Sculpture Inspired Musical Composition

Francisco Manuel Alves Pereira Braga
francisco.braga@tecnico.ulisboa.pt

Instituto Superior Técnico, Lisboa, Portugal

November 2019

Abstract

Creativity is one of the most beautiful wonders of humanity. It is an essential factor for areas from the arts to science. Another concept, intrinsic to creativity, is inspiration. Inspiration is something that arises in us and motivates us to do all kinds of tasks. It can come from within us, or from the world around us.

In this thesis, we have tried to make an inspirational system that takes as its source of inspiration a sculpture to compose music. What we present is only a possible approach to the problem of mapping different domains. Our approach does not consider the semantics of the sculpture but instead looks at it more abstractly.

The result of our system is three different compositions: *Primal*, the direct result of the application of our approach; *Tonified*, a version of the first in which out-of-scale notes are altered so as to have only in-scale notes; *Genetic*, the result of an evolutionary process that uses a Genetic Algorithm.

The *Genetic* composition was evaluated through questionnaires. In these questionnaires, the musical quality of the composition and the association with the sculpture were evaluated. The results were promising, with the majority of the participants having given a high classification (4 out of 5) to the preferred interpretations of the compositions and relating them to the respective sculpture. Much work can still be done on this topic, nevertheless, we consider that what has been done in this thesis is a valid, interesting, and successful approach to this topic.

Keywords: Computational Creativity, Inspiration, Genetic Algorithm, Sculpture, Musical Composition

1. Introduction

Creativity is one of the pillars of several areas, such as arts or science and engineering, where it is mostly used when trying to solve a particular problem [6]. Many would consider trying to reproduce this unique skill an impossible task, or even immoral. Nevertheless, machines capable of mimicking creative behavior could have countless benefits, either helping in problem-solving or in being responsible for the creation or co-creation of works of art. From an artistic point of view, this would be very controversial since art is mainly considered a human activity. Still, systems able to apply creative processes are interesting both as a scientific project and in the world of art.

Computational Creativity is the field of Artificial Intelligence that deals with this type of research. In this field, not only computational aspects are investigated, but also other areas such as cognitive psychology and philosophy. As such, it deals both with practical and theoretical issues concerning creativity.

One can easily state that something is creative, either it being a thought, product, or something else than to define creativity itself. Still, there is some consensus regarding what creativity involves: a person, process, press, and product, i.e., "four P's of creativity" [8]. Although the four "P's" play an essential role in creativity, since we are working in the Computation Creativity field, we aimed at partially taking the person out of the equation. One element to consider in our case is that our goal was to create an inspirational system. As such, much importance is given, in our case, to the press aspect of creativity.

Inspiration is a fascinating topic by itself. We search for inspiration when attempting to create a certain product, regardless of the field. Inspiration can come from inside of us or from the world around. Focusing on creativity inspired by other creative artifacts, existing ideas inspire new ones in all sort of areas, such engineering, where new solutions to problems can be significantly influenced by already existing solutions, or in product design, where inspiration may come from shapes in nature [2], or art, where already existing artworks greatly

help one to create original works or styles [7]. Even though in these three cases inspiration comes from the same domain as the initial problem, it can also come from different domains. Here, there is a much more significant challenge because there is no direct connection. This forces the person to reconsider his/her framework since the inspirational artifact does not directly fit in it, leading to the construction of new frameworks and perspectives [7].

In the computational realm, inspiration can be seen as the mapping between two different artifacts. This mapping is made so that the new artifact gets the essence of the already existing one. Several authors were successful in this matter, by making systems capable of doing this mapping, but still, not many examples of inspirational systems are found in this field [10]. The big problem here is how to relate two different domains.

In our work, we aimed at making an inspirational system that composes music inspired by sculptures. We decided to draw the inspiration from existing human-made artworks since, this way, we were able to maintain the human hand in the equation. Three goals have driven our research:

- First we aim at building a system that both acts creatively, and generates a product that is considered creative.
- Second, we want to compose a piece of music that is inspired by a sculpture. As such, the music should, in a certain way, be associated with it.
- Third and last, we aim at building a system that composes music and not just sound. Although the exact difference between sound and music has not yet been defined, it is essential that the music can be considered aesthetically pleasing.

Our work has been divided into several stages. The first stage was to explore the two domains involved in our work, sculpture and music. In the next section we describe the background on these domains necessary to fully understand our system.

2. Background

2.1. Sculpture

When first looking at a sculpture, several characteristics come to our attention. These can be divided into two groups: shape features, regarding the shape of the sculpture, and texture features, regarding the color or texture of the sculpture. Moreover, we decided not to address the semantic meaning from sculptures, since this would be a complex task both computationally and in terms of interpretation, since each individual may have its own perception of a sculpture. So, a general approach that does not address semantics was chosen, leading to a more abstract interpretation.

When dealing with sculptures computationally (or any form of data), the first problem is the representation. There are many possibilities to obtain a sculpture in a computer, such as retrieving it from images, but the one we followed was using 3D objects since all the information regarding both shape and texture can be more easily obtained.

