

Evolved Aesthetic Analogies to Improve Artistic Experience

Aidan Breen^(✉), Colm O’Riordan, and Jerome Sheahan

National University of Ireland, Galway, Republic of Ireland
{a.breen2,colm.oriordan,jerome.sheahan}@nuigalway.ie

Abstract. It has been demonstrated that computational evolution can be utilised in the creation of aesthetic analogies between two artistic domains by the use of mapping expressions. When given an artistic input these mapping expressions can be used to guide the generation of content in a separate domain. For example, a piece of music can be used to create an analogous visual display. In this paper we examine the implementation and performance of such a system. We explore the practical implementation of real-time evaluation of evolved mapping expressions, possible musical input and visual output approaches, and the challenges faced therein. We also present the results of an exploratory study testing the hypothesis that an evolved mapping expression between the measurable attributes of musical and visual harmony will produce an improved aesthetic experience compared to a random mapping expression. Expressions of various fitness values were used and the participants were surveyed on their enjoyment, interest, and fatigue. The results of this study indicate that further work is necessary to produce a strong aesthetic response. Finally, we present possible approaches to improve the performance and artistic merit of the system.

Keywords: Computational evolution · Genetic algorithms · Genetic Programming · Grammatical evolution computational analogy · Aesthetic analogy

1 Introduction

Analogy making is a cognitive tool that humans begin to use from an early age with children as young as six demonstrating a clear understanding and use of spacial analogy in problem solving tasks [7, 8, 10, 22]. It has been said that analogy may even be “the most important cognitive mechanism” [10] that we use to make sense of the world around us. Computer science researchers, recognising the value of analogies, have explored the possibilities of computational analogy making in problem solving for quite some time. However, problem solving is not the only application of analogy. Analogy is used heavily in the artistic world as both an inspiration and subject, as we discuss in more detail in Sect. 2.2 below. It is the combination of computational analogy making and artistic analogy that we explore in this paper.

Recent work we have conducted demonstrates the design of a system capable of making aesthetic analogies between two artistic domains. The proposed system, described in detail in Sect. 3.1 below, makes use of *mapping expressions* to make artistic analogies. We have shown that these *mapping expressions* can be successfully evolved using Grammatical Evolution, discussed in Sect. 2.4, to estimate a mapping between two empirically gathered aesthetic data sets. While gathering aesthetic data sets in a reliable manner, and evolving *mapping expressions* is not trivial, implementing the proposed system to generate real-time visual displays based on a live music input is also a challenging task. The work presented in this paper describes an implementation of that system.

Beyond this implementation, an exploratory study was conducted to test the effectiveness of this approach and to guide the development of similar systems in the future. A number of music and visual displays were generated in real time, and recorded for reliable presentation to subjects. Subjects watched the displays and gave feedback on their enjoyment, the interestingness of each display, and their fatigue or boredom over time.

1.1 Contribution and Layout

The primary contribution of this paper is a description of the implementation of an aesthetic analogy system using *mapping expressions* evolved using Grammatical Evolution. This implementation demonstrates the value of a *mapping expression* as a real-time artistic tool and the validity of the system design proposed in previous work.

This primary contribution is composed of the following secondary contributions. First, we explore the use of live audio as an input to *mapping expressions* in order to generate visual displays. Second, we explore the use of MIDI messages in the place of a live audio signal as input. Third, we demonstrate that evolved *mapping expressions* of varying complexities can be evaluated in real-time using both audio signals and MIDI messages as input demonstrating the value of *mapping expressions* as artistic tools in a live performance. Finally, we demonstrate how a visual display can be generated using the output of a *mapping expression* with acceptable time delays.

The layout of the paper is as follows. First we present and briefly discuss related work in Sect. 2. In Sect. 3 we discuss the development of a system capable of creating real-time aesthetic analogies beginning with the evolution of *mapping expressions* using Grammatical Evolution in Subsect. 3.1 followed by the real-time system implementation in Sect. 4. In Sect. 5 we present the methodology of a study conducted to test the effect of the implemented system on the enjoyment, interest and fatigue of subjects. We present the results of this study in Sect. 6. Finally, we present a discussion and conclusion in Sects. 7 and 8.

2 Related Work

2.1 Computational Analogy

Computational Analogy (CA) focuses on the use of analogical problem solving as a computational approach to artificial intelligence. CA systems have been proposed since the 1960's [5]. As the field matured, analogical systems began to fit broadly into three main categories; Symbolic, Connectionist and Hybrid systems [6]. Symbolic systems made use of symbolic logic, means-ends analysis and classical logical techniques [1, 5, 21]. Connectionist systems made use of networks, with spreading-activation and back propagation building networks of similarity between domains [4, 12]. Hybrid models often combined other models and made use of an agent based, distributed structure [16, 17].

