



影像辨識鋼琴譜生成器

PIANO SHEET GENERATOR

指導教授:郭天穎 組員:葉宏洋、蕭方凱、呂紹愷、吳宥霖

目錄

01

使用工具

02

研究動機

03

研究內容

04

成果

05

未來展望



使用工具 | 01

使用工具



VSCODE



OPENCV



LILYPOND



研究動機 | 02

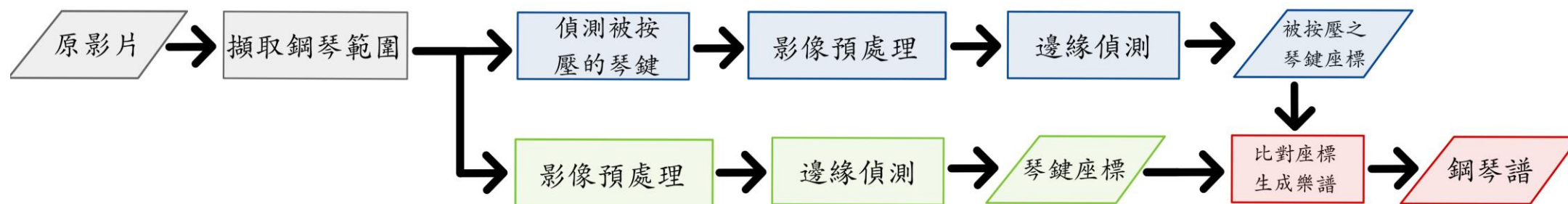
研究動機

隨著串流影音技術的成熟，網路上有非常多的鋼琴彈奏影片，但多數都不會附上免費樂譜，多半只能自己聽、自己寫譜，而有些影片是無聲的，或者是走音的鋼琴，導致我們難以單靠聲音來辨別音符，於是本專題利用影像辨識，透過影像生成五線譜，既能省下時間，又能補足聲音辨識的不足，帶給使用者更好的體驗。