Considering the 3D objects domain, there are several different possible representations. In our work, we decided to use Polygonal Meshes, since it is a simple representation, yet it has all the information we needed. In this representation, the object is represented by its boundary surface, composed of several planar shapes (or *facets*), which are defined by a series of *vertices* connected by *edges*. The precision of this representation depends on the number of facets and vertices. It is more efficient computationally, since manipulating planar shapes is relatively simple (linear algebra), but in high precision requirements it may need a large amount of data [1]. The most commonly used shapes are triangles (*triangular* meshes). Having specified our choice of representation, we can move to the extracted feature.

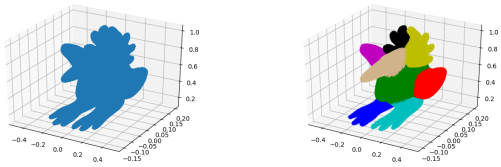
First considering shape, different features can be identified, some more intuitive than others. From all the features that could be extracted, three main ones were chosen as the most relevant for our work: curvature, angles, and segments.

Starting with the curvature, there are many possibilities to measure this feature. From all the possibilities, such as the two classical measures Gaussian and mean curvature [5], we used the mean curvature. Each vertex in the mesh has its own mean curvature.

Regarding the angles, these are measured using the normal vector of the faces. To obtain each vertex angle, the angles between all its faces are measured, and only the maximum is considered. One important aspect to note is that, since the angle is obtained from the normal vector, if the faces form a plane surface, the angle is $0°$, while two overlapping faces have an angle of $180°$. Using the angles, we also obtained another measure, which we called zero angle predominance. This measure verifies if the vertices with $0°$ angles in an object are inserted in a plane surface, or if they occur separately. If they are inserted in a plane surface, than our focus goes towards what arises from it, or to limits of that surface.

Finally, the segmentation is performed using the spectral clustering algorithm. This algorithm fits our purpose since its roots are on graph theory, and a mesh can be seen as a graph. It receives the adjacency matrix of the object's vertices and the num-

ber of clusters. It returns a label for each vertex. Using this, we divide the mesh into segments. In Figure 1, an example of segmentation of a 3D Mesh using Spectral Clustering can be seen.

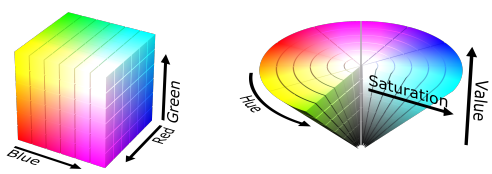


(a) Original (b) Segmented

Figure 1: 3D Mesh Segmentation example using Spectral Clustering

Moving to the texture, although we are extracting features from a 3D object, when it comes to this it is referring to a 2D surface. A simple analogy is one of a gift wrap. The gift itself is a 3D object, but the wrapping alone is a 2D surface. In order to do the mapping of the texture in the 3D object to 2D, there are several processes that can be can be applied, being the most common the UV mapping, that simply takes the coordinates of a 3D object, *(x,y,z)*, and converts it to 2D coordinates, *(U,V)*. Since we are dealing with a 2D image, the feature extraction process will be done using the same feature extraction algorithms as in a 2D image. As such, it is required to look further into color, starting by its possible representations.

Of the different available options, the ones more relevant are RGB and HSV. On the one hand, in RGB, any color is described as a combination of values of the three principal additive colors: Red, Green, and Blue. This model can be represented as a cube, as seen in Figure 2(a). On the other hand, HSV refers to a combination of Hue, Saturation, and Value. Hue refers to the color itself, saturation to the intensity (purity) of a particular hue, and value to the lightness or darkness of that color. This model can be represented as a cone, as seen in Figure 2(b).



(a) RGB Cube (b) HSV Cone

Figure 2: RGB and HSV representations

There are several different color features, being the two main categories histogram-based methods and color statistics [11]:

- Color histogram-based methods - These methods mainly represent the distribution of colors in the image. Although there are several ways to do this, the most common and useful way is the histogram of the distribution over the color model (3D in the case of RGB and HSV) or each channel of the color model.
- Color Statistics - These are statistical measures, such as the value, standard deviation, median, percentiles, among others.

Aside from these two categories, several features can be obtained by processing the image. From those, the two more relevant to our work are:

- Most Common Colors - To obtain these colors, a simple counting approach was used. First, nine equally spaced values were defined for each channel of the RGB model (thus a total of 729 distinct colors). Having the image's pixels represented in the same model, for each pixel, the closest value for each channel is chosen, and that color's count is incremented.
- Perceived Brightness - This can be seen as a conversion from the original colors to a greyscale, based on the brightness of each color. This feature can be obtained in several ways, for instance, a simple average of the three channels in the RGB model. In our work, to obtain a better brightness descriptor, a more complex formula was used. In this formula, different brightness weights are attributed to each of the RGB channels [3]:

$$PB = \sqrt{0.299 * R^2 + 0.587 * G^2 + 0.114 * B^2}$$

2.2. Musical Elements and Concepts

Music is a form of art which consists of organized sound throughout time. The source of that sound can be vocal, instrumental, among others, and it is usually divided into three parts: melody, rhythm, and harmony. These parts and their existence vary from culture to culture, being these three the most commonly used in Western music. Several existing elements in music govern each of the already mentioned parts, such as notes, chords, scales, and modes.

