
Reproducing Musicality: Immediate Human-like Musicality Through Machine Learning and Passing the Turing Test

Aran Samson, Andrei Coronel

Department of Information Systems and Computer Science, Ateneo de Manila University, Quezon City, Philippines

Email address:

avsamson@ateneo.edu (A. Samson), acoronel@ateneo.edu (A. Coronel)

To cite this article:

Aran Samson, Andrei Coronel. Reproducing Musicality: Immediate Human-like Musicality Through Machine Learning and Passing the Turing Test. *American Journal of Artificial Intelligence*. Vol. 5, No. 1, 2021, pp. 38-45. doi: 10.11648/j.ajai.20210501.13

Received: April 17, 2021; Accepted: May 12, 2021; Published: May 26, 2021

Abstract: Musicology is a growing focus in computer science. Past research has had success in automatically generating music through learning-based agents that make use of neural networks and through model and rule-based approaches. These methods require a significant amount of information, either in the form of a large dataset for learning or a comprehensive set of rules based on musical concepts. This paper explores a model in which a minimal amount of musical information is needed to compose a desired style of music. This paper takes from two concepts, objectness, and evolutionary computation. The concept of objectness, an idea directly derived from imagery and pattern recognition, was used to extract specific musical objects from single musical inputs which are then used as the foundation to algorithmically produce musical pieces that are similar in style to the original inputs. These musical pieces are the product of evolutionary algorithms which implement a sequential evolution approach wherein a generated output may or may not yet be fully within the fitness thresholds of the input pieces. This method eliminates the need for a large amount of pre-provided data as well as the need for long processing times that are commonly associated with machine-learned art-pieces. This study aims to show a proof of concept of the implementation of the described model.

Keywords: Artificial Intelligence, Computer Music, Turing Testing

1. Introduction

Artificial intelligence and machine learning has had great strides in recent years in terms of musicality, specifically in human-like algorithmic composition, so much progress has been made, in fact, that there have already been discussions if music can be used as a valid metric in satisfying the Turing Test [1, 14, 26].

This study concedes that these previous studies and techniques are able to emulate human musicality by studying a large corpus of music or rulesets. This study takes a different approach in musical composition totally eliminating the requirement for large datasets and rather, focusing on a method dubbed "immediate learning" wherein an immediate singular input of music is used as the basis for composition of human-like music. While it is natural for a composer or musician to study a specific style or a specific composer's music and be able to produce a musical piece based on this study, it is also possible for musicians and composers listen to a single piece of music and take elements from just this

single piece and then create a new piece of music that is inspired by just this single piece of music. Musicians can do this without copying the music directly and by taking specific stylistic elements from the original.

This study serves as a proof of concept of a model in which this seemingly human-only skill of being able to quickly compose music after hearing only one song is emulated in algorithmic composition. This study investigates how the model can generate similarly styled musical pieces based off a specific single input piece. This study is heavily based on the concept of an instrument solo, improvisation or the concept of cadenza. The author personally associates this with guitar solos wherein a guitarist may start playing an instrumental solo, then queues or signals another guitarist to continue the instrumental solo. The second guitarist is now in a position wherein he has to continue the solo in a similar style but not be a direct copy of the first solo. This study focuses primarily on improvised musicality with a primary objective in algorithmic composition with real-time or instantaneous output.

This study borrows technical concepts from imagery and pattern recognition, specifically, the concept of objectness. Objectness as a concept is described as, within a specific image, objects can be detected even without prior training by an image or pattern recognition system [2, 3, 27-28]. As this concept does not require training, and therefore does not require large datasets, the application of this concept becomes an essential component in this study's objective of music composition that does not require large datasets.

This study takes musical objects and uses them as foundation for algorithmic music composition based on evolutionary algorithms. In this study, extracted musical objects are considered important style objects in the evolutionary algorithm. This study has designed the algorithm to consider specific detected musical objects as important stylistic choices by the human composer and attempts to re-apply these stylistic objects into its own computer-generated compositions. This helps ensure that the new piece still stays inspired by the original input music. As an example, a certain guitarist may like the use specific hammer-on and pull-off techniques which other musicians who aim to emulate his style, will incorporate into their own. To assess the similarity of the overall style of the output musical pieces, musical features are extracted from both the original piece and the algorithmically composed musical pieces and distance is compared between feature values. This method has been used in prior research in musical composition [4].

