Music Visualization

For this assignment you will produce a two-dimensional visualization of major and minor chords that places similar chords near each other and dissimilar chords far away from each other.

Theory

The pitch of a musical segment in the Echo Nest track analysis consists of 12 numbers which are the relative amounts of each of the 12 tones regardless of octave. If we wanted to produce audio that “sounded like” the pitch vector, we could use Shepard tones. Check out this auditory illusion of constantly increasing pitch using Shepard tones. Instead of listening to them, we want to visualize them. The Infinite Jukebox uses Euclidean distance to compare pitches in different segments. So, it seems natural that similar pitches should be closer to each other and dissimilar pitches should be farther apart in the Euclidean sense. The trouble is that this is happening in 12-dimensional space which is hard to see. Fortunately, there are dimensionality reduction techniques that transform the input dimensions into a smaller set of dimensions attempting to preserve the distance information. In particular, there is a known relationship between the major and minor keys that you will visualize in this assignment.

Music theory already gives us some clues about what keys are similar. The circle of fifths places the major and minor keys in a specific order that places similar keys next to each other (circle of fifths.svg). Here’s a look at the ordering of notes on the piano:

C	C#	D	Eb	E	F	F#	G	Ab	A	Bb	B

The most similar keys are those that are separated by five steps. For example, C’s neighbors are five steps up (F) and five steps down (G). Stepping through the sequence this way produces the circle of fifths:

Major:	C	G	D	A	E	B	F#	C#	Ab	Eb	Bb	F	C	…
Minor:	A	E	B	F#	C#	G#	Eb	Bb	F	C	G	D	A	…

Data

Simple representation (S-patterns)

Marc Leman describes an *S-representation* that indicates which Shepard tones are present in a chord:

S-patterns													
Abbrev.	Name	C	C#	D	Eb	E	F	F#	G	Ab	A	Bb	B
C	Single note	1	0	0	0	0	0	0	0	0	0	0	0
CM	Major triad	1	0	0	0	1	0	0	1	0	0	0	0
Cm	Minor triad	1	0	0	1	0	0	0	1	0	0	0	0

Residue representations (R-patterns)

Because musical instruments usually contain harmonics (not pure frequencies), when a C note is played its harmonics also contribute to the pitches present. For example, when a C note is played, Leman gives C a weight of 1.00, its fifth (F) will contribute about half as much (0.50), its major third (Ab) contributes about 0.33 [For details see the section below on “How do you get these weights”]. For a C note, Leman uses the following weights:

Subharmonic-sum Table							
Abbrev.	Name	Octave (C)	Fifth (F)	Major third (Ab)	Minor seventh (D)	Major second (Bb)	Minor third (A)
C	Single note	1.00	0.50	0.33	0.25	0.20	0.10

Given an S-pattern, we can produces something like an Echo Nest pitch vector by summing up the contributions for each note that is present. For example see the R-patterns below:

R-patterns													
Abbrev.	Name	C	C#	D	Eb	E	F	F#	G	Ab	A	Bb	B
C	Single note	1.00	0.00	0.25	0.00	0.00	0.50	0.00	0.00	0.33	0.10	0.20	0.00
CM	Major triad	1.83	0.10	0.45	0.33	1.10	0.70	0.25	1.00	0.33	0.85	0.20	0.00
Cm	Minor triad	1.60	0.20	0.25	1.33	0.10	0.95	0.00	1.00	0.83	0.35	0.20	0.33

Normalized representation (O-patterns)

Notice that these representations are not normalized so they don’t look very much like the Echo Nest pitch vectors that have a maximum value of 1.00. In order to compare pitch vectors that could have large magnitude differences, Leman normalizes each vector using the following equation:

$$O_i = \frac{R_i}{R_{\max}} \times \frac{1}{\sqrt{\sum_{i=1}^{12} \frac{R_i}{R_{\max}}}}$$

O-patterns													
Abbrev.	Name	C	C#	D	Eb	E	F	F#	G	Ab	A	Bb	B
C	Single note	0.65	0.00	0.16	0.00	0.00	0.32	0.00	0.00	0.21	0.06	0.13	0.00
CM	Major triad	0.51	0.03	0.12	0.09	0.30	0.19	0.07	0.28	0.09	0.24	0.06	0.00
Cm	Minor triad	0.47	0.06	0.07	0.39	0.03	0.28	0.00	0.30	0.25	0.10	0.06	0.10

Now any pitch vector can be normalized to be comparable to the O-patterns.

Input

You will visualize the major triad and minor triad for every pitch (24 total O-patterns). You can form the patterns for pitches other than C by circularly shifting the values.