Starting with the most basic element, a note represents a sound. The two main characteristics of a note are its pitch and duration. The interval between pitches is measured by tones, being the smallest interval in western music the semitone. There are twelve pitch classes in western music that are represented with a letter from A to G. They may also have a symbol called accidental. For our purpose, the accidental may be either flat (♭) or sharp (♯). The flat accidental translates into a semitone bellow, while the sharp translates into a semitone above. For instance, having a G, if we introduce a flat symbol, it becomes G♭, i.e., a semitone bellow G, and if we introduce a sharp, it becomes G♯, i.e.,

a semitone above G. It is important to note that in standard pitch, the pitch class repeats when we double a certain frequency. For example, the pitch class A right above the middle C in the piano equals to 440 Hz. If we double that (880 Hz), we get another A one octave above. The duration of a note is the time the note lasts. This duration is represented as a fraction of what is called a whole note. For our purposes, not all possible durations were used

A scale is a repeating cycle of organized notes that usually repeats every octave. The notes of a scale can be organized in ascending or descending order, and, commonly, the step between neighbor notes is either a tone or a semitone. Scales may have a tonal center, which is usually the first degree of a scale, that is called the tonic. Considering the major scale, the steps separating each of the notes are T-T-S-T-T-T-S (T - tone, S - semitone). The natural minor scale is simply the major scale starting on the sixth note, thus creating the pattern: T-S-T-T-S-T-T. The chromatic scale is a scale with twelve pitches, and the step between pitches is always a semitone. Considering that this scale includes all twelve pitches used in western music, it has no tonal center, meaning it has no tonic.

Before moving to modes, it is most relevant to look into the common names given to intervals. These names are given according to the number of semitones that separate two pitches, and can be seen in Table 1. What is called the modern modes

N. of semitones	**Name**	**Abbreviation**
0	Perfect Unison	P1 (1)
1	Minor Second	m2 (♭2)
2	Major Second	M2 (2)
3	Minor Third	m3 (♭3)
4	Major Third	M3 (3)
5	Perfect Fourth	P4 (4)
6	Augmented Fourth	A4 (♯4)
	Tritone	TT
	Diminished Fifth	d5 (♭5)
7	Perfect Fifth	P5 (5)
8	Minor Sixth	m6 (♭6)
9	Major Sixth	M6 (6)
10	Minor Seventh	m7 (♭7)
11	Major Seventh	M7 (7)
12	Perfect Octave	P8 (8)

Table 1: Common interval names and abbreviations

are seven scales that derive from the major scale, i.e., that use the same set of notes. The difference between the major scale and the modes is that modes start in one of the seven notes of the scale, making it the tonic, thus having different intervals. It is also important to mark that each mode has one or more characteristic notes, which are notes that better describe the mode (concerning the other modes). In Table 2, all the seven modes, and their characteristics can be seen.

Mode Name	**Tonic relative to major scale**	**Intervals Name**
Ionian	I	P1 M2 M3 **P4** P5 M6 **M7**
Dorian	II	P1 M2 m3 P4 P5 **M6** m7
Phrygian	III	P1 **m2** m3 P4 P5 m6 m7
Lydian	IV	P1 M2 M3 **A4** P5 M6 M7
Mixolydian	V	P1 M2 M3 P4 P5 M6 **m7**
Aeolian	VI	P1 M2 m3 P4 P5 **m6** m7
Locrian	VII	P1 m2 m3 P4 **d5** m6 m7

Table 2: Modes (characteristics notes are in bold according to interval)

A common concept when talking about modes is their brightness [9]. This notion is quite simple: a mode with raised degrees, i.e., more sharps, is brighter than a mode with lowered degrees, i.e., more flats. The order of modes from brighter to darker is the following: Lydian (Brightest), Ionian, Mixolydian, Dorian, Aeolian, Phrygian, and Locrain (Darkest).

A chord is a set of multiple notes played simultaneously. The number of notes in a chord may vary, but the most common in western music are chords with three notes, called triads, and chords with four notes, called tetrads.

Triads are chords that commonly contain three specific notes: the root note, and the intervals (above the root) third, which can be major or minor, and fifth, which, for our purpose, can be perfect or diminished. A tetrad usually contains the same notes as the respective triad plus the seventh, which may be major or minor. By positioning the root note at each degree of the major scale, the existing chords of the major scale are obtained. Each chord can also be associated with the mode formed from the respective scale degree. In Table 3, we can see the chord associated with each mode, its characteristics, and its representation as a triad and a tetrad.