This study serves as a proof of concept of a model in which this seemingly human-only skill of being able to quickly compose music after hearing only one song is emulated in algorithmic composition. This study does not claim to perfectly reproduce the musical style of a specific input musical piece but to rather quickly approximate the style, outputting as input has ended, "with immediacy."

2. Methodology

2.1. Methodology Explained

This study undertook the following four steps in achieving the end goal of immediate computer composition: 1. Construct and define the concept of musical objectness. Determine how to extract these musical objects from given melodies or composition. 2. Define quantifiable musical distance or closeness. Similarity and distance were the primary focus of this step. 3. Develop an automatic or algorithmic method that can accept musical objects as input and will output musical compositions. 4. Conduct a test based on the Turing Test to determine if human-like algorithmic composition has indeed been achieved.

2.2. Musical Objectness

There were two types of musical objectness determined in this study, Large Musical Objects and Micro Musical Objects. The concept of large musical objects takes directly from visual concepts of pattern recognition and objectness [2, 3] while

micro musical objectness takes from musicality and musical hooks [5, 6]. The following function illustrates the heuristic of extracting large musical objects from musical input:

```

1 function MusicObjectExtraction (i);
   Input : Source MIDI file i
   Output: Musical note segment
2 g ← global pitch averages and rhythm averages of i;
3 t ← first four notes of i;
4 tx ← rhythm and pitch profile of t;
5 repeat
6   p ← next four notes of i;
7   px ← rhythm and pitch profile of p;
8   determine if px is closer to g or tx;
9   if px is closer to tx then
10    if px is within similarity threshold to tx then
11      append p to t;
12    end
13  else
14    t ← p;
15  end
16 end
17 until end of file;
    
```

Figure 1. The Algorithm for Large Object Extraction.

Micro musical object extraction is based on string-finding and string searching algorithms, specifically suffix trees as implemented by Ukkonen. This was specifically used as it explicitly runs in linear time, an important requirement in minimizing processing time [7, 8, 23-25]. Figure 2 illustrates the suffix tree being constructed from a short string of musical notes.

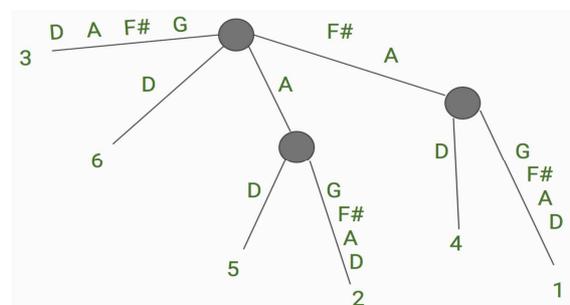


Figure 2. Suffix Tree Construction.

2.3. Musical Mathematical Distance

Musical mathematical distance was needed to be determined

in order to confirm and assess the relative human-ness of computer generated music. A statistical information extraction was used for input musical pieces. The tool used was Cory Mckay’s jSymbolic which extracts statistical information from input music pieces [9, 15]. The output statistical data for this was then used as input in the mathematical concepts of Taxicab Geometry and Manhattan distance. The following formula illustrates Manhattan distance:

$$d_1(p, q) = ||p - q||_1 = \sum_{i=1}^n |p_i - q_i| \quad (1)$$

The variables p and q were treated as input vectors based on features extracted by jSymbolic. The Manhattan distance was used as the basis in a normalized distance metric of any two musical pieces [10, 18-22].

Another concept explored in determining musical distance was a novel approach in the use of Density Degree Theory. The original author of Density Degree Theory, Dr. Orlando Legname, asserts that there is a mathematical complexity that is proportional to two notes’ consonance or dissonance to each other [11]. This complexity can be represented by drawing the Lissajous curve of any two notes’ frequency calculation. As illustrated below on Figure 3 and Figure 4, there is a distinct difference in mathematical complexity in the graphical representation of the relationship between two consonant notes such as the root and the 5th, versus two dissonant notes such as the root and its minor 2nd.

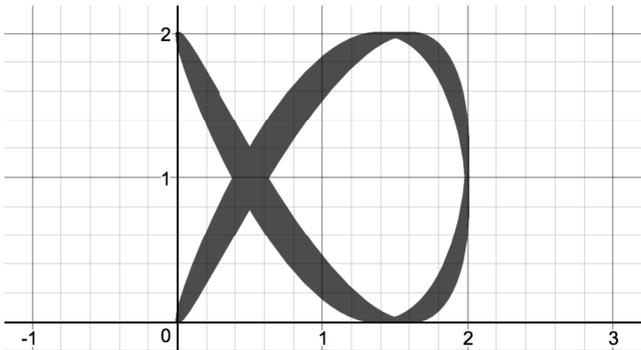


Figure 3. Lissajous Curve of Root and 5th.