Self-organizing maps

An overview of self-organizing maps is provided here. For this assignment, you will implement the following pseudocode to build your SOM:

```
1. Initialize W with values from a uniform distribution on [0, 1)
2. for s = 1 to λ
     a. for i = 1 to m
          i.  Select a random input vector, D(t)
         ii.  Use the Euclidean distance to find the node with the
              closest weights to D(t), the so-called best matching
              unit, u.
        iii.  Update the nodes in the neighborhood of the BMU with
              the following equation:
```

$$W(v) = W(v) + \theta(u, v, s)\alpha(s)(D(t) - W(v))$$

Symbol	Name	Value
m	Number of inputs	24
n	Size of a side of the square map	20
W	The map	3d array size = (n, n, 12)
W(v)	The weight vector of node v	Any of the 400 12-dimensional weight vectors
t	Index of the target input vector	1 to m
D(t)	Target input vector	Any of the m input vectors
s	Current iteration	1 to λ
λ	Iteration limit	360
c	Neighborhood radius decrement period	20
r	Initial neighborhood radius	n – 1
v	Index of a node in the map	$\{(i, j) \mid 1 \le i, j \le m\}$
u	Index of the best matching unit	$\{(i, j) \mid 1 \le i, j \le m\}$
d(u,v)	Distance between the (i,j) locations of u and v on the surface of a torus.	The distance is computed on a torus, so the top row connects to the bottom row and the left column connects to the right column.

θ(u,v,s)	Neighborhood function	$\theta(u,v,s)=e^{-\frac{d(u,v)^2}{2\sigma(s)^2}}$
α(s)	Learning rate	0.02
σ(s)	Standard deviation for neighborhood function	$\sigma(s)=\frac{1}{3}\left(r-\frac{s}{c}\right)$

What are all the S-patterns?

S-patterns													
Abbrev.	Name	C	C#	D	Eb	E	F	F#	G	Ab	A	Bb	B
CM	Major triad	1	0	0	0	1	0	0	1	0	0	0	0
Cm	Minor triad	1	0	0	1	0	0	0	1	0	0	0	0
Co	Diminished triad	1	0	0	1	0	0	1	0	0	0	0	0
C+	Augmented triad	1	0	0	0	1	0	0	0	1	0	0	0
CM7	Major seventh chord	1	0	0	0	1	0	0	1	0	0	0	1
Cm7	Minor seventh chord	1	0	0	1	0	0	0	1	0	0	1	0
Cx7	Dominant seventh chord	1	0	0	0	1	0	0	1	0	0	1	0
C07	Half diminished seventh chord	1	0	0	1	0	0	1	0	0	0	1	0
C+7	Augmented seventh chord	1	0	0	0	1	0	0	0	1	0	0	1
Cm-7	Minor with major seventh chord	1	0	0	1	0	0	0	1	0	0	0	1
Co7	Diminished seventh chord	1	0	0	1	0	0	1	0	0	1	0	0
Note: other keys can be obtained by circular shifting the S-representation (115 unique chords)													

How do you get these weights?

Explanation of weights from Marc Leman, “Music and Schema Theory,” Springer Series on Information Sciences, 1995, page 47.

3 The weights have been adapted to be in agreement with [5.13]. According to R. Parncutt, the minor third can support the root because its third and fifth harmonics (the perfect fifth and major third) are octave-equivalent to the root's seventh and third harmonics (minor seventh and perfect fifth). More recently [5.15], this reasoning is not followed anymore since it applies also for other intervals such as the major sixth. Actually, different weights are conceivable but the above set is also related to the one used in TAM. In TAM the weights W of the subharmonics are inverse proportional to the harmonic number ($W_i = 1/i$, for $i = 1$ to 15). In *SAM*, the weights that occur at different octaves are added. The octave-intervals then become $1/1 + 1/2 + 1/4 + 1/8 = 15/8$. Similarly, the total weight of components at fifth-intervals is $1/3 + 1/6 = 1/2$, while the components at major third-intervals is $1/5 + 1/10 = 3/10$. The obtained values are: 15/8 (added octaves), 1/2 (added fifths), 3/10 (added major thirds), 1/7 (only one minor seventh), and 1/9 (only one major second). Normalized (divide by 15/8) this gives: 1.0, 0.27, 0.16, 0.08, and 0.06 respectively. According to [5.13], an intuitively more reasonable set of values is obtained by raising the numbers to the power of 0.55. One then obtains the values depicted in Table 5.1. The weights

5.13 R. Parncutt: Revision of Terhardt's psychoacoustical model of the roots of a musical chord. *Music perception* **6**, 65-94 (1988)

5.15 R. Parncutt: Template –matching mdoels of musical pitch and rhythm perception. *J. New Music Research* **23**, 145-167 (1994)