MEASURING THE IMPACT OF SUBJECTIVE AND TRANSIENT AESTHETICS IN THE GENERATION AND APPRECIATION OF 3D VIRTUAL ARTWORK

Edward William Easton

Doctor of Philosophy

ASTON UNIVERSITY

September 2023

Abstract

Measuring the impact of subjective and transient aesthetics in the generation and appreciation of 3D virtual artwork
Edward William Easton
Doctor of Philosophy
2023

The topic of aesthetics is a vast puzzle with many contributing parts. It often seems incomprehensible; however, through decomposition into smaller problems, a solution may be within reach. It impacts almost every aspect of life, including romantic and platonic relationships, architecture, decoration, writing, theatre, mathematical proofs, algorithms, product design and artwork.

This thesis aims to partially deconstruct the individual aesthetic judgement of sculptures into its component parts through the steps of understanding these parts at both individual and collective levels, introducing formalised measures which calculate how well a specific sculpture exhibits a particular aesthetic feature and then using these measures to provide a novel, extendable and flexible model of individual aesthetic judgement.

This project represents a step forward in the computational research of aesthetics and aesthetic judgement through original Virtual Reality Environments and an art generation process capable of creating interesting 2D and 3D artwork. Experiments reveal some key aspects that contribute to human aesthetic judgement, both positively and negatively. The experiments also resulted in clear rules to follow for the generation of sculptures according to personal preference.

New aspects are formalised, calculating how connected, calm, or friendly a sculpture appears; all three are found to be closely aligned with the human understanding of the term. Finally, a theoretical model of individual aesthetic preference is presented. The model is based on three categories of aesthetic aspects: constant, subjective, and transient; these allow the generation of sculptures which are unique to an individual. Importantly, the models can be applied to replicating a specified style of artwork, meaning that new items can be generated that exhibit close relationships to the original style whilst still being novel. Constituting a multi-pronged approach, the contributions presented in the thesis move us closer to decoding the puzzle of aesthetic appreciation.

Keywords:
Aesthetic judgement, Aesthetic modelling, Virtual Reality, 3D Art Generation, Evolutionary Art, Gamification, Computational creativity

To everyone I've met

You are all a part of me

And now a part of this thesis

This is for you

Personal Acknowledgements

The main thing is to thank Dr Ulysses Bernardet and Prof. Anikó Ekárt, without you both I wouldn't be here, a lot has changed over the 6 years we have known each other, but you both have consistently pushed me forward and this work is really the triumph of us all. I have never been good at remembering names, and I will save the reader from me trying to describe how everyone who has contributed looks, (ironic, considering the content of the thesis). However, I can assure you that if you have had any interaction with me, from a brief chat, giving anonymous feedback on submitted papers to spending time helping me overcome a problem, you have all had a major impact and I will be perpetually grateful.

List of Publications

Publications Arising from this Thesis:

Easton et al. (2021) E. Easton, A. Ekárt, and U. Bernardet. Axial Generation: A Concretism-Inspired Method for Synthesizing Highly Varied Artworks. In *10th International Conference on Artificial Intelligence in Music, Sound, Art and Design, EvoMUSART 2021 held as Part of EvoStar 2021*, pages 115–130. Springer, 2021.

Easton et al. (2022) E. Easton, A. Ekárt, and U. Bernardet. Axial Generation: Mixing Colour and Shapes to Automatically Form Diverse Digital Sculptures. *SN Computer Science*, 3(6):505, 2022.

Easton et al. (2023a) E. Easton, U. Bernardet, and A. Ekárt. Is Beauty in the Age of the Beholder? In *Artificial Intelligence in Music, Sound, Art and Design: 12th International Conference, EvoMUSART 2023, Held as Part of EvoStar 2023, Brno, Czech Republic, April 12–14, 2023, Proceedings*, pages 84–99. Springer, 2023a.

Easton et al. (2023b) E. Easton, U. Bernardet, and A. Ekárt. Contributors to the aesthetic judgement of 3D virtual sculptures. In *2023 Third International Conference on Digital Creation in Arts, Media and Technology (ARTeFACTo)*. IEEE, 2023b.

Easton et al. (2024) E. Easton, U. Bernardet, and A. Ekárt. Modelling individual aesthetic preferences of 3D sculptures. In *Artificial Intelligence in Music, Sound, Art and Design: 13th International Conference, EvoMUSART 2024, Held as Part of EvoStar 2024, Aberystwyth, United Kingdom, April 3–5, 2024, Proceedings*. Springer, 2024.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

This thesis covers the expansive topics of aesthetics, looking to further scientific understanding of which aspects contribute to the aesthetic judgement of 3D virtual sculptures. This is achieved by determining how important these aspects are to the judgement process, how preferences of the aspects change between individuals, how the aspects contribute positively or negatively to an aesthetic judgement and finally, formalising new measures for use within an auto-generation context, ultimately aiming to expand the knowledge of these areas and allowing more personal artwork to be automatically generated.

Art is much more than lines on a canvas or chisel marks in stone, the starting point for the majority of artwork is a concept, something which builds the scene, that will later be translated onto the chosen medium. At its core, artwork is a form of storytelling, similar to academic writing, beginning by addressing the concept and setting the scene.

Figure 1.1: Going to work by L. S. Lowry. http://www.iwm.org.uk/collections/item/object/17026 [1], Public Domain, https://commons.wikimedia.org/w/index.php?curid=86205426

Modern-day life can be tedious, the endless repetition of day-to-day tasks is heavily linked to a rise in mental health issues WHO (2023). Through the lens of depression, the world can become very grey, the colour leaves along with the joy. Life becomes a waking LS Lowry painting (Figure 1.1), except the pollution-churning chimney stacks are replaced by pollution-churning cars, every person becomes a faceless, aimless, lifeless figure and even the ones who pass close by seem out of reach by an impossible distance. Once caged in that place, it requires courage to restore the colour to life, but something which can help along the way is finding something or someone that can contribute to the transition. The contradiction between the bleakness and joy of life routinely finds its way into artwork. The ironic depiction of life as told by Andy Warhol combines the two garish colour mixed with the drudgery of modern-day life (Figure 1.2). Mental health, creativity and art are closely linked, for the latter two are an almost essential part of modern life, helping to show why art and aesthetic appreciation has long been an important part of philosophy, psychology and computer science. It will continue to be a big part of these disciplines due to the ever-changing nature of technology, which forces new questions to be asked and new problems to be solved.

[Image removed from open access version of thesis; see link]

Figure 1.2: Campbell Soup by A. Warhol. Fair use, https://en.wikipedia.org/w/index.php?curid=4268566

Most humans are given the ability to judge artwork for its aesthetic appeal; theories surrounding this have been considered by philosophers such as Kant (Crawford, 2001) and Hume (Shelley, 2005). Like aesthetic appreciation, creativity is an inherent part of life, an underlying principle allowing us to extend existing rules (Boden et al., 2004). With art and aesthetic appreciation being so heavily human-centric, many things need to be considered when using a computer as the primary instrument creating it: can it be called art, who is the artist and does computer-generated art devalue art? These questions have been considered in many contexts, including within a computational space (Leymarie, 2021); however, to answer the first question, one needs to look to philosophers who have considered the question for millennia. Plato who required an ethical consideration for something to be considered art (Gaut, 2013); Sartre believed art needed an emphasis on imagination (Sartre, 2010). Irrespective of the opinion, creativity is the underlying principle which guides art and artists. Therefore, considering whether a computer can make art depends on *how* creative we believe a computer can be. A similar conundrum exists when considering who the artist of AI-generated art is: should the AI itself receive partial or even full credit? Unfortunately, there is no clear answer to this question, the credit for the machine can range from none to full, however, the vital point to note is that it is a human decision what their chosen level of credit is before the remainder can be apportioned out. The final question relating to the devaluing of art is a question which has been asked in many different contexts throughout history. Inventions like the printing press were deemed dangerous for similar reasons (Bell, 2010) and even modern-day technological advances have put this in the spotlight. The use of AI writing systems such as GPT-4[1] has proliferated, leading to significant increases in the amount of AI-generated fiction available online, due to this, AI-generated content is being banned at an unprecedented rate. This seemingly provides an answer to the question: the devaluation is happening and causing people to push back. However, this is a very glib interpretation of the issues AI poses, while mass production is not necessarily a goal when using AI, examples of singly generated images are still being banned from art competitions. This leads to the assumption that the devaluation is not necessarily related to the artwork but more to the experience it provides, which has long been a guiding principle of art. Taking humanity full circle back to the ideas purported by Plato where the "human-ness" of a piece of art is what is valued by other humans(Kieran, 2005), and so, for AI systems to be considered as creating art and being the artist, the system needs to consider and create art in a similar way to a human being, through partially formalising the aesthetic process, the area that this Thesis addresses.

Art is influential, it has long been documented (Daykin et al., 2008; Guetzkow, 2002; Pelowski et al., 2016) that living within an aesthetically pleasing environment can improve well-being, which is of paramount importance to everybody on the planet. Life allows many opportunities for aesthetically pleasing items to be viewed, most towns and cities have a

[1] https://openai.com/research/gpt-4

plethora of artwork available to view. However, 2D artwork (with some notable exceptions such as Street Art) resides primarily in museums. For various reasons, including lack of interest and even a lack of funds, these places are not accessible to everybody in equal measure. In contrast, free-to-access public artwork, often sculptures, are displayed in public places such as parks, several local examples of such sculptures are shown in Figure 1.3. Freestanding sculptures like these are often sought out and become part of regular routines, suggesting that sculptures can be a powerful type of artwork. The omnipresence of sculptures does not just exist for people who like them, it can also be seen in the media coverage, where the subject of sculptures in cities such as Cambridge and Bristol has come under severe scrutiny, (Farrer, 2020; Harris, 2022; Weale, 2021; Zayed, 2020). Whereas controversial 2D artwork does not attract the same level of scrutiny, partially due to sculptures being more commonly interacted with, and due to this, represent an important part of having artwork accessible by everybody.

However, this scrutiny indicates a flaw in providing public sculptures: due to the subjective nature of artwork, aesthetic judgement and art appreciation, it is impossible to create public artwork equally appreciated by everybody, for example, the sculpture shown in Figure 1.4, which depicts two women and a child with the title "A Real Birmingham Family", some commentators misinterpreted the sculpture to be supportive of the Gay rights movement, which led them to ask for the statue to be removed. These commentators were wrong and unaware of the artist's original intentions, which was that friends can become like family, and the statue depicted two friends helping each other. This shows that no matter the intention, all artwork is held to subjective standards, making the aim of creating personalised sculptures available to all an important endeavour. Many different types of sculpture exist, for example, Kinetic, Installation and Land Art; within this Thesis, the term is used to refer to Abstract Freestanding sculptures. This is one of the more accessible types of sculpture, often being found within public spaces, making it an ideal choice for the focus of study within this Thesis. Examples of Abstract Freestanding sculptures are shown in Figure 1.3.

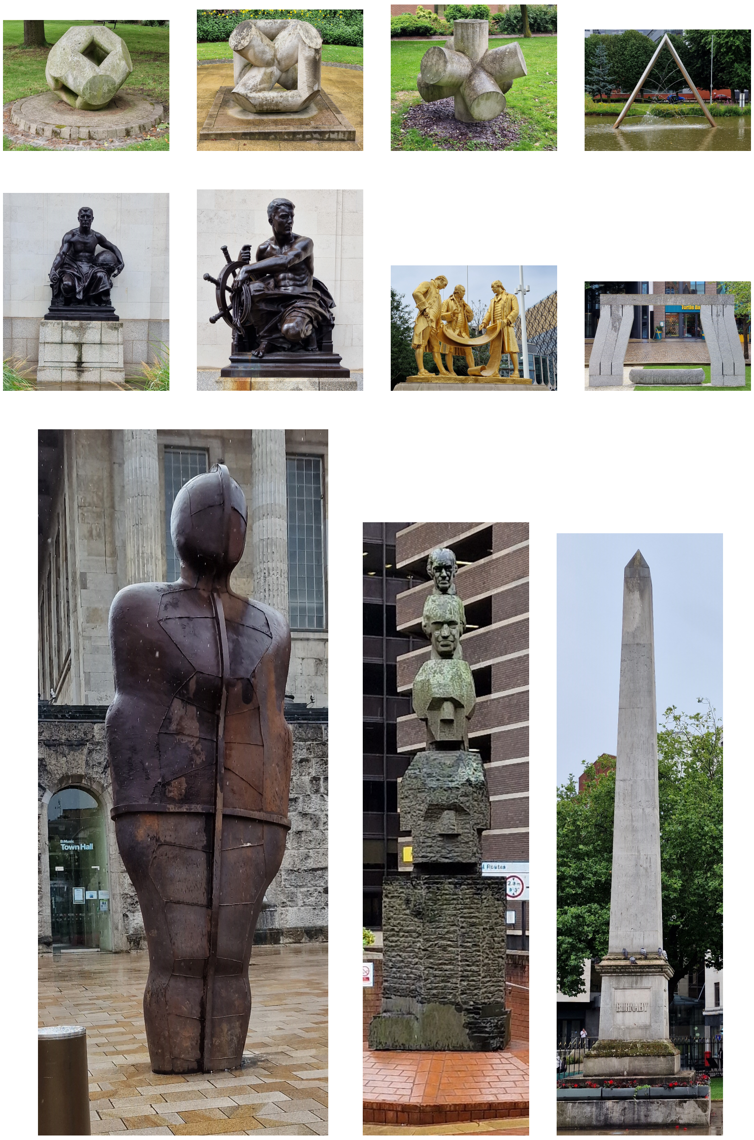


Figure 1.3: Sculptures found in and around Birmingham city centre. Taken by E. Easton

Figure 1.4: A Real Birmingham Family by Gillian Wearing. Taken by E. Easton

1.1 Aesthetic Appreciation and Computational Creativity

Whilst the terms are often synonymous, one important point to note is that aesthetic judgement is different from finding a piece of art beautiful. Many major artworks of the 20th Century have been explicitly created to be the opposite of beautiful, but the process of aesthetic evaluation still occurs, for example, Fountain by Duchamp (under the pseudonym R. Mutt), shown in Figure 1.5, was explicitly created to “test the openness of the society to artworks that did not conform to conventional aesthetic and moral standards without compromising the outcome or his relationships with board members, though at the expense of being able to avow that the work was his own.“(Tate, 2017). Considering the societal and contextual aspects is beyond the scope of this research, however, highlighting the difference between aesthetic judgement and beauty is important when discussing creating artwork that applies to everyone.

”Beauty is in the eye of the beholder”, this famous quote by Margaret Wolfe Hungerford (Hungerford, 1878) is often applied to discussing the beauty of individuals; however, it is also applicable to the judgement of some genres of artwork. The subjectivity of the platform is one of the few things considered an absolute truth about art, even as far back as Plato, who suggested that only highly educated individuals could truly understand art and, therefore, were the only ones who could create it. He held this view to such an extent that he believed *the Iliad* was not true art, an opinion not shared by others. More modern attempts to define the process consider a multitude of aspects internal and external to the artwork itself, such as the emotions invoked by the artwork in the viewer and the intended emotion conveyed by the artist. These numerous aspects interact with each other, suggesting that the appreciation of artwork can be considered a holistic process, where all aspects contribute in varying ratios. The potential that the appreciation of art considers a high number of largely unknown factors introduces three questions:

1. What are the contributors to the aesthetic appreciation of artwork?
2. How are different contributing aspects influenced by one another?
3. Can these aspects be effectively modelled to represent an individual's aesthetic preference and used to auto-generate artwork specific to them?

Figure 1.5: Fountain by R. Mutt. By Alfred Stieglitz - NPR arthistory.about.com, Public Domain, https://commons.wikimedia.org/w/index.php?curid=74693078

Aesthetic appreciation is not fully understood and is often considered to be a black box where making a judgement is easy, but describing the process and aspects which contribute to it is far more complex; this is mainly due to the extent that the judgement differs between individual people. Scholars and academics regularly disagree on which aspects contribute to an aesthetic judgement or struggle to find significant results when attempting to draw comparisons between people. These aspects have long been theorised about, however, limited concrete evidence has been found, introducing barriers when attempting to translate aesthetic judgement in a computational context, where the aspects are "understood" differently by machines and many aspects are also missed or implemented incorrectly due to their highly subjective nature. There are, however, some less accepted claims, often considered to be evolutionary traits, which are still part of our subconscious that contribute to how we formulate these judgements. One of the early philosophical writers on the nature of aesthetics and art was Plato (Janaway, 2013), who, throughout many writings, introduced the world to his view of aesthetic judgement. A large emphasis was placed on ethics and being wary of the potential damage various forms of art could cause. This ultimately led Plato to consider art only to be worthy if it was created by a philosopher, such as himself, regularly disparaging contemporary creatives under the guise that without a complete understanding and desire to search for truth, the art which is created should be discouraged, despite the merit of the actual piece in question. The danger, as Plato considered it, was that appealing to an undesirable part of the soul with drama and non-truthful imagery habituates someone to these notions and makes them more likely to act in this manner in future.

Whilst the formalising of artwork can be a means to an end, one of its primary purposes is to be used in other contexts, especially prolifically around auto-generating artwork. Recently, the concept of generating art has started encroaching into the general public's consciousness, thanks to systems like DALL-E (Radford et al., 2021), MidJourney and Stable Diffusion (Rombach et al., 2022). The potential power of these systems to generate realistic photographs from text prompts has brought about a wave of negative press about the impact of AI (Koidan, 2023; Plunkett, 2022; Taylor, 2022). Whilst these items have had enormous success, very quickly, the enthusiasm, along with the media coverage, has slowed down as people discovered their weaknesses. One glaring issue with systems like these is how they approach the subjectivity of artwork, these new approaches work from a much more generalised point, collecting huge amounts of data in order to train the system.

How the data is collected also becomes a problem, introducing a large bias into the results: BAME people are routinely excluded from simple prompts, and women can be portrayed in misogynistic ways. In addition to this, the systems understanding of abstract concepts, something very important within artwork, is minimal. All of these points are illustrated in Figure 1.6, showing images that were generated through the DALL-E system using the term "attractive". These images only showcase a narrow band of what humanity considers attractive, clearly indicating the bias introduced in these systems through the training data. It should be noted that the system can also generate more inclusive images, however, as these rely on the user inputting specific prompts, once again, the emphasis is back on the human to input the creative ideas and guide the system rather than the system generating art itself.

Figure 1.6: DALL-E 2 generated images for the prompt "Attractive". Generated by E. Easton

Due to this, these systems do not contribute to improving the fundamental understanding of aesthetic judgement. To further understand, individual aesthetic measures need to be identified, which look at specific aspects of a piece of art and provide a quantifiable value for this aspect. Identifying which aspects contribute to the aesthetic judgement of a piece of art will lead to comparison between artwork and a more explicit understanding of what contributes to artwork.

The state of these existing systems suggests that other approaches are required to auto-generate artwork that are more suitable to handle the subjective nature of artwork, and Evolutionary Computation is a technology capable of doing so. Generating art using Evolutionary Algorithms, more commonly known as Evo-Art, has been ongoing for decades, (Sims, 1991). Its significant benefit over other types of system is the generation can be determined at an individual level, however, introducing a model of aesthetic appreciation subjectivity requires thinking about how these aspects fit into aesthetic judgement.

1.2 Formalising and Modelling aesthetic judgement

Aesthetic judgement is a multi-faceted process with innumerable aspects contributing to the assessment (Hayn-Leichsenring and Chatterjee, 2019; Leder and Nadal, 2014; Pelowski et al., 2017; Ventura and Gates, 2018). Some aspects that contribute to aesthetic judgement are relatively well-known, such as the order and complexity (Birkhoff, 1933; den Heijer and Eiben, 2010b; Johnson et al., 2019). However, many other aspects are left in relative obscurity. Whilst identifying which aspects contribute to aesthetic judgement becomes easier with practice, a higher level of expertise in art can influence how items are judged (Leder et al., 2019; Monteiro et al., 2022). Experts can provide a wealth of information, however, without gaining further knowledge on which aspects contribute to aesthetic judgement, the aesthetic judgement process is challenging to model fully.

Whilst non-experts may find describing their judgement more difficult, aesthetic judgement still occurs, they will still like or dislike an object. Currently, experts are often involved to help model the aesthetic process, however, to truly understand the process, non-experts and everyone in between need to be included in the model, paving the way for computer systems which can generate artwork that appeals to a wider variety of people, expanding the reach, impact and potential uses these systems can have.

Formalising artwork is not a new endeavour, many examples of specific measures exist, such as the Global Contrast Factor (Matkovic et al., 2005), using Benford's Law (Acebo and Sbert, 2005) or measuring the symmetry (al Rifaie et al., 2017), however in several cases, these measures, by only considering a single aspect, do not cover the same aesthetic judgement, that we, as humans, perform. These quantitative aspects of aesthetic judgement form the basis of many attempts to model aesthetic preference, with a degree of success found with the approaches (Ekárt et al., 2012; Li et al., 2012), however, more is needed.

1.3 Aesthetic judgement

Many theories of aesthetic appreciation exist, in some, aesthetic aspects and their measures are considered constant values. However, in most theories, these values are often subjective and based on who is assessing the item (Korsmeyer, 2013). For example, when calculating a measure for the Global Contrast Factor (Matkovic et al., 2005), it was noted that during the small experiment run in order to determine the weights to apply to the contrast at various resolutions, there was a large number of contradictions between participants when they were asked to put a set of images in order from lowest to highest contrast. These contradictions were ignored in the final derived equation, however, the contradictions are a clear indication that aspects of assessing an image are very often subjective. The subjectivity has been investigated where aspects of aesthetic judgement have been split into private and shared taste have been determined from participants looking at abstract images and faces (Leder et al., 2016). However, other research (Leder et al., 2019; Monteiro et al., 2022), which has investigated which factors influence aesthetic judgement, has found a wide variety of factors seem to affect an individual's aesthetic preferences and judgement. This includes expertise with art (often determined by how often a person interacts with art), which does not easily seem to fit within the private and shared categories as whilst it is a subjective aspect of assessment, it will change with time. Interestingly, the change in level of expertise may also account for some of the effects of user fatigue, where initially the artwork created can seem exciting and attractive to the user; however, very quickly without the ability to introduce large levels of variety, the user becomes bored as they become more knowledgeable about the types of artwork they are looking at, indicating how influential novelty can be for aesthetic judgement.

Aspects that change over time should not be categorised as purely subjective aspects, as there is a high chance that even within a short time, an individual's preferences can change. Instead, they should be treated as volatile aspects of aesthetic judgement. This suggests a further category should be introduced over private and shared to appropriately handle these aspects, referred to here as Transient aesthetics. The private and shared categories will be referred to as subjective and constant, respectively, to align with the naming of the transient category. which better describes the aspects they contain, especially within a computational context. A full description of each category is shown below:

1. Constant Aesthetics - Universal aspects which apply to the majority of people regardless of characteristics such as age, culture or art experience.
2. Subjective Aesthetics - The aspects which will vary due to individual taste but are relatively constant within an individual
3. Transient Aesthetics - All aspects which vary due to an individual's trait, where the trait can be volatile and the judgement of these aspects is very likely to change, even at short notice.

One of the major benefits of splitting the aspects of aesthetic judgement into the three categories is that it allows different handling of each of the aspects within each category to help measure an individual's preferences. The subjective aspects will then depend on a small period of learning a specific individual's preferences, however, once they have been identified, they can be in place for the entirety of an experiment. Finally, the transient aspects are the most volatile and constant learning is required in order to properly represent an individual's preferences.

It also allows a distinction to be made when investigating multiple aspects of aesthetics when attempting to model an individual's preference, the clear categorisation allows a system to implement a clear set of rules for each category of aspects. For example, constant items can always be present, transient items can be updated regularly, and subjective items can sit in between.