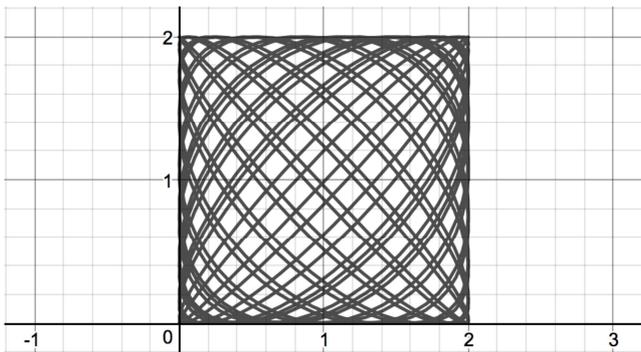


Figure 4. Lissajous Curve of Root and Minor 2nd.

The use of distance calculation based on the mathematical and statistical distance as well as the distance that can be extrapolated from the use of Density Degree Theory was important in assessing the closeness to human-like generated

computer pieces were.

2.4. Algorithms and Methods

The algorithmic approach in musical composition is taken directly from previous research done in evolutionary algorithmic musical composition [4, 16, 17]. The implementation in this study necessitated the use of the elements discussed in the previous section and was directly applied to the musical composition system in [4]. Previous research done [12] illustrates a proof of concept that the use of this method is promising in achieving the goals of this study.

For the purpose of this study, a simple application was constructed that functioned as the interface between a human musician and the computer algorithmic composer. Figure 5 shows the basic interface of the developed application.

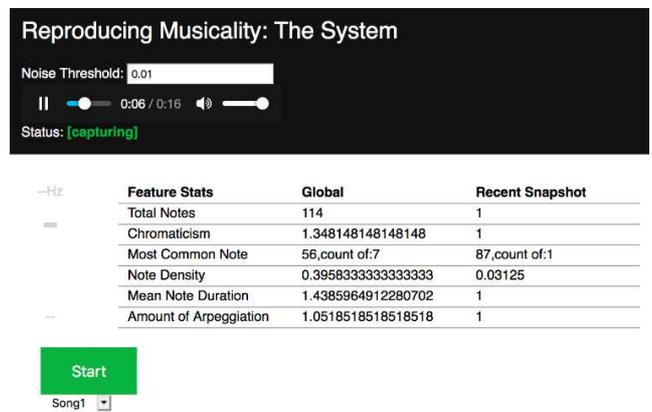


Figure 5. Developed application.

The process of data capture starts with a human musician playing music input such as a musical melody of any length. As the human musician starts his input (this is detected by audio input), the algorithms implemented immediately iterate to calculate and construct the output musical composition. When the musician finishes playing, the system automatically detects when the musician has finished and immediately outputs the audio of the composition. This computer composition is intended to be the closest-approximation to the style of the original input achievable in the timeframe before output is expected.

The outputs of this application were part of the answer to the original question of the ability of a computer to generate human-like compositions. The outputs were promising but a way to validate human-ness was needed to further give proof to the success of the computational composition model.

2.5. Validation and Human Testing

One of the goals of this study was to compose human-like algorithmic compositions. Music composition was proved indeed possible, but a qualification was needed to be conducted to confirm if the compositions were indeed human-like.

A qualification survey was constructed in three parts. All parts of this survey required the output of the musical system. Several computer-generated pieces were produced based on

the described music-composition methods described earlier. These were used as testing material for the qualification survey. The three parts of this survey are as follows:

2.5.1. Priming and Musicianship-Level-Metric

In this first part of the survey, respondents are tasked with listening to three pieces of music. The respondents are not informed that these musical pieces are composed by a computer and are instead made to believe that the musical pieces are composed by humans of varying proficiency levels. The respondents are asked to assess the probable number of years of musicianship the composer of each musical piece. The purpose of this section was to first determine the perceived level of proficiency of the algorithmic composer. The question at the end is to determine how believably human the composed pieces were. Respondents were not informed of the computer compositions as prescribed by previous research involving Turing-like indistinguishability tests [13].