Degree of major scale	**Mode**	**Notes**	**Triad**	**Tetrad**
I	Ionian	M3 P5 M7	I	I∆
II	Dorian	m3 P5 m7	IIm	IIm7
III	Phrygian	m3 P5 m7	IIIm	IIIm7
IV	Lydian	M3 P5 M7	IV	IV∆
V	Mixolydian	M3 P5 m7	V	V7
VI	Aeolian	m3 P5 m7	VIm	VIm7
VII	Locrain	m3 b5 m7	VII∘	VIIm7♭

Table 3: Chords from the major scale and corresponding mode

The concept of consonance and dissonance is a way to characterize notes played simultaneously or successively. Consonance can be associated with pleasant, satisfying, and stable, while dissonance with unpleasant, troubling, and unstable. For our

purpose, it is relevant to further look into this concept applied to notes played above a particular chord. There are many different ways to put these concepts in pratice, from which we chose the following. If the chord already includes that note, it can be considered as consonant, while if the chord does not include the note, a simple rule can be followed: if the note creates the interval m2 with any note from the chord, it is considered dissonant. The importance given to a dissonant note also depends on several other factors, being the two most important the note's duration and where it occurs in the bar. For the most common time signature in western music, $\frac{4}{4}$, we have fours bars. The first beat is the strongest, followed by the third beat, and finally the second and fourth beat. A dissonant note played in a weak beat is less critical than one played in a strong beat.

Another important musical element is a motif. It can be described as a short musical idea that often occurs in a piece of music. The motif does not need to be always equal, and may be subject to changes while maintaining its general idea. We considered four possible changes, called variations: Melodic, when the melody is altered in some way, but the rhythmic value is maintained, Rhythmic, when the rhythm is modified, i.e., the rhythmic value of notes is altered, maintaining the pitch, Retrograde,when the motif is reversed, i.e., the first note becomes the last, and vice-versa, and Inversion, when the motif is mirrored, i.e., all the intervals are maintained, but their direction is the opposite, usually considering the first note as the reference.

Finally, we consider the already mentioned parts of a music piece: melody, harmony, and rhythm. A melody consists of a linear sequence of notes. The harmony may consist of several elements, such as multiples distinct melodic lines. For our purpose, harmony consists of multiple sequenced chords that will define a chord progression. From the different types, we used modal harmony. The rhythm is associated with the tempo and the time signature of the piece. Having already explained the time signature, the concept of tempo in music is simply the pace of the given piece, and it is measured in beats per minute (BPM).

3. Approach and Implementation

Since we aim at composing sculpture inspired music, we started by creating an analogy from one domain to the other. However, before doing so, we needed to restrict our approach regarding the musical domain. Music is such a vast field with so many different styles and types that it becomes quite impossible not to restrict. As such, we decided to base our approach on modal music. This decision was made considering that with modal harmony, one can more easily pass on a specific sensation, associated with the respective mode. This decision also influences the type of chords used, which shall be later explained. Besides this, we also only used the most common time signature in western music, $\frac{4}{4}$. The first step towards creating the analogy was to map the sculpture's features into musical elements and concepts. From the start, we did not want a too precise approach, for example, where each point would be converted into a note. We intended to obtain the musical elements through the analysis of sculpture's features in a more general manner, by considering either the sculpture or each the sculpture's segments as a whole. With this in thought, let us start with a high-level association between the two groups of sculpture's features, shape and texture, and the parts of a musical piece, melody, harmony, and rhythm.

If we listen to a melody without any harmony (chords), we may perceive a particular sensation or emotion. Once we listen to it with harmony, this sensation or emotion may vary drastically from the first. Moving to the sculpture domain, if we think about a sculpture's shape, we will obtain mainly sensation associated with its characteristics, such as smooth or rough. However, once we add the texture/color, we get, once again, a much better-defined sensation, or even emotion. One can say that the same way the texture/color gives context to the shape, the harmony gives context to the melody. As such, our approach's foundations arise from the association of the sculpture's shape with melody, and the texture/color with harmony. It is important to note that this is not imperative, and some associations were made from texture to melody. There is still one part of the music piece left to map, the rhythm. As already stated, the rhythm is associated with the time signature and the tempo. Concerning time signature, as already stated, we decided to only use the most common time signature in western music, $\frac{4}{4}$. About the tempo, although initially we wanted both shape and texture to be considered, we ended up using only the texture. This decision was only made when confronted with the final results since one can only evaluate the tempo having the music piece. As such, it shall be later on explained.