While understanding the difference between constant and subjective aspects is intuitive and has been investigated previously (Leder et al., 2016), a problem exists with how to distinguish items between the subjective and transient categories. Transient aspects may present themselves as subjective, where a transient aspect may change too slowly to capture within the limited time frame of this PhD study, for example, several years. The exact categorisation of the aspects is out of the scope of this project, and it is expected that some aspects will be miscategorised. The purpose of this project is not to accurately identify which category an aspect may fall into but instead to use the categorisation process as a base to enable further elicitation of aspects which make up the aesthetic judgement of 3D sculptures and provide a formal definition for these aspects to allow their use to represent an individual's aesthetic preferences of 3D items. To achieve this, the following rules will be adhered to when determining the category for an aspect:

1. If the aspect displays a high level of commonality across multiple individuals regardless of age, gender and culture, the aspects will be placed into the constant category.

2. If the aspect does not display any commonality, or the commonality does not appear to transcend age, gender or culture, then this will place the item initially within the subjective category.

3. If a subjective aspect has the potential to change quickly, such as originality or where it is seen that ratings for similar sculptures throughout an experiment are changing, then this element will be considered transient.

An aspect may show properties of one or more categories, for example, it has been shown that symmetry is a prominent aspect of aesthetic preference across multiple different individuals, suggesting the constant category. However, the specific level can differ, for example, younger children can prefer higher levels of symmetry compared to older adults, thus suggesting the subjective category. These cases highlight the difference between considering the existence of a property and the specific value of that property, each of which should be considered separately.

As the project intends to provide a formal description and process for quantifying an aesthetic aspect within a sculpture, a few criteria have been identified that need to be met in order for a measure to be deemed successful. The first and most important will be that the measure accurately reflects human perception of the aspect. When trying to describe an individual's aesthetic preference and auto-generate artwork they may be interested in, this requires measures that can be used in conjunction with each other. For example, if an individual likes complex and symmetrical images, if the complexity measure is unable to rate highly symmetrical sculptures as complex, this would not be suitable for use for this individual. Therefore, the ability of the measures and investigated aspects within this project to work in conjunction with one another needs to be considered.

1.4 Research Questions

The extensive scope of the topic of aesthetics highlights that this area is impossible to research in its entirety. Due to this, this project will focus on a small, understudied part of the overall area, 3D aesthetics, broken down further with the overall aim of partially identifying and quantifying the aesthetics which contribute to an aesthetic judgement of a 3D sculpture by investigating four primary questions:

1. Is there a potential framework that can be developed to effectively study the aesthetic judgement of 3D sculptures at a collective and individual level?
2. Does a method exist that can create 3D sculptures which are visually varied and diverse enough to exhibit multiple aspects of aesthetic judgement?
3. How can different aspects of aesthetic appreciation be successfully formalised, allowing the automatic generation of 3D sculptures displaying a range of levels for these aesthetic aspects?
4. Can an individual's preference for different aesthetic aspects be modelled using the constant, subjective and transient categories, ultimately allowing the semi-automatic generation of 3D sculptures personalised to an individual?

Question 1, makes an important distinction between two different scopes of aesthetic judgement, collective and individual. The collective scope considers how a particular aspect contributes in general to aesthetic judgement and what other aspects may interact with the aspect. This provides important generalised information about how the aspect should be handled; however, it cannot as easily be applied to individuals, where their subjectivity can contradict these global, generalised rules. This introduces the need to consider both the collective and individual scope. This question is considered within Chapter 3, where two reusable applications will be introduced, designed explicitly to collect data about aesthetic judgement at the collective and individual levels. These applications form the basis for this project, allowing the collection of data in order to answer Questions 3 and 4, plus form the basis on how the project can be continued.

The main goal of Question 2 is to create an algorithm for generating 3D sculptures, which, in conjunction with the framework systems from Question 1, enables the investigation of 3D aesthetics. The generation process must adhere to several restrictions: be quick to run, the generated sculptures must also have the following attributes: they must be visually varied, exhibit a range of values across multiple aspects of aesthetic judgement, should be abstract and not intentionally represent real-world objects and should be usable within a VR environment with respect to the performance, ability to render the item in a straightforward manner. None of the existing generation methods, covered in more detail in Chapter 2, meet the criteria outlined above, indicating that a brand-new algorithm needs to be generated.

For Question 3, the aesthetic judgement of 3D sculptures will be investigated, with the aim to ascertain which aspects contribute to their assessment and whether they can be successfully formalised, allowing their use as an objective function as the basis of the sculpture generation method identified in Question 2 to generate sculptures which display one or more aspects. It is not expected that every possible aspect of aesthetic judgement will be identified and investigated, however, the most influential or common aspects will be identified and the set of attributes which contribute to the aesthetic judgement of 3D sculptures will be obtained through direct assessment of 3D sculptures, using the framework developed for Question 1 and the sculpture generation method from Question 2. In addition to obtaining a set of applicable aspects, their relative importance will also be established with respect to the other aspects and how positive or negative they are. Once the most common/important aspects have been identified, measures will be formalised to allow sculptures to be generated that exhibit these aspects. A measure will be considered successful if a sculpture generated using the measure exhibits the intended features. For example, if a sculpture is generated based on a specific emotion, if this emotion is identified by people when viewing the artwork, then the measure generated will be considered successful. This will be ascertained by conducting a human-subject experiment, using the human understanding of the aspect to confirm whether the measure is successful. A small number of aesthetics will be formalised due to the time constraints and the potential difficulty of the endeavour. In order to provide an answer to this question, each of the aspects must be reliably and successfully described by a measure. The number of measures which can be formalised will be limited due to the scope and time constraints on the project, however, a significant number of novel measures will be identified and presented, enabling the theoretical model from Question 4 to be built.

Question 4 will optionally be considered part of the project, depending on time constraints and the success of the measures as observed in preceding research questions. The main aim of this question will be to utilise the defined measures from Question 3 to automatically generate sculptures personalised to an individual based on that individual's preferences. Ultimately, we will study how combinations of the different aspects of aesthetics combine to represent an individual's aesthetic judgement.

Answering these challenging questions presents a unique opportunity to enhance academic knowledge in Computational Aesthetics, Evolutionary Computation, Evolutionary Art and Artificial Intelligence, all the areas which are required in order to answer all proposed questions. In addition, the research has potential applications to industry, for example, to help design new products, create artwork which can be 3D printed and auto-generate content for computer games or computer animation.

1.5 Contributions of this Thesis

Completing the set research questions allows a greater understanding of 3D aesthetics, and the aesthetic judgement of 3D items and answering each research question will yield major contributions:

1. The experimental framework, resulting from answering Question 1: the system presented in this Thesis allows for fully remote, interactive custom systems to be used within an experimental context, allowing the collection of data specifically relating to aesthetics at the two different scopes, split further into
   (a) The Virtual Reality Environment (VRE) that is specifically designed to collect data about aesthetic judgement at the global level
   (b) The VRE that is specifically designed to collect data about aesthetic judgement at an individual level.
2. A new artwork generation method, developed for Question 2, explicitly designed to allow the testing of aesthetic relevance and preference within both 2D and 3D artwork, allowing conversion and comparison between both dimensionalities.
   (a) The system allowing the generation of high-resolution, visually diverse and interesting 2D artwork
   (b) The system allowing the generation of high-resolution, visually diverse and interesting 3D artwork

3. New measures forming part of aesthetic judgement in 3D sculptures will be formalised to answer Question 3. These can be utilised individually or in combination for the analysis or automatic generation of 3D sculptures, considering the following aesthetic aspects, which have been identified as common and important aspects of the judgement of 3D sculptures:
   (a) Connected
   (b) Friendly
   (c) Calm
4. A theoretical model of aesthetic preference, designed to take into account the constant, subjective and transient nature of aesthetic judgement, to be utilised within an AI-based auto-generation system

Contribution 1, described in Chapter 3, includes these components: a fully remote experimental system allowing for the running of VR experiments even when in-person contact is not possible, and this provides many benefits such as the reduction of bias and the parallel collection of data. In addition, the two main VREs are utilised to investigate distinct areas of the aesthetic process. The first VRE allows the comparison of multiple pairs of pieces of artwork to allow the extraction of accurate aesthetic preferences across all ages, used in Chapter 5, to investigate the individual preference of symmetry within sculptures. The second allows ad-hoc grouping of sculptures within a VR environment, providing the ability for comparison, ranking and choosing of sculptures, allowing the extraction of a generalised overview of the aesthetic process or the direct comparison between human understanding of aesthetic terms and the formalised version of the term. These systems allow the collection of numerous types of data from participants, such as gaze information or positional details, which can contribute to further understanding how a person aesthetically judges a sculpture.

Contribution 2 will take the form of a versatile abstract art generation method which is capable of creating a wide range of both 2D and 3D pieces of art, allowing conversion and comparison between both, called the Axial Generation Process (AGP), full details are presented in Chapter 4.

Contributions 1 and 2 will be the basis for Contribution 3. With this approach, a ranked list of aesthetic terms was elicited, along with metadata around these terms, such as how much the term contributed positively or negatively to the aesthetic judgement of sculptures. This represents a step forward in understanding how people judge abstract sculptures and what they like and dislike, as well as how different aspects interact and are used in conjunction with each other. In addition, the most common/important measures were extracted and converted into formalised measures based on the data collected and utilised to generate a series of sculptures where the exemplification of the term closely matches the human understanding of the term.

Contribution 4 involves proposing a new model of aesthetic appreciation based on the above items, which allows a system to model the aesthetic preferences of a user and use this to generate sculptures. The model splits the process into three separate categories (constant, subjective and transient), and tracking the aspects and their values in each category allows the model to continually be refined and updated based on user choices.

1.6 Structure of this Thesis

The remainder of this Thesis is organised into eight chapters, with the following structure. Chapter 2 considers how aesthetic judgement has been thought of throughout history along with the latest thinking of the current understanding of what aesthetics are and which aspects contribute to the judgement, as well as describing how the understanding of the topic is currently being investigated across multiple different fields, revealing the opportunity of how the understanding of aesthetics can be improved, seized upon within this Thesis. Chapter 3 looks at how technologies such as VR can be utilised to develop a framework which can be followed to extract details about aesthetic preference on an individual and a general level. This is followed by Chapter 4, which discusses some potential sculpture generation methods looking to fulfil the unique requirements this project has and presents a fully realised custom sculpture generation algorithm creating artwork which can be used within the VR framework to further understand aesthetics and the aesthetic judgement, along with a comparison between the presented technique and other existing 3D generation algorithms. Chapters 5 and 6 combine the work from the previous two chapters, utilising the novel work in order to identify and assess the importance and positivity of a wide range of aesthetic terms when judging 3D sculptures within a VR environment, a second study is also presented here that considers the use of the second VR system to extract individual preferences for a stalwart of aesthetic judgement, symmetry. Overall, combining to deliver a picture of how well the VR framework can be used for the given task and identifying noteworthy patterns about aesthetic judgement, which can be utilised both within and without an auto-generation context. The information from these two studies is combined to create three formalised measures for the most commonly applied aesthetic tags in Chapter 7. The process of formalising and validating the measures is presented along with a further experiment which considers how closely the formalised measures match the human understanding of the same terms. These measures are then utilised within the novel model of aesthetic appreciation attempting to represent an individual's aesthetic preferences whilst taking into consideration the subjective and changing nature of aesthetic appreciation, detailed in Chapter 8. Finally, Chapter 9 closes out the Thesis by presenting the conclusions and potential future work and research focus arising from the work completed.

Chapter 2

Art as Aesthetics

Many theories have been proposed throughout history, emphasising that there is no correct way to judge art. However, this plethora of ideas can make the topic hard to grasp, but the angle through which art is viewed can help to reduce this complexity; philosophy considers art's role and impact on humanity, psychology considers art at a similar level but within the context of individuals, neuroscience and computer science consider the low level contributing aspects that combine to make art, the former through the mind's response and the latter the visual aspects that can combine to make art. Even though each context has a different understanding of art, a large amount of commonality exists. One way this commonality can be seen is to consider art in terms of aesthetics. These aspects which contribute to the judgement of artwork and form the basis of an aesthetic judgement, a concept which appears in all contexts. This chapter considers different aesthetic theories across the contexts, showing that even aspects identified millennia ago are still pervasive in current-day thinking, using these details to identify gaps in the existing research upon which this thesis is based.

2.1 Common Aesthetics

Aristotle introduced one of the first formal definitions of aesthetics and listed several aspects that allowed plays to be judged as a comedy or a tragedy. To make something a tragedy, it needed to have both action and moral significance, among other aspects. Whilst these aspects were defined, they were still highly subjective and required significant interpretation, for example, what contributed to moral significance. However, potentially recognising this subjectivity, Aristotle reduced the amount of potential external factors when judging art, not considering it in terms of political, religious or social contexts, providing an initial answer to the question about whether the study of artwork is best done in the social and material context it originates from or attempting to abstract it from that. This conundrum was also thought about by Frank Sibley, who believed that art needed to be judged within an appropriate context (Sibley, 1959), and Kant, who suggested that beauty can only be understood if the object does not have a purpose other than to be enjoyed as artwork (Haskins, 1989), indicating that art should not be judged in any other context than its own. This has been taken forward in other research fields, such as computer science and psychology, where abstract artwork is often studied to deliberately remove these external influencing factors from the judgement.

Each distinct aesthetic theory has introduced their own focuses and aesthetic aspects through which art can be judged; Pragmatist aestheticians such as John Dewey (Dewey, 2008) and artists such as Jackson Pollock only considered the aesthetic experience in its entirety, introducing several aspects about what combines to make art in this aesthetic experience, such as vitality and democratic meliorism (the purpose of art being solely to improve and understand art). Parallels can be drawn between this and auto-generating art, where not only is a specific aspect (or set of aspects) improved throughout the process, but also the novelty of the presented items is maintained to ensure that the same items are not generated on multiple occasions, overall, improving the knowledge and understanding of the type of art being generated. The concept of novelty was also investigated by Schiller, who referred to it as freedom, a form of rule-breaking in terms of the medium and style being used. The novelty also has close links to the concept of creativity (Boden et al., 2004), an underlying factor of generating artwork.

Expressivists considered art as the expression of feeling or emotion, where the art itself cannot be identified by its constituent parts, e.g. paint on a canvas. More recently, emotional content has been introduced into generating descriptions of images (Yang et al., 2019). The success of this approach suggests that emotional content may be a potential future method for generating artwork. However, translating emotions to specific aspects of art is exceptionally difficult, and similar to the Expressivists, may require considering the artwork as a whole rather than looking at low-level aspects.

Phenomenology is another area which introduced numerous aesthetic aspects, contemplating capturing and articulating the original meanings of the art experience, suggesting key notions such as imagination, negation, nothingness, freedom and commitment.

This expansive range of aspects only tells part of the story; how the aspects combine to form judgement is another important aspect of the process. Heidegger suggested a holistic approach where all aspects contributing to the judgement are equally essential to the judgement (Wartenberg, 2013), it was also important to Heidegger that art revealed the truth, similar to the back story and how we interpret it and what the artwork reveals. Whilst every theory introduced new aspects and areas for consideration, several aspects have been common throughout.

2.1.1 Ethics

Similar to how Aristotle considered tragedies to need to have moral significance, Plato considered art to have an essential role within society for the express purpose of teaching younger generations about what he considered fine and graceful (Allen, 2002). One of his major teachings was that artwork was only valid if it fit within his ideals, and actively sought to censor works which did not fit them. This ethical aesthetic of artwork is something which is a regular feature throughout many theories of judgement. Medieval philosophers such as Pseudo-Dionysius had strong religious ethics, disregarding artwork which had a pagan influence (Bredin, 2002), this religious aspect continued through modern-day thinking, Hegel used it as a basis to define art, linking all art to Christianity (Gaiger, 2002); and similarly to Plato believed art should be used to promote morality.

Even Plato's ethical considerations are still an important part of art judgement, often surging into popular thinking throughout history. Comparisons can be made between Platonic thinking and attempting to place blame for behaviour on many creative pursuits, such as Rock music damaging teenagers in the 1960s, video games such as Grand Theft Auto causing a rise in crime rates in those who played it, the seminal album *Straight Outta Compton* by N.W.A and most recently, the release of Trap music encouraging young people to commit acts of violence, even though it is regularly proven that these have had minimal effect on people's behaviour. The inclusion of ethics as a significant part of aesthetic judgement raises multiple interesting points: ethical values are extremely subjective and highly dependent on an individual's experiences.

The impact of ethical considerations on aesthetic judgement also suggests that the ethical judgement takes place in conjunction with an aesthetic judgement and has a higher cognitive weight, meaning that it can be used to override an individual's objective judgement of a piece of art. The extent to which this influences the judgement of artwork leads to the assumption that ethical judgement is a significant part of aesthetic judgement, as well as being an internal process governing other aspects of an individual's personality. As discussed in Chapter 1 with the sculpture "A Real Birmingham Family", the ethical judgement can also be tied to the backstory behind a piece of art; however, whether it is possible to assess an individual's ethical preferences regarding artwork and present them with artwork which reflects their preferences but has been considered in terms of the automatic generation of artwork (Brown and Ventura, 2022; Ventura and Gates, 2018).

2.1.2 Beauty

Beauty also holds a very prevalent place; for Plato, beauty created other aspects such as love (Grube, 1927), for Aquinas, beauty was defined in terms of religion, emphasising light, colours, harmony, symmetry and concordance, aspects which represented God and by extension beauty (Margolis, 2005). These concepts of order and harmony are common through many different interpretations of beauty, Hutcheson agreed and believed beauty could be investigated by assessing aspects such as order, harmony and design (Shelley, 2005). However, instead of being a set combination of these aspects, he thought that beauty was a subjective concept that related as much to the humans experiencing it as the objects which were creating it (Matthews, 1998).

This pure subjectivity approach was opposed by Hume (Winegar, 2011), who believed that objects themselves could be beautiful, and due to the shared taste of beauty, could be used to judge individual tastes. He introduced multiple stages in the judgement of beauty: perceptual, where the qualities of an item are observed and affective, where the sentiments are felt (Shelley, 2005). These two stages are similar to the shared and universal concepts used by Leder et al. (2016) or the persistent and conditional aspects proposed by Schiller (Jacquette, 2013). However, whilst recognising the changeability, one area which is overlooked is how volatile these different aspects can be, with some aspects potentially changing more regularly than others, making fitting the aspects into these categories a complicated process.

Kant did believe beauty to be subjective however, he did categorise what was required for an object to be called beautiful: quantity, quality, relation and modality (Hayn-Leichsenring and Chatterjee, 2019). These four aspects accounted for the different understandings of artwork, such as how the same objects are judged as beautiful by multiple people and that beauty is based on spatial and temporal relationships. Kant also started the distinction between beauty and aesthetics by introducing the concept of the Sublime, which stood in opposition to beauty and related to natural phenomena. This relationship between art and nature has often been used for formalised measures used to judge artwork, such as the Fractal Dimension (Spehar et al., 2003) or Benford's Law (Acebo and Sbert, 2005). Whilst these measures do not quantify Kant's idea of the Sublime (and do not intend to), they provide the ability to formally measure and compare artwork in terms of these specific aspects. Measures such as these are the backbone of Computational aesthetics.

2.2 Formal aesthetics

Formal aesthetics can be traced back to Clive Bell, who rejected most notions of previous theories, stating that the subject of art was the arrangement of lines, colours, shapes and space (Bell, 2005). Part of the reason for this approach was to search for an explicit definition of art while highlighting the collective nature of art where beauty is recognised by multiple people, something that computer scientists and mathematicians have sought to continue. One notable attempt at providing formal measures is from the mathematician Birkhoff (1933), who proposed that aesthetic judgment of paintings was related to many factors, but similar to scholars such as Hutchinson, who thought the uniformity and variation of art was an indication of its beauty, and as previously mentioned Aquinas (Margolis, 2005), quantified aesthetic value as the ratio between order and complexity of a piece of art.

The inclusion of order throughout many theories of art indicates the importance it has in art judgement, this can, in part, be due to its central role in processing visual information, where the more ordered an item is, the easier it is for our brains to process (Bertamini et al., 2018). The term "order" can still be very subjective and mean different things in different contexts; this has led to extensive research being performed, resulting in different methods being suggested for quantifying it. Most attempts to define the order relate it to the symmetry of the artwork, for reasons summed up in "Symmetry is one of the ideas by which man through the ages has tried to comprehend and create order, beauty and perfection" (Weyl, 2015). Symmetry also holds near universal importance across multiple applications, such as architecture (Zeki, 2019; Salingaros, 2020; McDermott et al., 2012), inter-human attraction (Rhodes, 2006), user interface design (Bauerly and Liu, 2006) and within artwork (al Rifaie et al., 2017; Bergen and Ross, 2013; den Heijer and Eiben, 2010b; den Heijer, 2012; O'Reilly and Hemberg, 2007; Osborne, 1986; Vinhas et al., 2016). With such a high level of evidence suggesting its importance, it is a safe assumption that symmetry plays a key role in the evaluation of artwork. However, the extent to which it contributes to aesthetic judgement is far less well defined, considering it a part of judgement and the multiple roles an aspect can have, e.g. shared or individual, limits the effectiveness of using these well-defined measures to quantify order. Instead, to achieve this, it needs to be understood how symmetry fits within the wider context of art appreciation, especially when there are a variety of factors which can influence an individual's symmetrical preference (Jones et al., 2007; Leder et al., 2019; Weichselbaum et al., 2018). To correctly place symmetry, evidence needs to be collected which considers how people judge artwork when different levels of symmetry are present, an approach taken in Chapter 6. The same holds true for the vast majority of formal measures; the quantification of the aspect only represents a small part of the aesthetic judgement process, and to further understand, the process needs to be investigated alongside any measures which make up part of it. Other examples of measures include using an image's 2D power spectra to determine how closely it matches a target image (Gircys and Ross, 2019a) or providing different measures to fit within the equation proposed by Birkhoff, such as using Shannon's Entropy and Kolmogorov Complexity (Rigau et al., 2007).

Nietzsche shared some commonality with Expressivism and theorised about the aesthetic experience of looking at art where the deepest and most horrible truth of the world is glimpsed (Nietzsche, 2017). He also believed that the aspects highlighted by an artist reveals their evaluative stance towards life and that these aspects could be understood in terms of dichotomous pairs ranging between life-affirming and life-denying, for example, strength versus weakness, health versus sickness. This gives an insight into how formal aesthetics could fit within the aesthetic judgement. Suggesting that judgement could be quantified by sets of measures, each ranging from one extreme of a high-level concept to the other. This can also account for subjectivity, where a person sits at a specific point in a vector of multiple aesthetic aspects, and this vector differs between different people. This dichotomy is already present in other methods of measuring aesthetic value, such as within the order/complexity ratio proposed by Birkhoff.

Birkhoff's formula aims to represent aesthetic judgement and has spawned multiple experiments, including the assessment of geometric objects (Katz, 2002) and abstract images (Osborne and Farley, 1970). Some of these studies suggested that the order and complexity are partially responsible for the aesthetic judgement of an item. However, such a simple formula fails to take into account the many other factors identified by the other theories and therefore would not be able to identify aesthetic judgement, especially considering it takes no account of any external factors such as age and gender (Salkind and Salkind, 1997), stipulating that this approach is limited when trying to fully model aesthetic appreciation.