2.5.2. Computer Assisted Composition

The second part of the survey tasked respondents with listening to three musical samples composed in-part by a human and in-part by computer. Respondents assessed the composition to determine how much of it they perceived to be composed by a computer and how much of it was composed by a human. The expected answers to this section of the survey were a percentage between 0 to 100 of how much the composition was constructed by a computer. All samples used in this survey were made 50% by a computer and 50% by a human and respondents were not informed of this equal composition. This section was constructed with the aim of determining if there is a noticeable bias to either computer or human when all samples are equally composed by computers and humans.

2.5.3. Human-Like Composition Believability

In the third part of the survey, respondents are tasked with listening to musical samples composed by either a human or computer. Respondents assessed if each given piece is composed by a human or automatically generated by a computer. This part of the survey aims to see the general believability of human-ness of the composed pieces by asking respondents if the musical piece is composed entirely by a human or computer. Of all the musical samples presented, only one was composed by an actual human, all the rest were computer generated.

2.5.4. Respondent Profile and Comments

Finally, the respondents were simply asked to state the number of years of musicianship they have, as well as the number of years of formal music education they have undergone. The respondents were asked for their comments, methods, and insights into how they are able to determine if a piece is composed by a human or computer. Overall, the survey was intended to be the validation method as well as the basis for success metrics for this study.

3. Results and Analysis

Overall, the results of creating the computer composer were promising and yielded compositions. Figure 6 and Figure 7 below are examples of computer compositions generated by the algorithms and systems that were developed throughout this study.



Figure 6. Computer Composition Example 1.



Figure 7. Computer Composition Example 2.

To validate the extent of success that we have achieved human-like composition, a survey was administered to test the level of perceived human-likeness of the computer compositions. Table 1 shows the variables that were retrieved from the survey and Table 2 shows the results that were retrieved from each of these variables as well as relevant averages for some of the items on the survey.

Table 1. Description of Each Variable*

Variable	Description
A1	Years of Musicianship (Instrument)
A2	Years of Formal Music Education
B	Average years of perceived computer composer musicianship (section 1 of survey)
C	Human-ness believability of specimens in section 1
D	Average perceived percentage of specimens composed by computer
E1	Perceived human-ness of computer-composed piece 1
E2	Perceived human-ness of computer-composed piece 2
E3	Perceived human-ness of human-composed piece
E4	Perceived human-ness of computer-composed piece 3
E5	Average Perceived human-ness of computer-composed piece 1-3

*these sections and variables are described previously in section 2.5

It was discovered through a one-way analysis of variance (ANOVA) that there is no significant statistical difference in the results of the survey regarding groups with greater years of musicianship and less years of musicianship (variable A1), as well there is no difference with more years of formal music education and less (variable A2). This now leads to an analysis of the entire body of respondents as a whole rather than initially intended to be a segmented analysis based on the proficiency and years of musical exposure of the respondents.

Table 2. Results of Survey.

Variables	N	Range	M	SD
A1	155	1-10	4.9	4.05
A2	155	1-10	2.08	2.92
B	155	1-10	3.16	2.02
C	155	1-10	5.79	2.83
D	155	1-100	57.51	16.39
E1	155	1-5	3.1	1.12
E2	155	1-5	3.23	1.32
E3	155	1-5	2.28	1.12
E4	155	1-5	3.27	1.11
E5	155	1-5	3.2	0.72

3.1. Perceived Average Number of Years of Musicianship

In this first part of the survey, respondents are tasked with listening to three pieces of music. The respondents are not informed that these musical pieces are composed by a computer and are instead made to believe that the musical pieces are composed by humans of varying proficiency levels. The respondents are asked to assess the probable number of

years of musicianship the composer of each musical piece. After the questions on this part but before the second survey part, the respondents are informed that the compositions were composed by a computer. The respondents are then asked to answer a question detailing if they believed that the pieces were composed by a human, prior to finding out that the pieces were composed by a computer.

The general average years of perceived musicianship of the pieces were calculated to be at the mean of 3.16 years. This means that the average perception of the participant group of the computer-composer was that it was a composer of around 3.16 years of experience in music. If taking into consideration that the algorithmic composition could only study the direct input of music, it was expected that this section would yield a smaller average number of years of perceived musicianship. It was a welcome surprise that the average perception was 3.16 years. It may be of importance to note that 28 of the 155 participants had rated the computer-composer as having 0 years of musical experience.

This section was designed without informing of the perceiver of the existence of the computer composer. This was intended to pattern the tests conducted by Colby, Hilf and Weber [13] wherein they assert that the Turing test be conducted without informing the perceivers of the computer. In this section, it was determined that the average of 3.16 is indicative that the perceivers were not able to discern that the composer for these pieces were not human.

