From Music to a Sculpture

Joana Silva Teodoro
Instituto Superior Técnico
Lisboa, Portugal
joanasilvateodoro@tecnico.ulisboa.pt

ABSTRACT

Creativity is a phenomenon that helps people express an emotional state or solve a problem. With the application of technology in daily tasks, Computational Creativity was born. This fields' aim is to develop systems that can reproduce and help the creative process.

In this work, we developed a system that uses music as inspiration and generates sculpture, using common techniques from this field. We created an analogy between the two domains - an association derived from the authors' perception that translates music information into sculpture information.

To evaluate the system, we carried online surveys to find out if people liked the sculptures, how they would describe them, and if they associated them with a given song. The results were positive - the participants liked the sculptures and, in general, associated the sculptures with the music from which they were inspired. This way, we believe that we contributed to the field of CC.

Author Keywords

Computational Creativity; Inspiration; Musical Analysis; Sculpture Generation; Evolutionary Algorithms

INTRODUCTION AND MOTIVATION

Creativity is part of us all. We have experienced a creative process when we have an idea, no matter if it is something simple, like improvising a recipe, or if it is something complex, like creating a lamp. As human beings, we evolve and shape the way we think, we speak, we act, and, consequently, the way we create.

Creativity can occur in very different ways but there are two main drives to create [18]:

1. We need to express an emotional state.
2. We adapt to our needs, i.e., when we have a problem to solve.

Computational Creativity is the field that uses Artificial Intelligence (AI) to build systems capable of creating artefacts considered creative. In this field, there are already several systems that are autonomous and contain some creative behaviour.

With this work, we studied and contributed to the Computational Creativity field. Our main goal was to develop a system capable of performing a creative process, generating new and valuable sculptures that are an inspiration result of a music file. We used use the inspiration's definition of Thrash, T. and Elliot, A. (2003) [20]:

> ”Inspiration implies *motivation*, which is to say that it involves the energization and direction of behavior; inspiration is *evoked* rather than initiated directly through an act of will or arising without apparent cause; and inspiration involves *transcendence* of the ordinary preoccupations or limitations of human agency.”

This means the music piece will be the motive, making our system transcendentally inspired and, consequently, evoke sculptures from it.

The music pieces we are going to use are artefacts already considered creative. Moreover, the created sculptures are not the only possible result, meaning that other sculptures can derive from the same music piece.

RELATED WORK

The way we perceive things and appreciate what surrounds is different, which, consequently, can lead one to consider something as creative while others might not. Therefore, how does one classify something as creative?

Definitions

Sawyer (2011) [18] introduces the **“big C” Creativity** and **“little c” creativity** concepts. He considers the solution of an extremely difficult problem, or an important and genius work, as creative and these ideas get included in the “big C” creativity. Small activities, like finding a way to wrap a present when you do not have any gift paper, or modifying a recipe when you do not have all the necessary ingredients, are labelled as “little c” creativity. Moreover, the author states that creativity cannot exist without **appropriateness**. The natural and social entities that can generate novelty are:

- **Nature:** Nature is the creator - Natural Creativity.
- **Group:** The group creates a product. This kind of creativity involves distributed cognition (each member has a weighted contribution to the product) - Group creativity.
- **Society:** Society is the big influence when creating. To better explain this type of creativity, we need to consider the following factors: economical and political strength, cultural values of the community, system of patronage of the wealthy, and apprenticeship systems for the new emerging artists - Societal Creativity.

According to Boden (2007) [11], there is a difference between something creative and something new. To present something creative, that something must be not only new, but also surprising and valuable.

In 1961, Rhodes [16] announces that creativity has four P's that influence the creative act:

- **Person:** Information about personality (likes, dislikes, habits, and moral values).
- **Process:** Motivation to create, perception of things, way of thinking, and communicating.
- **Press:** Relationship with others, social environment and how they influence the creator.
- **Product:** The created artefact and how it is communicated to others.

With these four concepts, we can define creativity as "A *process* executed by a *person*, *pressed* by his/her environment, by which a *product* is generated" [19].

In 2009 [12], Boden defines **H-creativity** and **P-creativity**. These are two different types of creativity, based on historical creativity (H-creativity) and psychological creativity (P-creativity). H-creativity is concerned with human history ideas. It means that no one has come up with that idea before, and P-creativity deals with an idea that needs to be valuable and surprising to the person who creates it [11].

Newell et al. (1962) [15] classifies creativity as problem solving. To find a way to solve a problem is to be creative, and four aspects must be present:

- Having a novel and a valuable idea.
- Having an unconventional idea.
- Having motivation to solve the problem.
- Having clarification of the problem [19].

Sawyer (2011) [18] presented the four stages of creative process the following way:

- **Preparation:** Initial phase for collecting data, information and search for related ideas.
- **Incubation:** Moment between preparation and the insight, where the information collected is organized.
- **Insight:** The "eureka" moment. where the idea arises.
- **Verification:** This stage is based on two substages:
    1. Evaluation of the idea.
    2. Elaboration of the idea.