Order is not the only aspect which has been formalised; a wide variety of aspects have been specified within quantitative measures, often being derived from real-world data such as examining existing artwork and extracting different features of the images (Acebo and Sbert, 2005; Alamino, 2015; Machado and Cardoso, 1998; Matkovic et al., 2005; Rigau et al., 2007; Ross et al., 2006; Spehar et al., 2003; Tinio and Leder, 2009; Vázquez et al., 2001). The approach of providing quantifiable formula in order to assess an aspect of aesthetic appreciation is an integral part of Computational Creativity, investigating aesthetics and the judgement of artwork, its main focus is to create systems which are capable of generating creative output. These systems can be split into three main subsets; autonomous systems, tools which support the creative process and systems which work with human users to generate output (Karimi et al., 2018). In order to allow a system to produce any kind of creative output, it often uses formal measures representing aesthetic aspects. Due to the difficulty in quantifying and measuring aesthetic judgement, generally, these focus on single measures, including how interesting an image is, by measuring its contrast (Matkovic et al., 2005), using Benford's Law to assess how natural an image may appear (Acebo and Sbert, 2005), whether colour gradients found in an image match those used in existing paintings (Ross et al., 2006), the relation between image and processing complexity (Machado and Cardoso, 1998) and how symmetrical an image is (al Rifaie et al., 2017).

Within recent years, subjective themes have slowly started to be introduced, linking knowledge between computational creativity and psychology (Johnson et al., 2019; Lamb et al., 2018). As well as specific topics such as how it may be possible to quantify emotion based on high- and low-level features (Rodrigues et al., 2018) and considering how important aspects such as perceived story are to the assessment of an image (Colton, 2008b). These works represent the first steps of researching subjective aesthetics, however, there has not yet been any attempt to fully quantify the details of these aspects. Without these subjective aspects, any creative system will struggle to produce content on a similar level to a human. Introducing measures which formalise these aspects will greatly improve the efficacy of these systems.

One of the best approaches for automatically generating artwork would combine these systems and formal measures in such a way that accurately reflects a person's aesthetic preferences. This approach has previously been investigated for 2D aesthetics by observing individual preferences for a period of time and modelling that as a multi-objective optimisation (Ekárt et al., 2012), as a machine learning model (Li et al., 2012) or through creating artificial art critics (Machado et al., 2021; Romero et al., 2003). These systems show some good results, however, they were not found to represent a full aesthetic judgement. In addition, these models can be limited with little room for expanding based on new knowledge or changing preferences from an individual, this is a required aspect when moving forward with trying to model aesthetic preference.

Some of the difficulty of formalising the aspects lies with the differences between the computational understanding of these terms compared to how they are understood by humans. Computational approaches generally look at attributes which are easily understood by computers. These will not necessarily reflect the human understanding of the same term and do not account for any ambiguity in terms. For example, the term complexity could relate to a piece of artwork with complex emotional content or complex subject matter as well as visual complexity such as multiple colours, lines and shapes (McCormack et al., 2021). This indicates why a disparity exists between the human judgement and the analogue followed by auto-generation systems. Attempting to solve this disparity forms an important part of the auto-generation of aesthetically pleasing artwork. Without it, it will be difficult for any generation system to reliably create aesthetically pleasing artwork.

Not all theories of art work well with this formal approach; one common argument made against formalist theories is the idea that all artwork does not share common features which can be measured and compared. However, this view is limited and assumes many things about artwork. For example, similar aspects may be present, but the combination and interplay between these aspects may be highly complex. This interplay was identified by Derrida (Marriner, 2002), who thought we have systems of signs where each sign only gets its meaning in relation to other signs. This is interesting as it implies a relationship between how familiar someone is with art affecting how they interpret and judge art and refers back to the subjective aspects.

2.3 Subjectivity of art

As previously discussed with the concept of order, a lot of the aspects have a high level of subjectivity. To identify these aspects, different methods have been used such as analysing critiques of artwork and identifying common terms used to judge artwork (Sibley, 1959) and looking at specific attributes about an individual such as their level of art expertise (Chatterjee et al., 2010), equating physical manifestations of emotion to link to the aesthetic experience (Pelowski, 2015) or how the aesthetic appreciation of visual media changes over time (Isik and Vessel, 2019).

This subjectivity has been identified within most of the previously discussed aspects, how varied the explanations for aspects such as ethics and beauty are, and theories such as Expressionism all highlight the high level of subjectivity in art. There are many reasons which may account for the high variance in experiences people have when viewing artwork, such as education and expertise. Even the experience itself is known not to be the result of a single judgement (Nadal and Chatterjee, 2019). This also seems to be reinforced by research establishing that multiple pathways of the brain are used when making aesthetic judgements, including pathways not linked to the physical properties of the object (Skov, 2019) aligning closely with the theories proposed by Kant. The relationship between different types of aesthetics has been directly investigated by (Leder et al., 2016), looking at the contribution of private and shared taste when judging artwork, closely linked to the concepts of subjective and constant aesthetics introduced in this project, showing the complicated link between types of aesthetics and how they all contribute to an aesthetic judgement. Similarly, (Pugach et al., 2017) has shown that the aesthetic preference of an individual changes over time, indicating that transient aesthetics form a major part of the judgement, something which goes deeper than just subjective versus constant. This demonstrates the need to focus specifically on these aspects as factors such as expertise and context knowledge change more quickly than individual preferences of aspects like colour preference.

Between aspects such as emotion (Rodrigues et al., 2018), perceived back story (Colton, 2008b) and more stable aspects such as the fractal dimension (Spehar et al., 2003), it suggests that all aspects can be split into the three categories described in Chapter 1.

The role of expertise has long held high importance for judging artwork, Plato considered many of his ideas as things which non-philosophers would be unable to grasp and therefore restricted the creation and understanding of artwork to highly educated philosophers (Allen, 2002). This influence can also be seen in the teachings by Hume, who believed that art could only be truly understood by people with the appropriate level of good sense and delicacy (Hume, 2017). This restriction on the availability of art knowledge is something which can still be seen today where expertise can be viewed as a pre-requisite for recognising aesthetic beauty, and it has been shown that expertise is an influencing factor on the judgement of artwork (Leder et al., 2019). However, this restriction limits the effectiveness of any research, directly contradicting the universality of art indicated by the importance of aesthetic judgement for everyone on a daily basis.

2.4 Understanding art in terms of language

Understanding which aspects contribute to aesthetic judgement can be complicated, as it can be challenging for individuals to articulate their thoughts when judging a piece of art. This is a difficult process for many people, however, everyone uses a common tool in order to discuss their aesthetic judgement: natural language. This formed the basis of the approach used by Nelson Goodman, whose views considered aspects like fluidity and freshness (Carter, 2005; Shusterman, 2013), moving the dialogue of artwork beyond emotion and formal aspects and helping to emphasise the context within which artwork is judged (Wollheim, 1970). This ideology was also followed by Frank Sibley, who used analytical aesthetics to compile a list of high-level terms used to describe artwork, known as aesthetic concepts. These concepts have been a starting point for formalising some of the aspects of artwork, e.g. complexity and their usage within critical works suggest they may be based on repeatable rules, e.g. the presence of features such as colour gives an indication of the aesthetic properties such as using bright colours precludes the use of terms such as Garish. This suggests that aesthetic judgements are objective to a certain degree where we have a communal sense of what is good in the same sense as we have a communal sense of what each colour is. It has been shown that people are more consistent when judging art others find aesthetically pleasing (Leder et al., 2016).

2.5 Art Generation

The process of automatically generating artwork is not a new process, even manual algorithmic methods have been used to generate physical artwork. Over time the rise of evolutionary art (Sims, 1991; Ekárt et al., 2011; den Heijer and Eiben, 2010b; McCormack and Lomas, 2020; Secretan et al., 2008; Tweraser et al., 2018; Vinhas et al., 2016; Machado et al., 2008; Galanter, 2012) has produced a wide variety of different generation techniques. A clear dividing line between these systems is whether the system generates 2D or 3D artwork.

Among the many methods used within systems capable of creating creative works, one of the most popular is Evolutionary Algorithms. These use an abstracted version of Darwinian Evolution to effectively search through a high number of possible items and find potential solutions to problems (Eiben and Smith, 2003). However, unlike other applications based on Evolutionary Algorithms, they often do not have a specific problem to solve and can continue to generate content indefinitely (Cook and Colton, 2018). Their ability in being able to generate creative content has been shown in a variety of different mediums such as dance (Antunes and Leymarie, 2012), fashion (Lourenço et al., 2017), music (Bountouridis et al., 2017) and most commonly artwork (den Heijer and Eiben, 2012) known as Evolutionary Art systems. These were initially introduced by (Sims, 1991) and provided an effective way of generating artwork, working either in unison with human users (Takagi, 2001) or working in a completely autonomous manner (den Heijer and Eiben, 2010a).

The issue of assessing the appeal of a generated item and the inability of existing aesthetic measures to do so limits the effectiveness of these systems. This has led to the majority of Evolutionary Art systems having a human user who assesses the generated artwork, for example, in the PicBreeder system (Secretan et al., 2008). In these systems, the user works in conjunction with the system, providing guidance to help it traverse the search space (Machado and Cardoso, 2002). This ensures that the system does not have to assess the items for artistic merit, greatly reducing its complexity. Unfortunately, involving humans to judge the output has some major drawbacks that can prevent the system from generating aesthetically appealing pieces of art. The most detrimental issue is User Fatigue (Takagi, 2001), which has been credited as the main reason for these systems being unable to generate suitably complex and interesting pieces of art. User fatigue is the phenomenon where a user will only perform a small number of generations, it is attributed to many aspects, such as the user becoming bored or not seeing significant changes in the generated images.

In order to negate this issue, many workarounds have been proposed to neutralise its impact, split across two categories: adding in additional aspects the program must check such as favouring novel items (Machado and Cardoso, 2002; Vinhas et al., 2016), improving the diversity of the population (Lehman and Stanley, 2008) and secondly changing the context the program is used in, such as sharing the generation of items across multiple users (Secretan et al., 2008) or introducing a semi-automatic assessment by generating user profiles (Li et al., 2012; Ekárt et al., 2012; Hornby and Bongard, 2012). This approach represents a step forward with assessing aesthetics; however is generally limited in two main ways: they have only been studied with respect to two-dimensional images, and secondly, the models generated are often a black box, making extracting specific details of an individual's preference difficult or impossible. This highlights the importance of investigating features on an individual basis, allowing specific aesthetics to be used for content creation as well as application to modelling an individual's preference.

Aside from using different aesthetics, other methods have been introduced to help generate aesthetically pleasing content, involving using different representations of artwork. Some of these representations have been inspired by Natural Language systems, known as shape grammars (Colton, 2012) or expression trees, which represent an equation which is subsequently used to generate an image (Sims, 1991). Each of the representations can be used in different ways in order to generate the final artefact. Expression trees can be translated into an image by using their value to calculate the colour of each pixel (Ekárt et al., 2012; Secretan et al., 2008), or generating the path of multiple lines across a canvas (Colton et al., 2011).

These varied solutions generate interesting content; however, they are still limited by the number of users willing to generate the content and the time it can take to generate complex and aesthetically pleasing content. The most effective solution for these issues would be to create a fully autonomous system capable of assessing its own output before presenting the results. Examples of this style of system already exist, such as the one presented by Cohen et al. (2017), which tried to re-create artwork in the style of Mondrian. However, these systems face limitations on how the output is assessed and still often fail to generate the aesthetically pleasing content achieved by their human counterparts, indicating that the current level of knowledge surrounding aesthetics and the ability to formalise these aesthetics is insufficient to create an autonomous system capable of generating aesthetically pleasing content.

Two methods exist for navigating to the region containing someone's preferences: using formalised measures to automatically navigate or asking the person to direct the search. Formal measures are often limited in their effectiveness to accurately describe aesthetically pleasing artwork, and on the other hand, a human user, who, despite being instinctively more proficient at assessing the aesthetic appeal, hinders the search in other ways such as through user fatigue (Takagi, 2001). One possible explanation for this is due to the extensive size of the search space, the system is not capable of locating sufficiently interesting artwork sufficiently quickly to continually engage the user.

Attempts have been made to address these issues by giving the user more control over the direction of the search (Easton et al., 2019). However, these methods also have their limitations. The most effective method would be to constrain the search space to include only the items someone finds aesthetically pleasing. However, due to the lack of exact formalisms for individual aesthetic preferences or even a full definition of what aspects may contribute to an aesthetic judgement, this would not currently be possible.

2.5.1 Generation of 2D images

2D images represent the most popular generation type when referring to evolutionary art, however, as this project focuses on 3D abstract sculptures, an in-depth discussion is beyond the scope here. The generation of two-dimensional items has been widely studied, and the generation of images based on mathematical expressions is the most popular approach within evolutionary systems. These can be represented by lisp-style expressions (den Heijer and Eiben, 2010b; Sims, 1991) or in the form of expression trees (Colton et al., 2011; Ekárt et al., 2011; Li et al., 2012; Machado et al., 2015; Mills, 2016; Vinhas et al., 2016). The way these expressions are used to generate an image takes a variety of forms, the most common procedure is to use a value generated by the expression to set a property of each pixel in an image, such as the luminosity (Gircys and Ross, 2019b; Machado et al., 2015), using multiple expressions to calculate the RGB value of a pixel (Colton, 2012; Ekárt et al., 2011; Li et al., 2012) or mapping the value to a colour lookup table (den Heijer and Eiben, 2010b). Expressions have also been used to calculate the position and colour of a line over a series of time steps (Colton et al., 2011; Easton et al., 2019) and to create animated content (Mills, 2016; Tweraser et al., 2018). Other underlying data types are routinely used, such as Compositional Pattern Producing Networks (CPPNs) (Lehman and Stanley, 2008; Nguyen et al., 2015; Secretan et al., 2008; Tweraser et al., 2018); Context-Free Design Grammars (CFDGs) (Vinhas et al., 2016); encoding the parameters to use in another generation process (Davies et al., 2016), using SVG (den Heijer and Eiben, 2012), or maintaining a list of shapes to be placed on the canvas (Colton, 2008a). Entirely custom representations have also been designed to solve particular problems, for example, to efficiently hold the data to create a piece of art in the style of Mondrian (Cohen et al., 2017).

2.5.2 Generation of 3D artwork

In the wider context of visual art, aesthetics are not just applied to plays, two-dimensional paintings and images, as is the case with the majority of output generated by existing evolutionary art systems. Other mediums exist, such as computer games and 3D art, where aesthetics are of paramount importance (McCormack et al., 2020; Niedenthal, 2009). Some research has been completed which looks at the aesthetics of 3D items, for example, looking at whether the rendering method and resolution affects a user's judgment of different items (Dev et al., 2017), looking at the relative importance of resolution compared with frame rate (Claypool et al., 2006) or (Wages et al., 2004) who found that a higher level of realism in computer games led to a decrease in the believability of the environment. Even with this research being limited in comparison to the amount completed in other fields, it is clear that visual aesthetics significantly affect the enjoyment and engagement of 3D items.

Similar to 2D items, generating 3D artwork is not a new endeavour; many examples of 3D generative art exist (Lambert et al., 2013; Latham and Todd, 1992). Often, these systems had limitations due to available processing power; however, some better-known systems have been exhibited in public art galleries such as Mutator (Latham et al., 2021), Galapagos (Sims, 1997) and xTNZ (Antunes and Leymarie, 2008). 3D art, such as these examples, and particularly Sculptures, challenges the notions of what makes something artwork transcending both physical and digital media. Many theories, such as those from Kant, suggest that art can have no purpose beyond itself, however, even sculpting during

Plato's time was considered an art form. Sculptures have been created for many reasons throughout history, often promoting a mythical theme: to worship deities or to guard the resting places of the dead. Their ability to make the imagined real or to illustrate complex topics means they have had a powerful impact, exemplifying the old adage that a picture (or sculpture) can be worth a thousand words.

Within Evolutionary Art, 3D item generation does not have an extensive number of existing systems however, there is still a diverse range of representations and methods being used. This includes manipulating the colour and visibility of voxels using CPPNs (Hollingsworth and Schrum, 2019) or Context-Free Grammars (Bergen and Ross, 2013; McDermott et al., 2012; O'Reilly and Hemberg, 2007), using Graph Grammars to generate a set of points in 3D space (McDermott, 2012), using Shape Grammars to create 3D items (Byrne et al., 2013; Muehlbauer et al., 2017; O'Neill et al., 2010) and evolving parameters to use within an external generation system (McCormack and Lomas, 2020; Nicolau and Costelloe, 2011). Other non-direct generation methods have also been used, such as taking 2D content and adding the third dimension (Easton, 2018).

Due to the limited generation systems, it is perhaps no surprise that only a handful of examples of systems assessing or generating 3D items exist. These look at different outputs such as creating textures for 3D items (Reynolds, 2016), working within a 3D medium such as architecture (Bak et al., 2016), generating 2D artwork which can be viewed and explored within three dimensions (Easton, 2018) and generating 3D models either of representational objects (Bergen and Ross, 2013), abstract objects (Chu, 2021; McDermott, 2013; Tian, 2023) or as the basis for architectural designs (Muehlbauer et al., 2017). All of which delegated the assessment of the items to a human user to apply their own judgement. This could be due to many different reasons, such as keeping the systems simple, however, none of the existing aesthetic measures discussed previously apply to 3D items, making it impossible to currently apply automatic assessment to 3D items.

Chapter 3

Development of an experimental aesthetic framework

The work presented in this chapter has been adapted from the following publications:
E. Easton, U. Bernardet, and A. Ekárt. Is Beauty in the Age of the Beholder? In *Artificial Intelligence in Music, Sound, Art and Design: 12th International Conference, EvoMUSART 2023, Held as Part of EvoStar 2023, Brno, Czech Republic, April 12–14, 2023, Proceedings*, pages 84–99. Springer, 2023a.
E. Easton, U. Bernardet, and A. Ekárt. Contributors to the aesthetic judgement of 3D virtual sculptures. In *2023 Third International Conference on Digital Creation in Arts, Media and Technology (ARTeFACTo)*. IEEE, 2023b.

3.1 Introduction

This chapter introduces the experimental framework through which aspects of aesthetic judgement can be identified, providing an answer to Question 1 and discussing the contributions linked to the question in detail. Two VREs were explicitly created to collect: (1) individualised aesthetic data using a novel mine cart environment (Section 3.3) and (2) generalised data through a gallery environment (Section 3.7). Validation will be provided for each environment, indicating their suitability towards understanding aesthetic judgement. Through the use of these two environments and the registration process, it will be possible to collect data in the following chapters, which will be used to formalise measures of aesthetic aspects and be presented in a model of 3D aesthetic judgement.

Using Virtual Reality (VR) as the technology for conducting the experiments provides multiple benefits, such as being in control of the environment, being able to view the sculptures in 3D without having to create them physically and allowing specific data to be collected, such as positional information, which would be much more difficult using more traditional technologies such as a desktop monitor. The final aspect to address was how to present the sculptures to the participants; viewing the 3D artwork on a 2D monitor would be limiting and potentially affect the results, considering that the environment is an essential aspect of aesthetic appeal (Braun Janzen et al., 2023; Pelowski et al., 2017; Specker et al., 2017). In addition, VR provides additional benefits which can be utilised to enhance the user's experience further when viewing the artwork, such as being able to rotate the sculptures, something which would potentially be difficult if the sculptures were full-size physical items; another benefit of using VR is the ability to display 'impossible" sculptures, for example, ones where all the contained elements were not attached to one another. To avoid these limitations and take advantage of the benefits, VR was chosen, and a custom VR environments were created, which allowed the 3D artwork to be viewed within a 3D environment via a Head Mounted Display (HMD).

The concept of using VR environments for viewing artwork is not new and an ever-increasing number of art projects have utilised VR as a medium such as Osmose (Davies, 1995), Chalkroom (Anderson and Huang, 2017), A Show of Kindness (Cowles et al., 2018) and CAVE (Layng et al., 2019). As well as its increasing use within the art world, VR is also becoming a more popular choice for running experiments (Gulhan et al., 2022; Easton et al., 2019; Osimo et al., 2015) due, in part, to its ability to explicitly control the environment a participant is placed in and the ability to enhance the art viewing experience.

An important aspect of assessing aesthetic preferences in Virtual Reality is ensuring high presence. In this project, presence is considered to be when participants act in the same fashion in the Virtual Environment as they would in real life (Slater, 2003). Having a high presence ensures that the information obtained and presented in the project is valid when considering physical art as well as virtual art. Some known perception differences in Virtual Reality that need to be overcome, such as proprioception, can be altered in VR with participants judging distances to be closer than they are in a Virtual Environment (Valori et al., 2020). One way this can be overcome is to exaggerate smaller distances by using lighting to ensure they are as visible in the VRE.

3.2 Remote experimental setup

One of the challenging aspects to overcome for this project was the Covid-19 pandemic and the related global lockdowns which occurred afterwards. Due to the social distancing requirements in place, many plans for how experiments will be run had to be adjusted to consider social distancing rules. As the restrictions did not seem likely to change when the data was collected, the decision was made to run all the experiments online. However, running the experiments planned for this project completely online required different considerations than running in-person: the experimenter cannot be guaranteed to be present or available when the participant is using the system, any problems with the environment or experiment are likely to be less tolerated and are much more complicated to be rectified, involving having to implement formal versioning and release processes. Aiming the experiments at experienced VR users introduces several concerns about how the system will be judged and used, especially relating to the environment the participant is placed in as well as control methods which will potentially affect the data collected. Finally, all the data management processes need to be different due to the asynchronous nature of how the experiments will be run.

These additional considerations require a significant amount of work, however, conducting an experiment entirely online does introduce additional benefits compared to running an in-person experiment. It is easier for the data collected to reach a broader range of demographics as the experiment is no longer restricted by geographic and institutional constraints, making the potential participant pool extremely large, now including anyone who owns their own VR headset and finally, the experiments can be run and the data collected very quickly, for example, everyone registered can complete the experiment at the same time, without intervention from the researcher.

In order to cope with these changes, the problem can be considered as a Software Engineering problem, where we can use software to mitigate some of the issues raised and take full advantage of the benefits the remote experimental paradigm offers. Two main parts of the experimental process must be considered: the participant and data management and the experimental tasks.

In order to implement the data collection, the client-server architectural design was chosen. The client is the participant's headset, and the server is an API that allows data to be sent and received, along with other critical operations such as checking participant access to a specific experiment and checking which sculptures to display. A more conventional paradigm is used for participant management, a website is provided to allow the participant to register, read through all essential documentation and receive links to download the VR environments from the Meta Quest App Lab. Through utilising this approach, the entire process can be completed remotely with minimal, if any, support provided by the experimenter, aside from posting the initial recruitment notice on the participant recruitment platform and responding to any questions. The complete experiment process is shown below:

1. Participant reads the recruitment notice on the relevant online platform, e.g. Prolific and expresses an interest in participating, potentially including payment for participating, depending on the platform being used.

2. Participant is directed to the registration website, where they will be presented with the relevant Participant Information Sheet (PIS) containing all details about the experiment

3. After reading the PIS, if the participant is willing to continue, they are shown the full consent details and asked whether they agree. If they do, the participant enters their identifier (either email address or platform identifier, e.g. Prolific ID).

4. The participant is then given instructions on downloading the environment from the Meta Quest App Lab along with a five-character identifier used within the VRE.

5. Participant downloads the VRE and starts the application. They are asked to enter their five-character identifier and then start the experiment.

6. Once all tasks are complete within the VRE, the participant is prompted to return to the registration website where an exit questionnaire can be provided. Finally, the participant is shown completion instructions, including how to uninstall the VRE.

The system tracks each participant's progress, ensures that they cannot perform an experiment more than once and ensures that completion of an experiment is recorded to allow the right participants to be paid via their platforms.

3.3 Using VR to explore individual aesthetic preferences of 3D sculptures

This section presents a novel gamified approach to collecting individualised aesthetic preference details when judging artwork, focusing on one or more aesthetic aspects. The process uses a custom VRE to present multiple sets of sculptures a participant chooses between. After making ten choices, the participant's exact value for the investigated aesthetic aspect is ascertained.