2.2 Artistic Analogies

Perhaps the most obvious examples of analogy in art might be found in the work of Wassily Kandinsky. Kandinsky is credited as being one of the first purely abstract painters, with his accounts describing what we now call synaesthesia, a condition that involves the mixing of senses within the brain. The condition results in a person literally seeing sounds, or hearing colours. Kandinsky held a life-long obsession with the connections between music and painting and also published a number of books on this topic [14].

Paul Klee produced and published similar work. Indeed, Klee and Kandinsky were colleagues at the Bauhaus school of art, design and architecture. Klee's handbooks [15] are cited as influential to most of modern abstract art with explicit notes on the representation of sounds as visuals.

In a more contemporary setting, lighting design has become ubiquitous with stage performance. Live music performance, especially popular, electronic, and rock music, all rely heavily on stage lighting design that enhances the feel of a performance. This can be seen as a practical analogy, where music is the source domain and lighting is the target domain. For some time, media player applications have displayed automatically generated visuals when playing music. These systems use simplistic generation techniques, and are not generally regarded as artistic, or artificially intelligent systems.

Further research has been conducted on the emotional connections between music and visuals. In her thesis, Behravan proposes a system which makes use of the detectable emotional aspects of music and visuals using psychological models of emotion together with an "artificial ear" and "artificial eye" [2]. Using a feedback loop and weighting mechanism tuned using Genetic Programming the system matches the emotional aspects of input music and generated visuals. The proposed system is functionally similar to the system proposed in this work, however the approach proposed in this paper aims to avoid subjective emotional aspects in favour of objective physical attributes within the domains of music and visuals.

2.3 Aesthetic Models

There have been numerous attempts to define aesthetic experience and to model or predict aesthetic value. Birkhoff’s over-simplified “aesthetic measure” formula $M = O/C$, the ratio of order (O) to complexity (C) [3] has been often criticised but it did begin an important discussion on the definition of aesthetics. More recently, this has been tackled by Ramachandran who defined 8 “laws of artistic experience” [20]. Of the factors outlined, many are measurable, such as contrast and symmetry, but others remain abstract such as “grouping”. Recent discussion has continued in the area [9, 11, 19].

The work in this paper relies on the existence of aesthetic models which are used as the core of the fitness function in the evolution of *mapping expressions*. Our previous work uses models gathered from two similar studies surveying people on their aesthetic preferences for pairs of music notes and pairs of colours. We adopt the same aesthetic models here.

2.4 Grammatical Evolution

Grammatical Evolution (GE) as proposed by O’Neill and Ryan [18] is a computational evolution approach whereby a grammar is used in a genotype-phenotype mapping. The application of this genotype-phenotype mapping enforces a stronger resemblance to biological systems where DNA codons are used to create proteins of a particular shape. In both systems, a many-to-one relationship occurs where many genotypes may produce one particular phenotype, which introduces a natural robustness while still allowing crossover and mutation to take effect. By defining a grammar, we can take any genotype and guarantee a grammatically correct phenotype.

A gene consists of an array of integers, known as codons. The grammar defines a starting non-terminal expression, and a set of terminal and non-terminal expressions. Terminal expressions represent fixed pieces of the output *mapping expression*, such as an input variable, a constant value or a mathematical operator. Non-terminal expressions represent pieces of the output *mapping expression* that are recursively replaced by other terminal or non-terminal expressions. The specific replacement expression is dictated by successive codons, with legal replacement expressions as defined by the grammar, equally distributed across the potential values of the codon. The recursive replacement continues until either a complete legal expression (phenotype) is created, or a length threshold is reached. It is common for an expression to require more codons than are present in the gene array and in this case, we simply begin at the start of the array again.

GE provides a number of distinct advantages. Primarily, a GE system suits the creation of executable expressions which can be easily defined by a relatively small grammar. In comparison to other Evolutionary Programming systems, the implementation is relatively straightforward and simple to implement. The output of the GE system, in this case a lisp like s-expression, can be parsed and

executed simply and efficiently. This is of particular value in this work as a real-time execution of the expression is necessary when generating visuals with live music input. Using a grammar provides the ability to include useful ‘pre-baked’ expressions like \sin , \cos and \log functions, as well as application-specific expressions like *plus90*, used in this implementation to represent offsetting a variable by 90 degrees on the colour wheel. Finally, the output *mapping expression* is a textual, human readable expression which can be stored in a text file for later evaluation, analysis or debugging.