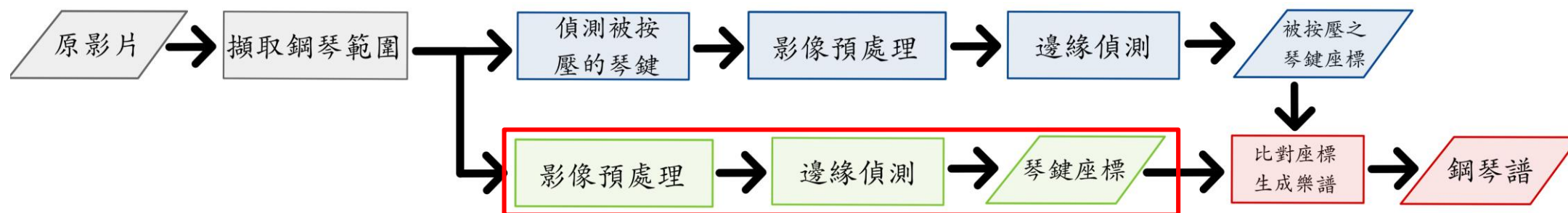


研究内容 | 03

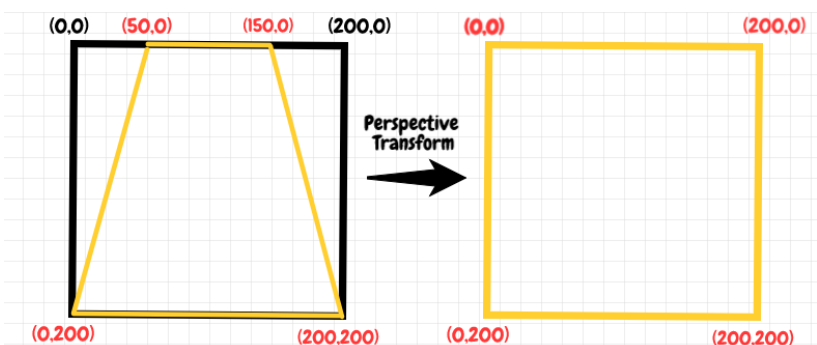
流程圖



流程圖



PERSPECTIVE TRANSFORMATION



如圖的四個點座標

$[(50,0), (150,0), (0,200), (200,200)]$

映射到另四個點座標

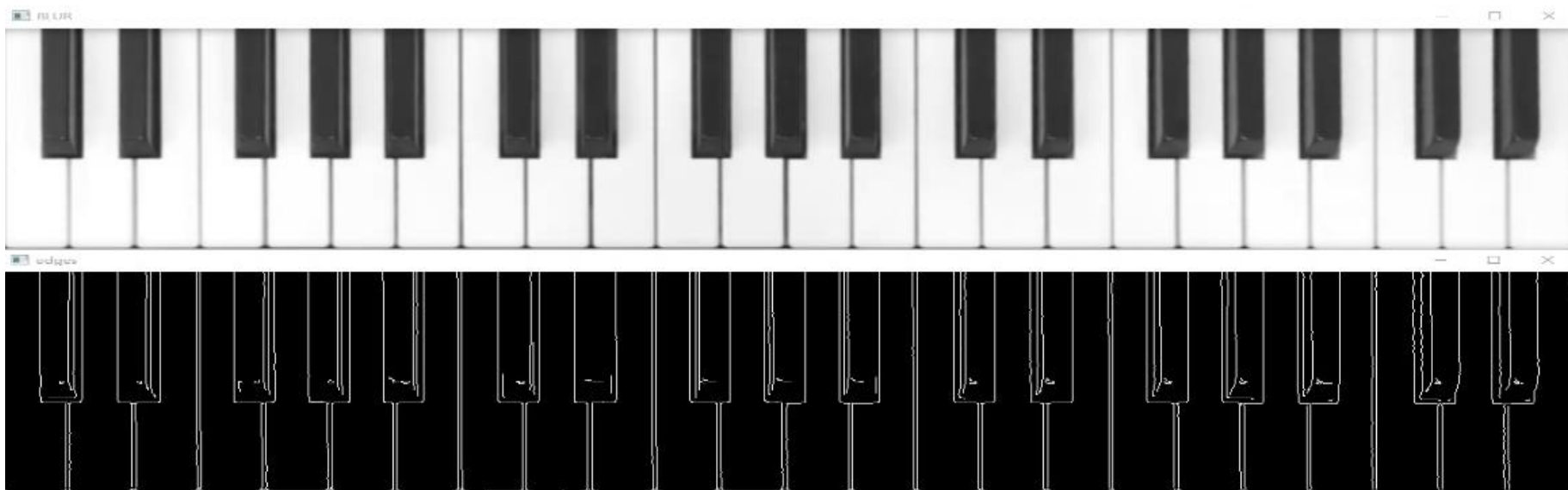
$[(0,0), (200,0), (0,200), (200,200)]$.

透過一變換矩陣 $\begin{pmatrix} 2 & 0.5 & -100 \\ 0 & 2 & 0 \\ 0 & 0.005 & 1 \end{pmatrix}$ 進行轉換。



邊緣偵測 (CANNY)

Canny演算法是一個複合性的邊緣偵測演算法，結合了Gaussian Filter、梯度偵測、非最大值抑制、判斷邊界四個演算法去實踐邊緣偵測，但相對的，Canny濾波器在實作上沒那麼靈活，無法像Sobel跟高斯可以搭配各種不同的影像處理演算法。