Figure 8 shows the distribution of the perception of the number of years of experience the computer has in music.

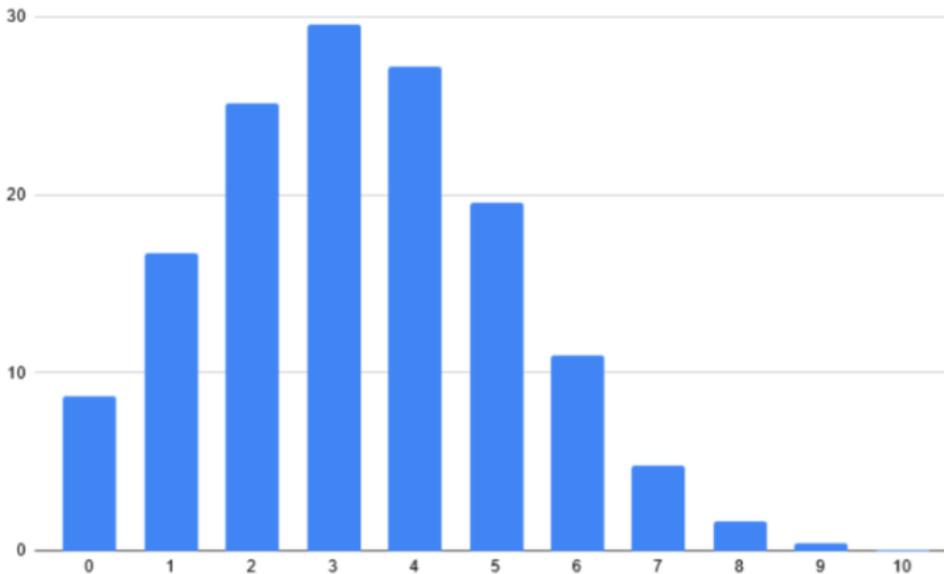


Figure 8. Distribution of results of section 1 of the survey.

3.2. Perceived Percentage of Computer

The second part of the survey tasked respondents with listening to three musical samples composed in-part by a human and in-part by computer. Respondents assessed the composition to determine how much of it they perceived to be composed by a computer and how much of it was

composed by a human.

The expected answers to this section of the survey was a percentage between 0 to 100, of how much the composition was constructed by a computer. All samples used in this survey were made 50% by a computer and 50% by a human and respondents were not informed of this equal composition. This section was constructed with the aim of determining if

there is a noticeable bias to either computer or human when all samples are equally composed by computers and humans.

In summary, an average that skews more towards 0 would mean that the general perception would be that the pieces were more composed by a computer and a score that skews more towards 100 would mean the perception is that the pieces were composed by a computer. 50 would mean a

general perception that they were in-fact 50-50 composed by a computer and human. The actual result of this section is 57.51, which means that it is close to the expected number of 50 with a slight skew in the perception that the pieces were composed by a computer. Figure 9 illustrates the results and skew of this section of the survey.

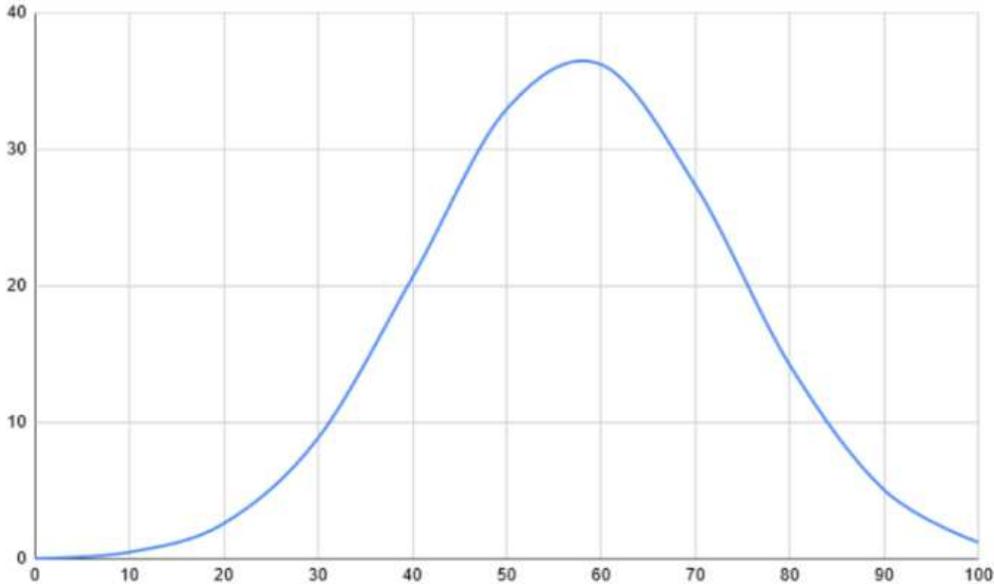


Figure 9. Distribution of results of section 2 of the survey.