3 System Methodology

3.1 Evolving Aesthetic Analogies

We have previously described an analogical approach based on a metalanguage constructed of *Mapping Expressions*. Each *Mapping Expression* maps one measurable aesthetic attribute in a source domain to a similar measurable aesthetic attribute in a separate target domain. An analogy may contain multiple *Mapping Expressions*, each mapping the value of some aesthetic attribute in the source domain to an attribute in the target domain. The structure of the metalanguage is illustrated in Fig. 1. In previous work, the musical harmony and colour harmony were used as measurable aesthetic attributes. We adopt this structure here also.

The GE system used to generate output *mapping expressions* is implemented as follows. The fitness of an individual is a measure of the similarity between the input musical harmony and output visual harmony. An individual contains a single gene. Genes are represented as an array of 8-bit positive integer (0–255) codons. Gene arrays of 60 codons were used within a population of size 50 and a population is seeded with entirely random genes. To begin the evolutionary process, the fitness of each individual is calculated. Tournament selection is then used to select individuals for evolution. Both single-point and double-point crossover are used to build the succeeding generation. Elitism is employed to maintain the peak population fitness from one generation to the next. Mutation is applied at the gene level where each codon may have its value randomly reassigned based on the mutation rate shown in Eq. 1 where α represents the number of generations since a new peak fitness was reached.

$$Mut_1 = \left(\frac{0.02}{70} \alpha \right) + 0.01 \quad (1)$$

A hyper-mutation rate is applied after a threshold to increase variation further. This approach ensures we allow adequate exploration of the genotype landscape while allowing local optimisation to occur for a short period. Evolution is halted after a threshold of generations where the peak fitness has not increased. The parameters used are summarised in Table 1.

Our previous work has shown that *Mapping Expressions* evolved with this procedure produce increasingly more accurate mappings between domains over time, with evolved expressions showing a distinctly higher fitness than randomly generated expressions.

Table 1. Genetic algorithm parameters.

| Parameter | Value |
|------------------------------------|-----------|
| Population size | 50 |
| Chromosome length | 60 |
| Crossover rate | 0.8 |
| Standard mutation rate (Mut_1) | See Eq. 1 |
| Hyper-mutation rate (Mut_2) | 1.0 |
| Mutation threshold | 100 |
| Halting threshold | 200 |

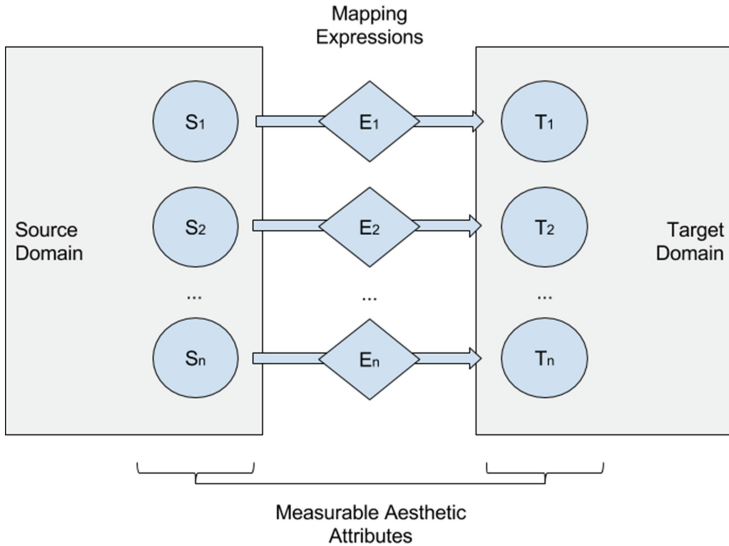


Fig. 1. Analogy structure overview with *Mapping Expressions* E_1 to E_n mapping source attributes S_1 to S_n to target attributes T_1 to T_n .

4 Real-Time Analogies

One of the major challenges of designing a system that receives live music input is making sense of the data being received. Our initial design intended to analyse live audio samples using signal processing techniques to detect salient information. While this approach has valuable applications, implementation is not reliable. Nevertheless, it still produces potentially useful results, presented in Sect. 4.1 below. The alternative approach, outlined in Sect. 4.2, uses a digital MIDI signal, removing most of the issues with live audio sampling in favour of a noise free, reliable input. This approach, however, is less flexible, limited to live input from digital instruments and synthesizers only.

4.1 Sampled Musical Input

A Fast Fourier Transform (FFT) was implemented to identify the frequencies being played in a live audio sample. In testing this approach we identified the strongest frequencies to be the fundamental frequencies of notes being played on an instrument. By taking the strongest N frequencies, the harmony of the audio could be calculated and sent on for *mapping expression* evaluation.