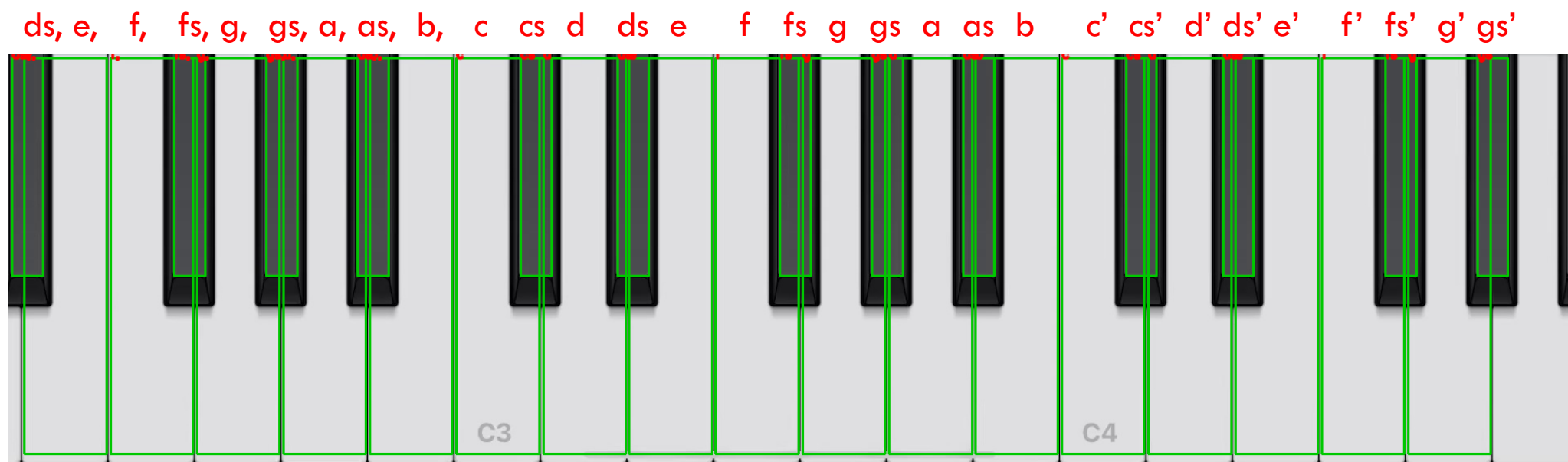


將琴鍵進行邊緣偵測

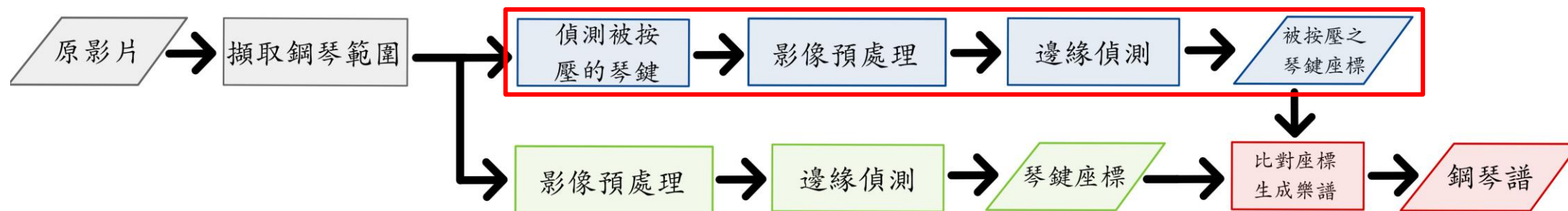
琴鍵座標

按鍵座標[按鍵序,按鍵左座標,按鍵右座標]:

```
[[0, 5, 43], [1, 20, 121], [2, 126, 227], [3, 204, 242], [4, 232, 333], [5, 317, 355], [6, 338, 439], [7, 429, 467], [8, 444, 545], [9, 550, 652],  
[10, 629, 667], [11, 657, 758], [12, 748, 786], [13, 762, 864], [14, 868, 970], [15, 947, 985], [16, 975, 1076], [17, 1060, 1098], [18, 1081, 1183], [19, 1172, 1210], [20, 1187, 1288], [21, 1293, 1394], [22, 1372, 1409], [23, 1399, 1501], [24, 1491, 1528], [25, 1506, 1607], [26, 1612, 1713], [27, 1690, 1728], [28, 1718, 1819], [29, 1803, 1840]]
```



流程圖



前景/背景分離(ABS.DIFF)

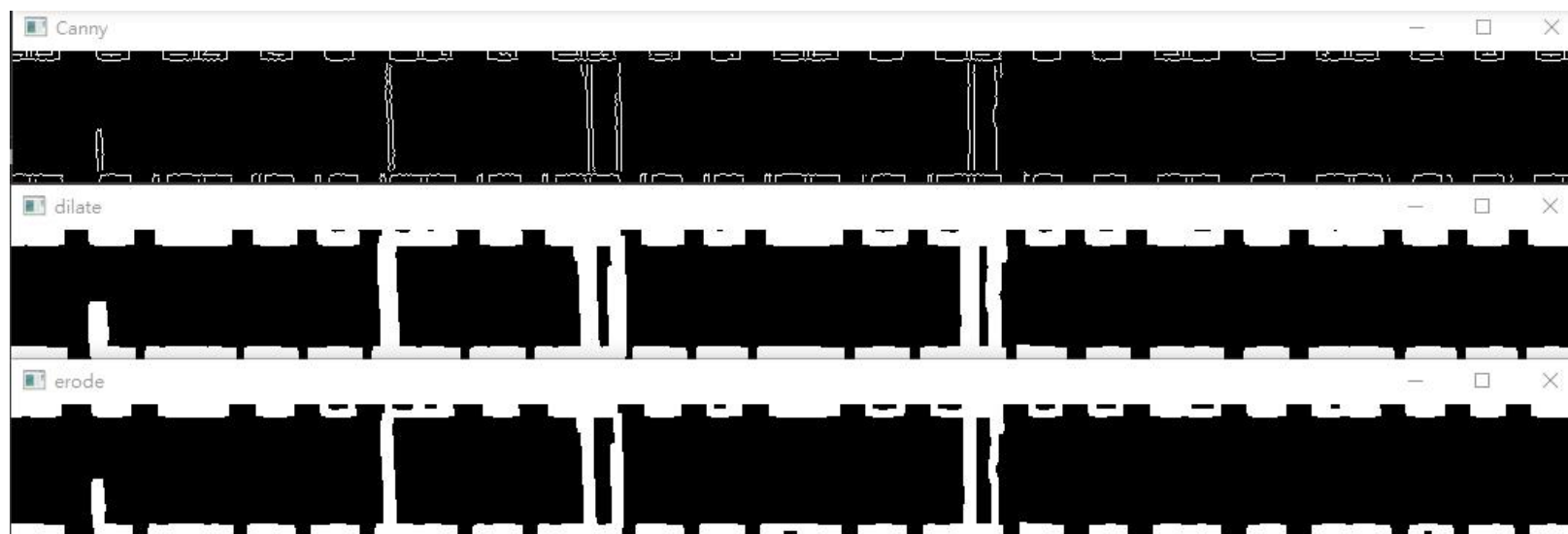
前景/背景分離技術可以分離出移動中的前景，有許多種方法可以做到前/後景分離，本專題選擇使用abs.diff去進行前景/背景分離，它的功能是将兩張照片去做相減後取絕對值，即可找出兩張照片中的相異之處。