Starting with the melody, this element consists of a linear sequence of notes, while each note consists of pitch and duration. Accordingly, we tried to associate these two characteristics with shape features. We used the sculpture's lines or curve to obtain the pitch of the notes. If the lines of the sculpture are smooth, the pitches of the melody would have to give this same sensation, as well as the contrary. To obtain the line's description, we used the angles. We started by obtaining a general representa-

Angles (degrees)	0°-15°	15°-30°	30°-45°	45°-60°	60°-75°	75°-90°	90°-105°	105°-120°	120°-135°	135°-150°	150°-165°	165°-180°
Interval	M2	M3	m3	P1	M6	m6	P5	P4	m7	m2	M7	TT

Table 4: Mapping from angles to musical intervals.

tion of the angles, the histogram. The histogram is a representation of the distribution of data, where we can obtain the frequency of intervals of values. These intervals are called *bins*, which may or may not have the same width. Having this, we decided to relate the musical interval for the next note concerning the current note with each *bin*. Since in one octave there are twelve possible intervals, the histogram was calculated using twelve *bins*. With the histogram, we created a probability distribution to obtain a musical interval, thus having the probability for the next note's interval in relation to the current note. The mapping made from angles to musical intervals can be seen in Table 4

One issue with this approach is the fact that if one particular *bin* does not contain any occurrences, that interval will never be used. We acknowledged that an interval of angles could not only be mapped to the respective musical interval, but also the neighbors' musical intervals. With this in mind, after using the probability distribution to obtain the musical interval, we also apply a normal distribution to give the neighboring musical intervals a chance to be chosen. We used what is know as the standard normal distribution ($\mu = 0$ and $\sigma = 1$). Using a value obtain from the normal distribution, there are three possible cases: if the value is between -1 and 1 (probability of 68.3%), the original musical interval is chosen; If the value is below -1 (probability of 15.85%), the left neighbor musical interval is chosen, unless the original value is the first position, in this case, the original musical interval is chosen; If the value is above 1 (probability of 15.85%), the right neighbor musical interval is chosen, unless the original value is the last position, in this case, the original musical interval is chosen.

The only aspect left to decide is the direction of the interval, being the only musical element in the melody that we decided to map from the texture. Here, we decided to make use of the perceived brightness. Colors with a higher perceived brightness value would be related to upwards intervals, and colors with lower values with downwards intervals. From the histogram of perceived brightness values, we took the value that occurred more often. This value ranges from 0 to 1, and set the probability of the interval direction being upwards between 20% and 80%.

The duration of a note (or rhythmic value) will significantly influence the way we perceive a music piece. If we think of a sculpture with much variation concerning the curvature, one can easily associate that with a dense music piece, where the notes keep changing in a small fragment of time. The opposite can also be related. As such, we aimed at relating the variation of the sculpture's curvature with the note's duration. To obtain the curvature's variation, we used the autocorrelation. We gathered the curvature value of each point, grouped them by neighbors, and used autocorrelation in these groups. We decided to use the mean of all results, and map this value with what would become the most probable rhythmic value. The mapping can be seen in Table 5.

Mean of the Autocorrelation Results	0-0.05	0.05-0.1	0.1-0.4	0.4-0.7	0.7-1
Most Probable Rhythmic Value	𝅘𝅥𝅯	𝅘𝅥𝅮	𝅘𝅥	𝅗𝅥	𝅝

Table 5: Mapping the mean of the autocorrelation results to the most probable rhythmic value.

We did not want only one type of rhythmic value. We were looking for a general way to relate the autocorrelation results with the rhythmic value without making a direct relation. To achieve this, the Standard Normal Distribution was used once again. Using this tool, we can set the most probable rhythmic value while maintaining the possibility of using other rhythmic values. For each note, a rhythmic value is chosen based on the most probable one and a value obtained using the normal distribution.

Instead of composing all the melody at once, we decided to compose motifs that would later be joined, composing the whole melody. As seen before, a sculpture can be segmented. As such, we decided to compose each motif using features obtained from each segment. To obtain the size of the motif, we used each segment's size in relation to the whole sculpture.

Having the various motifs that will form the melody, we can now move to the harmony. It was decided to relate the harmony with the sculpture's texture/color. As such, we want to harmonize each motif according to the respective segment's color. For each segment, we obtained the most common colors, as explained before. Having the HSV color model in mind, for the same Hue, and the highest Saturation (100%), the Value channel will give us how dark or how pure that Hue is, ranging from black (Value close to 0%) to the pure Hue (Value close to 100%). Similarly, the musical modes can also be ordered from darker (Locrian) to brighter (Lydian).

As such, having the most common colors, we re-

lated the Value channel from the HSV color model to musical modes. A color with a lower value in the Value channel would be associated with a darker mode, while one with a higher value with a brighter mode. However, not all Hues should be matched with all modes, since, for the same Value, some colors are perceived as darker than others. As such, the Hues were divided in three groups, and for each group the Value channel maps to a range of modes. The division in groups can be seen in Figure 3.

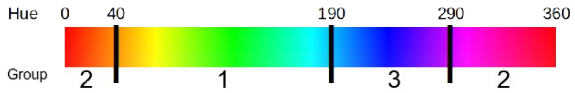


Figure 3: Hue value based groups

Besides these three groups, two special cases are not accounted for by the Hue values: If the color is black, or very close to black (Value channel below 20 %), the mode is set to Locrian (the darkest); if the color is white, or very close to white (Saturation channel below 10%), the mode is set to Lydian (the brightest). It is important to note that even if the Saturation is near 0% if the Value channel is close to 0%, the color will be close to black. As such, the first verification to be made must be if the color is close to black (Value ¡ 20%). The range of modes per color group association and the special cases can be seen in Figure 4.