3.4 Helping people release their inner art critic

At each step, the VR application expects participants to make binary decisions between two sculptures displaying two different levels of the investigated aspect to elicit the participants' aesthetic preferences. Each decision leads to updating the value of the aesthetic aspect displayed in the next pair of sculptures in the VR application, bringing the values of the displayed sculptures closer together converging on the participant's preferred value. The sculptures displayed in the first step should contrast each other, with one having the maximum value and the other having the minimum value of the chosen aspect. At each step, a sculpture with the same value as the last one the participant chose is displayed again in the next pair of sculptures, however, the other sculpture has its value amended by a fifth of the maximum value. The value of the sculpture is decreased if the participant chooses the lower-valued sculpture and increased if the participant chooses the higher-valued sculpture; thus, the participant's preference is established in five steps. To improve confidence in the recorded preference of the participant, the five steps, starting from sculptures with the maximum value and minimum value, through to convergence to one value, are repeated with different sculptures corresponding to the same set of values at each step, meaning the participant makes ten choices to ascertain their preference level. The tree displaying all possible paths of symmetry value pairs that the participant can follow is shown in Figure 3.1. Once all the choices have been made, the participant is presented with a (new) final sculpture, which displays the participant's chosen value calculated as the mean of the outcomes of the two sets of choices. At the end, the participant is asked to confirm how well the final sculpture matched their preference. This acts as a check to determine whether the participants' preference was captured.

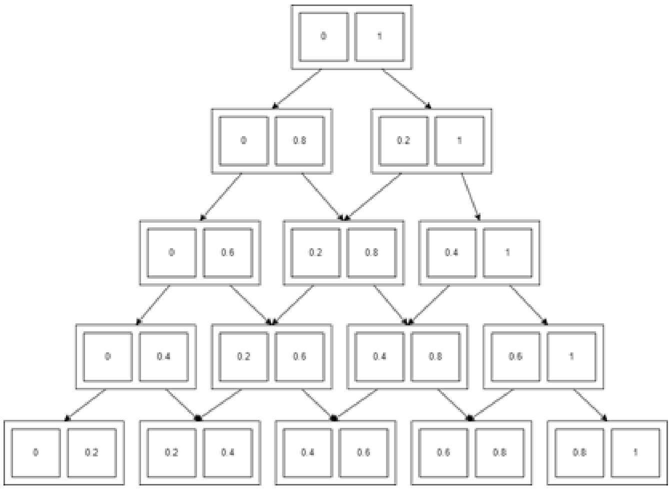


Figure 3.1: Decision tree showing all paths which could be followed by participants

3.5 Making the environment engaging

Due to the potential for the VRE to be reused across multiple experiments, the environment needs to be as engaging as possible to allow it to cater for a wide range of ages or if the same person participates in multiple experiments. To achieve this, a mine cart theme was chosen, where the participant would travel along a track through a series of caves using a simple control mechanism which does not require a VR controller, where a participant can lean in four directions to control their avatar. The pitch (forward/backwards) controls the speed of the cart and the Yaw (left/right) allows the participant to choose a sculpture or select an answer to a question, depending on the stage in the experiment.

Upon starting the experiment, the participant is given an introduction to what the experiment would consist of, i.e. choosing between sets of sculptures, followed by training on operating the cart, making sculpture choices and answering questions within the system. Once the training has been completed, the participant is then free to travel through the caves at their own pace. The participant completes the experiment by travelling through two types of cave, scenic caves and sculpture caves in an alternating manner. The scenic caves were intended to help keep the VR environment engaging by presenting a unique, fun theme in each one, Figure 3.3a. After going through the scenic cave, the participant found themselves in the sculpture cave and were presented with a choice between two sculptures,

Figure 3.2a, and asked to select the one they preferred before continuing along the track. To align with the theme of the VRE, the participant was not allowed to interact with the sculptures, e.g. by rotating them and only had the time from the entrance of the cave to when a decision was made to view them. This does provide the ability to view the sculptures from multiple angles, but not from all angles. This helps keep the controls simple, increasing the range of participants for which this system is suitable.

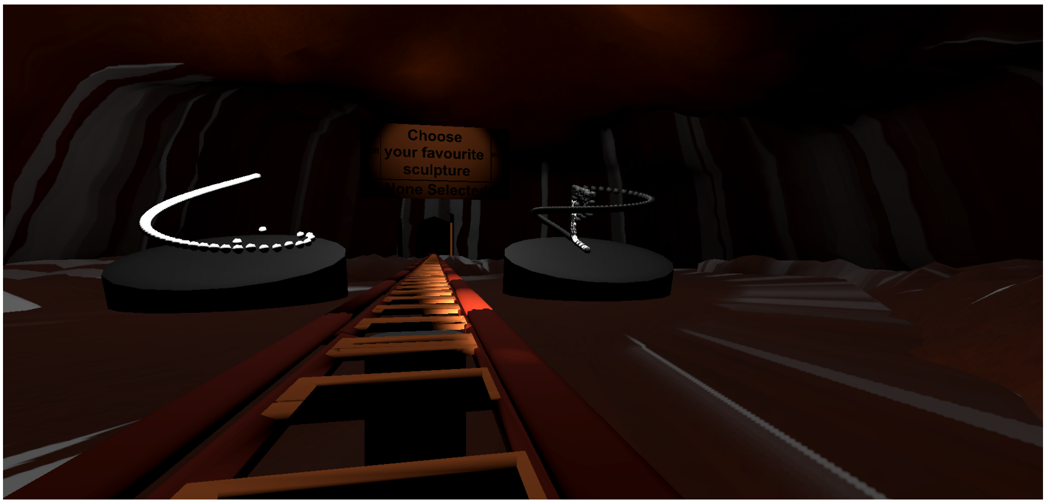
After the two converged values had been obtained from the two sets of choices, their average was taken and used to select and display the final sculpture to the participant, Figure 3.2b. Participants are given time to view the final sculpture and instructed to make a judgement on how much they liked it and are informed that the next cave would contain a set of questions about their experience. In this final cave, Figure 3.3b, several questions are presented to the participant covering different aspects about their time in the VR environment, such as whether they liked the final sculpture, providing ratings for the best and worst sculpture they had seen throughout the experience and whether they experienced any motion sickness. These questions can be tailored to the specific use of the VRE, this set was specifically created for this project.

3.6 Process validation

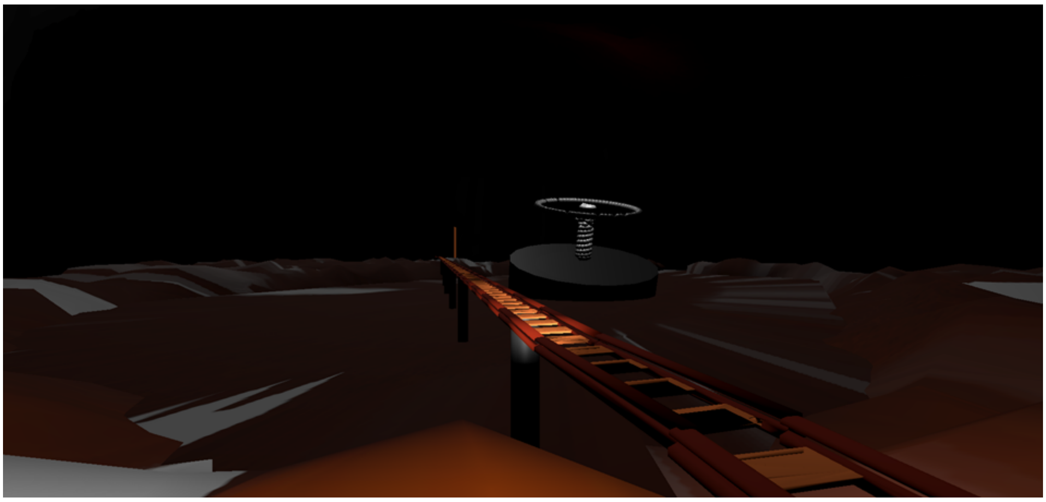
To validate this process for extracting aesthetic preferences from participants, a study on symmetry preference was conducted in Chapter 7.

Looking at the ratings for the final sculptures within the system gives an indication of whether the process was valid. It is argued in this thesis that if the system was able to correctly determine a preferred value for the investigated aesthetic aspect (in this case, symmetry), then the final sculptures will be rated higher than others. The results collected, Figure 3.4, show that whilst the final sculptures were not consistently rated as highly as the best sculptures the participant had seen during their time within the Virtual Environment, there is a clear bias towards higher ratings than the worst sculptures the participant had seen. This suggests that the process presented here did have a positive effect on how much the participant liked the sculptures, reflecting that a preferred level of symmetry was obtained for each participant.

Confidence can be had in the values collected for the symmetry preferences due to how



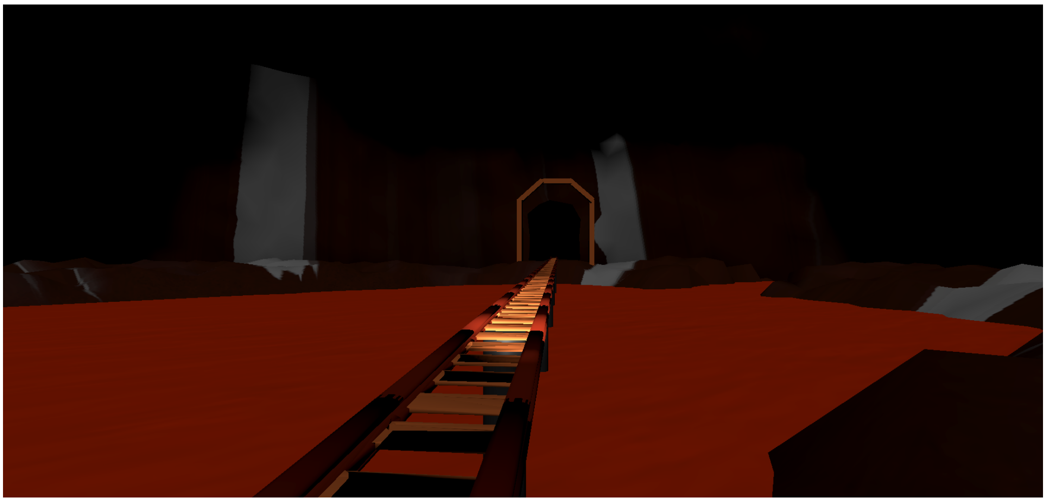
(a) Sculpture choice cave, presenting a pair of sculptures for the participant to chose between



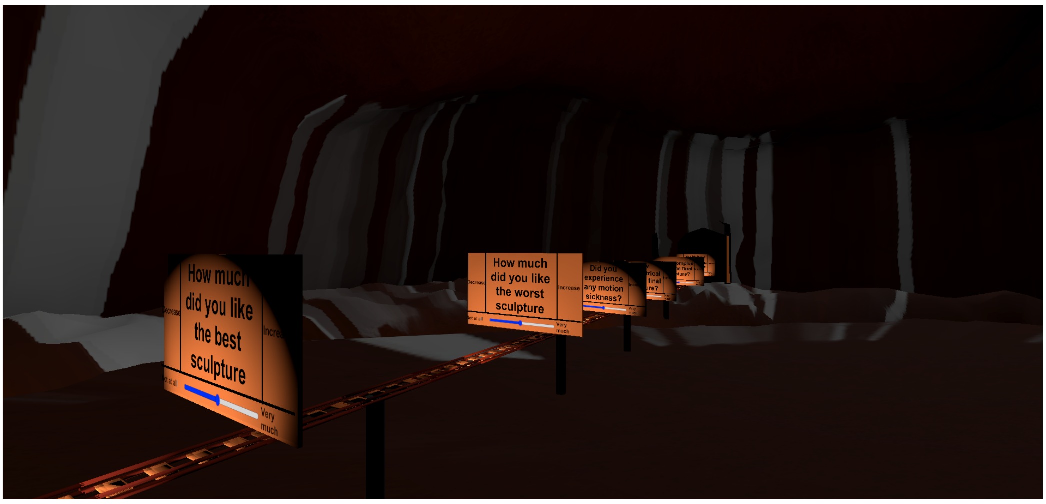
(b) Final sculpture Cave, showing an example of the sculpture generated from the participants choices

consistent the participants were in their choices between the two repetitions. The difference between the levels chosen by the participant at each stage of the experiment was found on average to be 0.267, indicating that the participants were consistent across their choices, being only a single step apart from the symmetry value obtained in the previous set of choices. The data for each group of participants are shown in Figure 3.5b.

These two statistics indicate that the process is a suitable method for extracting individual preferences of the aspects of aesthetic judgement, as study participants were able to select the level of the aesthetic aspects they preferred. As this is information that people may struggle to articulate, this system becomes a powerful tool for understanding aesthetic



(a) Scenic cave example with a lava theme



(b) Question Cave, presenting the final questionnaire to the participant

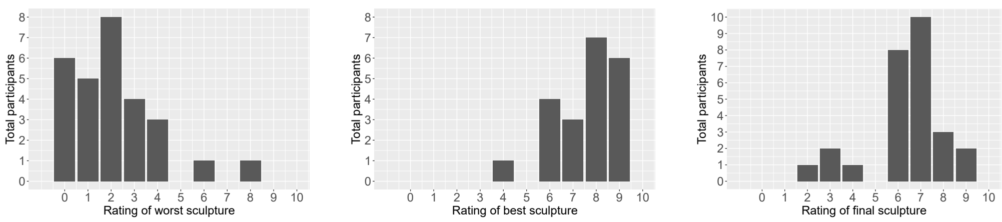
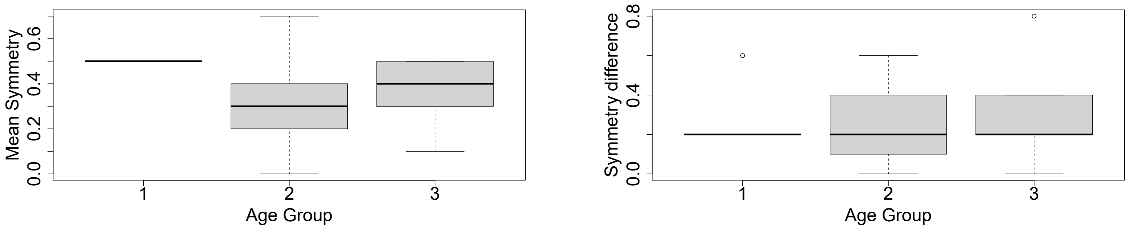


Figure 3.4: Overall aesthetic preference ratings provided for the sculptures in the system: (left) Ratings for the worst sculpture, (middle) Ratings for the best sculpture, (right) Ratings for the final sculpture

preferences.

The proposed process worked well at finding an individual's preferences, having only a



(a) **(b)**

Figure 3.5: a) Distribution of mean symmetry levels per age group, b) difference in the chosen symmetry levels by age group

slight difference in rating between the best and final sculptures shown to the participants representing their ideal value of the aesthetic aspect.

The use of VR itself contributes to the success of the application, through informal feedback provided about the app, it was suggested that the experience was engaging, encouraging the participant to use the app for longer. This, coupled with the success of measuring preferences, means that the system could be expanded to cover other facets of aesthetic judgement with few changes, such as using a different selection of items to view, which would allow the investigation of other aspects which may influence the symmetry judgement. This approach may also be an excellent candidate for measuring and assessing more complex variations for aesthetic preference. Including these additional aspects is expected to improve the viability of the final sculpture as it would be considering a more complete view of aesthetic judgement. The environment did not provide a full range of motion around the sculptures, this was deliberately excluded in order to fulfil the goal of making the environment engaging and entertaining whilst keeping the control mechanisms as simple as possible and limiting the potential for motion sickness, which other movement types such as flying can cause.

The exciting implication of this system as a whole is that it allows aesthetic preferences to be established without requiring people to explicitly verbalise or rate their preferences, circumventing the barrier of art being notoriously difficult to judge. This allows data to be collected from any individual and allows the system to be used without any skill or knowledge requirements, truly allowing judgement details to be collected from everyone.

Overall, the system was successful in its goal of extracting individual aesthetic preferences of 3D sculptures. Allowing this greater understanding of aesthetic judgement leads to more accurate methods of measuring aesthetic preference, enabling more accurate user modelling to be implemented and giving a greater ability to generate aesthetically pleasing sculptures.

3.7 Using VR to investigate generalised Aesthetic judgement details

As mentioned in Chapter 1, there are two sides to aesthetic judgement: the individual level details and the collective level details. These aspects inform different parts of aesthetic judgement, and both are equally important. Similar to collecting individual details, a novel VRE is presented in this section, which is designed to collect generalised aesthetic judgement details.

This VR environment places the participant in a room resembling a warehouse art gallery with the sculptures on a platform along the walls, see Figure 3.6. The centre of the room formed the main "play area", giving the participant ample space where they were able to inspect, move, group and assign aesthetic tags to each of the sculptures. The proposed layout of the VR environment was carefully designed to not give any perceived indication of preference or other attribute to any of the presented sculptures. All sculptures were spaced equally around the participant in a pre-determined random order. This avoids unintended bias when the participant decides which groups to place the sculptures in. In addition, it also avoids influencing the results by taking into account details such as the display of the sculptures affecting how the participant perceives them. The process of using a museum-like environment in VR has been shown to influence different factors such as the perception and visualisation of the presented objects, along with the recognition of these items are artworks (Gartus et al., 2015; Lopes, 2020; Kabassi and Maravelakis, 2015; Carliner, 2003; Post, 2021). This is an important consideration for the tasks being presented to the participants. The Gallery VR environment presents minimal distractions to the participant, allowing them to focus entirely on the presented artwork, similar in process to theories utilised for Modern Art galleries (Brieber et al., 2015).

The main task involved the participant choosing their own groups from the provided sculptures based on the minimal instruction of grouping similar sculptures, with the number and size of the groups being determined by the participant. Detecting the groups was

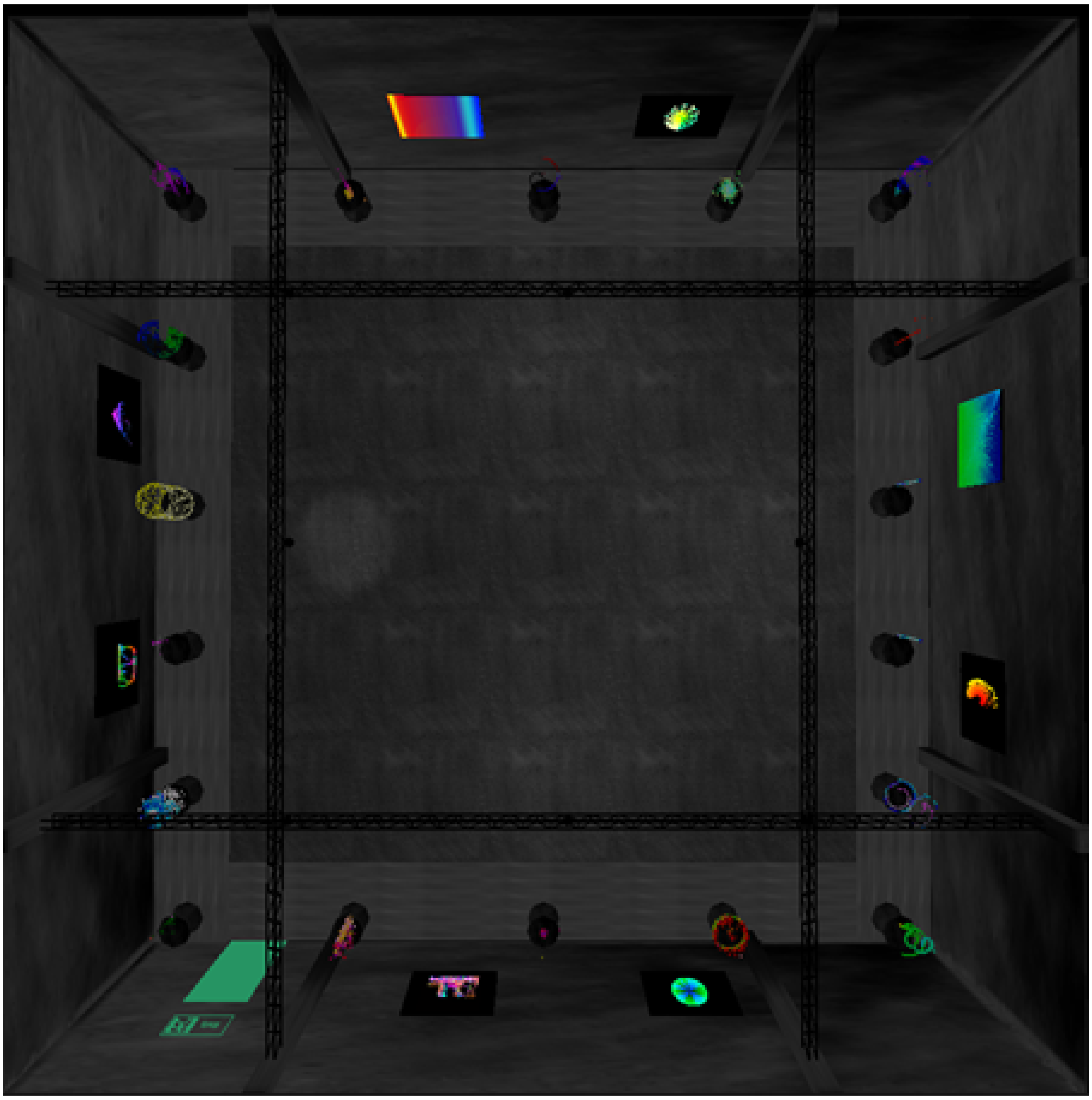


Figure 3.6: Top-down view showing sculpture starting positions and main play area

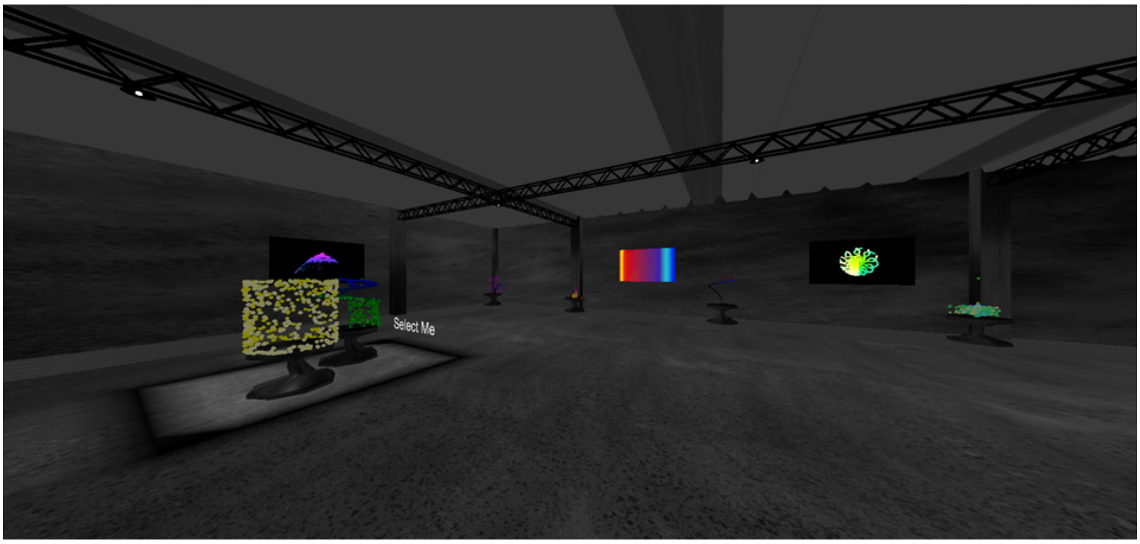


Figure 3.7: Environment with a detected region

handled by implementing the mean shift clustering algorithm (Derpanis, 2005), which ran in the background, identifying the clusters the participant had placed the sculptures within, allowing the real-time detection of any number of groups formed of any number of sculptures. Each cluster was visually indicated to the participant by highlighting the floor around each group, see Figure 3.7, allowing the group to be amended by adding or removing sculptures as necessary. As soon as a group was detected, tags could be chosen from the available list and applied to all sculptures within that group.