3.3. Human-like Composition Believability

In the third part of the survey, respondents are tasked with listening to musical samples composed by either a human or computer. Respondents assessed if each given piece is composed by a human or automatically generated by a computer. This part of the survey aims to see the general believability of human-ness of the composed pieces by asking respondents if the musical piece is composed entirely

by a human or computer. Of all the musical samples presented, only one was composed by an actual human, all the rest were computer generated. Item scores are from 1 to 5 where 1 is human and 5 is computer. Individual item score averages are listed below. It is important to note that the middle value is 3. Figure 10 illustrates the histogram and skew of the results of this section of the survey.

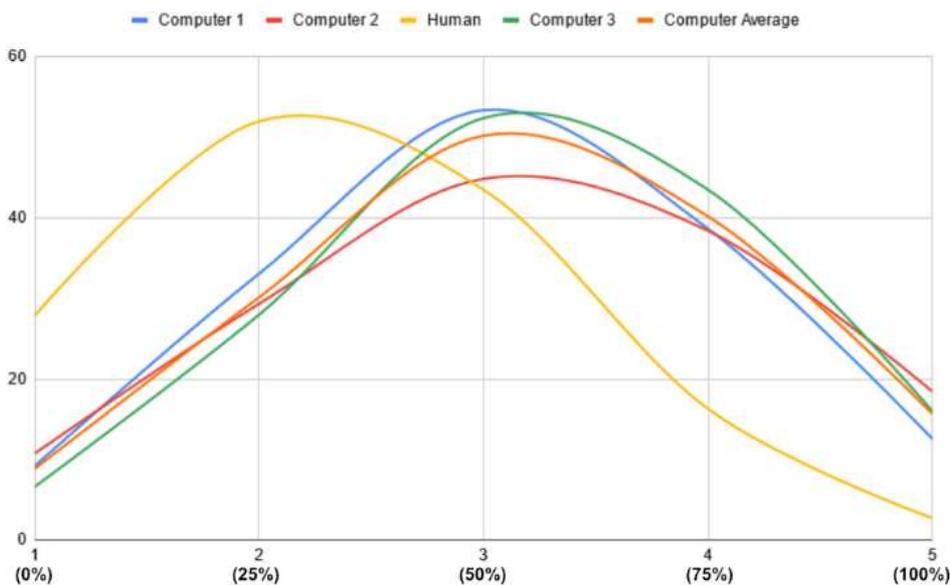


Figure 10. Distribution of results of section 3 of the survey.

This study interprets the values of each result under the context that a score of 1 is full believability that the piece is human-made (equivalent to 0% perceived artificial) and 5 is full-believability that the piece is made by a computer (equivalent to 100% perceived artificial) and 3 is interpreted as complete uncertainty (equivalent to 50% perceived artificial).

All individual scores of the computer-generated pieces were very similar in that they are very near the uncertainty point of 3 (50%). This is interesting to note as it means that statistically, for these specific three compositions, there is an uncertainty or vagueness in the perception of the human-ness of these compositions. The general average of the computer-generated pieces scored a 3.2 or an equivalent of 55% believability that it is artificially constructed.

A very clearly human-piece was also presented and even this human composed piece did not achieve a score closer to 1 and is closer to a middle ground between uncertainty and full-certainty. This further supports that there is still a range in the perception of if a piece is composed by a human or not.

With the statistical results of this section, it is interpreted that there is a difficulty in assessing if the pieces of music were composed by a human or computer. This vagueness or indistinguishability is the primary requirement in satisfying the Turing Test and it can then be safely said that statistically, in this study's scope, this system has passed the Turing Test.

3.4. Results Summary

The results of the study and its several parts has proven that while it does require a complex set of components to compose human-like music, it is indeed possible with a consistent level of success. The use of borrowing concepts from imagery and pattern recognition to determine and extract large and micro musical objects proved a reliable way of finding the defining features and sections of musical pieces. A mathematical descriptor of musical similarity developed for this study also showed that a similarity metric is useful in constructing algorithms that compose music that rely on comparing with a real-human composition.