There are some obvious drawbacks to this approach. Chord recognition is a complex and actively researched field. Simply taking the strongest frequencies of a FFT is an unreliable, naive approach. The fundamental frequency of a note being played is not guaranteed to be the strongest frequency and the usefulness of this approach is diminished further as more instrumentation is included and the frequency spectrum becomes more crowded with overtones and noise. Another issue with this approach is the speed at which the calculated harmony changes. In testing, we found the harmony to fluctuate wildly as the frequencies identified changed. To combat this, a smoothing window was used to find the

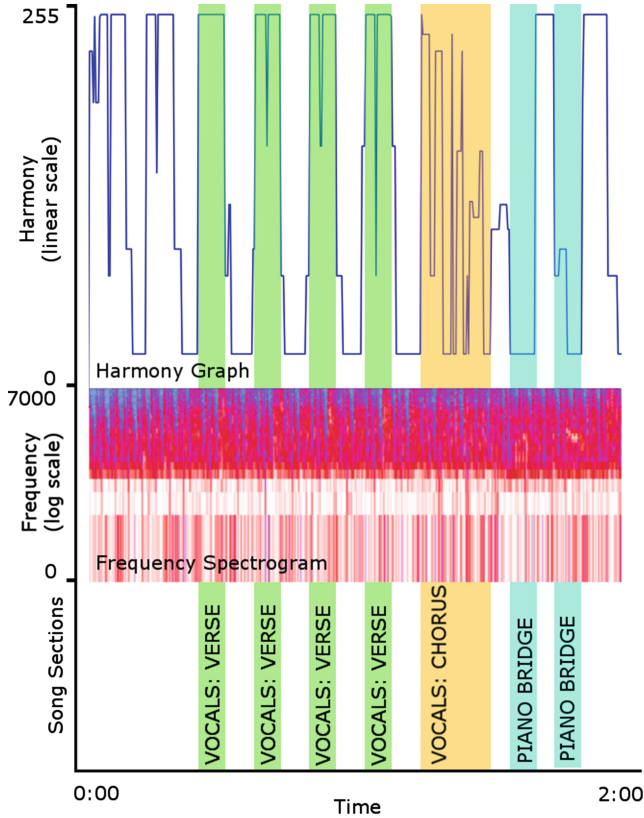


Fig. 2. Harmony calculated using a naive Fast Fourier Transform approach compared to the frequency spectrum and manually identified parts of a sample song.

mode harmony value. Using a short window of 10 samples was quite effective, improving the signal to noise ratio greatly without adding any significant delay.

Another major drawback of using a FFT with live audio is the timing of visual updates. Without a smoothing window, colour changes are rapid and distinctly unenjoyable. With a smoothing window, colours change at a more enjoyable pace, but do not seem to change in synchrony with any musical cadence.

Regardless of these drawbacks, using the FFT approach produced some interesting results as illustrated in Fig. 2. The figure compares the calculated harmony, the frequency spectrum and the manually identified sections of a sample song. Even using the naive FFT approach, it is clear that the harmony value is indicative of the part of the song, showing obvious differences between the verses, chorus and bridge.

4.2 MIDI Musical Input

To combat the downfalls of a FFT with live audio, we investigated the use of MIDI messages as a musical input.

MIDI (Musical Instrument Digital Interface) is a digital protocol originally designed to allow electronic controllers communicate with sound synthesizers in a modular fashion. The protocol was introduced in the 1980's but is still widely used today.

By using the MIDI protocol, we can ensure that we know exactly what notes are being played at any time without any noise or interference. This clearly restricts us to monitoring digital instruments in a live setting. However, for the purposes of this study, we can take a sample recording and score each instrument in MIDI manually. Using a Digital Audio Workstation we can then play the sample recording and the MIDI score in synchrony. This approach proved to be the most successful, producing a true representation of the music being played without any noise interference, and harmony values updating in synchrony with the music.

Interestingly, as illustrated in Fig. 3, the harmony value calculated using the MIDI approach is remarkably similar to the value calculated using the FFT approach, with parts of the section similarly distinguishable. This suggests that while using an FFT to detect the timing, notes and chords of live audio is unreliable, it may still be a useful approach to calculate harmony.

4.3 Mapping Expression Delay

With a reliable musical input available, the next step in generating a live visual output is evaluating a *mapping expression* with musical input. First, a chosen *mapping expression* is read from a file and parsed into memory as a set of nested sub-expressions. As we will be using only one *mapping expression* at a time, this can be done before any real time output is required, with no overhead. Evaluation is then a case of passing the parsed expression to an evaluation script which recursively evaluates each sub-expression.

