IDEAS IN AUTOMATIC EVALUATION METHODS FOR MELODIES IN ALGORITHMIC COMPOSITION

Alan R. R. de Freitas, Frederico G. Guimarães, Rogério V. Barbosa
Programa de Pós-Graduação em Engenharia Elétrica & Departamento de Teoria Geral da Música
Universidade Federal de Minas Gerais, Belo Horizonte MG 31270-901, Brasil
alandefreitas@gmail.com
frederico.g.guimaraes@gmail.com
rogeriovb@terra.com.br

ABSTRACT

Algorithmic Composition (AC) methods often depend on evaluation methods in order to define the probabilities that change operators have to be applied. However, the evaluation of music material involves the codification of aesthetic features, which is a very complex process if we want to outline automatic procedures that are able to compute the suitability of melodies. In this context, we offer in this paper a comprehensive investigation on numerous ideas to examine and evaluate melodies, some of them based on music theory. These ideas have been used in music analysis but have been usually neglected in many AC procedures. Those features are partitioned into ten categories. While there is still much research to do in this field, we intend to help computer-aided composers define more sophisticated and useful methods for evaluating music.

1. INTRODUCTION

Many AC procedures rely on a measure of quality to distinguish between suitable and inappropriate melodies. The quality of the melodies can be used to guide the process in many computer-aided composition methods, such as Markov Models, Generative Grammars or Artificial Neural Networks. The method that certainly depends the most on those measures of quality is Genetic Algorithms, at least when they are found in its simplest form. For instance, one way of computing the quality of a melody would be by directly comparing two tunes or using music theory to evaluate a song.

Given the open problems for evaluating music and art (Section 2), we present metrics that could be used to deal with the problem of defining an automatic evaluation function (Section 3) for melodies. Few of the metrics mentioned here have been used in earlier evaluation-oriented algorithms while most of them are extracted from an extensive musicology research, that might be useful to scientists interested in creating algorithms for generating melodies.

In that context, the contribution of this paper is to provide an extensive research on ways of analyzing melodies by means of:

1. Pitch (Section 4)
2. Tonality and dissonance (Section 4.1)
3. Intervals (Section 4.2)
4. Melodic expectation (Section 4.3)
5. Contour (Section 4.4)
6. Rhythm (Section 5)
7. Patterns (from Section 6)
8. Phrases (Section 6.1)
9. Originality (Section 6.2)
10. Second order analyses (Section 6.3)

These analyses can help the development of more sophisticated automatic evaluation functions for AC methods. Even though the focus of this paper is to analyze metrics that would be helpful in guiding AC methods, those metrics may also be useful from the point of view of Music Information Retrieval or Computational Musicology themselves.

We conclude in Section 7 that while there is still much research to do in this field, this work can surely help computer-aided composers define more sophisticated and useful methods for evaluating music generated by AC techniques.

2. EVALUATING MELODIES

As evaluating music includes the problem of codifying aesthetics, the definition of automatic methods for calculating the melodic features becomes a very complex task. For that reason, evaluating melodies is one of the biggest problems in generate-and-test AC methods because it is still not clear how aesthetic judgement can be expressed by an evaluation function [1].

Serious problems with interactive aesthetic selection in evolutionary computation [2] are that (i) the population size is limited to the ability of people to perform subjective comparisons; (ii) interactive comparison is a slow subjective process and it may take many hours to evaluate few generations; (iii) small changes in the genotype can radically change the phenotype and the quality of the individual; (iv) the complexity of the genotype is limited and complex pieces may take a long time to process, even when it is not possible to distinguish the rare ones that will result in something interesting.

A usual approach is to use a human mentor to evaluate solutions. At first sight, it seems to be a practical, efficient and elegant solution for the problem of evaluating art. However, in music, which is a temporal art, this approach rises the trade-off between the legitimacy of the evaluation and the size of the sample shown to the mentor. This problem is often referred to as the fitness bottleneck [3, 4].