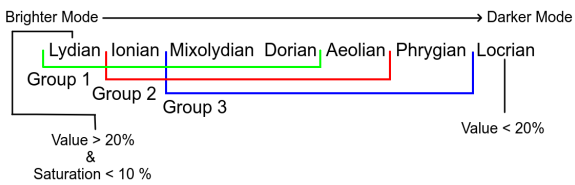


Figure 4: Range of Modes per Color Group and Special Cases

To obtain the chord, the closest tone of each motif is found, and then obtain the mode's chord in that tonality, as seen in Table 3. For each segment, we get the two most common colors, from there the two respective modes, and finally, the two chords. The chord obtained from the most common color is placed first in the motif and the one from the second most common chord second. The number of bars for the first chord is obtained by dividing the motif's number of bars by two and rounding up, while the remaining bars are left for the second chord.

To join the motifs in a logical order, we decided to get the probability distribution of intervals from the histogram of angles, although this time from the whole sculpture. From there, we calculate which of all possible permutations of the motifs' order best corresponds to the probability distribution of the intervals, being this the chosen order.

The only aspect considered left to map is the tempo. For this element, we used the variation of color among segments. To get this variation, first, we obtained the most common colors of the whole sculpture. Then, for each segment, we verified how many of the most common colors of that segment matched the ones from the whole sculpture. In the end, we obtained a ratio of the colors that matched the total colors verified that ranges from 0 to 1. The higher the ratio, the lower the tempo should be, as well as the contrary. For our purposes, we considered a tempo range of 80 BPM to 160 BPM.

Having the melody and the harmony from the ordered motifs, and the respective tempo, we obtain the first version of the music piece, we can call it the raw composition. The second version was obtained by correcting notes in the motifs that did not belong to the respective scale, changing them to the closest note found on the scale. This way, we obtain a version that respects each of the motif's scale. One might call it a "tonified" composition.

Besides these two versions, we decided to use a Genetic Algorithm. This algorithm provides a way to search for better results while maintaining a certain randomness that leads to exciting results.

This algorithm may vary from application to application, where the representation is can be seen as the basis. we defined each individual (candidate solution) as a music piece. As already stated, we considered a music piece to be the gathering of several motifs (melody and harmony), their ordering, and the music's tempo. Therefore, each individual being a music piece, its genes are motifs. There is a crucial property in each individual that is the ordering of the motifs (genes). The algorithm is divided in the following phases:

1. Initialization - This phase is the initial stage of the algorithm. We decided to have a population of 100 individuals. The number of motifs generated per segment is the chosen number of individuals. Each individual is created by randomly choosing a motif from each segment. The initial order of the motifs is randomly chosen.
2. Elitism - In this each generation, 10% of the individuals with the highest fitness are directly passed to the new generation. This is done to guarantee some level of quality.
3. Selection - A pair of individuals is selected. The probability of each individual to be chosen is based on the individual's fitness. The higher the fitness, the higher the probability of being chosen. As such, one individual may be selected several times or no times at all.
4. Crossover - Having a pair of selected individuals, there is a probability of 80% of suf-

			$weight <= 1$	$weight > 1$
Note belongs to the Mode	Characteristic	Consonant	$2 * weight$	
		Dissonant	$2 * weight$	$-weight$
	Not characteristic	Consonant	$weight$	
		Dissonant	$weight$	$-weight$
Note does not belong to the Mode			$weight$	$-weight$

Table 6: Mode definition Fitness measure for different types of notes

fering crossover. We defined two types of crossover, both equally probable: motif and note crossover. In the first case, half of the individual's motifs are randomly selected and crossed over. This crossover occurs between motifs generated from the same segment. Considering the second, the pitch of the individuals is swapped. Since one may be bigger than the other, among the proper starting points, one is randomly chosen.

5. Mutation - The individuals obtained have a probability of being mutated. This involves altering the individual's genes. Four types of mutation were used, each with an independent probability. The first is order mutation, with a probability of 10%. In this case, the order of the motifs is randomly altered. The second is the "tonify" mutation, with a probability of 10% for each motif. In this case, notes that do not belong to the motif's scale are corrected to the closest note found on the scale. The third and fourth are the Inversion and Retrograde mutations, each with a probability of 5% per motif. In these cases, the respective variation is applied.

6. Fitness - The fitness for all individuals is calculated at the end of each generation. This is obtained from a fitness function. In our case, there are three measure to obtain the global fitness: order, mode definition and range fitness. The order definition fitness evaluates how well the order matches the probability distribution of intervals obtained from the histogram of angles from the whole sculpture. This measure ranges from 10 (perfect fit) to -10 (no fit). The mode definition fitness evaluates how well the motif describes the respective mode. For this, each note has a weight according to its beat and its duration. Only the beat with the highest value is considered if the note includes more than on beat. If it is in the first beat, the weight is $4*duration$, the third $3*duration$, the second $2*duration$, and the fourth $1*duration$. The value for this measure is than calculated according to the type of note regarding the mode, as described in Table 6. The range measure penalizes -4 for pitches that are not in an acceptable range. We defined that acceptable range between 55 and 90 in MIDI.