Finally, the use of an evolutionary approach in designing the algorithm proved effective when considering the many modules and requirements that were determined needed in the development of a computer composer.

Statistical evidence showed that the computer composer was deemed near human-like with an ambiguous result when humans were asked to assess if the composed pieces were human or not. This, in addition to the average perception that the computer composer had 3 years of musical experience shows that the study was successful in producing a proof-of-concept of a computer composer that does not need a large dataset to learn human-like composition.

4. Conclusion

This study has proven that a computer is indeed capable of composing musical pieces in perceived real time or immediately after musical input. This study also serves to take

away some of the perceived importance of large data-sets for the purpose of musical modeling through machine-learning. Moreover, human-like composition was proven possible even when applied in a context of limited-composition time.

This study has succeeded in creating a foundational model for algorithmically composing music using only direct learning and without the need for big-datasets but which can also generate pieces quickly without sacrificing a human-ness that is often associated with music composition. This was done through the combination of multiple pre-existing technologies and concepts, as well as novel ideas and approaches that were directly created for the use in the development of this study. This study exhibits an approach of using the concepts from visual pattern recognition and objectness applied to sound and music. This study also shows a novel approach in musical distance and dissonance assessment using geometry and mathematical formulae.

This study has accomplished its set objectives of implementing a musical object extraction method and then using a similarity metric to compare similarity of these objects. This was accomplished using multiple methods and concepts together. Concepts of objectness and saliency and methods based on visual pattern recognition were used to extract musical objectness while a musical similarity was developed using mathematical concepts such as taxicab geometry as well as the novel Density Degree Theory. This study has also accomplished its objective of producing a proof-of-concept system that can algorithmically compose music based on these extracted music that can approximate the style of human-composed pieces. An application was able to be constructed that can accept a live musical input and then immediately respond with an algorithmic musical output. Overall, all objectives have been accomplished.

Results showed that algorithmically generating music in real-time may result in a musical pieces that pass the Turing test of ambiguity. The original test designed by Alan Turing only required the computer system to fool judges 30% of the time, a value that was arbitrarily selected. More modern tests have set this requirement to 50%. This study's results come close to this 50% believability in two sections of testing. One of the test sections were conducted with pieces of half-and-half music composed by a computer and human, and scored a 57.51% average or a 7.51% skew towards computer-like. Another test section conducted had participants gauge if a piece was human or computer-made which scored a 55%, skewing 5% towards computer-made.

The first section of testing also asked participants to rate the perceived number of years of musicianship the composer of select pieces were. These pieces were composed by a computer but participants were not informed of this until later in the text. This was intended to pattern the tests conducted by Colby, Hilf and Weber [13] and had results that showed an average of 3.16 years of perceived musical experience for the computer composer. This section of the study also showed only 28 (18%) out of the 155 participants scoring the computer-composer with 0 years of musical experience. From

the Colby et. al. study, it can be seen as a success that only 18 percent of participants consider the pieces as being composed by someone of no years of musical experience.

Overall, the scientific development and the results of this study show very promising results in this model for algorithmic musical computation. However, this study also does not claim to have perfect emulation of human-like composition as there are more dimensions that need to be explored for this to come close to perfect human-ness. This study concedes that models that make use of big datasets come closer to human emulation. It is important to state that the objective of this study is to develop a model that aims not to remove the requirement of big-data-based learning and composition but rather aims to compliment and supplement these models by providing a different perspective and method to musical composition, independent of big-data.

Lastly, it is hoped that the methods, results, and insights discussed in this study could provide knowledge that would be useful in and serve as basis for future algorithmic music generation studies.