An existing bolder approach to cope with the fitness bottleneck are fitness-free genetic algorithms [5]. In those algorithms, originality is allowed to emerge from a fitness-free environment as musicality is strongly controlled through the genetic operators and the initial population. Besides avoiding the study of all the nuances of the problem, this approach is limited in the sense that the algorithm does not "know" where it is heading to, and all we can do is to expect that the initial population is good and the genetic operators will not disrupt the melodies into something less acceptable than the original reference melodies. In order to guarantee that, the genetic operators will inevitably have to be more conservative, in contrast to the boldness of the initial idea, and they will not be allowed to strongly alter the solutions.

Another substitute for a human mentor is neural network evaluation, which has not yet achieved much success as the typical result is a network unable to generalize beyond the training sets or even evaluate the sets properly [6]. The problem is often based on the lack of information or lack of easily computable features in the musical material [7]. Other attempts to use such connectionist models were focused on evaluating simple music [8], images [9], and paintings [10].

Therefore, an open problem in the area is to formalize automatic evaluation functions [1]. Those functions must be machine representable, capable of measuring human aesthetic properties and practically computable. Evaluation functions should not only define what is more likely to occur on melodies but they should also allow creativity when considering all the different aesthetic objectives to generate ideas not imagined before [11]. Computational aesthetic evaluation is a distinctly non-trivial unsolved problem [1].

3. AUTOMATIC EVALUATION FUNCTIONS

Many examples of the use of at least simple formalized metrics for automated evaluation can be found in the literature. The automatic evaluation schemes are of three types [7]: heuristic features, rule-based and learned functions. Many heuristic features come from perceptions of the composer while other features are based on music theory.

There have been partial attempts to develop automated aesthetic measures of quality [1]. In an extensive research, Towsey [12] considers several simple features to model an evaluation function and analyze which of those impose strong constraints on melodies. With the lowest standard deviation among the features on a dataset, pitch variety appeared as a strong constraint on the melodies.

Aesthetic measures were developed to evolve four part Baroque harmonization for input melodies [13], and the results were restricted to the C major scale to reduce the search space. Some formalized automatic evaluation measures have also been used for the generation of jazz melodies [14]. In this work, there are several metrics such as intervals, patterns, suspensions, position and duration of notes, and contour.

Although we are describing characteristics to be explored by the fitness function, it is important to notice that other results [5, 15] suggest that some musical features have to be considered not only by the evaluation function but also in conscious manipulation methods that can maintain the solutions feasible. Otherwise, rule-based systems that do not iteratively evaluate solutions can be better suited for those tasks [15].

Besides measures of how people perceive art, comparisons to target values have also been used to evaluate melodies, as it has been done to generate synthesizer voices similar to sound targets [16]. Although target melodic features such as sound patterns can be used to evaluate melodies [17, 18], in the case of algorithmic composition the problem is very complex as it involves a necessity of originality.

Empirical measures can also be used to formulate evaluating concepts. For instance, the Zipf's law [19] and fractal analysis [20] have been used to evaluate music and art. There are also works that tackle aesthetics in general, and an aesthetic measure has been defined regarding the image-processing task of the brain [21], in which the aesthetic value of an artwork would be directly connected to image complexity and inversely connected to processing complexity.

As we can see from the literature, several metrics have been used individually to evaluate melodies in Evolutionary Computation. Very often, these Genetic Algorithms did not produce satisfying results because the whole context in which a melody is inserted was not considered. That motivates our effort in this work to consider a musicology research to define functions for evaluating melodic aesthetic criteria. We define here many measures for fitness evaluation divided into ten categories from Section 4 to Section 6.

It is important to mention that the applicability of these measures of fitness presented in this paper is dependent on the genre of music with which the composer intends to work. Some common criteria, such as the penalization of large vertical intervals, for instance, would not apply to many baroque melodies, which may have much larger vertical intervals than many Renaissance melodies. For other melodies, the tonality cannot be so easily implied and that makes the transposition of all melodies to a specific key less feasible. Thus, although some adaptation to the material the reader will be working with is expected, most observations in the text are adequate to a repertoire of classical or baroque tonal melodies and they also apply for a large spectrum of popular twentieth-century melodies.