Systems

Horn et al. (2015) [14] implemented a creative inspiration model named VIV that, through a 2D source image (given by the user), generates 3D-printable vases. For the image, VIV needs to check four aesthetic measures (based on the image's colours): Activity, Warmth, Weight, and Hardness. They performed an image evaluation on these measures, which creates a profile for the image that VIV uses, afterwards. After uploading an image online, the model provides the resulting 3D vase through genetic evolution (a population of 100 vases, over 100 generations). Using the image's colour profile, VIV proceeds to use an evolutionary algorithm, creating vases with a similar profile. The produced vases are ranked according to the aesthetic measures and the highest-ranked is the final printable vase returned to the user.

In 2020, Aleixo, L. et al. [10] proposed a possible approach for a cross-domain association between the musical and visual domains. The system attempts to exhibit creative behaviour by generating abstract images, inspired by musical artefacts. It identifies and extracts available features from a MIDI music file given as input, interpreting and associating them as a starting point to translate into several components of the visual domain. It generates three different outputs that were evaluated through online surveys:

- Random Image: results from a pure translation of the musical features into the visual features. The association between musical instruments and visual shapes is random.
- Associated Image: similar to the Random Image except that the association between each musical instrument and respective visual shape is predefined and deeply studied through one platform developed for the purpose.
- Genetic Image: results from the application of one Genetic Algorithm that uses the two previous versions to generate the initial population.

The system uses information extracted from melody, notes, harmony, chords, and rhythm of the music that influences the final image features (colour, shape, position and size - either in the background or foreground).

Rowland et al. (2002) [17] produced a genetic sculpture park. This park aims to create complex computer-graphic models of sculpture. In this park, a visitor can interact with a computer, so that the computer is able to generate a more aesthetically pleasing design for that visitor. Evolutionary algorithms were the base of this project and the park was a successor of other previous researches. When a visitor enters the park, eight random sculptures are depicted. The visitor can see them and interact with the park's scenario, to view each sculpture closely. If the visitor decides to click on one sculpture, it gets removed from the population. When there is only two or three left, the park learns the visitor's preference and another evolutionary cycle is repeated. The visitor is advised to leave only one or none. This process can be repeated as many times as the visitor wants. In this park, sculptures are graph structures and their nodes (13 in total) contain genetic information.

DOMAINS

To approach our challenges, we first need to get acquainted with the domains we are working with: the musical and visual domains. Therefore, we need to study their characteristics, so that we are able to work with them.

Musical Domain

Music is a composition of notes which produce sounds - vocal or instrumental - and silence. Depending on how these notes are combined, different emotional expressions may be associated with the resulting composition. Music pieces are labelled based on their properties such as rhythm, melody and harmony.

Starting with the basic element of music, we have **notes**. One sound defines a note, containing a certain duration and pitch. The duration defines how long we will listen to that sound and the pitch indicates how low or high the sound is. Pitch is a subjective evaluation of the listener (how high or low the listener perceives the note). Although it is not the same, we associate it with the note's frequency, which is measured scientifically (how high or low the note is in Hertz). We may perceive the same pitch differently, hence why sometimes we still can listen to higher-pitched notes while others can not.

Scales can be considered the artist's palette since they arrange notes in a certain manner helping musicians to understand some relationships and compose music. For the goal of this work, we decided to use the Circle of Fifths. The Circle of Fifths is a wheel that portraits the relationship of the twelve pitch classes of the Chromatic Scale. The Chromatic Scale uses twelve different pitches with an interval of a semitone between them (C, C♯, D, D♯, E, F, F♯, G, G♯, A, A♯, B).

In order to achieve some musical results, musicians sometimes play more than one note at a time. There are common combinations that are called **chords**. Chords can have three notes (named triads) or more that may guide the musician, allowing him to achieve an emotional or censorial goal [4].

Rhythm is the "time" element. When we analyze the rhythm we have three factors: duration (how long a sound or lack of it lasts), *tempo* (pace of the music piece), and meter (organization of beats into measures).

Melody is defined as the linear/horizontal presentation of pitch (highness or lowness of a musical note). Melodies can have several scales (families of pitches) like major or minor.

Harmony is the verticalization of pitch. It is the combination of pitch and chords. The chords are organized into sentence-like patterns and we call these chord progressions.

The **harmonic series** can be seen as a series of different frequencies (also called tones) of one original one, called the fundamental [4]. The fundamental frequency can be divided or multiplied by a number, producing other frequencies - higher or lower.

Dynamics translate the loudness or quietness of music. The transition between loudness to quietness is called *diminuendo* or *decrescendo*, and between quietness to loudness is called *crescendo*. *Accent* is the "punch" to a certain note to emphasize it.

Besides these, there are also other features, such as tone colour, texture, form, and style, but, for our work, we are going to focus on the ones mentioned above.

Visual Domain

The art of sculpture consists of creating a 3D object, usually carved or built with a certain material. Initially, these sculptured objects started as realistic natural human body forms and now sculpture evolved to more abstract models, which can be purchased and used as a house or garden adornment.