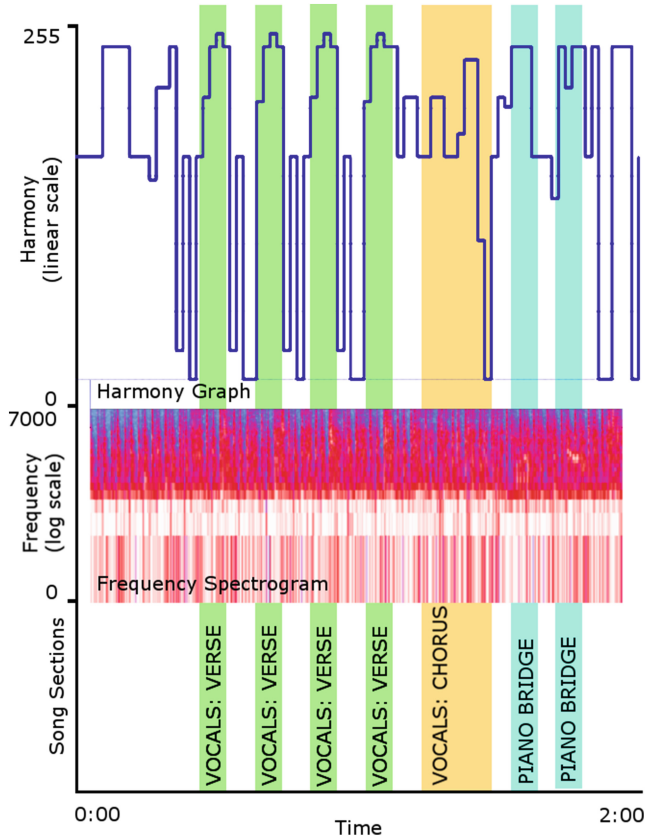


Fig. 3. Harmony calculated using a MIDI score compared to the frequency spectrum and manually identified parts of a sample song.

Evolved expressions vary in size between less than 10 sub-expressions, to over 7000, however, the vast majority of fit expressions contain less than 100 sub-expressions. The evaluation of larger expressions leads to a time delay. This is illustrated in Fig. 4, where we plot the average time taken to calculate the fitness of *mapping expressions*, by size, measured in number of sub-expressions. The figure displays the time taken to evaluate fitness rather than real-time evaluation times. This approach was chosen in order to obtain a fair evaluation delay across all input musical harmony values. The real-time evaluation delay can be accurately obtained by dividing the fitness evaluation time by 11, as shown in blue in Fig. 4.

4.4 Visual Display Generation

Finally, presenting and updating a visual display with minimal time delay is also a challenge. To achieve this, we implemented a *visualisation server*, a web

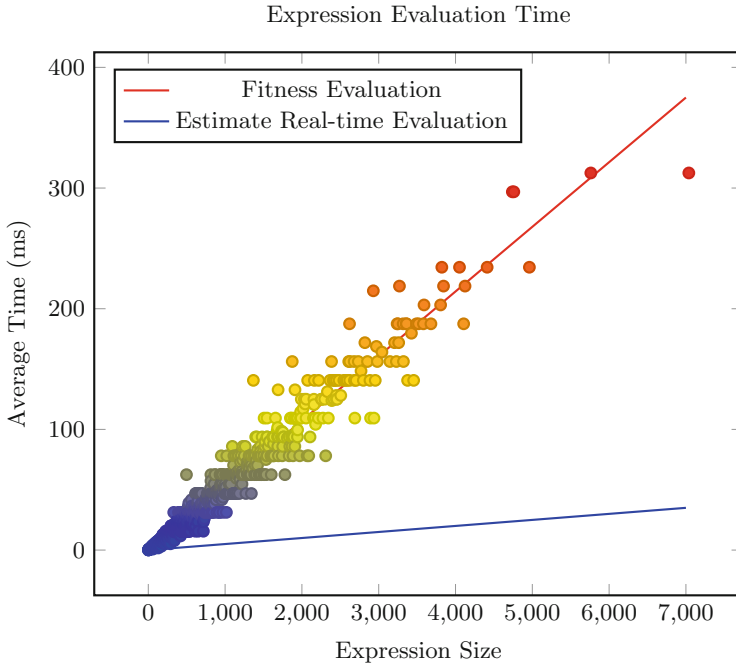


Fig. 4. Average execution time to evaluate the fitness of mapping expressions with an estimation of real-time evaluation. Mapping expression size is measured by number of sub-expressions.

server with API endpoints accepting colour update requests, and web pages served to render the display. Websockets were used to update the display as new colour update requests were received. The display could then be shown on a data projector, external monitors or any other display device. This approach was not ideal, but presented a cost effective and robust solution without the need for expensive stage lighting equipment and proprietary hardware for data transfer.

The overhead of evaluating a larger *mapping expression* as well as communicating with a visualisation tool could result in unacceptable delays, preventing suitable time synchronisation between audio and visuals. According to the ITU-R BT.1359-1 (1998) standard, a sound/vision timing difference of more than 90 ms is unacceptable, while anything less than 45 ms is undetectable. By plotting the time delay between a MIDI signal being received and a visual being rendered, as shown in Fig. 5, we can estimate this value. The figure displays a typical run with maximum delays of 79 ms, minimum delays of 11ms, and average delay of 24.357 ms indicating an acceptable delay.