4. PITCH METRICS OF MELODIC EVALUATION

Pitch is how humans discern sounds of different frequencies. Pitch values are probably the main aspects to be analyzed in a melody. A large musical corpora has been produced about pitch [22] to determine the perceptual relevance of different possibilities of pitch sets.

A first intuitive approach to analyze pitches would be a histogram of **pitch distribution**, which can reveal the most used notes for all melodies. When melodies come from tunes in different keys, it is probably more useful to shift the melodies to the same key before generating the histogram. The histogram of **shifted pitch distribution** can then give a hint of the scales being used the most in the melodies. Those histograms can also consider one or more octaves, in which case the same pitch would be counted separately if it is being performed in a different frequency.

It is also important to analyze the **pitch variety** in each melody as all the pitches revealed by the histogram may not be occurring in the same proportion in all melodies. A method of measuring pitch variety [12] would be dividing the number of distinct pitches by the number of possible pitches for the melody. Another aspect of pitch variety is **pitch range** [12], which is the difference between the highest and lowest pitches in a melody.

Melodies are divided into musical phrases, which also may have some tendency to begin or end with specific pitches. Thus, it would be important to study histograms of **beginning pitches** as well as **ending pitches**. When melodies are generated for specific chord sequences, it is particularly important to study the relation between those notes and the chords being played.

As used by some approaches, such as the key of a melody [23], it may be useful to make histograms of **weighted pitch distribution**, where the relevance of a pitch is weighted by its duration. That intuitively makes sense for many applications as we hear more the pitches with longer duration.

4.1 Tonality and Dissonance

In relation to melodies, it is important to define the **tonality of the melody** if that information is not available because it changes the notes that will be desirable when the intervals are decided. The K-S key-finding algorithm [23] uses key profiles to find the key of a song based on its pitches. The tonality of a melody can also be induced and projected on a **self-organizing map neural network of tonality** trained with 24 key profiles, as described by Toiviainen & Krumhansl [24]. The key of the melody can also be analyzed through the melody with smaller windows of time.

As we identify the key, we can usually recognize the occurrence of each **dissonance**, or infrequent notes, used in the melody and the **dissonance variety**. It is important to measure which dissonances are attracted to the closest note belonging to the appropriate scale or having another function.

Other features already employed [12] are **key centered**, the proportion of notes which are tonic or dominant, **non-scale notes** and **dissonant intervals**.

4.2 Intervals

Another class of melodic information relates to the intervals between pitches. Those intervals are usually measured in terms of scale degrees or semitones and are the basis of many AC algorithms based on Markov Chains. Thus, the first information we can extract in this category is a histogram of **interval distribution** and **interval size distribution**. A theoretical model of the occurrence of each interval [25] indicates that short intervals should happen more often than large intervals. Thus, another way of analyzing the intervals would be a comparison of the **difference between the intervals and the theoretical model**.

As described for the pitches, it is also important to study the **interval variety** when extracting characteristics from melodies. The interval variety can be calculated as the number of intervals divided by the number of possible intervals, as defined by a practical limit.

There is also evidence that short melodic intervals tend to be predominantly descenders as large ones are ascenders [22]. So a histogram of **ascenders and descenders** as well as the **proportion of ascending intervals** are good features to analyze the direction of the intervals.

In order to know the form of the melody, another metric is the **stability in relation to the last interval** [12]. That would represent how much the intervals tend to go in the same direction as some musical phrases aim to get to higher or lower pitches.

As very large intervals are usually rare, it is a common practice to **penalize very large intervals** in the evaluation of the solutions [14]. Although that is a common practice, we remind that this rule can be very genre-specific and it usually makes less sense when the interval leap is between phrases.

4.3 Melodic Expectation

When someone listens to a melody, it is common to have an expectation of what is going to come next. One important study on that subject is Narmour's Implication-Realization Model [26], which can predict melodic expectancy reasonably well. Although the whole model is very complex, some main points are mentioned here.