To allow the participant to complete the main task, four main interaction techniques were implemented to allow the participant to interact with the environment: Translation, Flying, Scaling and Rotation. The interaction techniques used both available controllers to allow the performing of multiple interactions simultaneously, for example flying whilst moving a sculpture. The translation interaction technique allowed the participant to move a sculpture from one position to another along the floor, flying allowed the participant to travel around the environment, moving the participant forward or backward at a consistent rate. The movement speed was chosen to avoid potential motion sickness and to allow the participant to accurately perform multiple actions whilst flying. Rotation allowed the participant to twist the sculpture, providing a shortcut to allow a sculpture to be seen from all sides without having to specifically fly around each side or physically move whilst wearing the HMD. This encouraged the participants to view the sculptures from each angle with minimal effort, meaning they could make an informed decision about the descriptive tags.

The final interaction technique was scaling, this allowed the participant to scale themselves down to a tenth of their original size and be placed at the base of a specific sculpture. The participants were able to then fly around the sculptures, providing a novel experience, uniquely available within a Virtual Reality environment to explore the sculpture from a new perspective, travelling in and out of any part of the sculpture. This afforded the ability to inspect all aspects of the sculpture closely and helped to provide the participant with an exciting VR experience.

To run the experiments using this VRE, it was substituted into stage four of the registration process, with the VRE downloaded from the Oculus Quest store [1].

3.7.1 Process

The VR app starts by asking the participant to enter their unique code, which is subsequently validated to ensure the participant has registered and confirmed acceptance of all essential documents and consent. This code is also used for the anonymous logging of the data collected within the environment. The next stage, comprised of multiple training tasks, enabled the participants to independently learn the four control mechanisms implemented within the environment and how to complete the environment-specific task. This was achieved by providing a task for each technique, such as collecting pellets, moving blocks into a goal, chasing ghosts and scaling to collect a key to open a door. The training section of the environment was critical due to the remote nature of running the experiments, so training videos and descriptions were available to the participant to help them learn the techniques without anyone being present. Once the training is completed, the participant completes the main experiment task.

Once the training had been completed, the participant completed the main experiment task and was shown the 18 sculptures placed around the edge of the room. The participants were given as long as they wanted to explore and interact with the sculptures whilst creating groups and assigning tags. This approach was chosen as it allowed the participant to change their mind about earlier groupings based on their changing opinions on the sculptures and the tags being assigned.

The approach also allowed the collection of data, which can be used to help guide the auto-generation of artwork, where common aspects, e.g. order and complexity, can

[1] https://www.oculus.com/experiences/quest/4632259860201837

be considered less important and therefore be less important when attempting to create aesthetically pleasing artwork.

Once the participant had confirmed they had finished the main task, they were prompted to return to the registration website, where they were requested to complete a final questionnaire asking questions relating to their experience. Upon completion of the questionnaire, the participants were given the uninstall instructions and a completion code, which allowed them to mark their submission on the Prolific platform as complete and collect the payment for participation, which was set as a minimum of £10 per hour, and based on the estimated time to complete the entire process.

Less validation of the gallery environment is needed as it is similar in design to other experimental paradigms. The key expectation of the environment is that it facilitates the participant's completion of all given tasks, which can be measured in terms of a few metrics: (1) whether all participants completed the given tasks, (2) how long the participant chose to stay within the environment and (3) whether the environment caused any motion sickness. This last metric is of high importance as a low value indicates that participants may participate for longer and that the environment is more appealing for novice VR users. The data is collected from the experiment presented in Chapter 5, where aesthetic terms are used to judge 3D sculptures and their relative importance using this VRE are investigated and Chapter 7 where the environment was used to validate the formalised measures, confirming whether the measures were capable of creating sculptures which displayed the intended attributes when considered by a human.

The average time the 58 participants spent in the environment was 27 minutes. Within this time, there were no sculptures which were left unassigned, indicating that the participants were able to complete the tasks. The average rating for motion sickness experiences was 2.3 out of 10 across all participants, once again indicating that the environment is fit for purpose.

Whilst this is a very brief and shallow look at validating the environment, looking at whether the environment is fit for use, the full validation of whether the environment is suitable for collecting generalised aesthetic data will be shown in Chapters 5 and 7.

3.8 Conclusion

Two VR systems and the online registration portal have been presented within this chapter, combining to answer Question 1 and providing a framework enabling the remote collection of individual and collective aesthetic judgement details. Each VR system is capable of being reused; e.g. the mine cart system can be used to investigate multiple aspects by amending the sculptures presented to the participant, and the Gallery VRE can be amended by updating the sculptures and available tags, ensuring that these systems are flexible and can be used across a wide range of contexts.

All systems have been custom-developed to achieve their stated goals including those related to the Covid-19 pandemic, including the fully remote conducting of the experiment. The two VREs represent novel approaches to the problems of investigating individual and collective aesthetic judgement. Both have been shown as capable of achieving the stated goals, i.e. the Mine cart VRE enabled a range of participants to consistently select their preferred level of symmetry (the chosen aesthetic aspect), which in turn generally improved the aesthetic rating of the sculptures shown to the participant when compared to the other sculptures. The gallery VRE provides a highly interactive environment where participants can use multiple interaction techniques to complete complex tasks, such as grouping sculptures based on the different aspects they display or by a value of a specified aspect. This allows generalised data to be collected from the participants, including positive overall validation of the formalised measures or details on which aspects contribute to the aesthetic judgement of virtual sculpture.

Due to the success of these three systems, it becomes possible to collect the data the data to provide answers for the final two research questions posed in this project, which relies on both individual and collective data to be collected and analysed.

Chapter 4

Creating a versatile 3D artwork generation algorithm

The work presented in this chapter has been adapted from the following publications:
E. Easton, A. Ekárt, and U. Bernardet. Axial Generation: A Concretism-Inspired Method for Synthesizing Highly Varied Artworks. In *10th International Conference on Artificial Intelligence in Music, Sound, Art and Design, EvoMUSART 2021 held as Part of EvoStar 2021*, pages 115–130. Springer, 2021.
E. Easton, A. Ekárt, and U. Bernardet. Axial Generation: Mixing Colour and Shapes to Automatically Form Diverse Digital Sculptures. *SN Computer Science*, 3(6):505, 2022.

A system is often only as strong as its weakest point, even with the systems described in the previous chapter, it would still not be possible to investigate aesthetic judgement of sculptures without having a sculpture generation method which is capable of creating sculptures that display a wide range of aesthetic values, one such generation method is presented here. Each implementation of 3D virtual sculptures introduces constraints on what needs to be produced and how to meet the unique needs of the implementation. For example, trying to re-create sculptures in the style of Roy Lichenstein would require different parameters and generation methods than for creating generalised sculptures. To generate 3D sculptures which are suitable for the aesthetic judgement investigation in this study, multiple constraints need to be adhered to. Some are common across all virtual sculptures, for example, due to their use in a dynamic VR environment, the generation process must be quick to run, and the sculptures must be renderable with high performance. In addition to these common constraints, the more specific constraints introduced by this project need to be considered, such as the generated sculptures should have the following attributes: they must be visually varied, exhibit a range of values across multiple aesthetics and should fit within the Concretism style. This chapter introduces the underlying mechanism for creating new sculptures, outlines a 3D generation algorithm, the Axial Generation Process (AGP), which meets all of the constraints and demonstrates its validity to the project by comparing it against other existing 2D and 3D generation methods, providing a complete answer to Question 2.

4.1 Evolutionary Computation

Evolutionary Computation (EC) is a wide-ranging term which covers multiple variant algorithms which find their root in the biological process of evolution (Eiben and Smith, 2003). All variants have found significant success in solving optimisation-based problems such as the Travelling Salesperson Problem (TSP) (Grefenstette et al., 1985). EC describes four core processes: Initialisation: the creation of an initial group of items, known as the population, which represent potential solutions to the problem being solved; Assessment: rating each member of the population by a scoring system known as a fitness function; Crossover: combining two or more members of the population to generate two or more distinct child items; Mutation: randomly altering a population member to provide additional variety.

Two of the major variants within EC are Genetic Algorithms, introduced by Holland (1975), which involves all of the core processes and focuses on using string values as the representation and Evolutionary Programming (Fogel et al., 1966), which use functions as the base representation, and the flavour of EC that is used within this project.

As with problems such as TSP, the generation and evaluation of artwork can also be considered an optimisation problem, where an image or sculpture can be measured in terms of aesthetic value, which can theoretically be maximised.

Within the assessment phase, the score is used to discriminate between different members of the population, identifying better solutions to the problem being solved, within the context of art generation, this often refers to the value of a piece of art, which, as discussed in Chapter 2, is a highly varied concept which contains, but is not limited to aesthetic value. Understanding and formally describing aesthetics and the aesthetic value of a piece of art is notoriously difficult, making assessing the items complex. Due to this, the assessment of the generated items is often delegated to a human user, who use their innate sense of aesthetic judgement to choose which items have higher value and should, therefore, be taken into the population for the next generation. This approach is known as Interactive Evolutionary Computation (IEC) and was introduced by Takagi (2001), and the approach mimicked when creating examples of the aesthetic model in Chapter 8.

4.2 Limitations of using EC

EC techniques do have disadvantages, such as premature convergence, however, these can be mitigated. Premature convergence can be a big issue when generating artwork as it affects the novelty of generated items. Novelty is often cited as a significant influence on the aesthetic judgement of an item (Boden et al., 2004; Vinhas et al., 2016). It can be mitigated using techniques such as maintaining a list of created individuals and then using this to filter subsequent population items.

One crucial aspect is how effective evolution is as a method of generating sculptures, evolution is known to be an effective and versatile process when creating biological life, however, the potential can be limited. This could be related to how progress is being measured. One school of thought states that evolution is greedy and only allows solutions which are effective at a single moment in time, even if another weaker solution would be more effective in the future, effectively limiting the usefulness of Evolutionary Search. This is often used as an example to explain why artificial concepts are not found in nature. The process of random mutation is relied on to escape local maxima, but this process can be ineffective and slow, especially where the optimal solution is changeable.

There are a few methods of working around these limitations, the first works by relying on the standard evolutionary processes to circumnavigate these issues. The issues can also be mitigated further by introducing the concept of novelty, where the algorithm prefers solutions that have not been seen before, a process used throughout this project. This is an effective solution and allows the navigation of different paths through the search space.

4.2.1 Benefits of using EC

Even with these concerns, Evolutionary Search Algorithms offer many benefits, especially in areas with a less defined goal, where they implicitly add constraints onto the process, ensuring that the process and output are valid. Within this project, to ensure that the data structure used was valid, a few different rules were used to prevent node configurations. This included looking at the arity of the function which formed a particular node, if the function required a single operand, for example, then only a single node would be created beneath it. This was managed for all arities of all available functions to ensure that each expression tree would be a valid function and produce a valid response.

The flexibility of EC algorithms is another aspect which is beneficial for generating artwork. The fitness assessment, which guides the process through all potential solutions, can rely on almost any rating method, either static or dynamic, allowing combination with other techniques to achieve the given task. Whilst the formal assessment of aesthetics is complex, some success has been found through training Neural Networks to estimate aesthetic ratings, which could be applied to an EC system or, as previously mentioned, the assessment can be delegated to a human user. This is in contrast to other AI techniques used with artistic intentions, which are often only capable of performing a single task and changing this task can be a complex and lengthy process. The ability to use one or more assessment measures of varying types is highly beneficial to this project and another reason why EC is utilised.

In addition, no ”correct” solution exists within the context of artwork, making the generation of multiple items a highly desirable property. Another important aspect is that, unlike other AI techniques, EC algorithms do not require large amounts of high-quality training data, making them easier to set up and potentially more ethical. Due to these reasons, Evolutionary Computation has been heavily applied to artwork generation, quickly becoming one of the main techniques utilised and, for the same reasons, applied within this project, requiring all standard EC processes to be implemented.

4.2.2 Evolving artwork

In order to create an Evolutionary Program to generate artwork, several decisions need to be made: how the items are going to represented internally in the program (genotype), how the items will be converted from the internal representation to a visible representation (phenotype), how the fitness of the items will be calculated, how the genotype of the item can be mutated, how crossover can be implemented, how items should be selected for crossover and which values to use for the varying parameters within the algorithm.

4.2.3 Representing artwork

A commonly used data representation for the genotype within Evo-Art systems is expressions, typically represented as Trees (den Heijer and Eiben, 2010b; Colton et al., 2011; Ekárt et al., 2011; Li et al., 2012; Machado et al., 2015; Mills, 2016; Vinhas et al., 2016). These are representations of mathematical functions which control one or more aspects of a generated item. These offer the benefits of being easy to implement across most programming languages, capable of generating a wide range of output and having well-known methods of implementing the standard EC processes. Expression trees are comprised of nodes representing arithmetic functions and terminal nodes representing constant values. These values usually are either random values or index values from the generation process. Each of the functions for a node requires a specified number of arguments, a value known as its arity, which defines the number of child nodes that each node requires. Increasing the size of the tree increases the complexity of the expression, which translates through to the generated item.

4.2.4 Combining maths and art

The list of functions available to create the expression trees within this system are shown in Appendix A. This includes additional functions compared to other systems within the Evo-Art field, which often only use basic arithmetic functions. The choice of functions was made due to three reasons. As the more complex functions are combinations of the simpler arithmetic functions, e.g. Matyas is a combination of constant values, addition and multiplication, these functions can be generated using the simpler functions, however, it would require larger expression trees for the same functions to be re-created. In the interest of the versatility of output, these functions have been added. The functions are also capable of generating interesting 3D shapes, which are translated into the items generated by the algorithm, some examples of the 3D forms the expressions create are shown in Figure 4.1. Including the functions also allows for a more expansive genotypic search space, again helping with the versatility and novelty of the system and allowing for complexity to be selectively added into the system at the genotype level. In addition to the functions used for the non-terminal nodes (Eiben and Smith, 2015), the expression trees also utilise specific terminal node values as the basis for building the expressions. These terminal nodes use random constants; a common example is using the X or Y coordinates of a 2D image. For this project, whilst it is possible to assign a coordinate system, as there is no universal coordinate system for 3D sculptures, implementing this would cause confusion and a lack of reproducibility. Instead, two values were used: the index, which represented the zero-based index of the element being placed in 3D space and an ephemeral constant between zero and one (Koza, 1990), as shown in Figure 4.2.

4.2.5 Evolving representations

When creating the initial population, the ramped half-and-half method was chosen to initialise the expression trees for each individual. This technique involves creating the population in two halves: the first half are created using the maximum allowed depth, and the second half are created to be half the size of the maximum depth (Koza, 1994). This introduces a good variety of items to start the evolutionary processes and balances many factors, such as the complexity of items, which can form a good starting place when searching for items with a varying level of aesthetic value. The exact process of how expression trees are used to create artwork (phenotype) for this project is explained in detail in Section 4.4, and an example of the expression trees which will be used is shown in Figure 4.2.

To perform cross-over, population members need to be selected as the parents, one of the most common methods of performing this selection is tournament selection (Eiben and Smith, 2003, pp. 86). This approach chooses a specified number of population members (k) at random, the fitness of each is obtained, and the item which scores the highest (across one or more measures) is chosen as the winner of the tournament and selected by the system. This process is repeated to obtain the multiple parents required. Figure 4.3 depicts how cross-over is implemented using expression trees, with the cross representing the cross-over point. This is a randomly chosen number based on the smallest number of nodes within the two trees chosen for cross-over. The node at this index and all sub-nodes are separated from both trees, the first child is then created by using the main tree of the first parent except the chosen node and sub-nodes, which comes from the second parent. The second child is the reversal of this process, taking the main tree from the second parent and the sub-tree from the first. Through this method, two new population members are obtained, which replace their parents in the population.

Mutation is a process where population members are chosen based on a specified rate, and

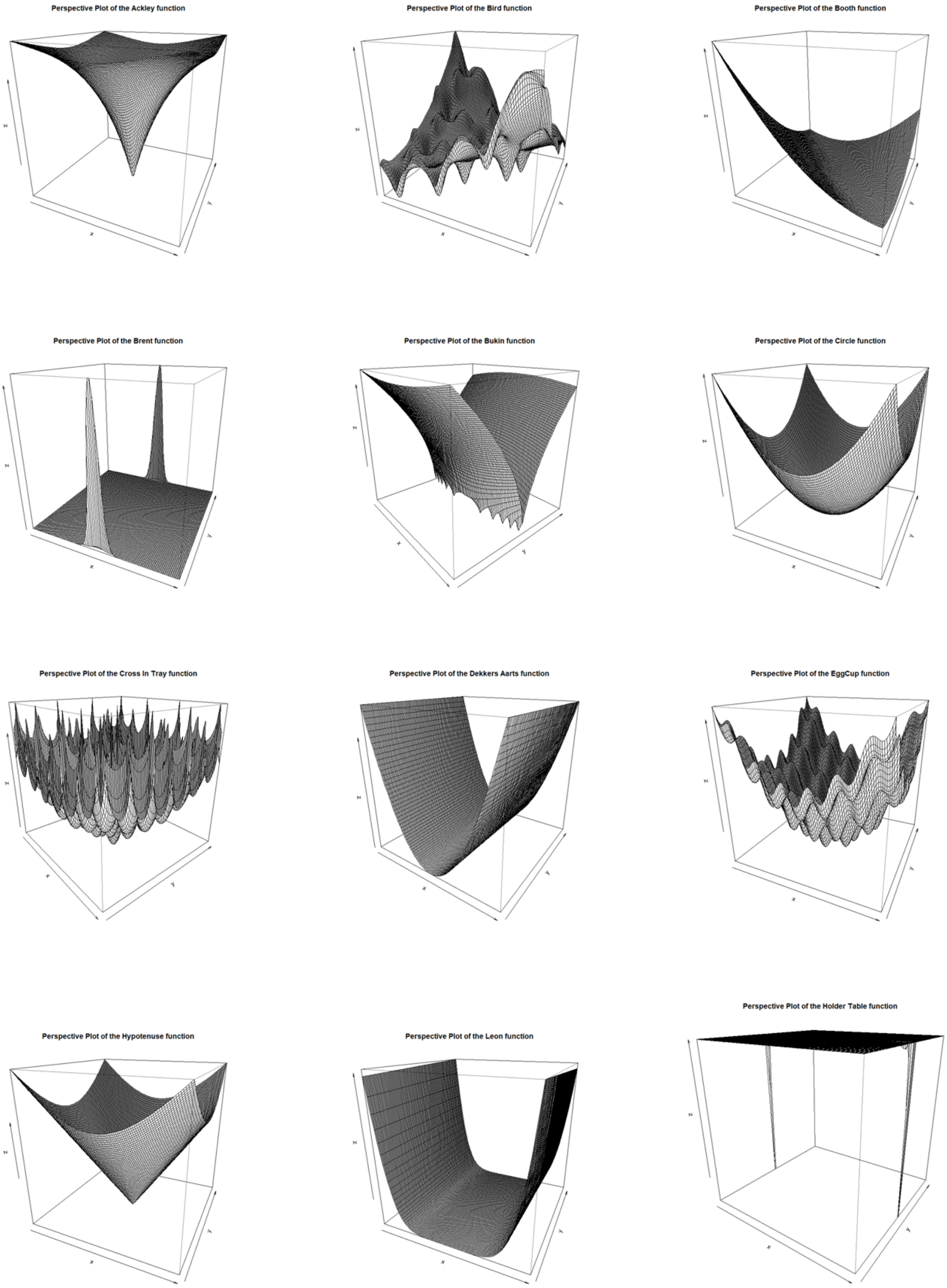


Figure 4.1: Example plots generated using the additional functions available within the Expression Trees

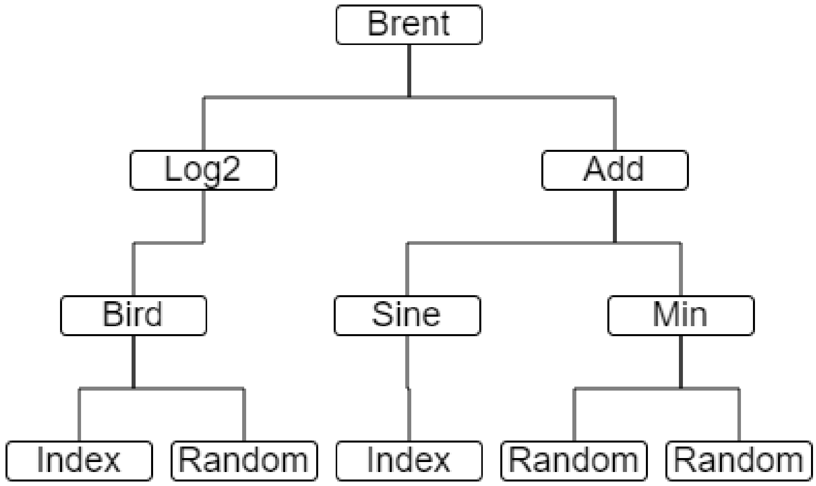


Figure 4.2: Example Expression Tree

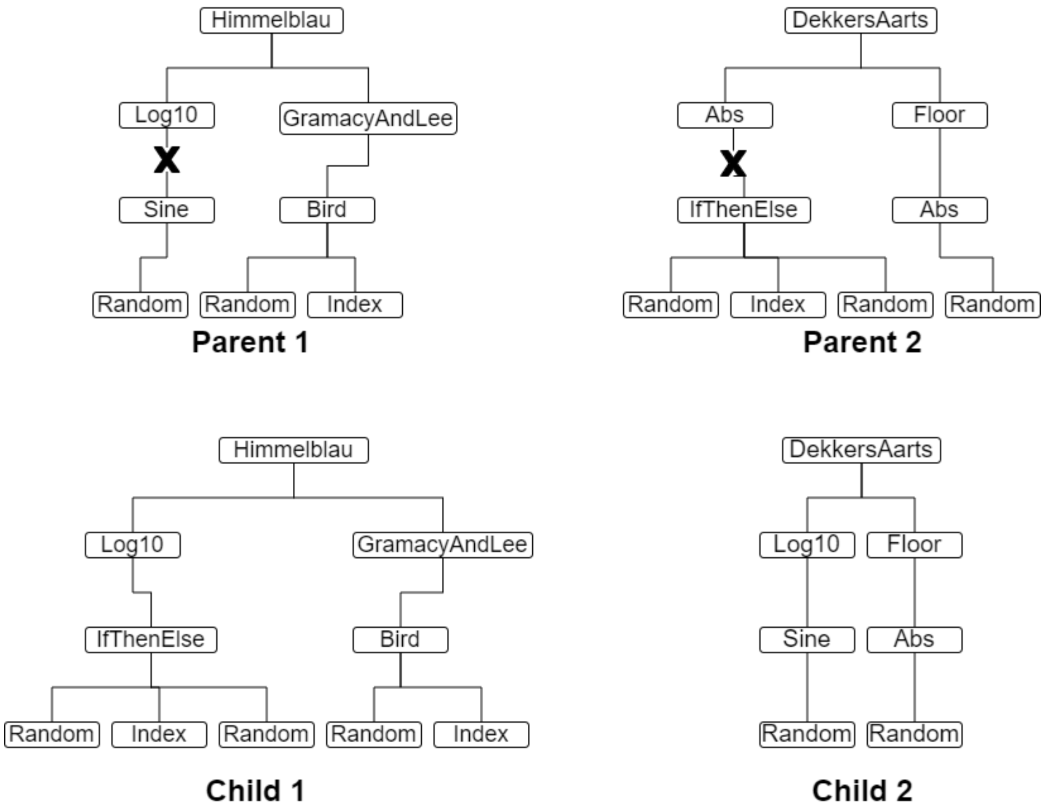


Figure 4.3: Example cross-over between two trees

then an aspect is changed about their genotype. Unless specified elsewhere, the mutation rate set for all experiments in this project is 0.7, meaning 70 percent of the population's expression trees will be mutated in some manner before the following fitness assessment. This improves the ability of the algorithm to escape local maxima and introduce a high level of randomness into the search, which is beneficial due to the focus being on creating a variety of items rather than looking for a specific solution. However, conversely, the value is not set any higher to avoid situations where the algorithm is too randomised and is unable to converge on any particular solution, especially when detailing the individual model in Chapter 8 will be beneficial. The chosen value is set to try and find a balance between these two factors. For this project, three methods of mutation were chosen: Adding a node into the tree, where a terminal node is replaced by a new expression, chosen at random from the available expressions, increasing the tree's depth; Deleting a random node (and all sub-nodes) from a tree and finally replacing a random node within the tree with another expression, these three methods of mutation are displayed in figure 4.4.