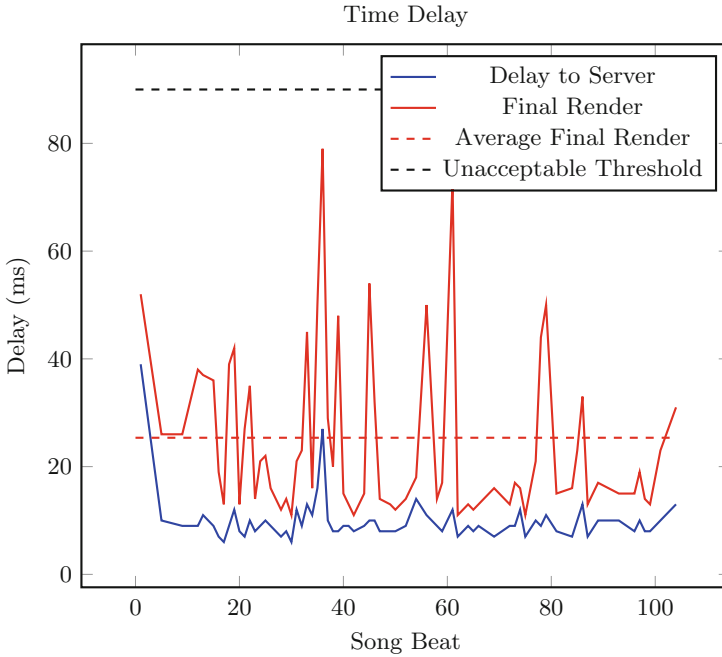


Fig. 5. Time delay in milliseconds from a MIDI message being sent to the Visualisation Server, and the visualisation being rendered for a single typical run

5 Survey Methodology

5.1 Experimental Setup

Using the approach outlined in Sect. 3.1, over 500 generations of *mapping expressions* were generated and stored. Three sample generations were selected and the lowest, median and highest fitness individual expressions were chosen for display. For each *mapping expression*, a visual display was generated using a sample music piece.

In the interest of displaying the same colours to all participants, and to avoid colour grading or exposure related loss of colour strength, colours were displayed on a screen which illuminated the room. The display consisted of a constant red background colour, and a varying foreground colour. The colours were arranged on screen as three vertical stripes, with the foreground colour in the center. This allowed the colours to illuminate the room distinctly.

The music piece, an arrangement of guitar, vocal harmony, piano, and light drums was used as it demonstrated a relatively strong periodic variation in musical harmony. The piece was composed by the authors to be enjoyable without focusing on any particularly polarising genre and also to avoid any intellectual property issues. An external speaker was used to ensure high audio quality throughout the study.

Table 2. Survey questions.

| Category | Question | Value |
|-----------|---|-------|
| Enjoyment | I enjoyed neither the music nor the colours | 0 |
| | I enjoyed the music but not the colours | 1 |
| | I enjoyed the colours but not the music | 2 |
| | I enjoyed both the music and the colours | 3 |
| Interest | I found neither the music nor the colours interesting | 0 |
| | I found the music but not the colours interesting | 1 |
| | I found the colours but not the music interesting | 2 |
| | I found both the music and the colours interesting | 3 |
| Fatigue | I feel tired or bored | |
| | I do not feel tired or bored | |

Table 3. Study participants summary.

| Attribute | Value | Percent (frequency) |
|--------------------|--------------------------|---------------------|
| Age | 15–24 | 37.5% (12) |
| | 25–44 | 43.75% (14) |
| | 45–64 | 18.75% (6) |
| Gender | Male | 56.25% (18) |
| | Female | 43.75% (14) |
| Musical background | No experience | 43.75% (14) |
| | Beginner or Intermediate | 43.75% (14) |
| | Expert | 12.5% (4) |

Each of the 9 videos were of 2 min in length and were presented in a random order for each viewing. After each video, participants were asked to provide one answer to each of 3 multiple choice questions as shown in Table 2. In total, 32 participants were surveyed. A summary of the age, gender and musical background distributions are shown in Table 3¹.

6 Results

Each participant was presented with nine visualisations, divided into three fitness levels: low, medium and high. Each of the four mutually exclusive responses to Enjoyment and Interest questions were given a value, 0 to 3 as shown in Table 2.