The first principle of the model is **registral direction** and it implies that small intervals tend to be followed by a continuation of pitch direction while large intervals give expectation of a change in direction. **Interval difference** is the principle that small intervals imply other intervals of similar size while large intervals imply smaller intervals. When a second interval leads to a pitch very close to the original pitch, we have the principle of **registral return**. Thus, listeners expect skips to return to a similar pitch. The principle of **proximity** simply defines that small intervals are more expected than large intervals. The last principle is **closure** and it occurs when melody changes direction or when a large interval is followed by a smaller interval.

The model has been quantified by Krumhansl [27]. In her work, she also proposes a new principle called **consonance**. This principle states that unisons, perfect fourths, fifths and octaves are favored intervals. However, this is a very context-specific assertion as melodies may have many more chromatic or dissonant intervals, depending on its genre.

Of course the melody key also influences the expectations of the listeners, so the values of expectation can be changed to conform more to the **tonal stability values** [28]. In addition to that, the current pitch being used in the melody can influence the size of the next interval as it may tend to be followed by a note that belongs to the key. As unstable tones tend to be attracted by stable tones, **melodic attraction** [29] can also be analyzed on the melodies.

Two reformulations in the model [30] are **tessitura**, which states that the following notes tend to the median pitch height, and **mobility**, which uses autocorrelation to predict if the next pitch is predictable in relation to the previous intervals. A quantifyed implementation of all these measures is available from Eerola and Toiviainen [31].

There are ways to use those expectations in an evaluation function: (i) using the model directly and penalizing melodies that do not follow the model or (ii) studying the **relation between a large corpora of melodies and the model** before penalizing melodies that do not have a similar relation.

4.4 Contour

Contour represents the sorts of movements being performed by the melodies. The contour of a melody is usually easier to remember than all the intervals [32]. There are various types of melodic contour and the occurrence of each **sort of contour** can be used to evaluate melodies. We can identify several sorts of contour [33] in melodies: (i) ascending, (ii) descending, (iii) undulating, (iv) pendulum, (v) cascading, (vi) arc or (vii) rise. Naturally, another feature to analyze would be the **variety of the type of contour**.

A form of controlling contour is to have the **average interval direction and size** and compare it to the melody [14]. Other contour metrics already employed [12] are the average **contour direction**, or the direction of the intervals, **contour stability**, how many intervals change their direction, **movement by step**, the amount of diatonic intervals, **leap returns**, large intervals not followed by a return interval, and **climax strength**, the inverse of the number of times a climatic note is repeated in the melody.

5. RHYTHM METRICS

The first feature to be analyzed in relation to rhythm is the **duration of notes** and **duration variety**. In order to do that, all possible durations have to be determined to a practical limit. Other features to be studied are the **duration of the first notes** and the **duration of the last notes** on each musical phrase. The **rhythmic proportion** or the **rhythmic difference** of the melodies, or the proportion or difference between the longest and the shortest note, is also something to be studied. Also as important as analyzing the occurrence of pitches in a melody is to perceive the occurrence of **silence** in the melodies as it can dramatically change its meaning.

Another characteristic which forms the rhythm is when the notes are played. Two important features are the **location of the notes in the bar** as well as the **location of the notes in the beat**. It has been revealed that a hierarchical grid of note locations may exist in the expectation of Western melodies [34]. This metric can also be useful to find syncopated rhythms in a database. It is also relevant to know the **location of the first notes** and the **location of the last notes**.

The data related to the location of the notes can also be compared to the data related to the **location of the notes weighted by their duration**, as the notes with longer duration are more notable to the listeners [35].

Other interesting metrics [12, 14] are **suspensions**, notes that lie across two consecutive chords; **notes at downbeats** and **notes at half-bar**, as those are significant beats in a bar; **long notes**, notes that last for too long; **note density**, the sparseness or "business" of a melody; **rest density**, the amount of silence; **rhythmic variety**, the durations used compared to the possible durations; **rhythmic range**, the rank difference from the longest to the shortest note, all divided by the number o possible durations.

