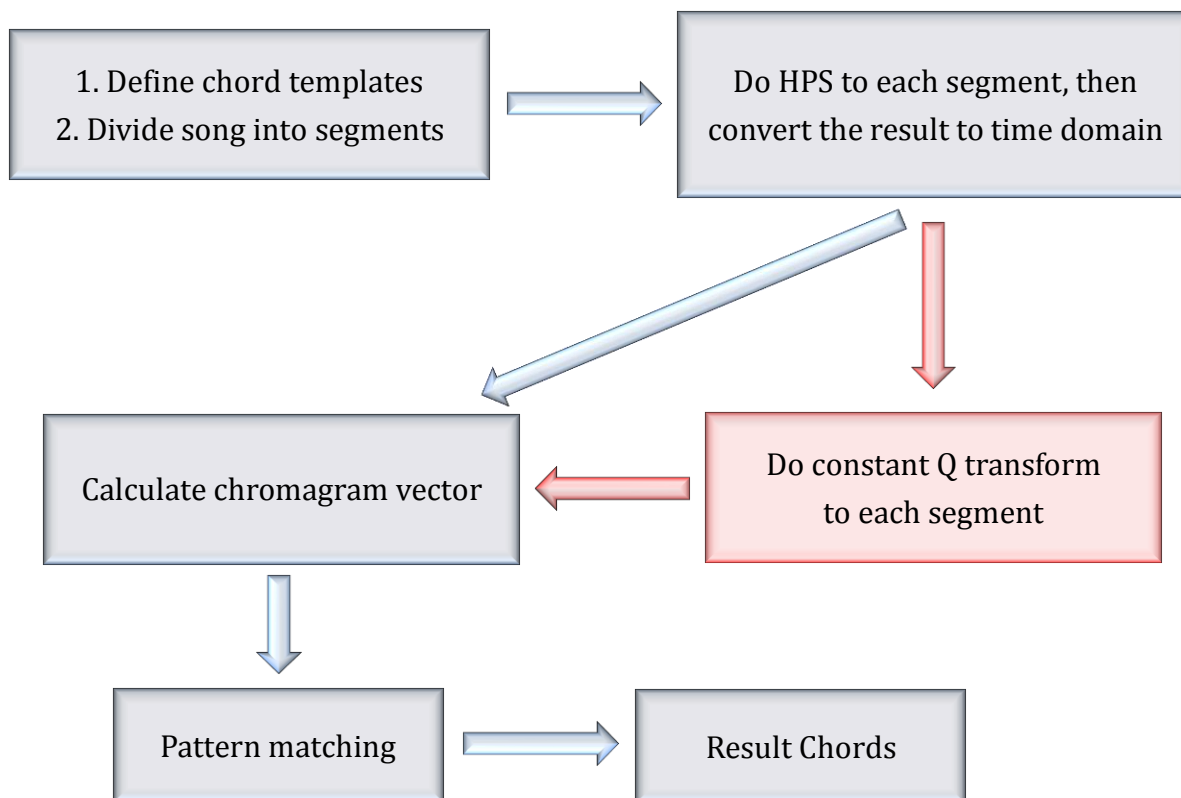


多媒體技術概論 Final Project Report

Automatic Chord Transcription

Implementation Procedure



Template Defining

我們定義的和弦種類為大三和弦(X)、小三和弦(Xmin)。

chordTemplate 為儲存和弦模板的矩陣，其內容大致如下：

	C	C#	D	Eb	E	F	F#	G	G#	A	Bb	B
C	1	0	0	0	1	0	0	1	0	0	0	0
C#~B											
Cmin	1	0	0	1	0	0	0	1	0	0	0	0
C#min~Bmin											