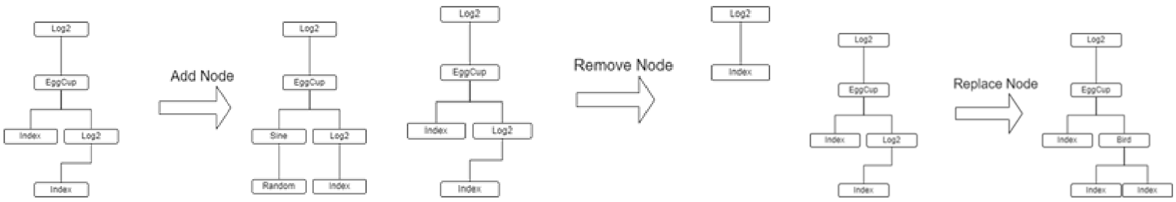


Figure 4.4: Each type of mutation

As this project requires high flexibility to cover all potential usage scenarios, the EC implementation is based on several additional parameters as described in Table 4.1.

4.2.6 Additional benefits

The basis of EC allows it to be run in different fashions: manually, semi-automatically and fully automatically. To run the system manually, the results are presented to the user after each generation, the user can choose their favourites, which are then used to create the population for the next generation. For semi-automatic evaluation, a specific number (greater than 1) of generations are run before returning the results to the user. Finally, fully automatic continuously completes generations until a specified value is found or if the percentage change between the average value of each item in the population is lower than a specified amount (indicating that the system has reached convergence and is unlikely to progress any further).

After each generation, the results are tracked, including various statistics and the changes made. This allows information to be generated, such as the genealogy of an algorithm run, as shown in Figure 4.5, which items were altered by mutation and, in addition, allows the user to return to a previous generation.

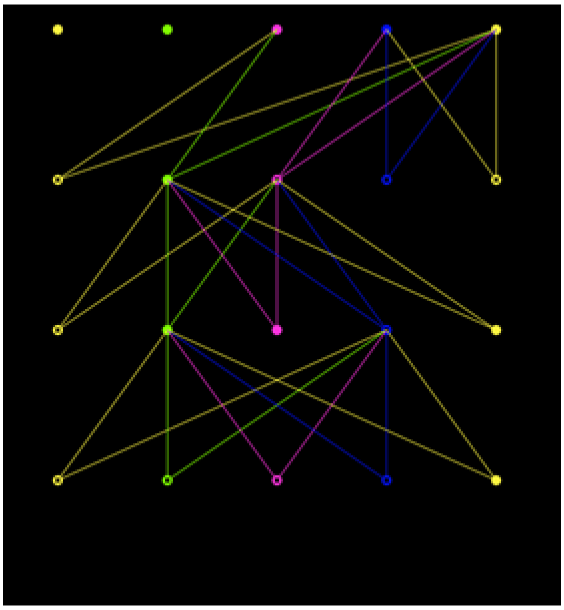


Figure 4.5: Example genealogy of an algorithm run

The utilisation of the EC system described here forms the basis of answering the project questions with enough flexibility to handle all potential situations and configurations which may arise.

4.3 Potential sculpture rendering methods

Four methods were initially considered for defining the shape, where each one determined a set of points which could then be rendered into a sculpture. Each of these methods is presented here with a brief evaluation of their suitability.

4.3.1 Growing out from an origin

One standard method of creating physical sculptures is combining multiple items to create the desired form and feel. A similar method for generating virtual sculptures creates sculptures by defining an origin point and then adding cubes (of 1x1m dimensions) depending on the values calculated from a set of expression trees. The process used three expression trees, which calculated an integer value for each of the X, Y and Z axes. The value obtained from each of the expression trees defined how many cubes should be placed along the specified axis, for example, if the process obtained a value of 1 for X, 2 for Y, and 1 for Z, then the system would place one cube along the X axis, two cubes along the Y axis and one along the Z axis. Once the initial set of values were obtained, the furthest point from the origin was selected, and the values were generated again. This process was repeated until a set number of cubes had been placed to create the sculpture. This algorithm created some interesting sculptures and had the benefit of always creating items which reassembled solid objects. The basis of expression trees allowed it to be easily used within an Evolutionary context, however, the main limitation of this process was that the items generated were constrained along the axes. This generally resulted in the sculptures looking very flat and platform-like, such as the sculpture shown in figure 4.6. This may mean that it is a method which is more suited for investigating architectural structures rather than aesthetic judgement of sculptures without significant effort to refine the process.

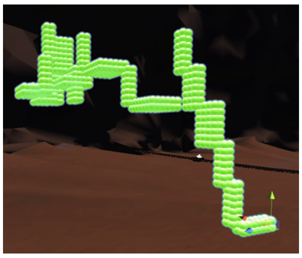


Figure 4.6: An example sculpture created by growing items from the origin

4.3.2 Adding space between items

Space is considered an important aspect of sculpting, and this process was designed to utilise the space between items in the sculpture to define patterns and shapes. This process can be considered similar to the method undertaken by a sculptor, removing the excess material from a block until the desired form is revealed. To generate sculptures with this method, initially, a 3D zone was defined by specifying a set number of metres along each axis, e.g. 5m along the x-axis, 5m along the y-axis and 5m along the z-axis, creating a 5x5x5 cube in which the sculpture could be generated. This defined space shape was filled by fitting in as many geometric shapes as possible. The system then maintained three expression trees, which represented the space that should be around each shape along each axis, e.g. if the three values generated for the first item were 1,3 and 5, this would mean that the first cube should not have anything within 1m along the x-axis, 3m along the y-axis and 5m along the z-axis. The process then worked in reverse to the previous one, where the value calculated for each tree represented the number of geometric items which should be removed on the corresponding axis. For example, if the values 1, 2 and 3 were obtained, the next item on the x-axis would be removed from the sculpture, the next two items along the y-axis would be removed, and the adjoining three items along the z-axis would be removed. The process would then loop through all items, ignoring any which have been removed, until the space around all the items within the sculpture had been determined, for example, Figure 4.7. However, similar to the first method, this process was unable to create a significant set of interesting shapes which were obvious to the viewer, with some features often being hidden from view even when the sculpture was rotated, making this process less suitable than the others proposed.

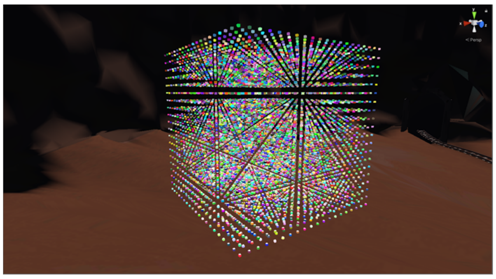


Figure 4.7: The sculpture block after items were removed

4.3.3 Adding items into space

This process was much more conventional and followed an approach used in other studies on generating 3D items (Hollingsworth and Schrum, 2019). This process maintained three expression trees, which determined the exact X, Y and Z coordinates to place a 3D object. This was the easiest method to implement and is comparable to several existing 2D as well as 3D generation methods. The sculptures mainly fit into the specified style, more often than not appearing as solid objects, and, similar to the other presented generation method, the use of expression trees allowed a straightforward application of EC Algorithms. The most attractive characteristic of this method was that this process was also more capable of generating a varied range of shapes and patterns, including curves, compared to the previous two methods. However, when compared to the other methods, the range of shapes generated were not as extensive, and the potential to extend the process was not as broad as the other methods, so this process was also discarded. An example of a sculpture generated using this method is shown in Figure 4.8.

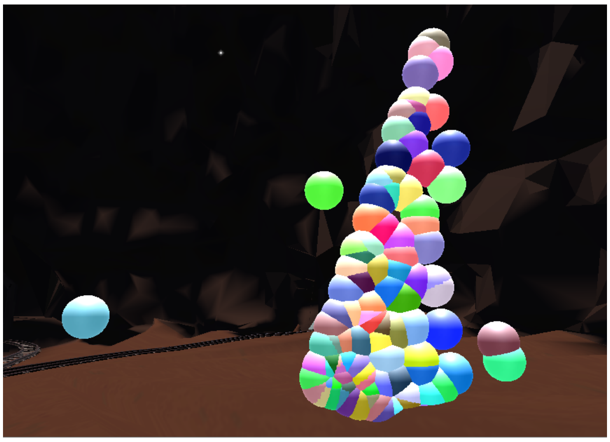


Figure 4.8: An example sculpture generated by placing items at calculated coordinates

4.4 The Axial Generation Process

The final proposed method uses three values to position geometric items around a central axis: the item's height on the axis, the angle around the axis where the item is placed and the horizontal distance the item is placed from the axis. This Axial Generation Process (AGP) maintains three expression trees which represent each of these values, for example, if the values we obtained from the expression trees were 1, 180 and -5, this item would be placed 1 unit up on the x-axis, and -5 units along the Z-axis. How the values are used to place the objects is shown in Figure 4.9. Algorithm 1 shows the process of converting the values from the expression trees into a 3D position. This process is capable of generating a wide range of visually different sculptures containing a wide range of patterns and offers several potential methods for extension, seemingly meeting the criteria outlined. However, an evaluation of how effective the process is at generating interesting content based on a variety of measures and these results are then compared to two existing systems in both 2D and 3D: a particle-based generation system (Colton et al., 2011; Easton et al., 2019) and a pixel/voxel-based generation system (Hollingsworth and Schrum, 2019; Colton, 2012; Ekárt et al., 2011; Li et al., 2012), to ensure the AGP objectively met the criteria.

This system is built around producing artwork within the Concretism style, either abstract images in two dimensions or digital sculptures in three dimensions and is capable of generating a wide range of visually varied artefacts of both types. As defined by Tate: "Concrete art is abstract art that is entirely free of any basis in observed reality and that has no symbolic meaning" (Tate, 2017). Concretism was introduced in 1930 by Theo van

Doesburg, who released the *Manifesto of Concrete Art* (Van Doesburgh, 1930), outlining the vision for Concrete art through six well-defined tenets describing how and why Concrete art should be created. Due to the style being disinterested in describing natural objects or sentiments, it fits nicely with the auto-generated nature of Evo-art systems. The similarities continued to grow through further investigation of the manifesto. The third tenet describes that concrete art "must be entirely built up with purely plastic elements, namely surfaces and colours" (Van Doesburgh, 1930) resembles the use of basic shapes to generate artwork, allowing a connection between Concretism and the AGP to be established.

Initially, the process focused solely on the composition of the created items and did not involve colour or different shapes to make up the items. However, this approach limited how the AGP could be assessed, as no formal measures relying on colour could be applied. The assessment was instead restricted to using only a small number of form-based 3D measures. This left several unanswered questions relating to the ability of the AGP to perform well across a wider range of measures and how the AGP compares to other generation systems at producing stimulating items. Answering these questions required the AGP to be substantially extended to include calculating shape and background colour, allowing the content to be created using additional shapes, allowing these shapes to change size dynamically and allowing the content to be rendered in two dimensions from multiple viewpoints with the extended process presented within this chapter. Through implementing these changes, the process is not only capable of generating a more comprehensive range of visually varied items but can also utilise a wider range of formalised measures to assess the output, allowing the AGP to be directly compared against other generation systems.

4.4.1 Choosing a style

The abstraction of real-world phenomena is a key element of art. The level of abstraction depends on many factors, such as the subject of the artwork, the medium the artwork uses and the style it is created in. The style of a piece of art often transcends across both 2D and 3D media and heavily influences the subject. A style can be thought of as the combination of aspects that can be used to identify and categorise pieces of art. These categories are used as a method of identifying and comparing both artists and their work. For paintings and sculptures, these styles are often well defined, Van Gogh is known as a Post-Impressionist, Andy Warhol for Pop Art, Banksy for Street Art, while Rodin and

Brâncuşi are known as modern sculptors. This categorisation process offers many benefits, such as making comparisons between artwork and artists more straightforward, for example, comparing Georgette Chen to Vincent Van Gogh is easier than to Banksy. Styles can also help people understand and describe their aesthetic preferences.

In art generation systems, the style in which the items are created is often not well defined due to the system being minimally constrained, searching through all possible permutations of colours on a canvas, meaning any style of artwork may be produced. A person's aesthetic preferences only represent a minute portion of the items which can potentially be created. The size of the search space compared to the region containing an individual's preferences makes finding aesthetically pleasing artwork difficult.

The process of generating art is often similar for creating 2D or 3D items, where the subject, real or not, is abstracted, losing superfluous details to align with the presented constraints for the medium. 3D media holds several benefits for representing real-world objects over 2D, such as allowing a closer representation of the item being modelled through the use of the extra dimension losing less information in the process. Modelling objects using computers has been around since their conception, with the unique ability of computers allowing the process to be much quicker and more versatile than manually creating physical items. It also holds the major benefit over physical media that the laws of physics do not necessarily need to be adhered to. As technology improves, so does the capacity to create items, leading to many technologically advanced artefacts, for example, the latest computer games, which often push the boundaries of what is possible, allowing a proliferation in the interest surrounding 3D art within Virtual Environments.

The Axial Generation Process (AGP) creates items by combining a specified number of geometric shapes within set bounds, three values are used to position the shapes around a central axis: the height, the angle and the radius. Figure 4.9 shows how these three values are used to represent a precise location in 3D space. In order to calculate these three positions for each shape, the process maintains three expression trees. The expression trees used differ slightly from other implementations, which commonly use position details and instead only use two possible values for the terminal nodes, the index of the shape being placed and a random value between 0 and 1, as the position data for each item is not known until after the expression has been evaluated. The process, shown in Algorithm 1, is used to position the shapes for 2D and 3D items.

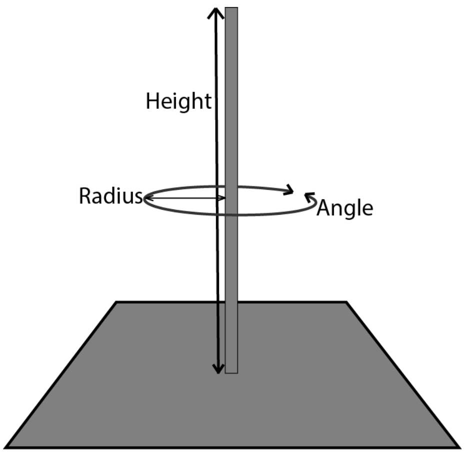


Figure 4.9: The use of the height, the radius and the angle to place the geometric shapes around the axis

---

Algorithm 1 Positioning process

$totalItems \leftarrow n$
$heightTree \leftarrow Tree$
$angleTree \leftarrow Tree$
$radiusTree \leftarrow Tree$
$Points[] \leftarrow [totalItems]$
for $k \leftarrow 1$ to $totalItems$ **do**
  $y \leftarrow heightTree.getValue(k)$
  $angle \leftarrow angleTree.getValue(k)$
  $radius \leftarrow radiusTree.getValue(k)$
  $x \leftarrow radius \times Sin(angle)$
  $z \leftarrow radius \times Cos(angle)$
  $Points[k] \leftarrow (x, y, z)$
end for

Eight parameters are used to control this process, described in Table 4.1. These include the number of shapes to be added, the bounds the items should be placed within and how the items should be rendered. Amending these parameters can change the visual structure of the generated items, as shown in Figure 4.10, which contains items generated using only the initial process.

The initial process served as a proof-of-concept, showing that the method can generate a wide range of items, both in 2D and 3D, making it a good candidate for further study. It also provides many areas that can be extended to further increase the range and diversity of items that can be generated whilst also extending the range of measures that can be used to assess the created content, ensuring that the project contributions are met. The

Total Items	The number of items which will be placed in the artwork
Maximum Height	The furthest distance away from the base the items can be placed
Maximum Radius	The furthest distance from the axis the items can be placed
Is 3D	Indicates whether 2D or 3D items should be generated
Element Size	The radius of the circle to be placed on the canvas (2D only)
Background Colour	The canvas colour to be used when drawing the image (2D only)
Canvas Width	The width of the bounds to render the items in (2D only)
Canvas Height	The height of the bounds to render the items in (2D only)

Table 4.1: Parameters used by the AGP

Calculate Item Colour	Whether the colour of the shapes should be calculated
Item Colour	The default colour used for the geometric shapes
Standardise Item Colour	Whether all shapes should have the same colour

Table 4.2: Parameters used by the AGP relating to calculating the colour

base method has been extended with four additional features detailed below: (1) colour, (2) dynamic sizing, (3) different base shapes, and (4) multiple viewpoints.

4.4.2 Colour

Many factors contribute to making an aesthetic judgement, one ubiquitous factor is the colour contained within the items. Colour is a major component of art and is the focus of many existing generation systems. Due to its importance to aesthetic judgement, it also forms the basis of a large number of existing measures which attempt to assess images (Acebo and Sbert, 2005; Matkovic et al., 2005; Ross et al., 2006). These look at aspects such as the relative positioning of the colours or which gradients are present. Due to the stalwart nature of colour as a part of aesthetic judgement, adding colour to the generated sculptures and images was necessary. Another benefit of utilising colour within the generation process is the potential for increasing the range of visually different artefacts generated by the AGP. As noted by Arnheim (2009), shape and form are highly dependent on differing areas of colour and brightness, and the inclusion of colour enhances visual differentiation. This is an important aspect as the sculptures generated using the AGP will be utilised to demonstrate a large area of the aesthetic concept space as colour is helpful when identifying the boundaries and features of objects, which improves the ability for different sculptures to be identified over one another. This will greatly improve the potential for the AGP to generate more visually diverse artwork. The colour calculations rely on three expression trees representing the R, G and B components of each shape and additional parameters to allow greater control over how the AGP handles the new components, as shown in Table 4.2. Figure 4.11 shows examples of axial artwork created in full colour.

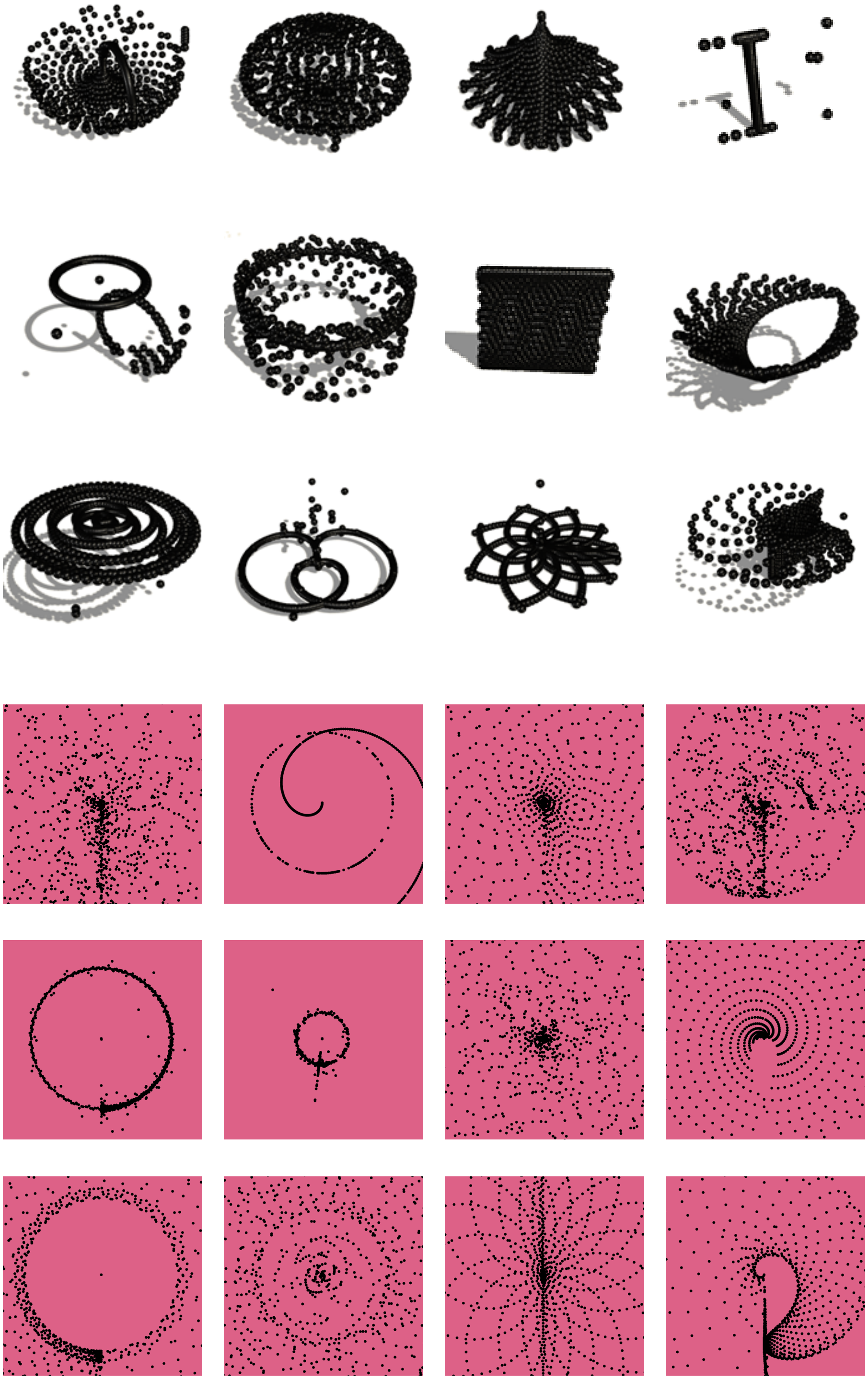


Figure 4.10: A variety of generated items using the base AGP

As well as item colour, the background colour is another important feature. It is also a significant factor when assessing 2D artwork, with the brain amending its response to seeing a colour based on the background it is presented against (Zeki, 2002). It can be

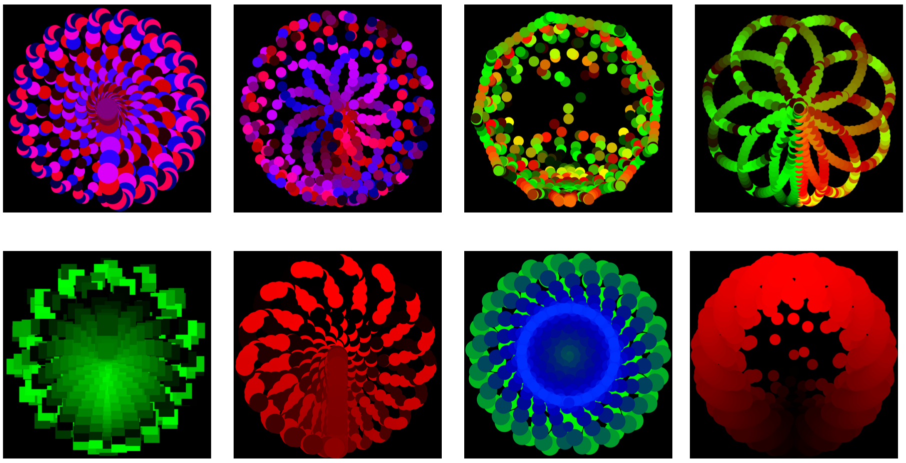


Figure 4.11: A variety of colour images generated using the AGP

Calculate Background Colour	Whether the background colour should be calculated

Table 4.3: Parameters used by the AGP relating to calculating colour

used to emphasise or hide different parts of the form of the artwork, which contributes to the expression of emotional content, dictating the importance of including the background colour in the AGP. The background colour is only implemented for 2D items, as it would be inappropriate to implement in 3D. The environment a 3D item is viewed in is context-specific and based on how the sculptures are being used, it also includes complex considerations such as lighting position, making it out of scope for this project. The background colour feature is demonstrated in Figure 4.12, where the same forms are shown against different backgrounds. The images in Figures 4.12a and 4.12e emphasise the form of the generated shape and the remaining images mask it, for example, through having the background colour match the colour of some of the shapes being placed, as seen in figures 4.12f, 4.12g and 4.12h or through using contrasting colours in the background. Figure 4.12d has a yellow background, making the items on the bottom left of the form not as noticeable as in image 4.12a. Similarly to shape colour, the background colour uses three expression trees representing the R, G and B values. It relies on one additional parameter shown in Table 4.3 and the background colour parameter from the base process.