For our work, we need to have a **computational representation** for sculpture. The four most common representations are:

- **Polygonal meshes:** Polygonal meshes have as a polygon as basic unit, which can be defined as a face. The face contains 3D vertices that are connected from edges and, consequently, build a certain shape [6].
- **Surface-based Models:** A Surface-based Model is a model defined mathematically through its surfaces [9], linked with edges [7].
- **Volumetric Representation:** Contrary to surface-based models, a volumetric representation allows to go inside the surface of the object and is displayed according to a voxel grid which measures pixels volumetrically [13]. This representation allows to define and describe the material content of the 3D object.
- **Point Cloud:** A point cloud is a collection of 3D points that, together, produce "a digital 3D representation of a physical object or space" [3]. Besides having the usual x, y and z coordinates, the point can also include RGB (Figure 1) colour data and intensity value.

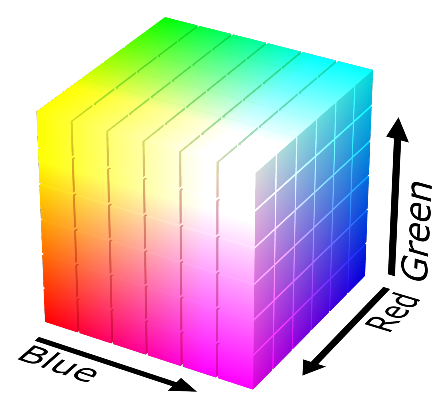

Figure 1. RGB colour model.

For our work, we decided to use a polygonal mesh representation due to its simplicity and flexibility. Moreover, in computer science, a 3D object is defined by three files: an object file, a material file and a texture file.

Focusing on the 3D world, **shapes** can be defined as an object having three dimensions: length, width, and height. The shape of a sculpture is essentially made with three elements: faces, edges and vertices [2], which relate to each other. According to [1], "most complex 3D models start as a simple

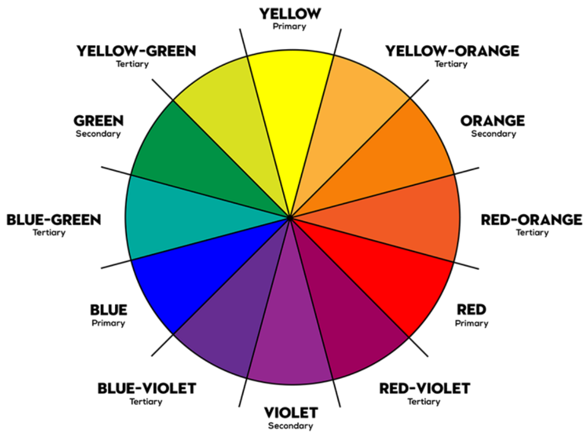

Figure 2. Colour Wheel.

geometric shape, like a cube, sphere, or cylinder", which are called object primitives. These primitives are later shaped according to the artist's final resulting aim. Our system uses primitives as well to create the final sculptures.

Texture can be defined as " a tactile quality of an object's surface" and can "evoke feelings of pleasure, discomfort, or familiarity" [8]. Artists tend to use a specific texture that seems fit to their artistic aim. Texture can simply be a colour or a material and, in a 3D object, there can exist more than one texture. Moreover, texture is real or implied [8].

Colour is an important element in all visual art forms. We convey a certain colour (or lack of it) when we want to express an emotion or a feel. The colour wheel is the main diagram used to include all colours in a spectrum. Isaac Newton created the first colour wheel (Figure 2) comprising primary, secondary and tertiary colours:

- **Primary colours:** Red, blue and yellow.
- **Secondary colours:** Violet, orange and green.
- **Tertiary colours:** Red-violet, blue-violet, blue-green, yellow-green, yellow-orange, and red-orange.

To represent colour in computer science we have colour models, such as **RGB (Red, Green, Blue)**, **HSV (Hue, Saturation, Value)**, or **HSB (Hue, Saturation, Brightness)**. For our work, we are going to use the RGB model.

APPROACH

Since the main goal of this work is to create a system capable of generating sculpture using music as inspiration, we need a bridge between music and sculpture.

Using some extracted music features (described in Section), we made associations that helped to convert the music features into sculpture characteristics. We used music dynamics, tempo, chroma, melody, harmonic series, and instruments. To obtain all these, we analyzed the same MP3 and MIDI files from several music pieces.

Before continuing, we want to clarify that the following associations are a result of our personal perception of these two art forms, meaning that other perspectives and, consequently, other analogies could be as valid as ours.

Starting by extracting the music dynamics, we decided to interpret it as the sculpture scale. When we listened to a quiet music piece, we saw a small sculpture, while when listening to something louder, we could envision a bigger sculpture.

Regarding the music's *tempo*, we perceived it as a texture indicator. When listening to a piece of music with a higher number of beats per minute, we associated it with something rougher. This means that, on the other side, something calmer, with fewer beats per minute, would be translated into something more smooth.

To add colour to the texture, we thought of using the music's chroma[1]. We created a default association between the Circle of Fifths and a Colour Wheel that chroma uses to obtain the three most predominant colours. This association links each element of the Circle of Fifths with a colour from the Colour Wheel. Since this colour analogy is especially subjective, we made a questionnaire that allows the user to choose to make his/her own or use the default analogy (ours).