將彈奏影片去做前景/背景分離可得到按壓鍵盤之邊緣

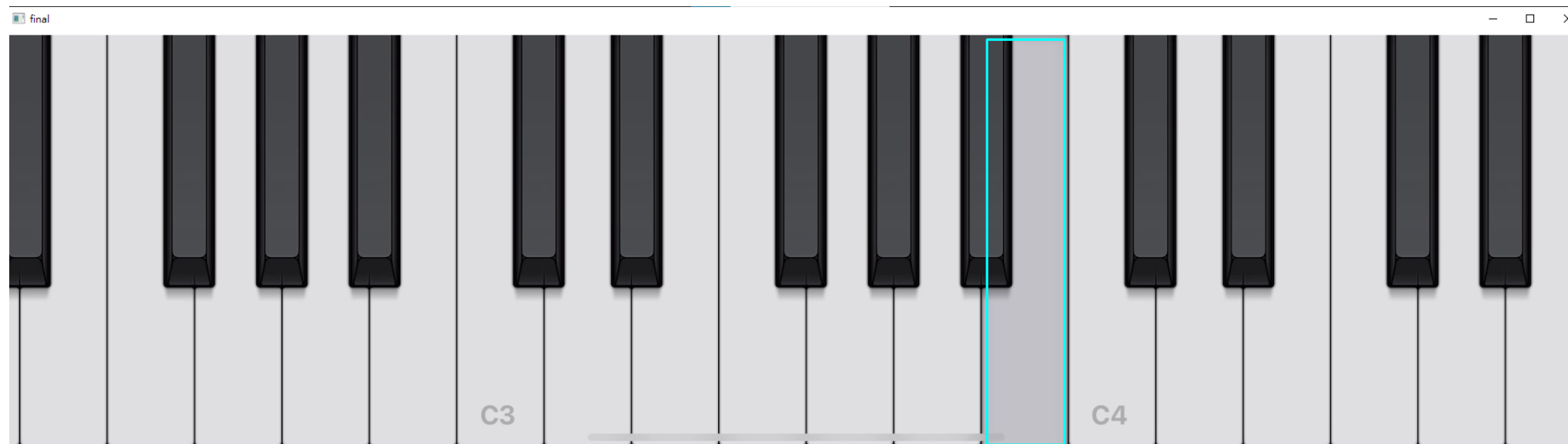
形態學(MORPHOLOGY)