先預訂好每個和弦種類的第一列再用迴圈去產生其他和弦的模板，並在最後把每一種和弦的模板合併成上表。

```

for family = 2:12
    for bin = 1:12
        i = bin-1;
        if i==0, i = 12; end
        X(family, bin) = X(family-1, i);
        Xmin(family, bin) = Xmin(family-1, i);
    end
end

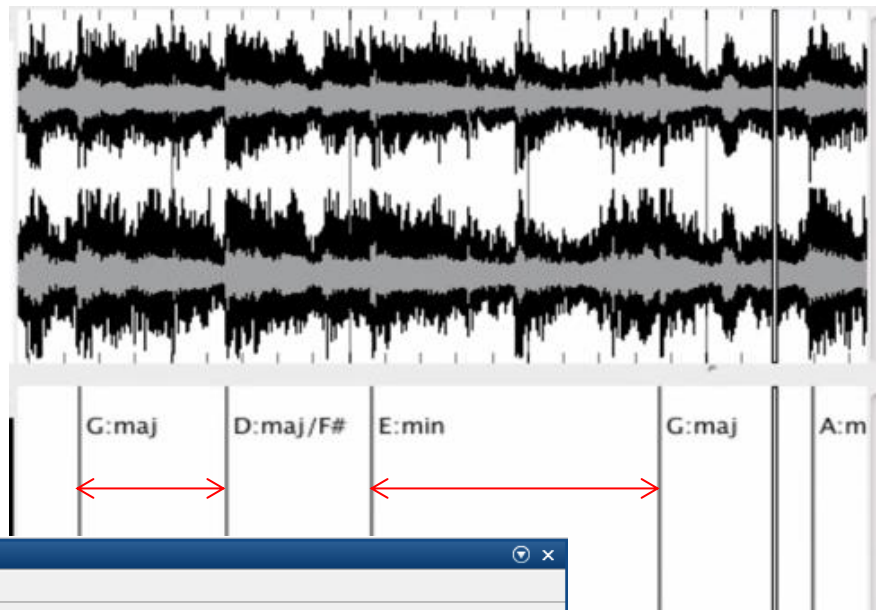
chordTemplate = [X;Xmin];
chordKey = {'C', 'C#', 'D', 'D#', 'E', 'F', 'F#', 'G', 'G#', 'A', 'A#', 'B'};
  
```

Measure Division

我們參考了 <http://schall-und-mauch.de/artificialmusicality> 切割歌曲的方式，以 $1/2$ 小節為單位將歌曲分割成好幾段，分析出每個小節可能的和弦，由於其基本單位為 $1/2$ 小節，故理論上和弦的分析能夠更為準確。此網站提供了完整的轉錄和弦的程式碼，我們只擷取了分割歌曲的程式碼片段，並進行修改及簡化(即為程式碼中 `beattracker` 資料夾)。

其分割方式為計算每個拍子的秒數、每個小節的秒數區間、每個小節的拍數，分析原理大概為依照音量、音量和過零率去抓取音量忽大後漸小的區間作為基本單位。

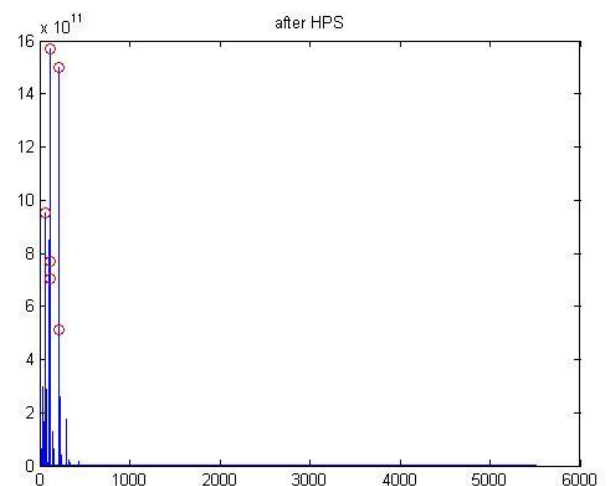
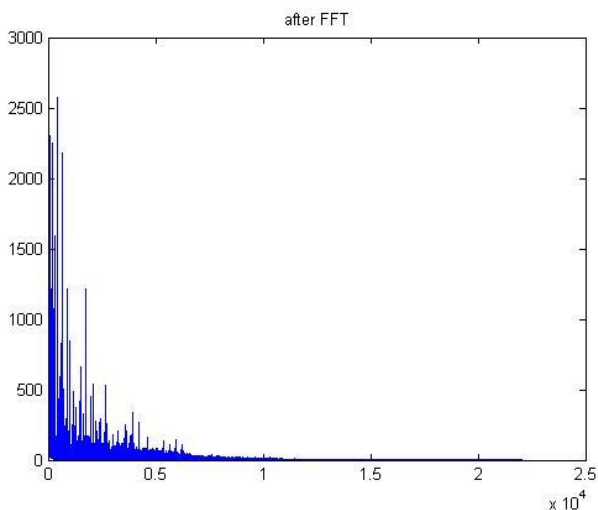
下表即為用此段程式碼得到的各小節的起始與結束秒數，可以依此為切割歌曲的基準進行和弦分析，則可得出各個小節可能的和弦。