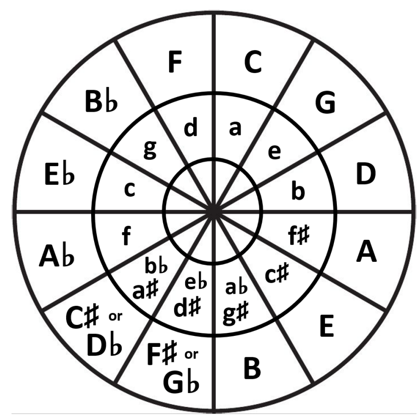

Figure 3. Circle of Fifths.

The melody and harmonic series were also used in the texture as a complement of chroma. Melody is the element that we most remember from music. This way, we decided to use it as the saturation complement of the colour. While higher frequencies translate into more vibrant colours, lower frequencies give colours a more faded and greyer tone. Besides complementing chroma, harmonic series can also help with perceiving melody. Therefore, we associated them with the colour brightness. Like harmony, harmonic series can impact the final emotional feel of the music. Similarly, as a minor harmony can produce feelings of sadness, nostalgia or melancholy, a minor harmonic progression can have the same result.

[1]Chroma captures the twelve profile pitches, which contain harmonic, melodic, timbre and instrumentation music characteristics.

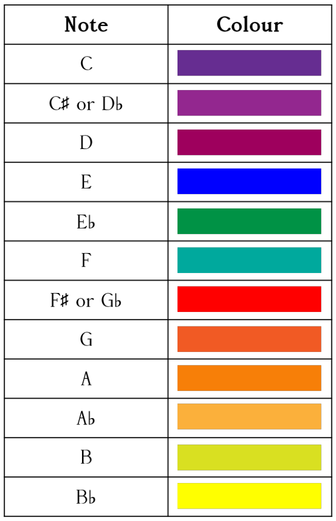

Figure 4. Default association between notes and colours.

On the other side, a major harmony leads to feelings of happiness, excitement and high energy, which can also be a product of a major harmonic progression. Having this in mind, we followed an association between harmonic series and brightness, which states that major harmonic progressions are associated with more brightness, and minor harmonic progressions with less brightness.

Using the extracted instruments, we obtained genre through a genre identifier that we created. This way, we could use the instruments' material along with others, such as acrylic, aluminium, clay, fabric, limestone, paper, and shell. Table 1 contains the associations between genre and the available materials.

Genre	**Material(s)**
Alternative	All materials
Blues	Brass, Acrylic, Bronze, Gold, Aluminum
Classical	Wood, Stone, Metal, Steel, Marble, Limestone
Country	Wood, Clay
Disco	Plastic, Bronze, Brass, Glass
Electronic	Metal, Acrylic, Plastic, Trash
Hip-hop	Copper, Brass, Paper, Bronze
Jazz	Gold, Brass, Steel, Velvet
Pop	Acrylic, Steel, Rubber, Fabric
Punk	Metal, Aluminum, Basalt, Trash
Rap	Copper, Brass, Bronze, Paper
Rock	Metal, Iron, Paper

Table 1. Association between genres and materials.

IMPLEMENTATION

For our system, we divided each challenge into a module:

- Have a module to extract and analyze music - **Music Analysis Module**.
- Have a module to translate the information of the music piece into sculpture features - **Cross-Domain Module**.
- Have a module to generate sculptures - **Sculpture Generation Module**.
- Have a module to improve and/or diversify the sculpture - **Genetic Algorithm Module**.

In the first module - Music Analysis Module - we analyze the input of both MP3 and MIDI files. It extracts each music's features and saves the information for the next module to use. We got features, such as dynamics, *tempo*, chroma, melody, harmonic series, and instruments (which were later used to obtain the music's possible genres). Using the knowledge of colleagues and people that had music theory knowledge, we created a dictionary for genres, each one containing instruments that usually are played in that genre.

Another feature we decided to extract was the music's sections (intro, verse, chorus and bridge). We performed this extraction with the music dynamics, previously obtained. We created a k-means cluster that classifies each section according to the dynamics. Using the values of the dynamics from every 22 seconds, we analyzed the values, according to a threshold of 0.03. Therefore, we obtained different groups. Later, we built a cosine similarity matrix that measures how similar these groups are, so that we could classify them into sections.

In the Cross-Domain Module, we followed the associations explained in the Approach section, where we link the musical and visual domains. We conceived several associations between intervals, so that, later, we could apply the first derivative, standard deviation, and obtain a continuous function from the scales that we are going to mention now.

With dynamics, we got the sculpture's scale. Values of dynamics between 0 and 0.16 (very quiet dynamics) represent a small sculpture, having 1 to 2 meters high.

For the sculpture's texture feel, we first defined how soft or rough the texture should be. Then, we determined that between 0 and 32 BPMs the texture feel should be very soft. Between 32.1 and 64 BPMs it is soft, between 64.1 and 120 BPMs it has a medium feel (neither soft nor rough), between 120.1 and 256 BPMs the texture is rough and above 256 BPMs we have a very rough texture feel. Having this previous translation, we used different parameters to obtain this texture feel: the noise scale, noise texture, a smooth factor, a smooth repeat value, and a subtraction factor. The noise scale is used to scale the size of the noise. The bigger the scale is, the smoother the texture will be. The smaller, the roughest it is since smaller sizes produce a grain effect. The noise texture is a value that determines whether a noise pattern is rough or smooth. With sharper peaks, obtained with smaller values, we created a rougher texture, while with rounder peaks we created a smoother texture (having bigger values).