在進行影像處理時，經常使用的一種方法，可以用來去除雜訊，也可以用來增強圖片所要傳達的訊息，主要分為兩大功能: 膨脹(Dilation)、侵蝕(Erosion)。

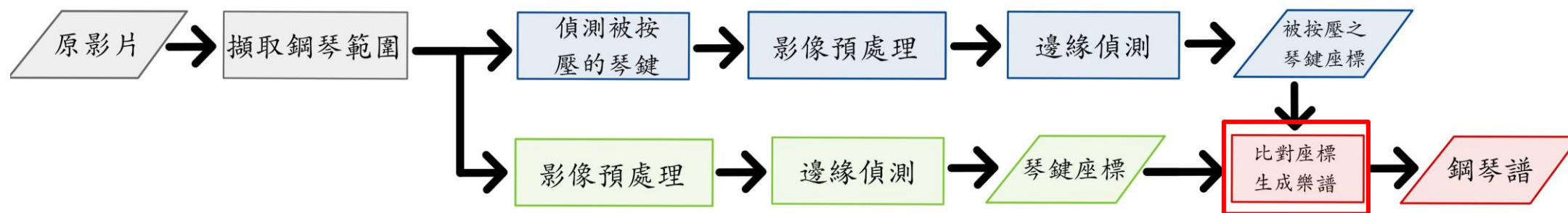


將琴鍵按壓變化去做Closing

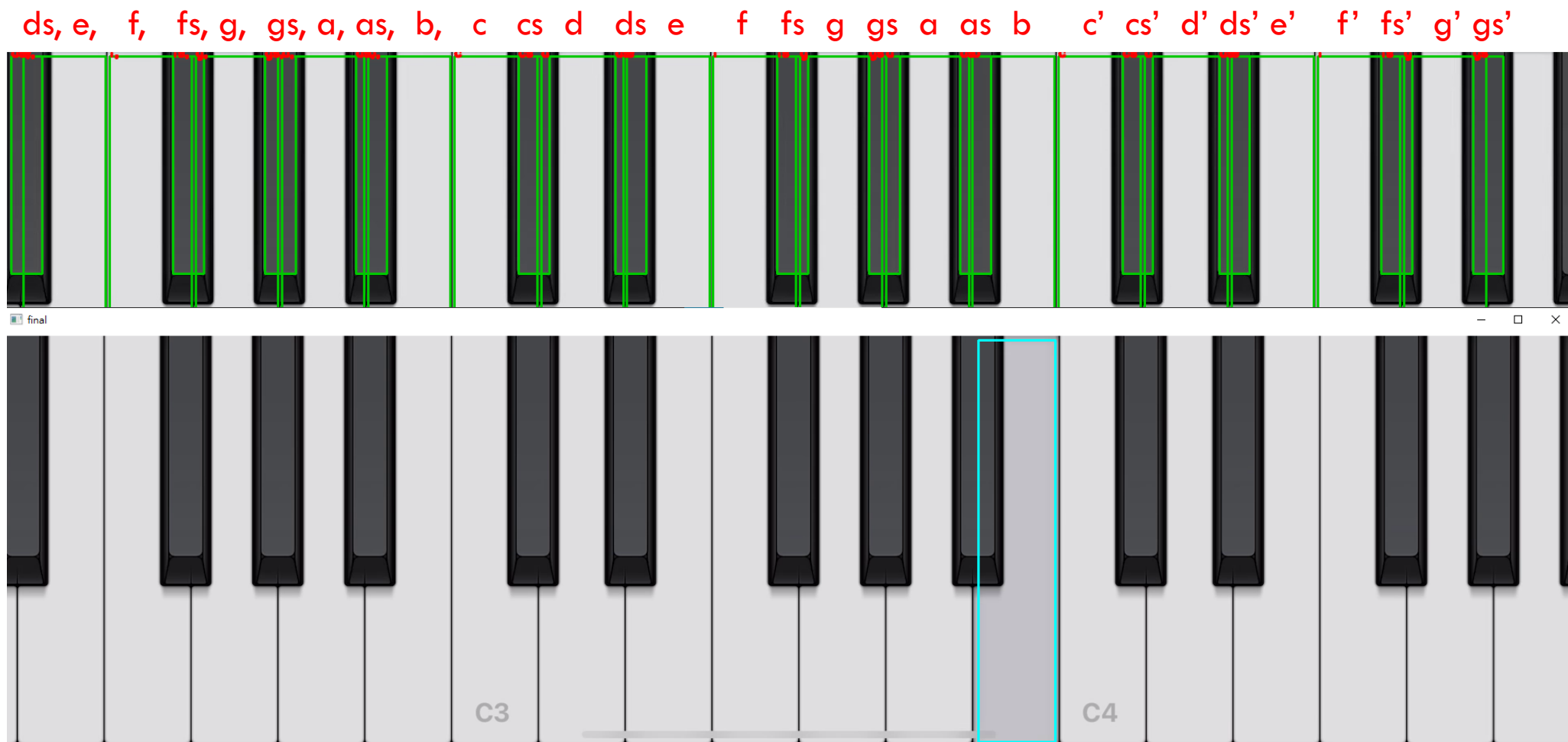
被按壓之琴鍵座標



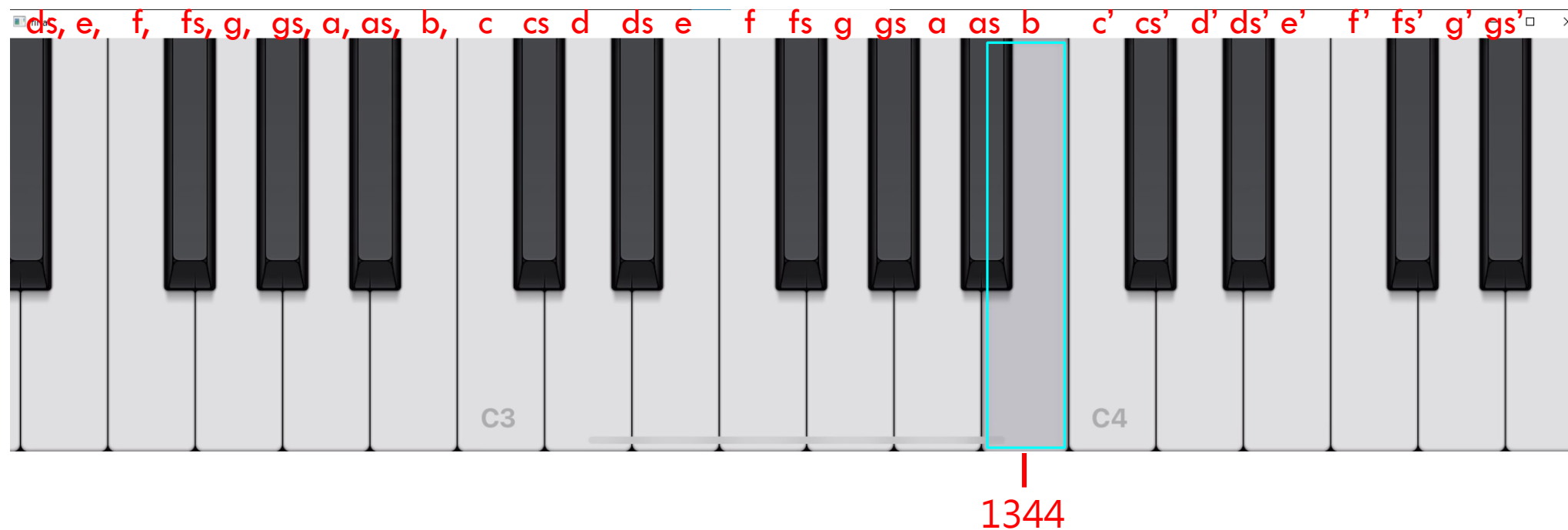
流程圖



比對座標



比對座標



按鍵座標[按鍵序,按鍵左座標,按鍵右座標]:

```
[[0, 5, 43], [1, 20, 121], [2, 126, 227], [3, 204, 242], [4, 232, 333], [5, 317, 355], [6, 338, 439], [7, 429, 467], [8, 444, 545], [9, 550, 652],  
[10, 629, 667], [11, 657, 758], [12, 748, 786], [13, 762, 864], [14, 868, 970], [15, 947, 985], [16, 975, 1076], [17, 1060, 1098], [18, 1081, 1183], [19, 1172, 1210], [20, 1187, 1288], [21, 1293, 1394], [22, 1372, 1409], [23, 1399, 1501], [24, 1491, 1528], [25, 1506, 1607], [26, 1612, 1713], [27, 1690, 1728], [28, 1718, 1819], [29, 1803, 1840]]
```