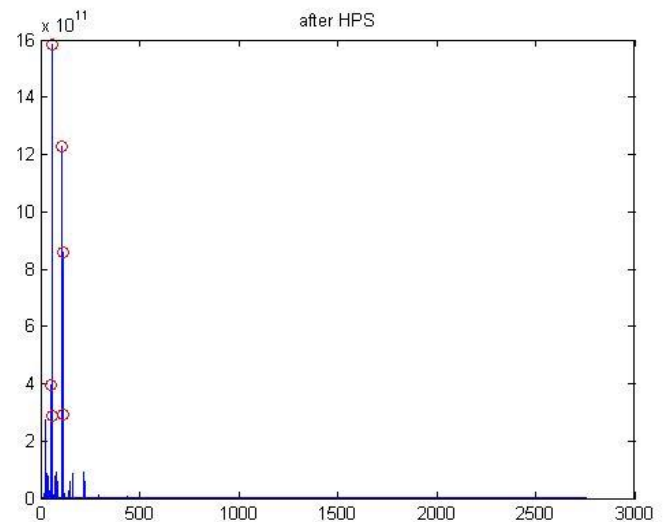
Variables - interval.measures										
interval.measures <1x199 double>										
	1	2	3	4	5	6	7	8	9	10
1	1.5209	2.6006	3.7152	4.7833	5.8746	6.9776	8.0689	9.1603	10.2516	11.34

Harmonic Product Spectrum (HPS)

HPS 的功能在於分析某段音訊中主要出現的頻率，將一段音訊轉換成頻域之後對其進行每 2、3、4 點向下取樣(downsampling)，即可得到壓縮 2、3、4 倍之資料，再將這些壓縮後的資料與原始資料相乘，就能得到此段音訊中最主要的幾個基頻。左下圖即為將音訊轉換為頻域後的結果，而右下圖則是經過 HPS 處理之後的結果，並用紅色圓圈標示六個振幅最大點。



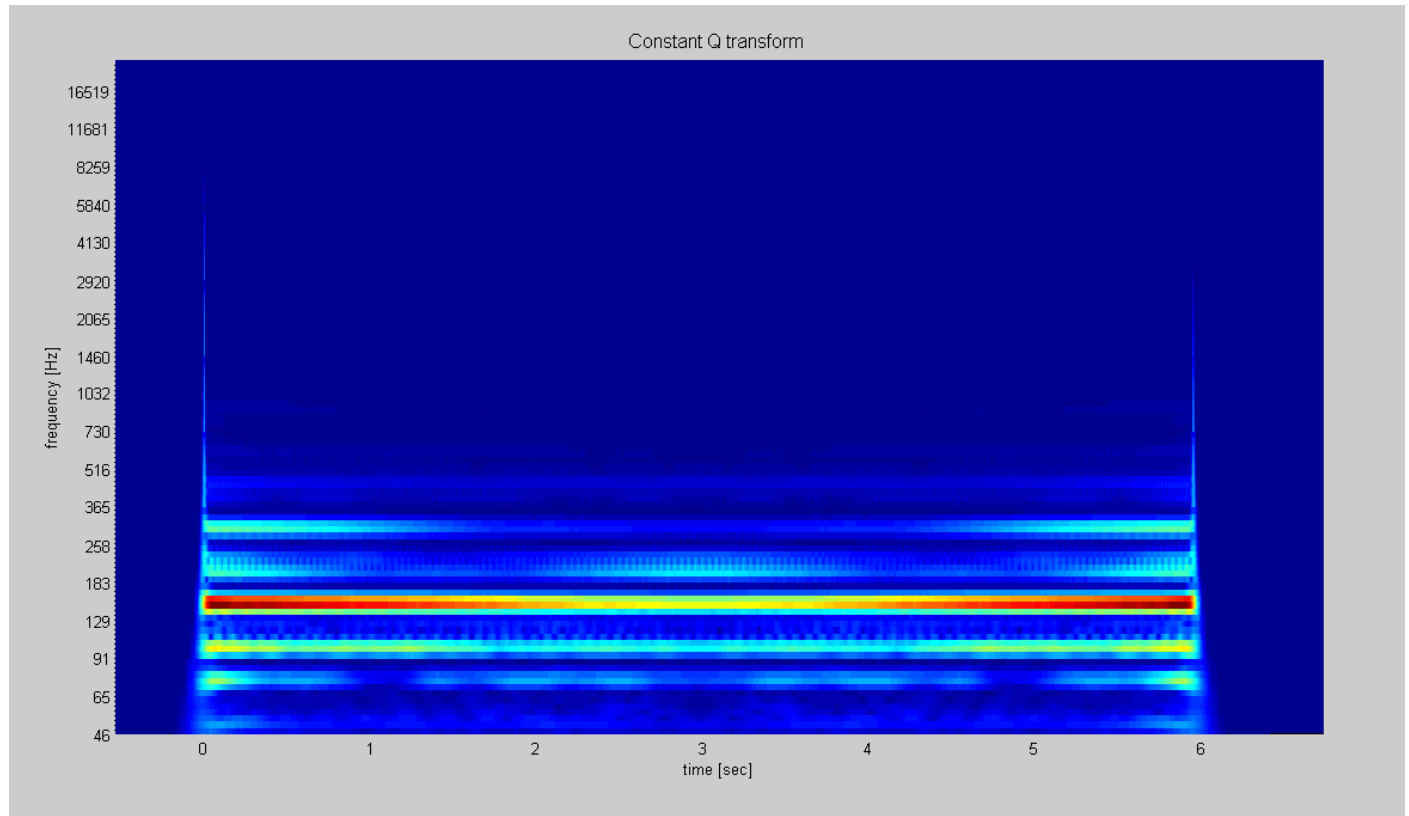
我們也嘗試了 2、4、8 的 downsampling rate，因為高八度的音為基頻的兩倍，即皆為二的次方倍數，以二的次方去做 downsampling 可以取得較精準，也比較可以除去一些不重要的頻率。我們最後決定用 2、4、8 的 downsampling rate。