The smooth factor is a value of a smooth filter we opted to apply, so that sculptures could have a more realistic look. This filter smooths the sculpture through its angles and adjacent faces. The bigger the smooth factor value, the smoother the sculpture will be. Besides the smooth factor, we also have the smooth repeat, which is the number of times the filter repeats.

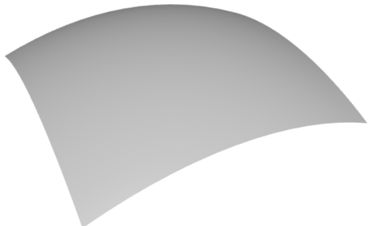

Figure 5. Round surface example.

Figure 6. Peak surface example.

Finally, we created a parameter called the subtraction factor. This factor determines the difference of movement between a point and its adjacent ones, i.e., how much the adjacent points move when one does.

Regarding melody, we first translated the values into a saturation level: very low, low, normal, high or very high. After doing so, we defined that when having a very low level of saturation, we need values next ranging from 0 to 0.8. Low means that we need values between 0.81 and 1.1. Above 1.11 and below 1.41 would result in the original saturation colours. The high level stays inside the interval of $[1.41, 1.7]$, and above 1.71 translates the very high saturation colours.

For the harmonic series, we followed the same line of thought where, firstly, we conceived levels of brightness: very low, low, normal, high or very high. Then, these levels were translated into intervals for brightness values. While experimenting with values, we decided to go with a scale ranging from 0 to 1.25. Values next to 0 or 0 mean that there is no brightness, i.e., we will obtain a full black result, values from 0.5 to 0.75 result in normal levels of brightness (not too dark and not too bright), and values above 1 mean that the results will be very bright, obtaining an almost total or total white result.

Moreover, with the first derivative results, we improved the values, where we added or subtracted a factor, according to how much ascending or descending periods existed. If more than half of the periods were ascending, the coordinates values are updated according to Equation 1, but if the majority of periods were descending (more than 50%), the coordinates are updated according to Equation 2. The variables pos and neg represent arrays of the positive and negative values of the ascending and descending periods, respectively.

$$newValue = oldValue + \frac{\sum pos}{length(pos)} \tag{1}$$

$$newValue = oldValue - \frac{\sum neg}{length(neg)} \tag{2}$$

The Sculpture Generation Module is responsible for generating the sculpture with all features previously saved. We created a script containing the commands to generate a sculpture from the initial primitive. We opted to have several available primitives: cylinder, sphere, cone, gem, torus, supertoroid, and torus knot. From these, the system chooses one according to what it finds best for the final result. We noticed that, with the point modulation, some textures will naturally be smoother, while others would be rougher due to the position of the initial points.

The sculpture is later scaled, according to the scale results. Then, with the texture information, the mesh points are transformed and moved. With the first module, we extracted music sections using music structure classification. This way, assigned several points to each section, which will be modelled according to that section. Each section has a percentage of points. Initially, we calculate the percentage for the number of sections the music has. If the music has an intro, verse, chorus, and bridge, it means each section will contain 25% of points. Then, these percentages are refined according to repetitiveness. While verses and chorus are often repeated more than once, intros and bridges have usually only one appearance. Therefore, we lowered the percentage for intros and bridges by half of the initial amount and increased the verse and chorus percentage with the taken amount. If the music has both verse and chorus, the percentage amount we took from the intro and/or bridge is divided by 2.

When creating a smooth surface, it is necessary to move one point and the adjacent ones. This allows a smooth transition between points. However, if we want to achieve the opposite effect - a rough surface - we can just pull one point, which will create a peak. To get a more round and smooth effect, we module the points with a value ranging from 0 to 1. This value is obtained according to a function created from a finite interval. Moreover, to create this effect, the adjacent points need to move too. The eight adjacent points (four behind and the front four) move with different values according to the subtraction value we created. This factor subtracts a value for the movement of the adjacent points. The first two adjacent points are subtracted once, the next two are subtracted twice, and so forth. For a medium texture, only two adjacent points move along with the original and, for a rough texture, only the original moves, which creates the peak effect. Additionally, if the texture is supposed to be smooth or very smooth, it is smoothed out with a filter, containing the values of repetition and factor (mentioned in the Cross-Domain Module).