6. PATTERN METRICS

The relation between pitch and rhythm can be used to make a psychologically motivated [36] identification of melodic patterns or phrases. An autocorrelation technique [37] can identify those patterns by correlating series with a delayed copy of the melody. This **autocorrelation** technique can be used to identify patterns in the contour, pitch or rhythm. Another idea for finding patterns in pitches is to use **pattern matching** [14] to make a histogram of the number of pitch patterns found for each melody size.

Other metrics [12] are the number of **repeated pitches** in relation to the number of intervals used and **repeated rhythmic values**. Extensions of this idea are the number of repeated pitch of rhythm patterns of n notes.

6.1 Phrases

Melodies usually consist of many musical phrases. By dividing the melody into phrases, we can analyze (i) the **phrase sizes** in the melodies as well as (ii) all the features mentioned so far in the paper individually for the phrases. The latter is nice to perceive if the extension of possibilities used in all the songs also apply for different phrases in the same song. In most cases, the information on the segmentation of the melodies is not available. In that case, there are rule-based approaches for melodic segmentation based on **Gestalt psychology** [38] or extensions of these ideas [39]. There are also **statistical approaches** based on the analysis of melodies [40].

Similarly to how we described the possibility to use autocorrelation to find patterns in the melody, once we have the phrases boundaries, we can also use correlation to detect **similarity between the phrases**. Dividing the melody into smaller units allows an analysis of **variance of pitch range between phrases**, which is an important resource to create melodic variety. Other features are the **distance between phrases** and **number of phrases per melody**.

6.2 Originality

Many methods have attempted to generate melodies that not only follow the rules but also are original and diverse [5]. AC algorithms cannot be considered efficient if they always return the same "best" melody. There is also a strong relation between the originality of the themes and their popularity [41]. The least original themes may be considered banal as the most original may be considered too complex.

Although the concept of originality in aesthetics is very complex, some originality of a melody in relation to many other compositions may be inferred by its **note transition probabilities** in relation to the average melody [42]. The **expectancy-based model** [43] uses melodic expectancy theories to define the complexity of a melody. Another way to measure the originality of a melody is to **monitor melodic parameters**, as any of the parameters mentioned herein. By monitoring those parameters, it is possible to detect a change in behavior in the melody as compared to the parameters used up to that point.

6.3 Second order analysis

As music is a contextual art, it makes sense to analyze the influence of the parameters into one another. For instance, it is possible that a particular pitch sequence may lead to expectations related to the duration of notes. In that case, it is important to have an **analysis of second, third or nth order** in a way that we can discover those interactions between the parameters.

Those analyses are usually not simple and lead to a task that is also important in first order analyses: **the metric importance**. We have mentioned so many metrics that we have to distinguish which ones are the most important, be it of first, second or third order. One proposed strategy to measure the importance of a feature is to analyze its **standard deviation** [12], as low standard deviations in a parameter indicate that it puts a strong constraint on the melodies.

7. CONCLUSION AND FUTURE WORK

As mentioned in this paper, the previous approaches to formalize an evaluation function for generating melodies have only been partial and the intention of this paper is to expand the pool of ideas on which computer-aided composers may be working on at the moment. Although there may be many more features to research, this work might expand the possibilities of feature categories of which composers have been thinking.

An AC procedure will not be able to handle all the features mentioned here. A simple weighted sum of objective functions or many objective optimization may not be the best answers to this problem and those may be in fact very naive solutions. Future studies must include how to put those metrics together to evaluate melodies.

In order to put the metrics together for use in a genetic algorithm, of course an analysis of which ones are the most important features as well as an analysis of the practical limits of those metrics have to be developed. Thus, future work should analyze many melodies and define which ones are strong constraints for a specific music genre or even to devise theoretical models. Naturally, the most important measures will be dependent on the style of the melodies analyzed. Besides defining which are the most important measures, a study of which measures overlap can help handling all the features mentioned here.

It is important to clarify that although we have mentioned many general metrics of automatic evaluation, there are always specific matters in each repertoire that must be taken into account for a good analysis of the songs.