The input variables include 3 categorical variables (or factors), namely Fitness, Gender, and Musical Background and 1 quantitative input (or covariate), Age. The response values within each fitness level were summed, obtaining

¹ Full results are available upon request. Please contact a.breen2@nuigalway.ie.

an ordered categorical response variable, taking values 0 to 9 representing the aggregate enjoyment or interest of a participant at each fitness level.

One complication in the analysis is that certain observations are correlated, namely those on the same participant at the different fitness levels.

In the software SPSS [13], suitable analysis can be conducted using Generalised Estimating Equations. This analysis provides no evidence of an effect of any of the three factors (Fitness, Gender, Musical Background), nor of the covariate (Age) on either of the responses (Enjoyment and Interest).

7 Discussion

The core goal of this work is to examine the possibility and practicality of implementing a system such as that proposed in our previous work. To this end, a system has been implemented that makes use of evolved *mapping expressions* and generated real-time aesthetic analogies by taking a live musical input, evaluating *mapping expressions* and generating a visual display. While this system was capable of creating aesthetic analogies, there is no evidence that these analogies have an effect on the aesthetic experience of a piece of music. While this is a negative aesthetic result, the implementation was successful. Given these results, a number of improvements could be made to improve the performance of the system, in terms of both optimizing the computational aspects of the system, and improving the aesthetic appeal of output visual displays.

7.1 Implementation

The grammatical evolution based system proved successful for evolving *mapping expressions*. There are obvious improvements that could be made, including increasing the gene length, increasing the size of the population and tweaking other parameters to obtain fitter individuals.

Obtaining a reliable musical input proved to be a major challenge in this work. The final system used a predefined MIDI score as musical input. While this approach might suit a live performance using digital instruments, it is clearly not practical for an acoustic performance. Without a predefined MIDI score, further work is needed for live audio input.

While live music performance may currently be beyond reach for analogue instruments, using a Fast Fourier Transform with a smoothing window does seem to be useful for identifying harmony with more granularity. This may have practical applications in song-part identification.

With a predefined MIDI input, it may be possible to look ahead at the score to dynamically change analogies. Meta-data MIDI messages may also be used for this purpose. Further, this may also be a practical use of the sampled music approach which may be used to identify the section of a song with more subtlety than the MIDI based approach.

The evaluation of *mapping expressions* is currently not a major factor in the input-output delay. However, if multiple attributes, longer gene length, or more

complex expression selection approaches were put in place, this may become a bottleneck. Fortunately, the design of the *mapping expression* allows for some optimisation to be done. Firstly, with expressions and sub-expressions taking on a tree like structure, expression sub-trees that do not contain any dynamic input variables like music input may be pruned. Using this approach, large sub-trees could be replaced with constant values, greatly reducing the processing required for each evaluation of the expression. Secondly, the use of multiple aesthetic attributes lends itself to multi-threading or distributed evaluation which may further reduce any processing bottle-neck that occurs. Finally, it may be possible to build a single mapping expression using sub-trees built from separate genes. This approach would allow the distribution of processing even for the evaluation of single expression, though this approach would clearly not be required without substantially larger expressions.

There are a number of drawbacks with the approach taken for displaying visuals. Notably, the use of a web browser to generate a display presented on a screen or data projector is clearly inferior to the use of professional stage lighting hardware. While the current implementation was cost effective and quick to build, professional stage lighting hardware would create a far better atmosphere. The use of dedicated lighting hardware would also greatly reduce the input-output delay with current delays being mainly the result of rendering within the browser.

7.2 Survey

Our hypothesis assumed that an evolved *mapping expression* between the measurable attributes of musical and visual harmony would produce an improved aesthetic experience compared to a random *mapping expression*. The survey conducted did not verify this hypothesis. We believe this is primarily due to the effect of musical and visual harmony being too weak to have an observable effect. To remedy this, we believe the use of more rudimentary analogies, such as musical loudness to visual brightness, will have more obvious observable effects.

A second factor may be the simplistic structure of the analogy with just one *mapping expression*. Similar to the use of musical loudness to visual brightness suggested above, the use of these expressions in combination may also create more obvious observable effects allow a more dynamic analogy to be made, perhaps improving the aesthetic quality of the analogy. Using the implementation outlined in this paper, the use of multiple attributes should be possible without major architectural changes. However, gathering the required aesthetic response data may be prohibitive.

Our previous work suggests the use of a supervised fitness method which may encourage the creation of analogies that generate more recognisable effects. Using the music-visual analogy as an example, this process might complement a normal rehearsal, allowing human supervision without inducing a great deal of fatigue.

We have mentioned that the use of predefined MIDI lookahead and the sampled music approach may allow the use of multiple analogies for separate parts of a song. This approach may also improve the aesthetic quality of the system.