For the texture application, a material file is conceived, containing characteristics such as: the material name, shininess exponent (Ns), ambient (Ka), diffuse (Kd), specular (Ks), emission (Ke), optical density (Ni), dissolve (d), and illumination values (illum). For the majority of these parameters, we used the recommended values since our main goal was to achieve the material feel with the texture file. We decided to use music features to obtain the final value to obtain the final value of shininess exponent. With the music's genre, we created a subjective scale from 0 to 10 of how shiny each genre seems like. For punk, the shiny value was zero, we gave to rap and hip-hop one, country is a two, rock is a three, alternative is four, pop is a five, classical is a six, jazz is a seven, blues is an eight, electronic is a nine, and finally disco got a ten. To improve this, we added the harmonic series to refine and obtain the final output. By using a histogram, we were able to identify the distribution of frequencies, i.e., where the frequency values were located most and least of the time in the scale. With this information, we perceived if the material should be more or less shiny. If the frequencies were present in high frequency intervals, the material is shinier or, on the other hand, if the majority of the frequencies is in lower intervals, the material is less shiny.

The application of colour was conceived through percentages that show how dominant each colour is. We opted to use the first more dominant colour in 20% of the material. The second more dominant colour is shown in 10% of the material and the third one appears in 5%. With the material's image size, we analyzed how many pixels each percentage would be. We created two versions of colours, trying to replicate colour "splashes". The first version has three different big splashes, each one with a different colour from the three most predominant ones. The second version is composed with smaller splashes, that can have different sizes. Both versions' splashes were painted over the image, using its size and pixels.

Our second generated sculpture, the altered version, was made using the raw's information, but the mesh composition is slightly different. Instead of using the classified sections (verse, chorus, etc), we are using the dynamic groups that were not classified. This means that, by evaluating the dynamics, we created parts when a 0.2 discrepancy was found between different intervals. Each part is later associated with a letter - A, B, C, or D. We named parts with as A (0 to 0.2), parts that contain dynamics above 0.2 and equal to or not bigger than 0.4, are B, C contains the interval 0.41 to 0.6, and parts above 0.6 are D. This way, the system creates sculpture sections with different sizes.

The last module - Genetic Algorithm Module - is based on natural selection, which create new elements for a certain population, in order to reproduce evolution as we know it in real life. In our work, this translates as evolving different sculptures according to some parameters. Our algorithm runs in 300 iterations through its several stages:

- **Initialization:** We created a population randomly, containing in total 30 individuals. Each individual of this population is a sculpture, having features (or genes) of scale, texture and colour. As for the number of individuals, we decided to go with 30, because according to [5] the best population sizes range between 20-30.

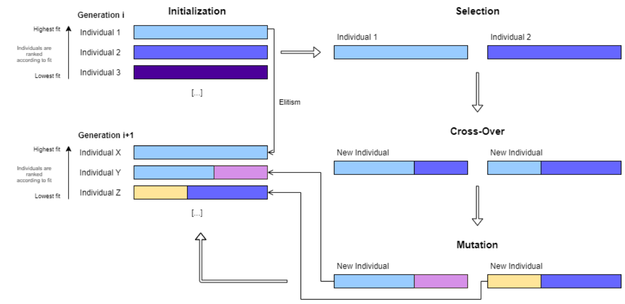

Figure 7. Genetic algorithm summary.

- **Selection:** All individuals start with the same fitness: 0. Then, according to a fitness function, the ones that show better features will be maintained and reproduced, while the ones that show poorer characteristics will die along the way, since they do not adapt to the current conditions.
- **Crossover:** We have a 90% of probability for crossover occurrence on features, such as scale dimensions, texture feel, and colour combination.
- **Mutation:** We opted for 1% of mutation probability on all possible mutations, since a lower value would result in almost no cases of mutation. These can occur in features, such as scale dimensions, texture feel, and colour combination.
- **Fitness Function:** Based on a sum. Good candidates will have positive points added, while bad ones they will have more negative points added. The bigger the score, the better the individual is. The scale and texture can add from $+8$ to -4, depending on the candidate's features. Colour can add from $+10$ to -10, depending on the candidate's colour combination.

The system's full architecture is depicted in Figure 8.

RESULTS & DISCUSSION

To evaluate our system, we decided to create three surveys that are divided according to an Era/Period: Contemporary, Modernism, and Classical. Each contained music pieces from that Era/Period, where the participants were asked several questions about music and sculpture artefacts.

In total, there were 113 participants. The big majority was between $18 - 29$ years old in all surveys. Regarding the attendance of expositions before and after the COVID-19 pandemic, the results show that, after the pandemic, people tend to go less to exhibitions and expositions. When asked about any educational background, the majority of our participants stated that had no education in sculpture.

From the obtained results, the overall preference was close in all three versions - Raw, Altered and Genetic (36,80%, 31,40%, 31,80%, respectively). This means we cannot state that there is preferred version, due to no major % difference.