該幀數正被按下的琴鍵之中點座標:

```
[[1344.0]]
```

該幀數正被按下的琴鍵之按鍵序:

```
[21]
```

LILYPOND

按鍵紀錄整理[音名,幾分音符,起始幀數,同時按鍵數]:

```
[["e'", 8, '19', 1], ["ds' e'", 16, '27', 2], ["e'", 8, '34', 1], ["ds' e'", 8, '42', 2], ["e'", 8, '52', 1], ['b', 16, '62', 1], ["d'", 8, '71', 1], ["c'", 8, '81', 1], ['a', 2, '92', 1], ['c', 8, '117', 1], ['e', 8, '125', 1], ['a', 8, '134', 1], ['b', 2, '143', 1], ['e', 8, '170', 1], ['gs', 8, '181', 1], ['b', 8, '189', 1], ["c'", 2, '199', 1], ['e', 16, '221', 1], ["e'", 8, '228', 1], ["ds' e'", 16, '238', 2], ["e'", 8, '246', 1], ["ds' e'", 16, '257', 2], ["e'", 8, '265', 1], ['b', 16, '275', 1], ["d'", 8, '282', 1], ["c'", 8, '290', 1], ['a', 2, '299', 1], ['c', 8, '319', 1], ['e', 16, '328', 1], ['a', 8, '336', 1], ['b', 4, '346', 1], ['e', 8, '366', 1], ["c'", 8, '372', 1], ['b', 8, '379', 1], ['a', 4, '388', 1], ['b', 8, '409', 1], ["c'", 8, '417', 1], ["d'", 8, '425', 1], ["e'", 4, '432', 1], ['g', 8, '451', 1], ['f', 8, '458', 1], ["e'", 8, '465', 1], ["d'", 4, '472', 1], ['f', 8, '490', 1], ["e'", 16, '496', 1], ["d'", 16, '502', 1], ["c'", 4, '509', 1], ['e', 16, '527', 1], ["d'", 8, '533', 1], ["c'", 16, '540', 1], ['b', 2, '548', 1], ["e'", 8, '572', 1], ["ds' e'", 16, '580', 2], ["e'", 16, '587', 1], ["ds' e'", 16, '594', 2], ["e'", 8, '600', 1], ['b', 8, '608', 1], ["d'", 8, '615', 1], ["c'", 8, '622', 1], ['a', 4, '630', 1], ['c', 8, '651', 1], ['e', 8, '658', 1], ['a', 16, '665', 1], ['b', 4, '670', 1], ['e', 16, '688', 1], ["c'", 8, '694', 1], ['b', 8, '701', 1], ['a', 2, '709', 1], ["ds' as, d g a e", 16, '739', 6], ['fs,', 16, '740', 1], ['gs', 16, '741', 1], ['cs ds', 16, '742', 2], ['fs, ds gs', 16, '745', 3], ['fs,', 16, '748', 1], ['ds,', 16, '749', 1], ['fs,', 16, '751', 1], ['d', 16, '753', 1], ['ds', 16, '759', 1], ['fs,', 1, '760', 1]]
```



按鍵整理(lilypond格式):






```
["e'4", "ds'8", "e'4", "ds'4", "e'4", 'b8', "d'4", "c'4", 'a1', 'c4', 'e4', 'a4', 'b1', 'e4', 'gs4', 'b4', "c'1", 'e8', "e'4", "ds'8", "e'4", "ds'8", "e'4", 'b8', "d'4", "c'4", 'a1', 'c4', 'e8', 'a4', 'b2', 'e4', "c'4", 'b4', 'a1', 'b4', "c'4", "d'4", "e'1", 'g4', "f'4", "e'4", "d'2", 'f4', "e'8", "d'8", "c'2", 'e8', "d'4", "c'8", 'b1', "e'4", "ds'8", "e'8", "ds'8", "e'4", 'b4', "d'4", "c'4", 'a1', 'c4', 'e4', 'a8', 'b2', 'e8', "c'4", 'b4', 'a1', "<f, ds' as, d f g a e b>16", 'fs,16', 'gs8', '<cs ds>16', '<fs, ds gs>16', 'fs,16', 'ds,16', 'fs,8', 'd16', 'ds16', 'fs,1']
```

LILYPOND

按鍵整理(lilypond格式):

```
["e'4", "ds'8", "e'4", "ds'4", "e'4", 'b8', "d'4", "c'4", 'a1', 'c4', 'e4', 'a4', 'b1', 'e4', 'gs4', 'b4', "c'1", 'e8', "e'4", "ds'8", "e'4", "ds'8", "e'4", 'b8', "d'4", "c'4", 'a1', 'c4', 'e8', 'a4', 'b2', 'e4', "c'4", 'b4', 'a1', 'b4', "c'4", "d'4", "e'1", 'g4', "f'4", "e'4", "d'2", 'f4', "e'8", "d'8", "c'2", 'e8', "d'4", "c'8", 'b1', "e'4", "ds'8", "e'8", "ds'8", "e'4", 'b4', "d'4", "c'4", 'a1', 'c4', 'e4', 'a8', 'b2', 'e8', "c'4", 'b4', 'a1', "<f, ds' as, d f g a e b>16", 'fs,16', 'gs8', '<cs ds>16', '<fs, ds gs>16', 'fs,16', 'ds,16', 'fs,8', 'd16', 'ds16', 'fs,1']
```