Acknowledgments

This work has been supported by the Brazilian agencies CAPES, CNPq, FAPEMIG.

8. REFERENCES

[1] P. Galanter, "The problem with evolutionary art is..." *Applications of Evolutionary Computation*, pp. 321–330, 2010.

[2] J. McCormack, "Open problems in evolutionary music and art," *Applications of Evolutionary Computing*, pp. 428–436, 2005.

[3] P. Todd and G. Werner, "Frankensteinian methods for evolutionary music," *Musical networks: parallel distributed perception and performace*, p. 313, 1999.

[4] J. Biles, "Genjam: A genetic algorithm for generating jazz solos," in *Proceedings of the International Computer Music Conference*. International Computer Music Association, 1994, pp. 131–131.

[5] A. Freitas and F. Guimarães, "Originality and diversity in the artificial evolution of melodies," in *Proceedings of the 13th annual conference on Genetic and evolutionary computation*. ACM, 2011, pp. 419–426.

[6] J. Biles, P. Anderson, and L. Loggi, "Neural network fitness functions for a musical iga," in *International ICSC Symposium on Intelligent Industrial Automation (IIA'96) and Soft Computing (SOCO'96)*. International Computing Sciences Conferences (ICSC), 1996, pp. B39–44.

[7] J. Biles, "Evolutionary computation for musical tasks," *Evolutionary Computer Music*, no. 2, pp. 28–51, 2007.

[8] S. Phon-Amnuaisuk, E. Law, and H. Kuan, "Evolving music generation with som-fitness genetic programming," *Applications of Evolutionary Computing*, pp. 557–566, 2007.

[9] T. Gedeon, "Neural network for modeling esthetic selection," in *Neural Information Processing*. Springer, 2008, pp. 666–674.

[10] P. Machado, J. Romero, and B. Manaris, "Experiments in computational aesthetics," *The Art of Artificial Evolution*, pp. 381–415, 2008.

[11] A. Freitas and F. Guimaraes, "Melody harmonization in evolutionary music using multiobjective genetic algorithms," in *Proceedings of the Sound and Music Computing Conference (SMC)*, 2011.

[12] M. Towsey, A. Brown, S. Wright, and J. Diederich, "Towards melodic extension using genetic algorithms," *Educational Technology & Society*, vol. 4, no. 2, pp. 54–65, 2001.

[13] R. McIntyre, "Bach in a box: The evolution of four part baroque harmony using the genetic algorithm," in *Evolutionary Computation, 1994. IEEE World Congress on Computational Intelligence., Proceedings of the First IEEE Conference on*. IEEE, 1994, pp. 852–857.

[14] G. Papadopoulos and G. Wiggins, "A genetic algorithm for the generation of jazz melodies," *Proceedings of STeP*, vol. 98, 1998.

[15] S. Phon-Amnuaisuk and G. Wiggins, "The four-part harmonisation problem: a comparison between genetic algorithms and a rule-based system," in *Proceedings of the AISB99 Symposium on Musical Creativity*, 1999, pp. 28–34.

[16] J. McDermott, N. Griffith, and M. O'Neill, "Toward user-directed evolution of sound synthesis parameters," *Applications of Evolutionary Computing*, pp. 517–526, 2005.

[17] Y. Khalifa and R. Foster, "A two-stage autonomous evolutionary music composer," *Applications of Evolutionary Computing*, pp. 717–721, 2006.

[18] A. Horner and D. Goldberg, "Genetic algorithms and computer-assisted music composition," *Urbana*, vol. 51, no. 61801, p. 14, 1991.

[19] B. Manaris, P. Machado, C. McCauley, J. Romero, and D. Krehbiel, "Developing fitness functions for pleasant music: Zipf's law and interactive evolution systems," *Applications of Evolutionary Computing*, pp. 498–507, 2005.

[20] T. Mori, Y. Endou, and A. Nakayama, "Fractal analysis and aesthetic evaluation of geometrically overlapping patterns," *Textile research journal*, vol. 66, no. 9, p. 581, 1996.