When analyzing the best and worst case scenarios, we see that the results were, overall, positive. We gave the participants a Likert scale to rate the sculptures, where values range

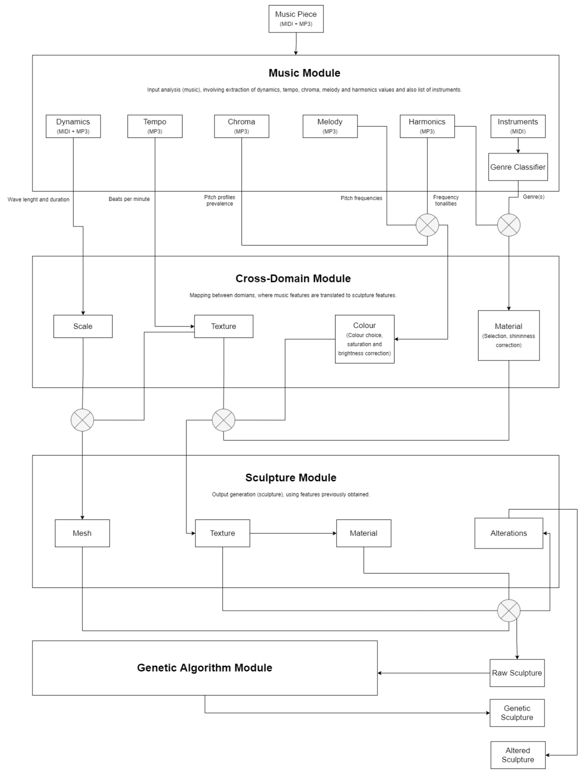

Figure 8. System's architecture.

from 1 (Do not like it) to 5 (Really like it). In the best case, when asked the participants if they liked the sculptures, the results of mean, median, and mode were all above 3 in all music pieces of the three surveys, meaning that the participants liked or really liked the sculptures. Looking at the worst results, we concluded that the majority of the participants neither like the sculptures nor dislike them, because the mean values slightly surpass 3, meaning that the participants, overall, were indifferent to the sculptures.

Regarding the relationship of music-sculpture, when asked "How do you think the previous sculptures are related to the music?", the results were positive, since the mean is above 3 in most cases. The scale goes from 1 to 5, where 1 means that the participant does not think there is a relationship at all, and 5 that the participant sees that the music and sculptures are totally related. This means that the overall feel corresponds to a sensation of relationship between the artefacts.

When analyzing the adjectives given to the music pieces and the sculptures obtained from them, we saw that the participants gave almost the same adjectives. Most of the participants gave the sculptures 4 out of the 5 adjectives that were present in the music. With the values of the common adjectives, we calculated that the total mean is 3.71, the median and mode are 4. This means that, overall, the participants gave 4 out of 5 adjectives, having a comparison success rate of 80%. However, when asked how the participants think the artefacts were related, the answers were more diverse, and the majority were not sure if one existed.

The last question of the survey asked the participants how much they agreed with the sentence "All the sculptures were created having the previous music as inspiration". The results show that for each survey the participants agree with the sentence since statistics point values above 3 in all means and mode. This means the participants agree with this sentence.

Lastly, we believe that our system achieved the results for the challenges we aimed to surpass, using a subjective association between two different domains.

CONCLUSIONS & FUTURE WORK

It is difficult to understand how creativity works. Even though we lay out some things as creative, the definition of creativity and how it occurs still needs to be deepened. Creativity is intrinsic and subjective. This is why it is hard to understand and define.

Like creativity, inspiration is also intrinsic and working with it scientifically is not the easiest task, since it involves personal experiences and personal emotions, which are also subjective.

With this work, we investigated the possibility of having a system capable of creating sculpture, using music as inspiration. This yields the field of CC, which intersects computer science with art, by having systems that work with AI techniques to develop artefacts usually considered as art.

Our work is based on the field of Computational Creativity and the belief that it is possible to have a system capable of generating new and valuable sculptures. Therefore, this was our main hypothesis.

Having this in mind, we, firstly, studied our problem, presenting in this work some previous related studies and systems done in this field, to help us understand the problem in our hands and possible solutions in depth. Then, we developed a system using music as a source of inspiration, that creates sculpture.

To do so, we linked the music and visual domains through an analogy, made from our perception of these art forms. By using this analogy, music information is translated to sculpture information, which is later used to create the final output. The final output consists in three different versions of sculpture: a Raw, an Altered and a Genetic sculptures. All three versions are sculptures created from our conceived analogy between the musical and visual domains.

Besides AI techniques, we combined math and a lot of experimentation. This work is the result of experimenting and going back and forward to achieve what seemed the best visual artefacts, possible in all three versions.

To evaluate the system, we created three online surveys which contain two different music eras: Contemporary and Classical. Inside the Classical Era, we divided the Modernism period and the Classic, and Romantic periods. This way, we are able to study the "Post-Great War" years separately from the more Classical times. Each survey has its own era/period, and, in each, we ask the participants to listen to two different music pieces from that era/period. You can see all the music pieces used and respective sculpture here: `https://frommusictoasculpture.wordpress.com/`. Later, the participants were asked to evaluate each music, describe it, and do the same for each resulting sculpture. With this information, we were able to analyze the results and conclude if the participants liked the artefacts, how they describe them, and if they find or do not find a relationship between them.

Considering the survey results, we can state that we followed a positive approach, where overall the participants liked all three versions. Moreover, they found a relationship between the music and respective sculpture, which is supported by the common adjectives between artefacts and directly asking it.

Our system has some limitations that could be improved or refined. The music structure classifier needs to be studied in a deeper level to be improved. Our classifier works with dynamic intervals, every 22 seconds. Two interval neighbours are analyzed to understand how similar they are, using a threshold. If they are not considered to be similar, they are labelled as different groups, which will then be classified into intro, verse, chorus or bridge. This last classification into a part is made from colleagues' perception of music that studied it, meaning it could be improved and done some different experimentation, to see how the results would be.

