Music generation with note similarity

Orges Leka[1]

[1] Germany, Limburg; orges.leka@gmail.com

Abstract

In this note we will describe a novel way to measure musical note similarity using a function between musical notes which tries to capture the perception of consonant sounding notes with a function. Having defined this similarity function between notes, we describe a potential algorithm for music generation, which uses this function.

Introduction

The use of the twelfth root of two in music is very well known in Western music (see for example [1]). It is known in music theory that two successive pitches a, b which sound "consonant" or "nice" if some ratio B/A is "simple". The notion of simplicity has not been defined precisely, and we will give a possible notion here:

Let $\alpha = 2^{\frac{1}{12}}$, $p_1 = \alpha^{k_1}, p_2 = \alpha^{k_2}$ where $0 \leq k_1, k_2 \leq 127$ are the midi pitches. We define the similarity between p_1 and p_2 to be:

$$K_p(k_1, k_2) = \frac{\gcd(a,b)^2}{ab}$$

where a = numerator of a rational approximation of $\alpha^{k_1 - k_2}$ and b = denominator of a rational approximation of $\alpha^{k_1-k_2}$. We argue that this similarity could capture when two pitches have a "simple" ratio and hence will sound "nice" together or when played in successive order. We look at the following matrix as an example:

$$\begin{pmatrix} \texttt{C4} & \texttt{C\#4} & \texttt{D4} & \texttt{E-4} & \texttt{E4} & \texttt{F4} & \texttt{F\#4} & \texttt{G4} & \texttt{G\#4} & \texttt{A4} & \texttt{B-4} & \texttt{B4} \\ 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ 1 & \frac{16}{15} & \frac{9}{8} & \frac{6}{5} & \frac{5}{4} & \frac{4}{3} & \frac{17}{12} & \frac{3}{2} & \frac{8}{5} & \frac{5}{3} & \frac{16}{9} & \frac{15}{8} \\ 1 & \frac{1}{240} & \frac{1}{72} & \frac{1}{30} & \frac{1}{20} & \frac{1}{12} & \frac{1}{204} & \frac{1}{6} & \frac{1}{40} & \frac{1}{15} & \frac{1}{144} & \frac{1}{120} \end{pmatrix}$$

In the first row are the pitch names and octaves. In the second row is the pitch difference $k_1 - k_2$. In the third row is a rational approximation of $\alpha^{k_1 - k_2}$. In the last row is the similarity measure $K_p(k_1, k_2)$ where $k_2 = 0$. This similarity measure goes from 0 to 1. A larger value indicates a larger similarity. We sort the matrix above by the last row, similarity measure, to look how the rational approximation changes:

$$\begin{pmatrix} \texttt{C4} & \texttt{G4} & \texttt{F4} & \texttt{A4} & \texttt{E4} & \texttt{E-4} & \texttt{G\#4} & \texttt{D4} & \texttt{B4} & \texttt{B-4} & \texttt{F\#4} & \texttt{C\#4} \\ 0 & 7 & 5 & 9 & 4 & 3 & 8 & 2 & 11 & 10 & 6 & 1 \\ 1 & \frac{3}{2} & \frac{4}{3} & \frac{5}{3} & \frac{5}{4} & \frac{6}{5} & \frac{8}{5} & \frac{9}{8} & \frac{15}{8} & \frac{16}{9} & \frac{17}{12} & \frac{16}{15} \\ 1 & \frac{1}{6} & \frac{1}{12} & \frac{1}{15} & \frac{1}{20} & \frac{1}{30} & \frac{1}{40} & \frac{1}{72} & \frac{1}{120} & \frac{1}{144} & \frac{1}{204} & \frac{1}{240} \end{pmatrix}$$

We see in the matrix above that 7-semitonse distance (for example from C4 to G4) are more "consonant" by the above definition than a 5-semitones distance (for example from C4 to F4) . The most "dissonant" is a semitone distance realised by 1 being the last number in the first row (for example from C4 to C#4) .

Since a pitch alone does not describe a note, we have also defined similarity measures for duration, volume and if it is a rest or not:

Herefore we make use of the Jaccard-similarity of two intervals:

$$J(A,B) = \frac{\mu(A \cap B)}{\mu(A \cup B)} = \frac{\min(a,b)}{\max(a,b)}$$

where $A = [0,a], B = [0,b]$ are closed intervals and $a, b > 0$ and $\mu([x,y]) = y - x$.

Using J we define the duration similarity:

$$K_d(d_1,d_2) = J([0,d_1],[0,d_2]) = \frac{\min(d_1,d_2)}{\max(d_1,d_2)}$$

for two durations d_1, d_2 given as multiple of quarter notes. And similarily we define the volume similarity as :

$$K_v(v_1,v_2) = J([0,v_1],[0,v_2]) = \frac{\min(v_1,v_2)}{\max(v_1,v_2)}$$

for $0 \leq v_1, v_2 \leq 127$ giving the volumes in midi notation. For rests we take the similarty = 0 if one is not a rest and the other is, or = 1 if both are no rests or both are rests.

Having two notes $n_1 = (p_1, d_1, v_1, r_1), n_2 = (p_2, d_2, v_2, r_2)$ we define a similarity between them as:

$$K(n_1,n_2) = \alpha_p K_P(p_1,p_2) + \alpha_d K_d(d_1,d_2) + \alpha_v K_v(v_1,v_2) + \alpha_r K_r(r_1,r_2)$$

where $\alpha_p + \alpha_d + \alpha_v + \alpha_r = 1$ and $0 < \alpha_x < 1$ are weights, which can be chosen by the composer.

The mathematical properties of this similarity measure are also nice and can be proven. We can use this similarity measure to define a distance between two notes:

$$d(n_1,n_2) = \sqrt{2(1 - K(n_1,n_2))}$$

This has the advantage of using the nearest neighbors algorithm in generating music. To capture similarities between fixed length sequences of notes, one could define the mean of the similarites:

$$K_S((n_1,\cdots,n_s),(N_1,\cdots,N_s)) = \frac{1}{s}\sum_{i=1}^{s} K(n_i,N_i)$$

This could be useful for measuring consonance of two melodies or so. The algorithm we propose starts with a single note for a voice and keeps adding nearest neighbor notes sorted by distance, with the last note in a sequence of neighbors, becoming again the first note etc.

We have implemented these ideas in the Python/Sagemath (code can be shared on request) language and could generate some consonant sounding piano solos as linked [2].

Summary and Conclusions

We proposed a new method for measuring note similarity based on knowledge from music-theory to generate consonant sounding Piano solo pieces.

Acknowledgements

I would like to thank Frank Farris for encouraging me to write these notes. In the musical side I would like to thank Tom Johnson and Stefaan Himpe for giving me advice in music which I lack due to missing training.

References

[1] Twelfth root of two "Wikipedia" https://en.wikipedia.org/wiki/Twelfth_root_of_two.

[2] Examples for Piano solo generated with the method above https://soundcloud.com/user-919775337/sets/waves-experiment.