4.4.3 Dynamic sizing

The AGP tries to focus on both the form and colour of the generated items. Part of creating different forms of sculptures depends on the primitive shapes placed within the

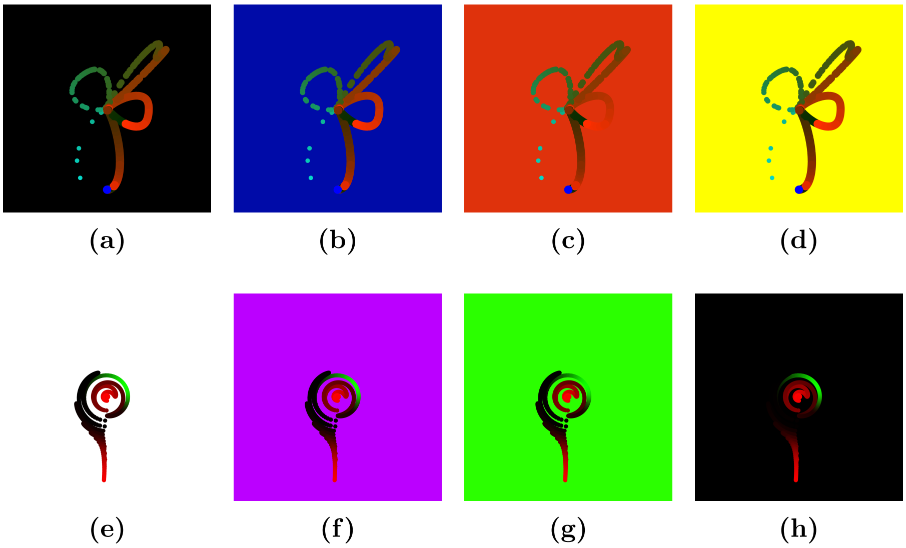


Figure 4.12: 2D images with various backgrounds generated using the AGP

Maximum Element Size	The maximum width or height a shape can have
Minimum Element Size	The minimum width or height a shape can have
Standardise Scale	Whether a shape should have the same scale across all axes

Table 4.4: Parameters used by the AGP relating to calculating colour

output. Within the prototype process, these shapes were limited to a single size, revealing another extension that can be implemented to further increase the range of forms the system can generate. This aspect was amended in two ways, the first is to introduce the dynamic sizing of the individual shapes within a specified maximum and minimum value. This uses three further expression trees, which calculate a value for the scale of the shape along the X, Y and Z axes. If the scale is standardised using the parameters, the scale value for the x-axis is applied across all axes, Table 4.4 shows the parameters that control the dynamic sizing. A selection of both 2D and 3D items, showing the capability of dynamic sizing, are shown in Figure 4.13, which uses standardised scaling and Figure 4.14, shows shapes to be scaled independently on all axes, indicating the influence this change can have on the final output.

4.4.4 Shapes: spheres and cubes

In addition to allowing the shapes to scale, the range of primitive shapes used to build the items was also increased to allow a sphere or a cube to be placed. These two elements were chosen as they are ubiquitous across all graphics software. To keep this feature consistent

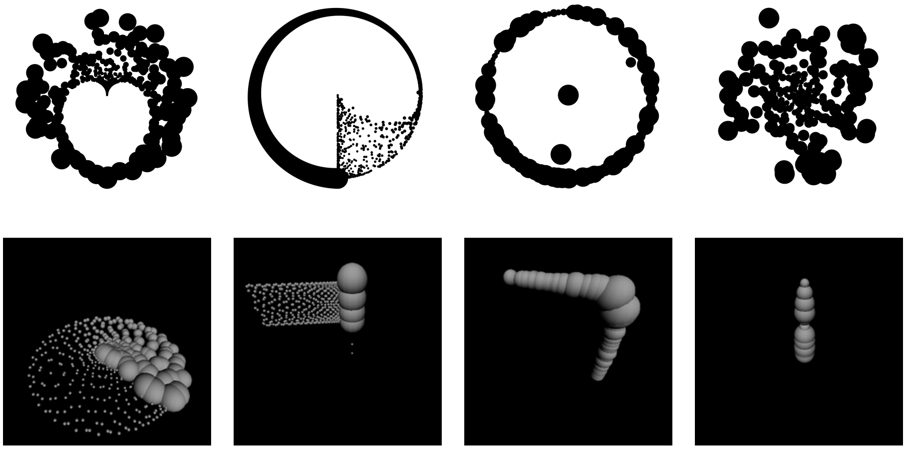


Figure 4.13: Generated content using standardised dynamic sizing of shapes

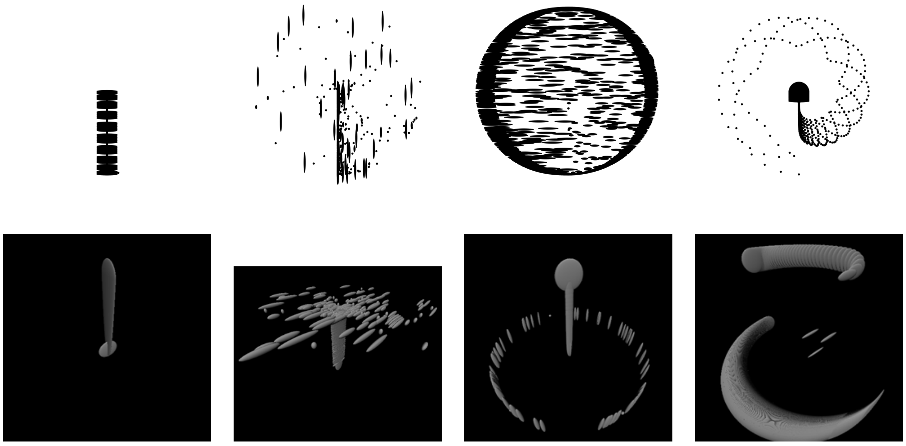


Figure 4.14: Generated content using non-standardised dynamic sizing of shapes

with the process and allow the shapes used to be dependent on the evolutionary processes, a single expression tree is used to determine which type of element is used to create the item. To keep the performance of the process stable, only a single shape per output item is allowed; however, this still considerably affects the look of the created items, as shown in Figure 4.15. As with all the form-related extensions, the automatic specification of the shape can be turned off, and a single element can be specified using the two parameters in Table 4.5.

Generate shape	Whether the shape should be automatically set
Base Geometry	The shape to use if not calculated automatically

Table 4.5: Parameters used by the AGP relating to setting the base element type

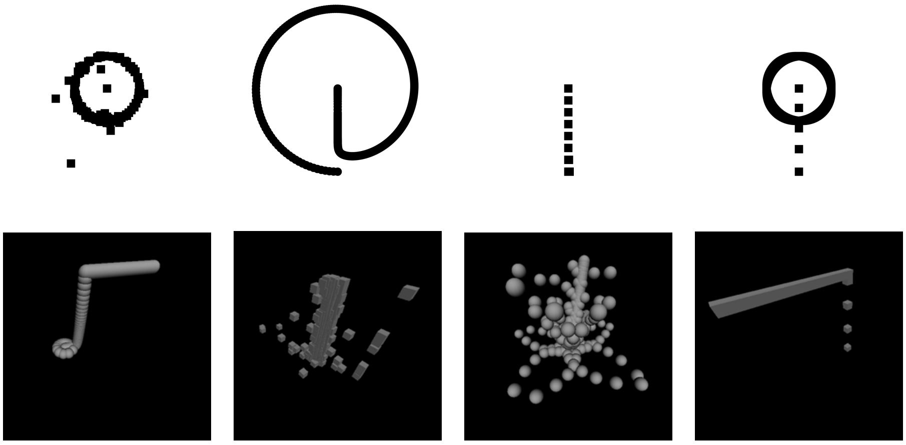


Figure 4.15: Generated content using different types of geometric element

4.4.5 Rendering the items from multiple viewpoints

The AGP abstracts the rendering process away from the generation method, allowing the same output to be rendered in either 2D or 3D. Rendering the 3D content is a simple process and involves placing the primitive shape at the specified location in 3D space. For 2D content, the rendering process works in a similar way to creating 2D art, where a 3D object is abstracted down onto a 2D plane. As both 2D and 3D content initially gets created as a 3D point cloud, this opens up a further extension for the process, allowing the 2D items to be rendered from multiple viewpoints.

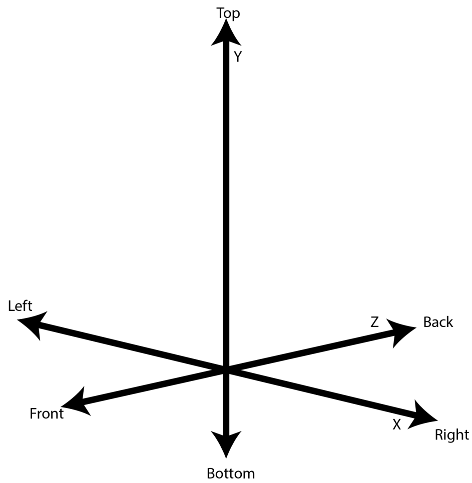
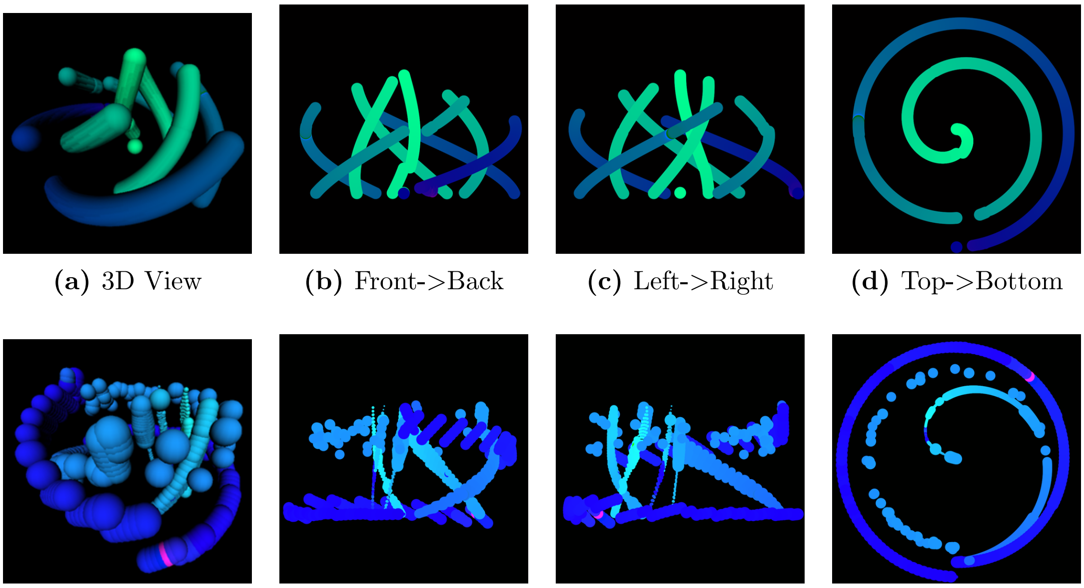


Figure 4.16: The directions for each viewpoint

Six additional viewpoints can be used to render a 2D item from the directions shown in Figure 4.16. The viewpoints are limited to six to find a balance between increasing the range of items that can be created and keeping the difficulty of using the process as low as possible. Allowing different viewpoints affords two significant benefits, which will be demonstrated further in section 4.5, where it increases the range of 2D items the system can create as well as allows 3D content to be assessed using 2D based measures, where a snapshot is taken from each viewpoint and assessed using the measure. The viewpoint is controlled by a single parameter allowing the viewpoint to be specified from the choices of Top->Bottom, Bottom->Top, Left->Right, Right->Left, Front->Back, Back->Front, figure 4.17 shows the same item from different viewpoints, indicating the importance of the viewing position on the aesthetic judgement of the sculpture. The chosen viewpoint can also dictate the overall shape of the form; viewpoints that use the Y-axis as the height create square forms rather than circular ones.



(a) 3D View **(b)** Front->Back **(c)** Left->Right **(d)** Top->Bottom

(e) 3D View **(f)** Front->Back **(g)** Left->Right **(h)** Top->Bottom

Figure 4.17: Different views generated for the same 2D item: (a) and (e) display the item in 3D and (b)-(d) and (f)-(h) the resulting views, respectively.

Figures 4.18 and 4.19 display a range of items created using either the prototype of the process (top half) or taking advantage of all the extensions (bottom half) in both 2D and 3D. As can be seen in the items which use the extended features, the forms are much less regular, this is related to the dynamic sizing of the shapes, which can allow different patterns to be expressed. The element colour also adds a significant amount of diversity where even sculptures with similar shape sizes can appear dramatically different. This highlights the effectiveness of these five extensions to the AGP, representing a significant improvement allowing the process to generate a wider range of more interesting items.

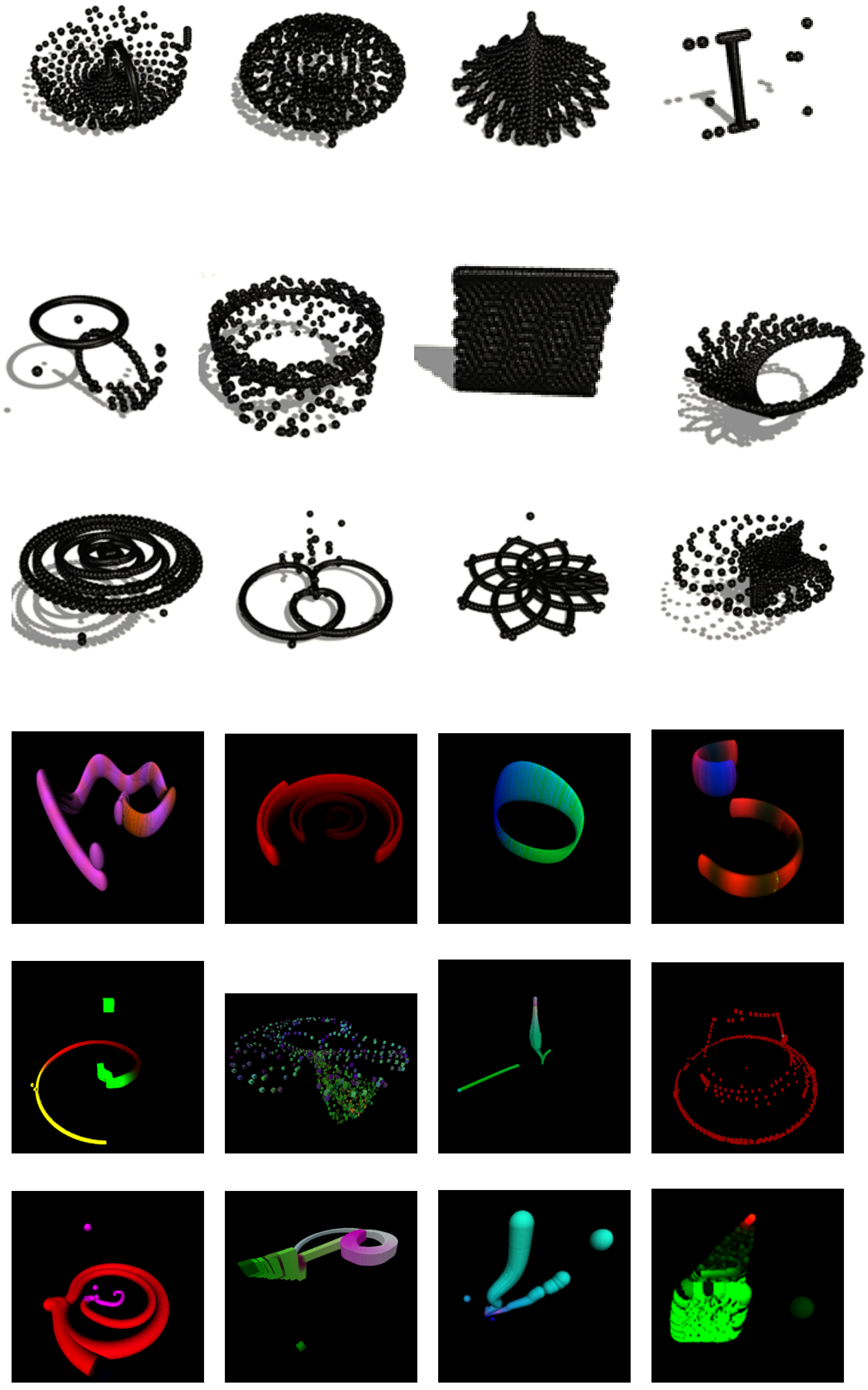


Figure 4.18: A variety of randomly generated sculptures using the AGP

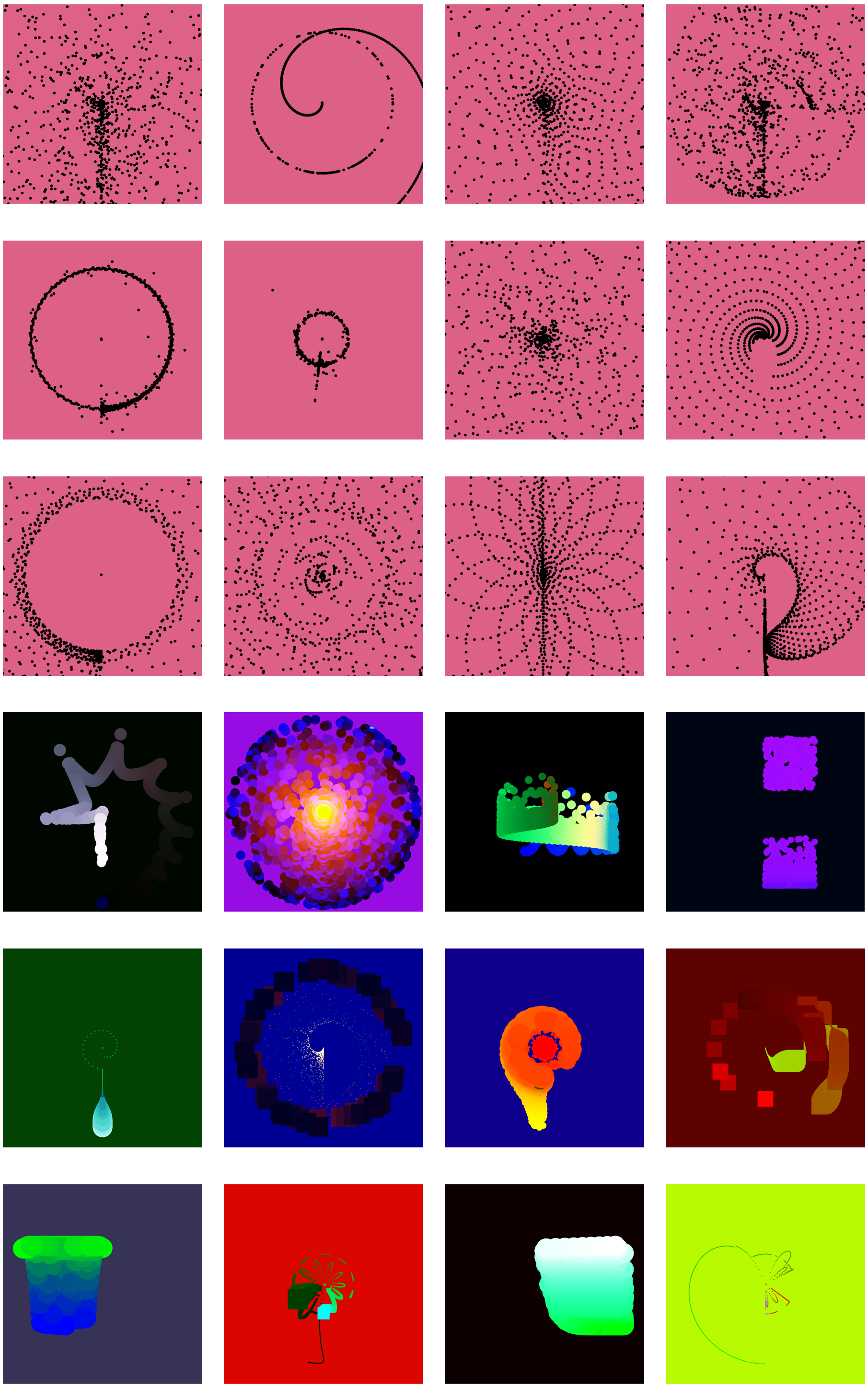


Figure 4.19: A variety of images generated using the AGP

4.5 Evaluation of the Axial Generation Process

For the AGP to be useful within an Evolutionary Art context, it must produce artefacts of interest. A common method is to classify artefacts based on their "novelty" and "value" (Boden et al., 2004; Canaan et al., 2018; Wiggins, 2006). However, novelty is a context-specific measure, depending on what items have previously been presented in the context they are viewed within and cannot be directly assessed without knowing how the output might be interpreted. Alternatively, the process can be evaluated for its versatility, i.e. the range of novel items that can be generated that show a variety of values across different measures. Many formal measures aim to describe the aesthetic appeal and can be used to evaluate these items. Due to the limitations of the prototype implementation of the AGP, the set of applicable measures to assess the output was limited. The chosen measures also had to be applicable to both 2D and 3D items.

Due to these restrictions, established measures such as the Global Contrast Factor (Matkovic et al., 2005), the colour gradients present within an image (Ross et al., 2006) or the naturalness of the image (Acebo and Sbert, 2005) were not considered. Instead, three canonical properties are used, all of which contribute to the aesthetic appeal of a generated item: the level of symmetry, complexity and compressibility were used to determine how suitable the prototype was.