✳Constant Q Transform

我們採用了 <http://www.eecs.qmul.ac.uk/~anssik/cqt> 的 toolbox，將做完 HPS 的資料用 ifft 從頻域轉回時域，然後做 constant Q transform 將其轉換成時頻圖如下，可以顯示在一段時間內各種頻率的強度，顏色越紅則強度越強，越偏藍則強度越弱，可以很直接看出此時間內主要頻率為何。

但我們實作的結果不甚理想，在計算 chromagram vector 的正確率極低，雖然 HPS 的幾個高峰值都是目標和弦的組成音，但在做完 constant Q transform 之後就錯了，我們猜測可能是在 ifft 轉換的時候出些問題，因為沒能解決，所以我們用另一個較土法煉鋼的方式取代 constant Q transform。



Chromagram Vector Calculation

因為 HPS 的結果是正確的，我們想要用從 HPS 得到的高峰值對應到的頻率範圍的強度總和，直接計算 chromagram vector。

因為頻率不會是單一值，而是在某個範圍內，所以要先訂定C的範圍、D的範圍等。C的頻率為16.352，而C#的頻率為17.324，所以我們取中間值作為分界：若頻率落在15.892~16.838則為C，若頻率落在16.838~17.839則為C#，下圖即為建立的分界表：

Variables - range													
range													
range <8x13 double>													
	1	2	3	4	5	6	7	8	9	10	11	12	13
1	15.8920	16.8380	17.8390	18.8990	20.0230	21.2140	22.4760	23.8120	25.2280	26.7280	28.3170	30	31.7850
2	31.7840	33.6760	35.6780	37.7980	40.0460	42.4280	44.9520	47.6240	50.4560	53.4560	56.6340	60	63.5700
3	63.5680	67.3520	71.3560	75.5960	80.0920	84.8560	89.9040	95.2480	100.9120	106.9120	113.2680	120	127.1400
4	127.1360	134.7040	142.7120	151.1920	160.1840	169.7120	179.8080	190.4960	201.8240	213.8240	226.5360	240	254.2800
5	254.2720	269.4080	285.4240	302.3840	320.3680	339.4240	359.6160	380.9920	403.6480	427.6480	453.0720	480	508.5600
6	508.5440	538.8160	570.8480	604.7680	640.7360	678.8480	719.2320	761.9840	807.2960	855.2960	906.1440	960	1.0171e+03
7	1.0171e+03	1.0776e+03	1.1417e+03	1.2095e+03	1.2815e+03	1.3577e+03	1.4385e+03	1.5240e+03	1.6146e+03	1.7106e+03	1.8123e+03	1920	2.0342e+03
8	2.0342e+03	2.1553e+03	2.2834e+03	2.4191e+03	2.5629e+03	2.7154e+03	2.8769e+03	3.0479e+03	3.2292e+03	3.4212e+03	3.6246e+03	3840	4.0685e+03