 VSLilyPond	2021/11/17 上午 08:07	檔案資料夾	
 final.py	2021/11/17 下午 01:09	Python 來源檔案	12 KB
 lily.log	2021/11/17 下午 12:56	文字文件	2 KB
 lily.ly	2021/11/17 下午 01:52	LY 檔案	1 KB
 lily.pdf	2021/11/17 下午 12:56	Microsoft Edge P...	61 KB

LILYPOND

```
\version "2.22.0" \n \language "english" \n \relative{\n \clef "treble_8"e'4  
ds'8  
e'4  
ds'4  
e'4  
b8  
d'4  
c'4  
a1  
c4  
e4  
a4  
b1  
e4  
gs4  
b4  
c'1
```

Lily.ly的文字檔





成果 | 04

成果

本專題將展示兩種情況下之成果：

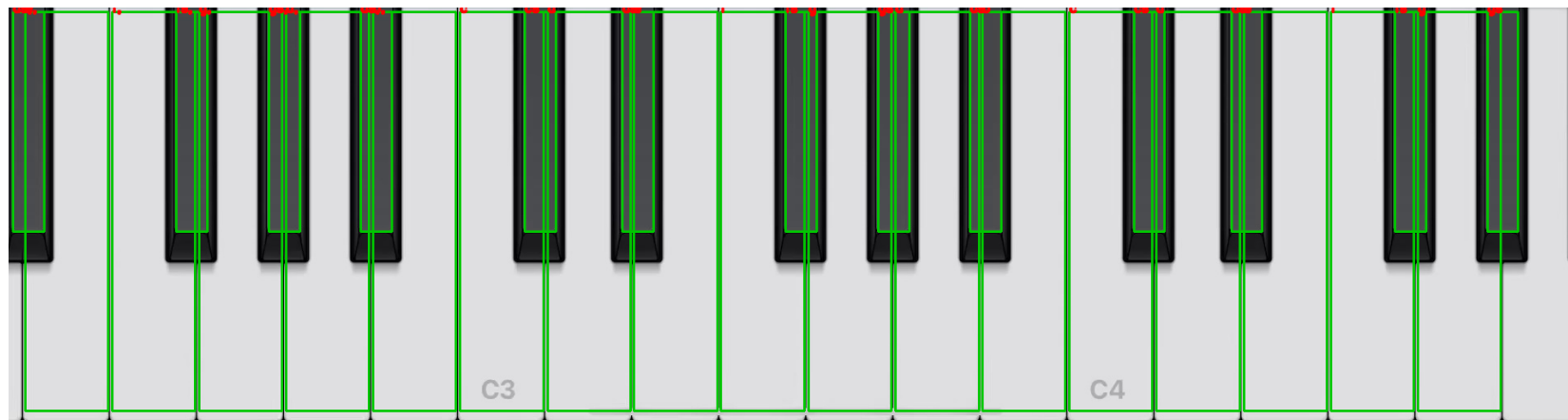
1. 有光影影響：