Another problem we faced was the genre classification. We concluded that the simplest approach was to use a dictionary, containing the most common instruments as a base for it. Even though this worked positively for our dataset, this is not a favourable approach for the problem itself.

The last limitation is the wrapping paper for the sculptures. Using a texture, we applied UV mapping to convert 3D coordinates into 2D coordinates. We used a dynamic UV mapping that creates a dynamic map, considering the manipulated mesh, to achieve the best result possible. However, UV mapping is a manual technique that 3D artists take a lot of time to do, meaning that, when performed by a computer, the task is harder. Consequently, UV mapping should be studied in more depth to seek a possible better approach when performing it.

To conclude this work, we believe in this field and its possibilities. Moreover, since this field works with creativity, it might boost other fields with art forms, and surround them with new and valuable ideas.

We believe we contributed to the CC field, and that our system provides a positive progress in it, as it regards some kind of creative behaviour in the visual domain, more specifically, inside the fields of computer science and sculpture.

ACKNOWLEDGMENTS

I would like to thank my supervisors H. Sofia Pinto and Nuno Correia for believing in me from the beginning and supporting my work. Secondly, I would like to thank my parents, Helena and João, which always supported me thick and thin at every level of my life. To all my friends and boyfriend, which have listened to me every day. Thank you. A special thanks to my very good friend, Leandro, which has been here since day one, listening to me, supporting me, and believing in me. I want to thank my grandparents, as well, for the comforting words I needed when I was tired or I thought I could not do it. Last but not least, I would like to thank professor Rui Jorge for being always available to help me. His insight was crucial for the conception of this work. To each and every one of you who has always believed in me – Thank you.

REFERENCES

1. 3d model components—vertices, edges, polygons more. `https://www.lifewire.com/3d-model-components-1952`. Accessed: 12/08/2021.
2. 3d shapes. `https://www.twinkl.pt/teaching-wiki/3d-shapes`. Accessed: 11/08/2021.
3. Everything you need to know about point clouds. `https://transport.itu.edu.tr/docs/librariesprovider99/dersnotlari/dersnotlarimak537e/notlar/7-wireframe-and-surface-modeling.pdf?sfvrsn=4`. Accessed: 08/08/2021.
4. Harmony explained: Progress towards a scientific theory of music. `https://arxiv.org/html/1202.4212`. Accessed: 04/08/2021.
5. Introduction to genetic algorithms: Recommendations. `https://www.obitko.com/tutorials/genetic-algorithms/recommendations.php`. Accessed: 31/08/2021.
6. Introduction to polygon meshess. `https://www.scratchapixel.com/lessons/3d-basic-rendering/introduction-polygon-mesh`. Accessed: 07/08/2021.
7. Surface modeling. `https://transport.itu.edu.tr/docs/librariesprovider99/dersnotlari/dersnotlarires112e/not/cadd-14.pdf?sfvrsn=4`. Accessed: 07/08/2021.
8. What is texture in art? `https://www.thoughtco.com/definition-of-texture-in-art-182468`. Accessed: 10/08/2021.
9. Wireframe and surface modeling. `https://transport.itu.edu.tr/docs/librariesprovider99/dersnotlari/dersnotlarimak537e/notlar/7-wireframe-and-surface-modeling.pdf?sfvrsn=4`. Accessed: 07/08/2021.
10. Aleixo, L., and Pinto, H. S. From music to image - a computational creativity approach. Master's thesis, Instituto Superior Técnico, Lisbon, Portugal, November 2020.
11. Boden, M. A. Creativity in a nutshell. *Think 5*, 15 (2007), 83–96.
12. Boden, M. A. Computer models of creativity. *AI Magazine 30*, 3 (2009), 23–23.
13. Braga, F., and Pinto, H. S. Sculpture inspired musical composition. Master's thesis, Instituto Superior Técnico, Lisbon, Portugal, November 2019.

14. Horn, B., Smith, G., Masri, R., and Stone, J. Visual information vases: Towards a framework for transmedia creative inspiration. In *Proceedings of the Sixth International Conference on Computational Creativity.* (2015), 182–188.

15. Newell, A., Shaw, J. C., and Simon, H. A. The processes of creative thinking. In *Contemporary Approaches to Creative Thinking, 1958, University of Colorado, CO, US; This paper was presented at the aforementioned symposium.*, Atherton Press. (1962).

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

17. Rowland, D., and Biocca, F. Evolutionary cooperative design methodology: The genetic sculpture park. *Leonardo 35*, 2 (2002), 193–196.

18. Sawyer, R. K. *Explaining Creativity: The Science of Human Innovation.* Oxford University Press., 2011.

19. Teixeira, J., and Pinto, H. S. Cross-domain analogy: From image to music. In *Proceedings of the Fifth International Workshop on Musical Metacreation* (2017).

20. Thrash, T. M., and Elliot, A. J. Inspiration as a psychological construct. *Journal of Personality and Social Psychology. 84*, 4 (2003), 871.