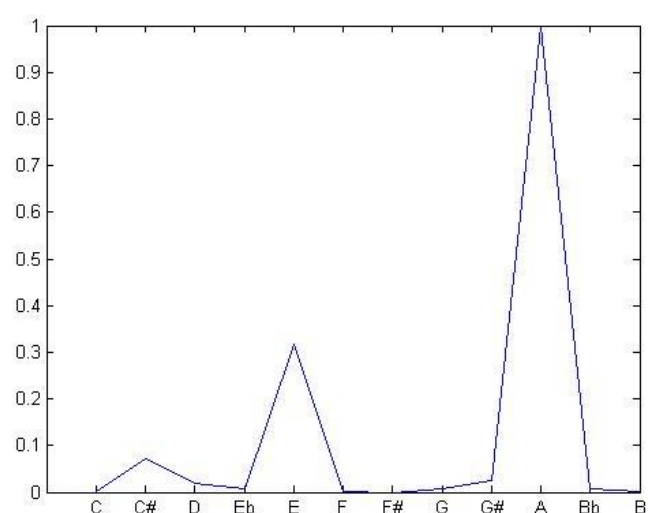
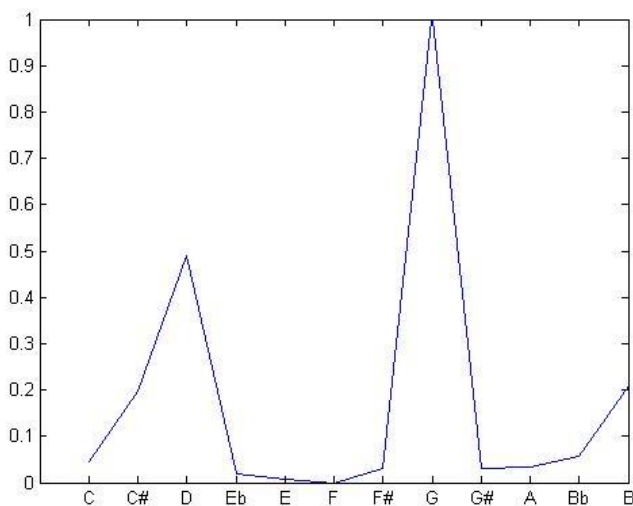
由此表就可以從HPS的結果中找到屬於C的所有頻率的振幅，並全部加總存入CH(1)，直到CH(1)~Ch(12)都建立，即完成chromagram vector。

```

for b = 1:12
    CH(b) = 0;
    for m = 1:8
        target = find(freq>=range(m, b) & freq<=range(m, b+1));
        CH(b) = CH(b) + sum(amp(target))/size(target, 2);
    end
end

```

左下圖為G和弦的chromagram，右下圖則為A和弦的chromagram。



Pattern Matching

將上一步得出的CH vector先normalize到0~1之間，然後去跟chordTemplate中的每一個模板做correlation，結果值最大的即為所求，再依照出現最大值的模板的index去輸出結果。

```
CH = (CH-min(CH))./(max(CH)-min(CH));

%figure, plot(CH);
%set(gca, 'XTick', 1:12);
%set(gca, 'XTickLabel', chordKey);

for i=1:24
    Chord(i) = CH*chordTemplate(i, :);
end

[~, maxInd] = max(Chord);
outNum(c)=maxInd;
```

```
outChord = '';
outChord{1} = 'N';
for c=2:N-1
    if mod(outNum(c), 12)==0, outChord{c} = chordKey{12};
    else outChord{c} = chordKey{mod(outNum(c), 12)};
    end
    if( outNum(c)>12 ) outChord{c} = strcat(outChord{c},':min');
    else outChord{c} = strcat(outChord{c},':maj');
    end
end
outChord{N} = 'N';
```

最後輸出如右圖：

```
outChord =

Columns 1 through 10

    'N'    'D:min'    'G:maj'    'D:min'    'G:min'    'A:maj'

Columns 11 through 19

    'G:maj'    'D:maj'    'G:min'    'A:maj'    'G:maj'    'D:min'
```