References

- [1] Belgum, E., Roads, C., Chadabe, J., Tobenfeld, T. E., & Spiegel, L. (1988). A Turing Test for "Musical Intelligence". *Computer Music Journal*, 12 (4), 7-9.
- [2] Shah, S. A. A., Bennamoun, M., Boussaid, F., & El-Sallam, A. A. (2013, February). Automatic object detection using objectness measure. In *Communications, Signal Processing, and their Applications (ICCSPA), 2013 1st International Conference on* (pp. 1-6). IEEE.
- [3] D. T. Bogdan A. and V. Ferrari, "Measuring the objectness of image windows", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 99, no. PrePrints, 2012.
- [4] Samson, A. V., & Coronel, A. D. (2016, November). Evolutionary algorithm-based composition of hybrid-genre melodies using selected feature sets. In *Computational Intelligence and Applications (IWCIA), 2016 IEEE 9th International Workshop on* (pp. 51-56). IEEE.
- [5] G. Burns: "A Typology of 'Hooks' in Popular Records", *Popular Music*, Vol. 6, No. 1, pp. 1?20, 1987.
- [6] Kronengold, C. (2005). "Accidents, hooks and theory." *Popular Music*, 24 (3), 381.
- [7] Tarhio, Jorma, and Esko Ukkonen. "Approximate boyer-moore string matching." *SIAM Journal on Computing* 22.2 (1993): 243-260.
- [8] Ukkonen, Esko. "On-line construction of suffix trees." *Algorithmica* 14.3 (1995): 249-260.
- [9] C. McKay and I. Fujinaga, "jSymbolic: A feature extractor for MIDI files", 2006.
- [10] Samson, A. V., & Coronel, A. D. (2018, February). Estimating Note Phrase Aesthetic Similarity using Feature-based Taxicab Geometry. In *Digital Arts, Media and Technology (ICDAMT), International Conference on*. IEEE.
- [11] Legname, O. (1998) Density Degree Theory. SUNY Oneonta, www.oneonta.edu/faculty/legnamo/theorist/density/density.html.
- [12] Samson, A. V., & Coronel, A. D. (2019, February). Reproducing Musicality: Detecting Musical Objects and Emulating Musicality through Partial Evolution. In *Intelligence in Information and Communication (ICAIC)*.
- [13] Colby, K. M.; Hilf, F. D.; Weber, S.; Kraemer, H. (1972), "Turing-like indistinguishability tests for the validation of a computer simulation of paranoid processes", *Artificial Intelligence*, 3: 199-221.
- [14] M. Boden and M. Bishop, *The Turing test and artistic creativity*, Kybernetes.
- [15] McKay, C., Cumming, J., & Fujinaga, I. (2018). jSymbolic 2.2: Extracting features from symbolic music for use in musicological and MIR research. *Proceedings of the 19th International Society for Music Information Retrieval Conference*, 348-354.
- [16] Horner, A., & Goldberg, D. E. (1991). Genetic algorithms and computer assisted music composition. *Urbana*, 51 (61801), 437-441.
- [17] Cope, David, et al. *Virtual Music: Computer Synthesis of Musical Style*. MIT Press, 2004.
- [18] Yujian, L., & Bo, L. (2007). A normalized Levenshtein distance metric. *IEEE transactions on pattern analysis and machine intelligence*, 29 (6), 1091-1095.
- [19] Aggarwal, C. C., Hinneburg, A., & Keim, D. A. (2001, January). On the surprising behavior of distance metrics in high dimensional spaces. In *ICDT (Vol. 1, pp. 420-434)*.
- [20] Lavoie, T., & Merlo, E. (2012, June). An accurate estimation of the levenshtein distance using metric trees and manhattan distance. In *Software Clones (IWSC), 2012 6th International Workshop on* (pp. 1-7). IEEE.
- [21] Aucouturier, J. J., & Pachet, F. (2002, October). Music similarity measures: What's the use? In *ISMIR* (pp. 13-17).
- [22] Wongsaroj, C., Prompoon, N., & Surarerks, A. (2014, May). A music similarity measure based on chord progression and song segmentation analysis. In *Digital Information and Communication Technology and it's Applications (DICTAP), 2014 Fourth International Conference on* (pp. 158-163). IEEE.
- [23] Gusfield, Dan. "Linear-time construction of suffix trees." *Algorithms on Strings, Trees and Sequences: Computer Science and Computational Biology* (1997).
- [24] Miller, B. (2020). "All of the rules of jazz": Stylistic Models and Algorithmic Creativity in Human-Computer Improvisation. *Music Theory Online*, 26 (3).
- [25] Hsu, C. C., Tsai, Y. H., Lin, Y. Y., & Yang, M. H. (2020, August). Every Pixel Matters: Center-aware Feature Alignment for Domain Adaptive Object Detector. In *European Conference on Computer Vision* (pp. 733-748). Springer, Cham.
- [26] Pang, Y., Wu, Y., Wu, C., & Zhang, M. (2020). Salient object detection via effective background prior and novel graph. *Multimedia Tools and Applications*, 79 (35), 25679-25695.