7. Termination - To terminate the algorithm, we decided to have a fixed number of generations. Regarding the final music piece, we could obtain it from the fittest individual in the final generation. However, our interest has moved towards an approach that considers the *evolutionary process as artwork* [4]. In order to do this, we decided to obtain the fittest individual in each fixed number of generations. In the end, the individuals are joined and form the final music piece. We decided to terminate the algorithm once it reached 300 generations. To obtain the final results, we obtain the fittest individual every 100 generations.

4. System Architecture

The approach above explained was implemented through the architecture seen in Figure 5.

At first, our system receives a sculpture. The first module encountered is the **Sculpture Module**. In this module, the sculpture (3D object) is processed and analyzed to obtain the necessary features. The results of this module are the sculpture's shape features (segments, angles, and mean curvature) and texture features (perceived brightness and colors).

Having the necessary features, we pass them on to the **Mapping Module**. This module is responsible for the mapping of the sculpture's features to elements of the musical domain, or mechanisms to obtain them. One can see it as the middleware between the sculpture part and the musical part. The result of this model is a set of properties or values to be used in the musical domain. These include the interval probability distribution and the interval direction probability for the pitch generation, the rhythmic value to be used as the most common, and the mode to be used. One important aspect to note is that different values are obtained for each of the sculpture's segments. The Tempo and the general interval probability distribution (from whole sculpture) are also outputs of this module.

The values and properties obtained in the mapping module, except for the Tempo that shall be directly applied in the music piece, are passed on to the **Composer Module**. Here, with all the necessary information, motifs are composed. The melody is obtained using the interval probability distribution, the interval direction probability, and the rhythmic value for the common case. The harmony is composed using the modes obtained, and

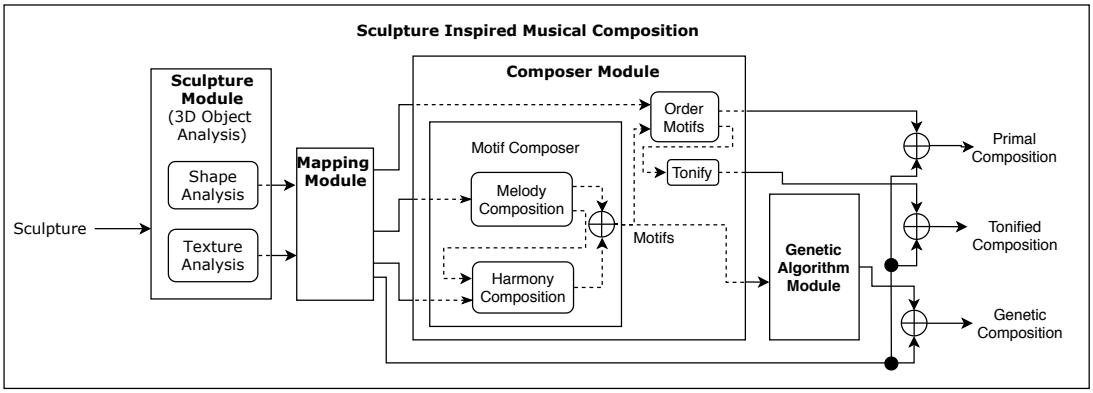


Figure 5: System Architecture

considering the melody. Following this module, there are two possible scenarios:

- The motifs are ordered using the general interval probability distribution - In this case, we obtain two compositions after adding the Tempo: the Primal Composition, and, after being "tonified", the Tonified Composition.
- The motifs are passed on to the **Genetic Algorithm Module** - In this case, the composer module generates a fixed number of motifs from each segment, instead of only one. These motifs are used in the genetic algorithm, and, after adding the Tempo, the Genetic Composition is obtained.

5. Evaluation

Six 3D Models of sculptures were used as the dataset. These had two sources: models found online, and models captured and processed by the author. One example can be seen in Figure 6. For the

Figure 6: Galo de Barcelos[1]

six sculptures, the genetic composition was generated. This was the only composition considered for the evaluation. Having the six compositions, three possible interpretations of each were generated using virtual instruments. One closer to Jazz, one to Rock, and one to Electronic/Space Rock.

To evaluate our system, the most important aspects to consider are our objectives. Having this in mind, we wanted to verify if the music pieces composed are associated with the sculpture and if the music is considered aesthetically pleasing. Although one of our goals was to make a system that acts creatively and generates creative products, this aspect was not evaluated directly. Instead, we considered that the evaluation of the other two goals would lead to this goal's evaluation.

We decided to use online surveys to perform the evaluation. Two surveys were made, each containing three compositions. Having three interpretations for each composition, the main part of the survey was done concerning the preferred interpretation. For each, we evaluated the quality of of the music piece and the association with the sculpture. The least preferred is also evaluated, to establish a ground floor for interpretation. The results are explained in more detail in the full version of the thesis. In here, we them very briefly.

