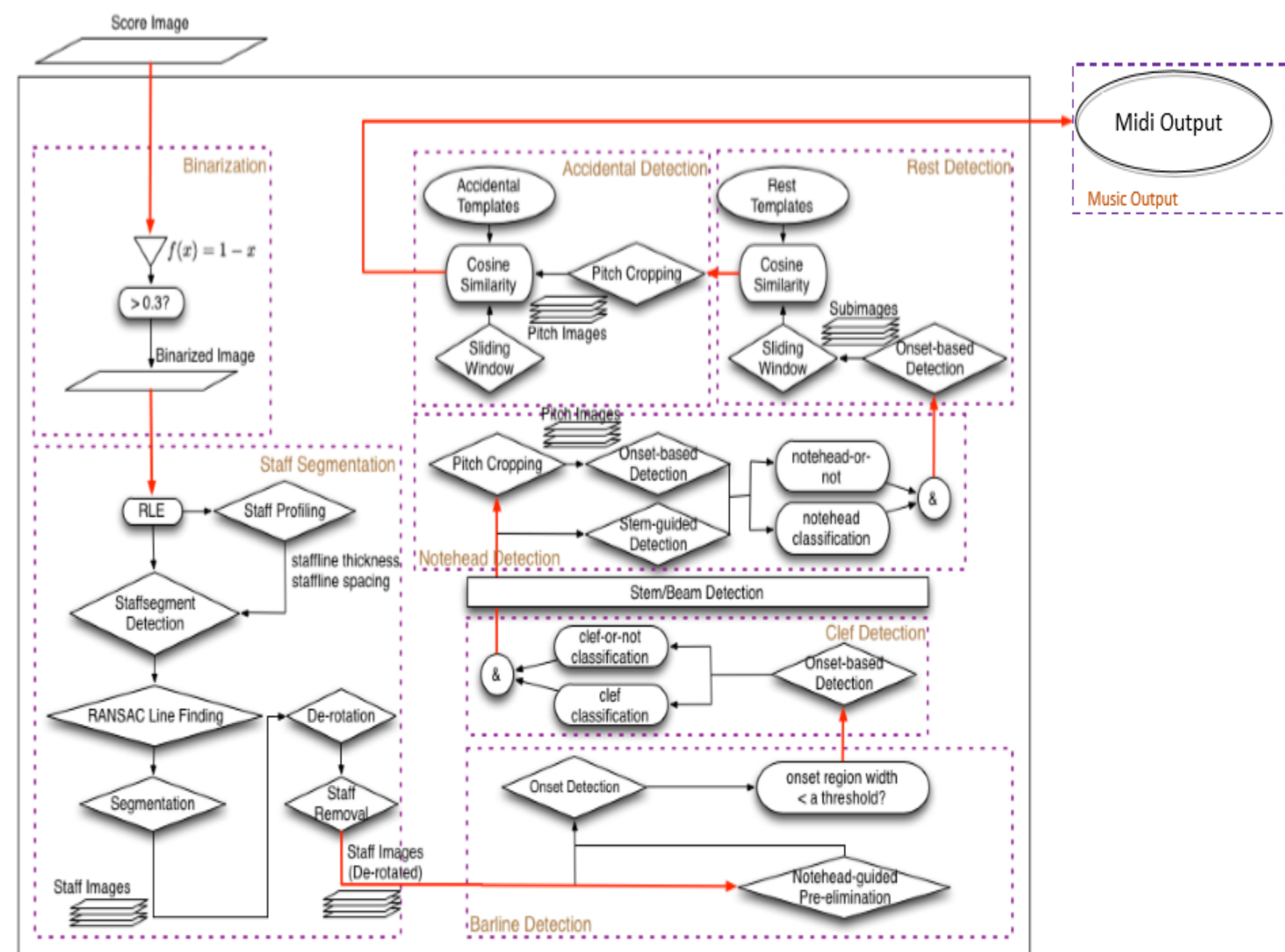
### Abstract

音樂在我們生活中是不可或缺的一部份，樂譜的發明是為了以圖像的方式記錄一段音樂資訊，將聽覺視覺化。對於初學者而言，最先遇到的瓶頸是無法單純從閱讀圖像式的樂譜準確地判斷旋律、節奏、升降等樂曲的特色，以至於對樂曲的演繹方式無從下手。