[21] P. Machado and A. Cardoso, "Computing aesthetics," *Advances in Artificial Intelligence*, pp. 105–119, 1998.

[22] P. Vos and J. Troost, "Ascending and descending melodic intervals: Statistical findings and their perceptual relevance," *Music Perception*, pp. 383–396, 1989.

[23] C. Krumhansl, *Cognitive foundations of musical pitch*. Oxford University Press, USA, 2001, no. 17.

[24] P. Toiviainen, C. Krumhansl *et al.*, "Measuring and modeling real-time responses to music: The dynamics of tonality induction," *PERCEPTION-LONDON-*, vol. 32, no. 6, pp. 741–766, 2003.

[25] W. Dowling and D. Harwood, *Music cognition*. Academic Press New York, 1986, vol. 19986.

[26] E. Narmour, *The analysis and cognition of melodic complexity: The implication-realization model*. University of Chicago Press, 1992.

[27] C. Krumhansl, "Music psychology and music theory: Problems and prospects," *Music Theory Spectrum*, pp. 53–80, 1995.

[28] C. Krumhansl and E. Kessler, "Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys." *Psychological Review*, vol. 89, no. 4, p. 334, 1982.

[29] F. Lerdahl, "Calculating tonal tension," *Music Perception*, pp. 319–363, 1996.

[30] P. Von Hippel, "Redefining pitch proximity: Tessitura and mobility as constraints on melodic intervals," *Music Perception*, pp. 315–327, 2000.

[31] T. Eerola and P. Toiviainen, *MIDI Toolbox: MATLAB Tools for Music Research*. Jyväskylä, Finland: University of Jyväskylä, 2004. [Online]. Available: www.jyu.fi/musica/miditoolbox/

[32] W. Dowling, "Scale and contour: Two components of a theory of memory for melodies." *Psychological Review*, vol. 85, no. 4, p. 341, 1978.

[33] B. Nettl, *Music in primitive culture.* Harvard University Press, 1956.

[34] C. Palmer and C. Krumhansl, "Mental representations for musical meter." *Journal of Experimental Psychology: Human Perception and Performance; Journal of Experimental Psychology: Human Perception and Performance*, vol. 16, no. 4, p. 728, 1990.

[35] W. Thompson, "Sensitivity to combinations of musical parameters: Pitch with duration, and pitch pattern with durational pattern," *Attention, Perception, & Psychophysics*, vol. 56, no. 3, pp. 363–374, 1994.

[36] M. Leman, *Music, gestalt, and computing: Studies in cognitive and systematic musicology*. Springer Verlag, 1997, vol. 1.

[37] T. Eerola, T. Himberg, P. Toiviainen, and J. Louhivuori, "Perceived complexity of western and african folk melodies by western and african listeners," *Psychology of Music*, vol. 34, no. 3, pp. 337–371, 2006.

[38] J. Tenney and L. Polansky, "Temporal gestalt perception in music," *Journal of Music Theory*, vol. 24, no. 2, pp. 205–241, 1980.

[39] E. Cambouropoulos, "Musical rhythm: A formal model for determining local boundaries, accents and metre in a melodic surface," *Music, gestalt, and computing*, pp. 277–293, 1997.

[40] R. Bod, “Memory-based models of melodic analysis: Challenging the gestalt principles,” *Journal of New Music Research*, vol. 31, no. 1, pp. 27–36, 2002.

[41] D. Simonton, “Computer content analysis of melodic structure: Classical composers and their compositions,” *Psychology of Music*, vol. 22, no. 1, pp. 31–43, 1994.

[42] ——, “Melodic structure and note transition probabilities: A content analysis of 15,618 classical themes,” *Psychology of Music*, vol. 12, no. 1, pp. 3–16, 1984.

[43] T. Eerola and A. North, “Expectancy-based model of melodic complexity,” in *Proceedings of the Sixth International Conference on Music Perception and Cognition. Keele, Staffordshire, UK: Department of Psychology. CD-ROM*, 2000.