有光影環境為自行錄製或擷取youtube上彈奏鋼琴(視角為俯視)影片

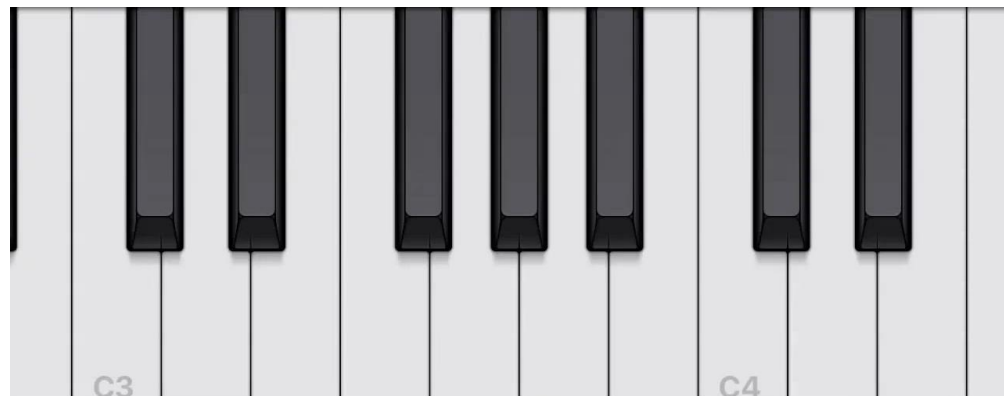
2. 無光影影響：

無光影環境則是用Garageband軟體進行人工彈奏並螢幕錄製擷取。

無光影環境—琴鍵邊緣偵測



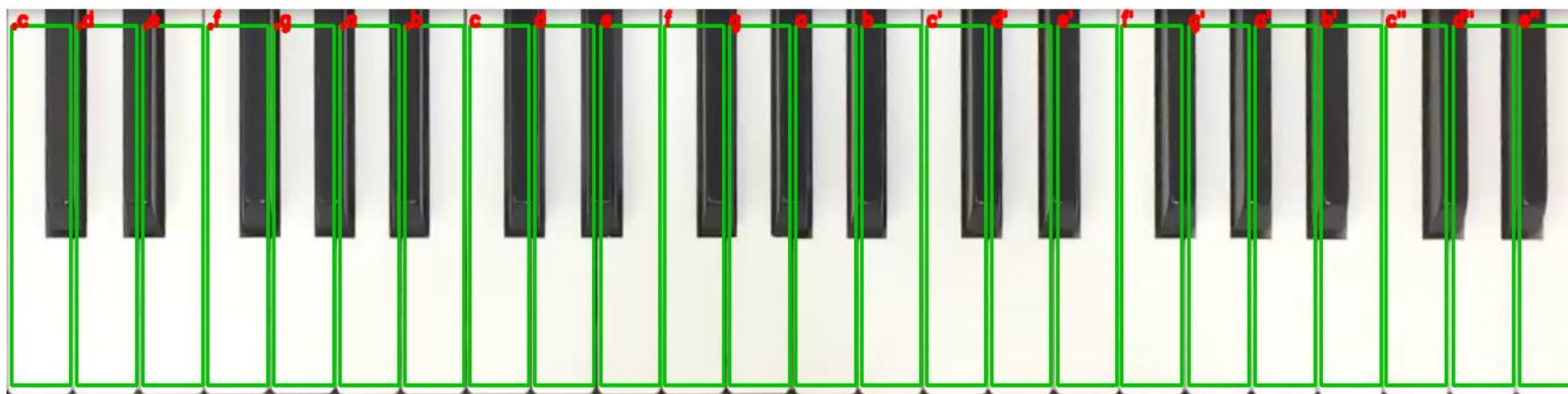
無光影環境—琴鍵按壓偵測



無光影環境-輸出樂譜



有光影環境-琴鍵邊緣偵測



有光影環境—琴鍵按壓偵測

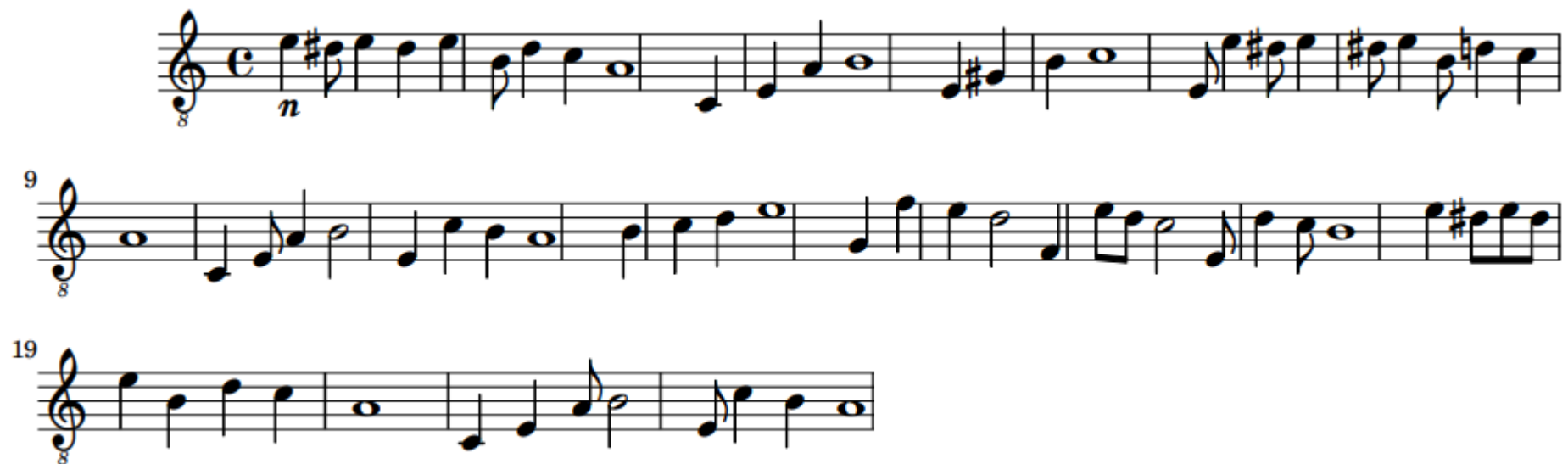


有光影環境-輸出樂譜



黒鍵偵測







未來展望 | 05

未來展望

本專題因光影問題，無法在真實鋼琴或youtube影片上準確判斷每個按壓琴鍵位置，若能夠突破這方面的問題，將能更廣泛運用在多種情況下。

另外在模擬鋼琴鍵盤app上，本專題採用螢幕錄製方法並人工彈奏曲子，放進程式判斷按壓琴鍵位置，雖然判斷準確，卻有節奏微小誤差的問題(人工錄製)，若可以找出演算法協助彌補這類人工誤差，將可使準確率提高至95%以上。

Q&A

Thank you