此專題用於設計一套演算流程，讓電腦也能夠自動辨認出原先給人閱讀的樂譜進而將音樂進行輸出。一般而言，樂譜被存為電子檔的格式多數為圖片檔。因此光學辨識的主要目的在於從一張圖片的每個像素中，取得其音樂資訊。

而我們的專題主要做的部份則是建立在李豪章學長的碩士論文的基礎上，同時提升樂譜辨識的精準度，進而將辨識出的樂譜進行輸出。

### System Structure

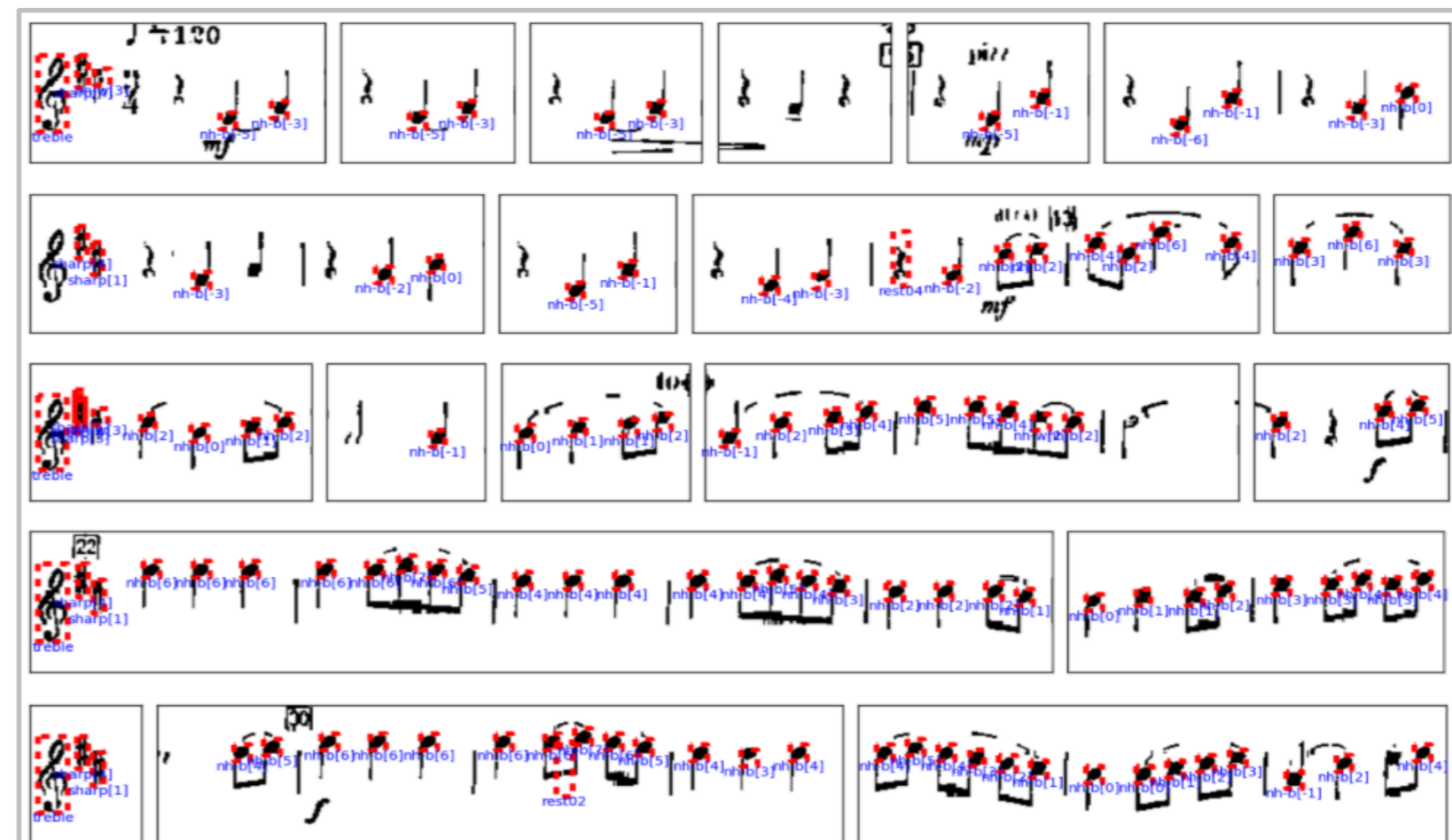


### Procedure and Result of the Experiment

在輸入端的樂譜我們採用，96dpi以上、每行五線譜上下高度60像素以上、左右寬度600像素以上的JPG或PNG圖檔為主，以確實配合原程式設計。

而為了能準確分析實驗結果，我們的實驗分成兩大部份：第一部份從李豪章學長的資料庫圖檔，作為輸入端圖片來源；第二部份則是使用自行在網路上找到的樂譜圖片檔來做測試及分析。

圖片經過系統的過程，首先在前置作業會針對原圖片的顏色及角度做處理，依各個像素的數值做二值化後，最後只留下黑白兩色，並利用每列五線譜中每條線相同的間距，估計出各列五線譜可能的中心點，描繪最高機率的中心線。接著以每條線為中心上下延伸切割出垂直長度，分割及扶正各列五線譜。之後偵測各列五線譜中的五條橫線並加以移除，再根據所辨識出的小節線，對各個小節加以分割。最後由系統對所剩符號進行辨識後，由紅色虛線方框標繪出所偵測到的音符範圍，並以藍色文字標示音符內容及其音高。輸出端的圖片結果與所有實驗測試資料分析如右圖及表格所示。



	total	miss	error	note	error rate (%)
treble	97	1	4	2(bass) + 2(nh-w)	5.15
bass	19	3	5	3(treble)	42.11