Regarding the quality of the music, the statistics for all compositions can be seen in Table 7. In this particular question, the participants rated the music for the preferred and the least preferred interpretations. The responses obtained were very good for the preferred version. On a scale of 1 to 5, considering the median of 4, we can state that at least 50% of the responses rated the pieces as 4 or higher. These results are the verification for the musical quality of the composed music pieces. The least preferred interpretation had lower results, as expected. However, the rating lowered about one step. These results were expected, and we considered them to be good since this was the least preferred interpretation of the three, and still got a median and mode of 3.

How would you rate this music?	Preferred	Least Preferred
Mean	3.55	2.79
Median	4	3
Mode	4	3
Std. Deviation	0.92	1.03

Table 7: How would you rate this music? Statistics

Considering the sculpture-music association, the results vary more from case to case. Two special cases were encountered. The first case, where the semantics of the sculpture led to a perspective drastically different from the one obtained through our approach. In this case, the results were obviously bad. In this case, the description given by the par-

ticipants for the sculpture was very distant from the composition. The composition was not considered related to the sculpture since, for both the preferred and the least preferred interpretations, about 70% of the participants rated the relation as either 1 or 2 on a scale of 1 to 5. The second special case is encountered in compositions inspired by very subjective sculptures. In this case, there is much variation on what is perceived from the sculpture. Despite this variation, for the preferred composition, at least 50% of the participants rated the relation as 3 or higher. For the least preferred interpretation, 51.6% of the participants rated the relation either as 1 or 2. For the typical case, the results were good, since most people related the music to the sculpture. Regarding the description, there was a clear connection between the sculpture and the music. Considering this case, for three out of the four sculptures, more than 50% of the participants rated the relation for the preferred interpretation as 4 or higher. For the least preferred, the values were about one step lower than for the preferred, which is expected.

6. Conclusions

The main objective of our thesis was create a system that composes music inspired in sculptures. Our approach involved relating the shape of the sculpture with the music's harmony, and the texture with melody. This association was not a strict association, yet it was our basis.

We believe that we created an inspirational system with an interesting approach that balances the freedom associated with inspiration with a logical mapping from the sculpture domain to the musical one. During the process, we imposed many limitations and restrictions that are both associated with any approach, but special to ours due to the vast musical knowledge used to build the system.

It is imperative to reinforce that what we have done is one possible approach. As such, for future work, there are other possible approaches that could be explored. Starting with the musical realm, more elements and concepts could be used. For example, instead of using a modal approach, use a tonal or atonal approach or maybe even consider the three, limiting the system much less. The use of different time signatures could also be implemented since we only used one. As for the association approach, others should be tested. The shape-harmony and texture-melody association was made at the beginning, and exploring other possibilities could be interesting. Other approaches to the genetic algorithm could also be investigated. One possibility would be to give more freedom to the algorithm by not limiting the crossover between motifs generated from the same segment of the sculpture. This could be compensated by adding a fitness measure that would favor individuals where all the segments are represented. A more concrete aspect to could be worked on is the evaluation since only one of the three output compositions was evaluated. The three could be assessed to verify, for example, which one is preferred and which one is considered to be more related to the sculpture. Besides this, having human interpretations of the composition could significantly improve the evaluation. The quality of music played by real musicians may be much better than interpretations generated using a computer.

References

[1] J.-L. Dugelay, A. Baskurt, and M. Daoudi. *3D Object Processing: Compression, Indexing and Watermarking*. John Wiley & Sons, 2008.

[2] C. Eckert and M. Stacey. Fortune favours only the prepared mind: Why sources of inspiration are essential for continuing creativity. *Creativity and Innovation Management*, 7(1):9–16, 1998.

[3] D. R. Finley. HSP Color Model - Alternative to HSV (HSB) and HSL. `http://alienryderflex.com/hsp.html`.

[4] C. G. Johnson. Fitness in evolutionary art and music: what has been used and what could be used? In *International Conference on Evolutionary and Biologically Inspired Music and Art*, pages 129–140. Springer, 2012.

[5] J. J. Koenderink and A. J. Van Doorn. Surface shape and curvature scales. *Image and vision computing*, 10(8):557–564, 1992.

[6] A. Koestler. *The Act of Creation*. Hutchinson, 1964.

[7] T. Okada and K. Ishibashi. Imitation, inspiration, and creation: Cognitive process of creative drawing by copying others' artworks. *Cognitive science*, 41(7):1804–1837, 2017.

[8] M. Rhodes. An analysis of creativity. *The Phi Delta Kappan*, 42(7):305–310, 1961.

[9] M. Ron. Modal jazz composition & harmony. *Advance Music, Rottenburg*, 1996.

[10] J. Teixeira. Cross-Domain Analogy: From Image to Music. Master's thesis, Instituto Superior Técnico, Lisbon, Portugal, 2017.

[11] M. J. Álvarez, E. González, F. Bianconi, J. Armesto, and A. Fernández. Colour and texture features for image retrieval in granite industry. *Dyna*, 77:121 – 130, 2010.