Finally, we hope to further investigate the consistency of feedback between judges using an output created by hand. Our survey is based on the assumption that an aesthetically pleasing visual can be created using the defined apparatus. It is possible that such an output is not possible or detectable given the constraints of the system and the inconsistency of responses.

8 Conclusion

It has been demonstrated that a system may be implemented to make use of Grammatical Evolution to make real-time aesthetic analogies by taking a live musical input, evaluating evolved *mapping expressions* and generating a visual display. *Mapping expressions* have been shown to have evaluation times fast enough to allow real time music to visual mapping. Furthermore, this mapping enables real time mapping within acceptable limits (90 ms), even considering the time delay resulting from input audio preprocessing and output visual rendering.

Statistical analysis of a survey conducted to investigate the effect of the system showed no evidence of any effect of any factors (Fitness of expressions, Gender, Musical Background) or the covariate (Age) on the enjoyment or interest of participants in a music and visual display.

Finally, a number of possible improvements have been proposed which may impact the aesthetic and artistic value of the system. The main proposed improvements are the use of multiple *mapping expressions*, the use of multiple measurable aesthetic attributes, and the use of a human supervised fitness method.

References

1. Becker, J.D.: The modeling of simple analogic and inductive processes in a semantic memory system. In: Proceedings of IJCAI-1969, pp. 655–668, Washington (1969). <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.77.8346&rep=rep1&type=pdf>
2. Behravan, R.: Automatic mapping of emotion in music to abstract visual arts. Ph.D. thesis, University College London (2007)
3. Birkhoff, G.: Aesthetic Measure. Cambridge University Press, Cambridge (1933)
4. Eliasmith, C., Thagard, P.: Integrating structure and meaning: a distributed model of analogical mapping. *Cogn. Sci.* **25**(2), 245–286 (2001)
5. Evans, T.G.: A program for the solution of a class of geometric-analogy intelligence-test questions. Technical report, Air Force Cambridge Research Labs LG Hanscom Field (1964)
6. French, R.M.: The computational modeling of analogy-making. *Trends Cogn. Sci.* **6**(5), 200–205 (2002). <http://www.ncbi.nlm.nih.gov/pubmed/11983582>
7. Gentner, D.: Children’s performance on a spatial analogies task. *Child Dev.* **48**, 1034–1039 (1977)

8. Gentner, D., Forbus, K.D.: Computational models of analogy. *Wiley Interdisc. Rev. Cogn. Sci.* **2**(3), 266–276 (2011). <http://doi.wiley.com/10.1002/wcs.105>
9. Goguen, J.A.: Art and the brain: editorial introduction. *J. Conscious. Stud.* **6**(6), 5–14 (1999)
10. Hofstadter, D.R.: Analogy as the core of cognition. In: Gentner, D., Holyoak, K., Kokinov, B. (eds.) *The Analogical Mind: Perspectives from Cognitive Science*, pp. 499–538. MIT Press, Cambridge (2001)
11. Huang, M.: The neuroscience of art. *Stanford J. Neurosci.* **2**(1), 24–26 (2009). <http://www.stanford.edu/group/co-sign/huang.pdf>
12. Hummel, J.E., Holyoak, K.J.: Distributed representations of structure: a theory of analogical access and mapping. *Psychol. Rev.* **104**(3), 427 (1997)
13. IBM Corp: IBM SPSS Statistics for Windows (2013)
14. Kandinsky, W., Rebay, H.: *Point and Line to Plane*. Courier Corporation, North Chelmsford (1947)
15. Klee, P.: *Pedagogical Sketchbook*. Praeger Publishers, New York (1925). http://monoskop.org/images/3/30/Klee_Paul_Pedagogical_Sketchbook_1960.pdf
16. Marshall, J.B., Hofstadter, D.R.: The metacat project: a self-watching model of analogy-making. *Cogn. Stud. Bull. Japn. Cogn. Sci. Soc.* **4**(4), 57–71 (1997)
17. Mitchell, M.: *Analogy-Making as Perception: A Computer Model*. MIT Press, Cambridge (1993)
18. O’Neil, M., Ryan, C.: *Grammatical Evolution*, pp. 33–47. Springer, Heidelberg (2003)
19. Palmer, S.E., Schloss, K.B., Sammartino, J.: Visual aesthetics and human preference. *Ann. Rev. Psychol.* **64**, 77–107 (2013)
20. Ramachandran, V.S., Hirstein, W.: The science of art: a neurological theory of aesthetic experience. *J. Conscious. Stud.* **6**(6), 15–35 (1999)
21. Reitman, W.R.: *Cognition and Thought: An Information Processing Approach*. Wiley, New York (1965)
22. Thibaut, J.P., French, R., Vezneva, M.: The development of analogy making in children: cognitive load and executive functions. *J. Exp. Child Psychol.* **106**(1), 1–19 (2010)