flat	62	4	0		6.45
sharp	47	1	12	1(bass) + 6(double sharp) + 1(nh-b) + 4(nh-w)	27.66
natural	2	0	0		0.00
rest01	5	0	0		0.00
rest02	16	9	0		56.25
rest04	150	78	14	14(sharp)	61.33
rest08	95	22	0		23.16
rest16					
rest32					
rest64					
nh-b	1128	60	7	1(flat) + 6(nh-w)	5.94
nh-w	306	35	13	13(nh-b)	15.69
<b>total</b>	<b>1927</b>	<b>213</b>	<b>55</b>		<b>13.91</b>

### Conclusion

這次的專題研究中，我們所完成的部分為：校正傾斜的樂譜、將大多數常用的符號辨識出來、最後將樂譜用MIDI音樂檔進行輸出。

實驗結果分析可以看出來，我們所測試的這幾分樂譜中，錯誤率平均為13.91%，還有很大的修正空間。我們歸納出的原因是：二值化的過程會使樂譜部分失真，出現一些汙點，而這些汙點會造成後續執行上的失真，從切割小節開始，到辨識符頭、符桿、休止符、升降記號等都會出現較難辨識的狀況。至於MIDI輸出的部分，目前只能做到單行輸出，如果樂譜是屬於有不同分部或和聲的話，便無法同時輸出多行。

### Future

可以將此專題從原來只能辨識的單一音符的樂譜，提升為可以辨識和弦的系統，並且將MIDI輸出的部分優化為可以同時輸出多個分部。

### Reference

Hao-wei, Lee, "Automatic Music Score Recognition", National Tsing Hua University, 2016