Implementation Result

除了我們自身演算法的誤差，以及只定義了基本模板之外，主要影響正確率的因素皆來自答案格式的不一致性：

1. 有些大三和弦並無輸出maj => LetLoveBeYourEnergy
2. 有些和弦查不到其結構 => LoveCallingEarth (Cmaj/b7)
3. 有些和弦是相通的但並沒有標準輸出 => D^b=C[#]
4. 有些和弦以不合理的格式出現 => EgoAGoGo (C^b)

以下為兩大組測資之結果：

	Result2002.txt	Result2005.txt
1	The 2002-Escapology/01-How PeculiarGTChords.txt	current rate:69.148936 %
2	The 2002-Escapology/02-FeelGTChords.txt	current rate:78.813559 %
3	The 2002-Escapology/03-Something BeautifulGTChords.txt	current rate:70.491803 %
4	The 2002-Escapology/04-MonsoonGTChords.txt	current rate:93.750000 %
5	The 2002-Escapology/05-Sexed UpGTChords.txt	current rate:75.652174 %
6	The 2002-Escapology/06-Love SomebodyGTChords.txt	current rate:64.912281 %
7	The 2002-Escapology/07-RevolutionGTChords.txt	current rate:56.896552 %
8	The 2002-Escapology/08-Handsome ManGTChords.txt	current rate:13.669065 %
9	The 2002-Escapology/09-Come UndoneGTChords.txt	current rate:98.809524 %
10	The 2002-Escapology/10-Me and My MonkeyGTChords.txt	current rate:75.675676 %

	Result2002.txt	Result2005.txt
1	The 2005-Intensive Care (Special Edition)/01-GhostsGTChords.txt current rate:86.000000 %	
2	The 2005-Intensive Care (Special Edition)/02-TrippingGTChords.txt current rate:74.766355 %	
3	The 2005-Intensive Care (Special Edition)/03-Make Me PureGTChords.txt current rate:86.153846 %	
4	The 2005-Intensive Care (Special Edition)/04-Spread Your WingsGTChords.txt current rate:55.813953 %	
5	The 2005-Intensive Care (Special Edition)/05-Advertising SpaceGTChords.txt current rate:95.945946 %	
6	The 2005-Intensive Care (Special Edition)/06-Please Don't DieGTChords.txt current rate:75.000000 %	
7	The 2005-Intensive Care (Special Edition)/07-Your Gay FriendGTChords.txt current rate:74.218750 %	
8	The 2005-Intensive Care (Special Edition)/08-Sin Sin SinGTChords.txt current rate:81.012658 %	
9	The 2005-Intensive Care (Special Edition)/09-Random Acts Of KindnessGTChords.txt current rate:47.586207 %	
10	The 2005-Intensive Care (Special Edition)/10-The Trouble With MeGTChords.txt current rate:72.173913 %	

Program Execution

1. 將beattracker及cqttoolbox以及其子目錄加入目前路徑

執行指定測試資料：

開啟finalProject.m，將InputAudioGTChords.txt換成要進行找尋和弦程式的時間分割檔，如: GhostsGTChords.txt，並且將InputAudio.mp3換成要進行找尋和弦程式的音檔，如:Ghosts.mp3，以及檔案最下方的InputAudio.txt改成輸出和弦的檔案名稱，如: Ghosts.txt。

執行其他測試資料：

開啟finalProject.m，將15~19註解，將11~12、21~25行去掉註解，並在audioread及getmeasures2裡面輸入欲分析的音檔。

2. 改好上述的地方後，執行程式，會輸出一個txt檔，如: Ghosts.txt，裡頭資料便是經由程式找尋到的和弦。

Reference

<http://stackoverflow.com/questions/19765486/matlab-code-for-harmonic-product-spectrum>

<http://cnx.org/content/m11714/latest/>

<http://chur.chu.edu.tw/bitstream/987654321/577/1/GM094020080.pdf>

http://en.wikipedia.org/wiki/Harmonic_pitch_class_profiles

<http://schall-und-mauch.de/artificialmusicality/2014/04/matlab-chord-transcription-code-dbn/>

http://lac.linuxaudio.org/2009/cdm/Saturday/21_Fugal/21.pdf

<http://hans.fugal.net/research/cq-octave/cq.m>

<http://www.ee.columbia.edu/~dpwe/resources/matlab/sgram/>

<http://robotics.ee.uwa.edu.au/theses/2012-MusicTranscription-Gouws.pdf>

<http://www.eecs.qmul.ac.uk/~anssik/cq>

<http://home.deib.polimi.it/zanoni/Site/Research.html>

<https://ccrma.stanford.edu/~kglee/pubs/klee-icmc06.pdf>