Firstly, symmetry is known to be an important factor in aesthetic judgement (Tinio and Leder, 2009), the AGP places items around a central axis, often yielding highly symmetrical items. From the many available methods to calculate symmetry, the measure presented in (Ecins et al., 2017) is used as it calculates the level of symmetry for point clouds, and therefore, is applicable for both 2D and 3D items. This measure works by generating multiple candidate symmetry planes at even intervals in a sphere around the object and then calculating a score for each point reflected in each of the initial planes based on how close the nearest neighbour is to the original point after being reflected across the plane. This score is used to refine the position of each plane by calculating an error value based on how close each point is to a corresponding neighbour after being reflected in the plane, then using the Levenberg Marquardt algorithm to minimise this error. If the error is lower than a specified threshold, the item is considered symmetrical around it. The score calculated for each item is the average distance error across all detected planes for an item. With the placement of geometric items around a central axis, other forms of symmetry could have been used, such as rotational symmetry; however, due to the AGP placing items around a central axis, this may give the AGP an unfair advantage compared to other systems. Reflectional symmetry was chosen in order to avoid this issue. In addition to this, reflectional symmetry is considered one of the easiest types of symmetry to be recognised by humans. It has been linked to aesthetic appreciation (Giannouli, 2013; Thornhill and Gangestad, 1999) and used in previous experiments to generate abstract stimuli for investigating aesthetic judgement (Jacobsen and Höfel, 2002; Leder et al., 2019).

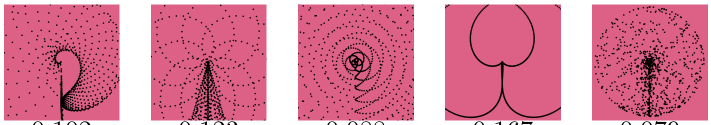
Secondly, the complexity of an item is also known to impact its aesthetic appeal (Birkhoff, 1933; Machado and Cardoso, 1998). The method employed here is based on the measure utilised by Birkhoff (1933), involving counting the vertices in an item. Instead of fully rendering each item and counting the vertices, the value was estimated by merging items with similar positions, using a distance threshold of 0.2. The remaining geometric items in the artwork are then counted and divided by the original total, the intention is that if items have many shapes in the same position, they will be less complex.

Finally, we use the Global Normalised Kolmogorov Complexity (Rigau et al., 2007), shown in expression (4.1), calculating the compressibility by encoding all points into text and then compressing this string using ZIP compression, similar to the process used in (Ekárt et al., 2011). This measure was chosen due to its ubiquity as a representation of complexity (Aaronson et al., 2014; Ekárt et al., 2011; Machado and Cardoso, 1998; Rigau et al., 2007), in addition, it has been shown that algorithmic complexity is a good proxy for estimating appreciation of images (Lakhal et al., 2020), which is an important factor for any art generation system, with the same effect expected for sculptures.

To perform the assessment, 500 items in both 2D and 3D were generated using the AGP. A selection of generated items and their calculated values are shown in Figure 4.20.

$$\frac{originalSize - compressedSize}{originalSize} \tag{4.1}$$

As shown in Figure 4.21, both 2D and 3D items were generated within a wide spread across the ranges for symmetry and complexity; however, the process could not generate items that simultaneously had a high level of complexity and symmetry. Additionally, there were no compressibility values lower than 0.6067 for 3D items and 0.6747 for 2D items, indicating that the lower end of the spectrum could not easily be reached. These two



Symmetry	0.102	0.123	0.088	0.167	0.079
Complexity	0.989	1	0.99	0.695	0.978
Compressibility	0.685	0.681	0.680	0.682	0.676

Symmetry	0.309	0.285	0.305	0.199	0.282
Complexity	0.904	0.953	0.897	0.775	0.692
Compressibility	0.616	0.609	0.623	0.658	0.633

Figure 4.20: Example output and associated values

Total Generations	30
Mutation Probability	0.7
Initialisation Method	Full
Selection Method	Tournament (k=3)
Population Size	150

Table 4.6: Evolutionary Algorithm parameters

gaps potentially represent a limitation with the base process when attempting to generate novel items, as these regions of the search space may not contain a large number of artefacts.

4.5.1 Analysing the gaps in the aesthetic search space

Four evolutionary algorithms were used to determine whether items exist within the gaps in the search space (i.e., no items with high complexity and symmetry and no items with low compressibility). These attempted to maximise both symmetry and complexity as two objectives and minimise the compressibility as a single objective for both 2D and 3D items. The evolutionary parameters are outlined in Table 4.6. A visual inspection of the resulting distributions, displayed in Figure 4.22, shows that the evolutionary optimisation was not able to fill the gaps in the measured values. The compressibility measure did not reduce further than the original values, although more items fell within the lowest band, and the symmetry and complexity could not be simultaneously improved. There are two possible explanations for this: the prototype was potentially limited and unable to produce both complex and symmetrical items, or the measures themselves do not accurately represent the details of the generated artwork, highlighting the potential need for more appropriate formalised measures to be created.

To ensure that the process was not the limiting factor, the extensions outlined above were introduced, and further analysis was performed. To determine whether the extensions

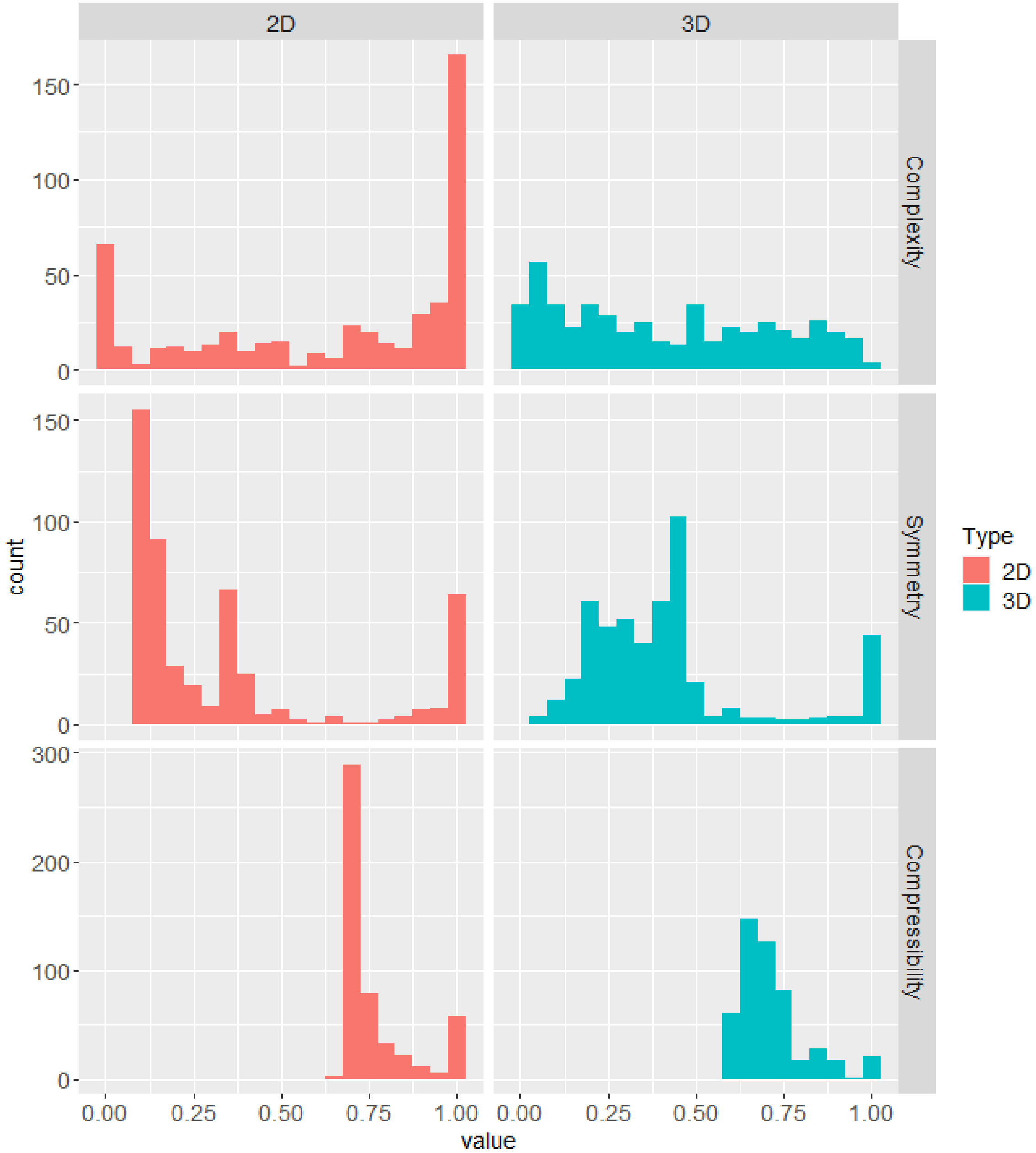


Figure 4.21: The distribution of the randomly generated items

helped to generate interesting and novel items, the generated content was compared to two well-established systems.

To compare 2D and 3D, two 2D processes were chosen that were not originally designed for 3D but only required minor changes to allow 3D items to be generated using their distinct features. The first additional method introduced is a pixel-based method that calculates the colour of each pixel within an image for 2D items. To amend this to work

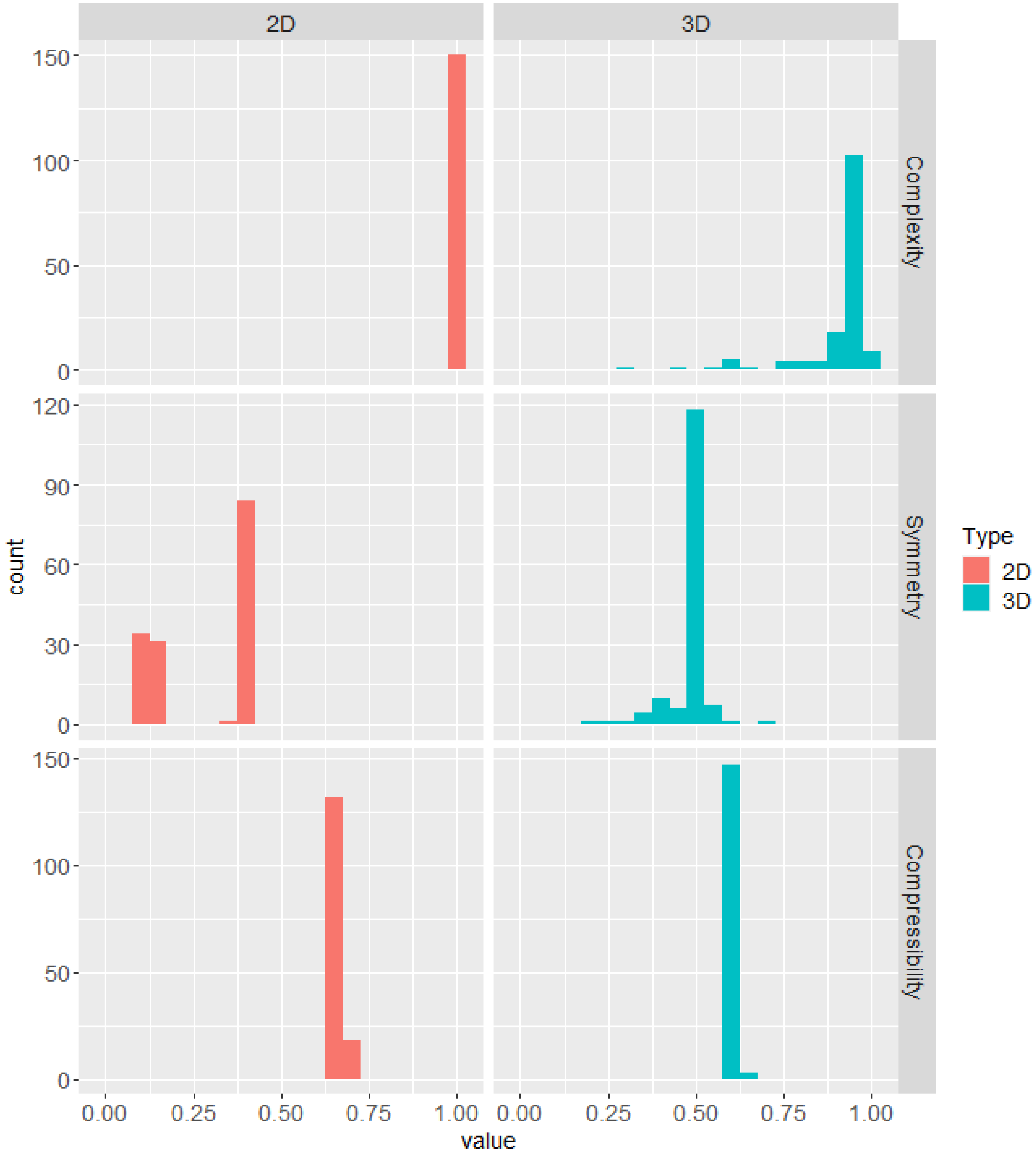


Figure 4.22: Distribution of the items after the Evolutionary run

with 3D items, an additional depth parameter was added, allowing 3D items to be created, similar to the implementation in (Hollingsworth and Schrum, 2019), where voxels are used to build sculptures with variations in form achieved by calculating the opacity of each voxel. Examples of items generated using this process in both 2D and 3D are shown in Figure 4.23.

The second method is an updated version of the process used in (Colton et al., 2011), this process tracks the position of a specified number of particles across a canvas over a

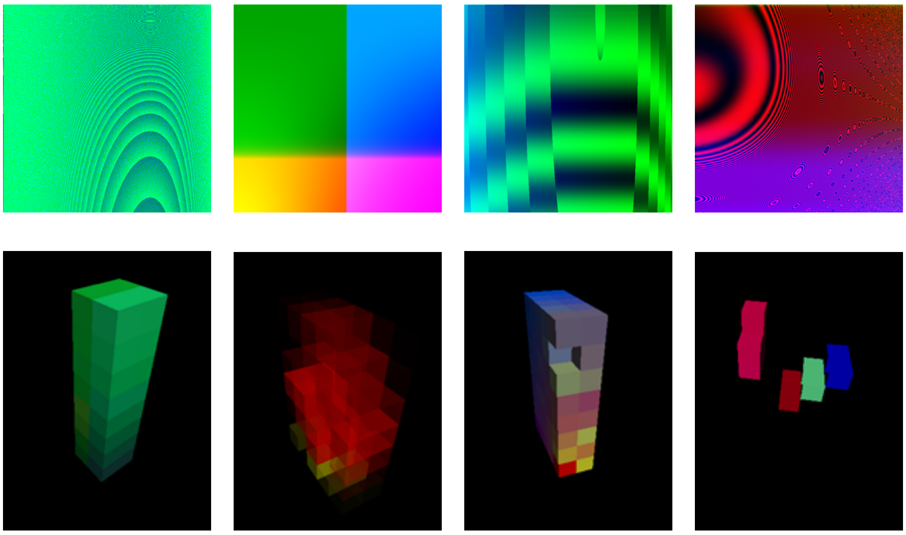


Figure 4.23: Pixel/voxel-based generated content

specified number of time steps, drawing a coloured line between the start and end position for each time step and then blurring the entire image. This process has been used within 3D before (Easton, 2018); however, this implementation focused on allowing a user to explore artwork in Virtual Reality with the 3D aspects being constant values based on the particle index. For this analysis, we have implemented the depth via a new parameter added onto the genotype, allowing each particle's path to be traced in three dimensions. Due to the extension of the process from 2D to 3D, the blurring that originally happened after each time step was no longer included as it would be overly complex for 3D items to implement in a similar manner to 2D items. Figure 4.24 shows generated content using the particle-based process.

Different measures are implemented to further assess the items generated by the AGP and the two contrasting processes. These follow the same themes as the earlier analysis, considering the symmetry, complexity and compressibility. The new measures were selected as they are sensitive to colour and form, allowing the extensions in the AGP to be taken into account and allowing a complete comparison between all three processes. For 2D items, the final generated image was used for the evaluation, considering the varying shapes and colours present in items generated through all three processes. To assess the complexity of the 2D items, Shannon's Entropy (Rigau et al., 2007) was used, which calculates how much colour information is present in the image. The symmetry was calculated using the SYM4

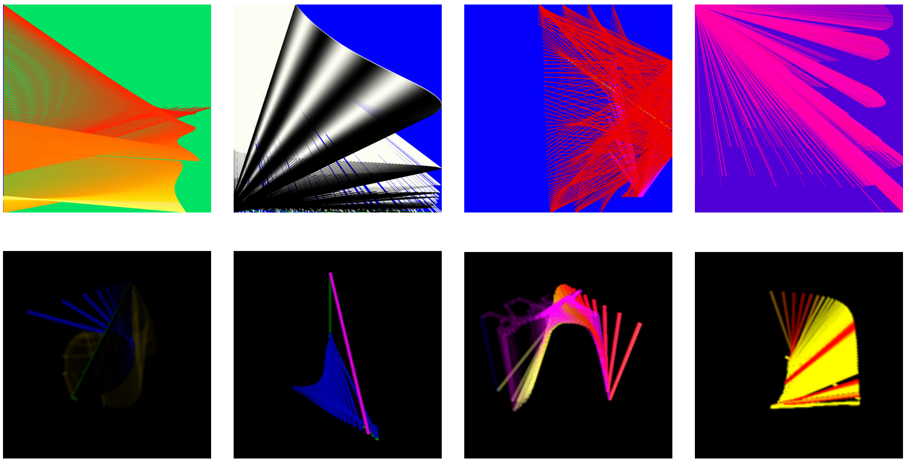


Figure 4.24: Particle-based generated content

method (den Heijer, 2012), which compares the luminance values of each pixel across four different quadrants of an image and finally, the Global Normalised Kolmogorov Complexity (Rigau et al., 2007) was calculated by first encoding the item into a bitmap and then applying ZIP compression.

The implementation of these measures differs slightly from the original, which considers every pixel in the image as having the potential to contain information. For this assessment, the background of the images was set as transparent and only pixels with a calculated colour value were used to normalise the values obtained. This decision was made for two reasons: firstly, the pixel-based systems do not have the concept of a background, making the comparison less fair between the three generation methods, and secondly, measures such as SYM4 can have their values artificially increased if large blocks of a single colour are present. 500 2D items from each process were randomly generated, using a max tree depth of 5 with the distribution of the values shown in Figure 4.25 (please note the logarithmic scale) and examples of the generated content shown in Figure 4.26, indicating the differences in blank space between each generation method and how the compressibility generally occurs on a limited range.

All three generation processes created content exhibiting a wide range of values for both the complexity and symmetry measures, with all processes showing a relatively consistent spread for both measures. The AGP produced content that generally had higher values for the symmetry, this is expected due to the rendering process of the AGP placing items around a central axis. In addition, the AGP has a slightly more uniform spread of values for the complexity measure. Compressibility is the only measure where significant differences exist between the processes. The pixel-based method covers a much wider range of values than the Particle or Axial processes. This is potentially due to differences in the generation process where both the Particle and Axial processes have a focus on generating the form of the item within the bounds of a canvas, which often leads to larger amounts of blank space being present in these images than there is within the Pixel process.

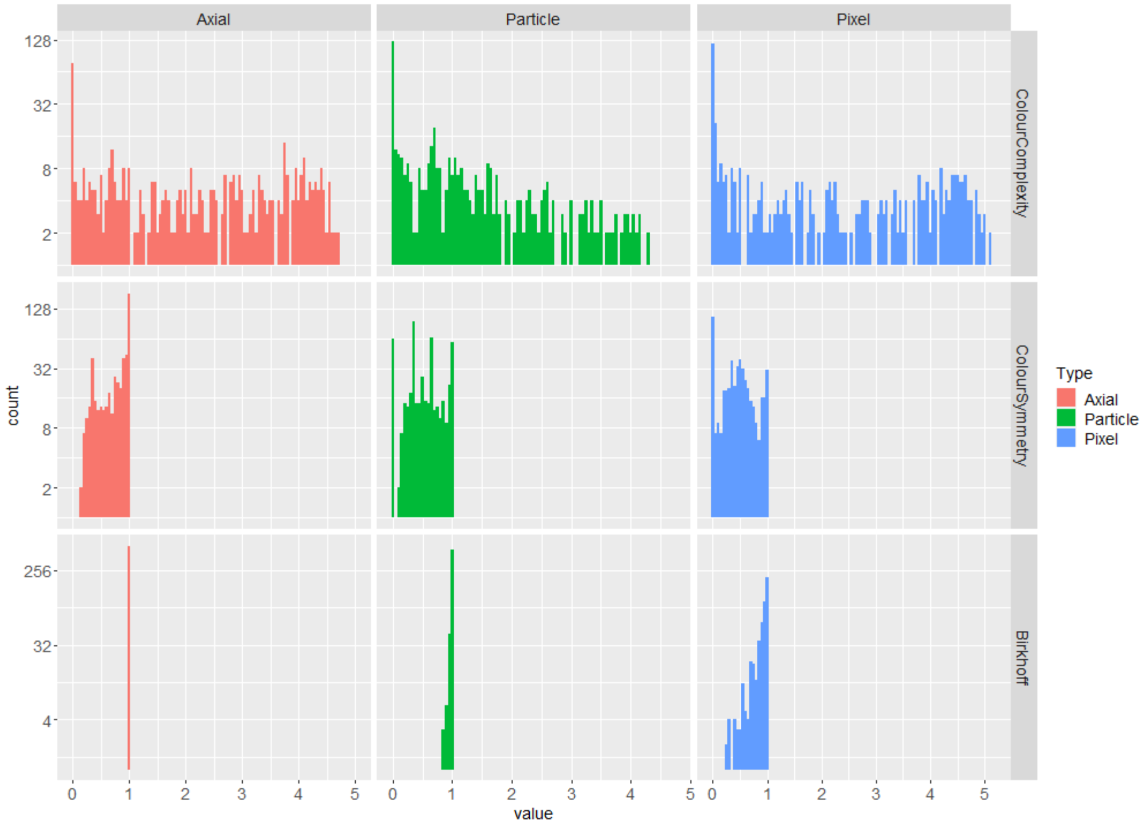
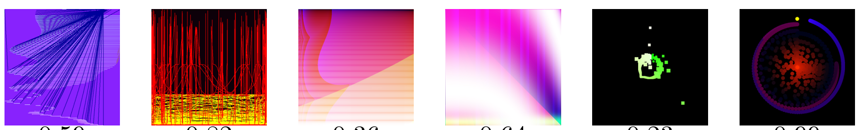


Figure 4.25: The distribution of the 2D randomly generated items across all generation types

For the 3D comparison, the symmetry and complexity for both the colour and form of the items were calculated, complemented by the compressibility as calculated in the base process evaluation. To assess the form of the generated items, the symmetry method



Symmetry	0.59	0.82	0.36	0.64	0.23	0.99
Complexity	1.66	0	4.69	3.56	3.96	4.14
Compressibility	0.97	0.96	0.81	0.73	0.99	0.98

Figure 4.26: Example 2D output across all processes

proposed by Ecins et al. (2017) was used, and the form complexity was calculated using the viewpoint entropy (Castelló et al., 2006). The viewpoint entropy is used to determine the best viewpoint of a 3D scene by calculating how much of the scene is visible from each viewpoint. For this evaluation, the six viewpoints available in the system are used to calculate the value, and the average is provided across all six viewpoints, indicating how much of the sculpture is visible.

For the colour-based measures, the symmetry was calculated using a similar method to the form symmetry, except the final score depended on how close the intensity of the colour was for each matched point across the plane of symmetry. An additional constraint was also added when detecting the original planes of symmetry: if matched item's colours were more than a specified threshold away from each other, this plane was discarded. The average colour error across each detected plane of symmetry is the colour symmetry value presented. The colour complexity follows the same process as the viewpoint entropy, except Shannon's entropy is calculated for each viewpoint and the average value provided. As with the 2D items and due to the background colour not being calculated for 3D items, the background was again ignored for calculating Shannon's Entropy; only pixels containing a set colour were included in the calculations. In addition to this, due to the scales of the 2D and 3D items not being the same (2D distances are measured in pixels, and 3D are measured in metres), the 3D items needed to be scaled to ensure they entirely covered the 2D canvas without losing any information. 500 3D items for each of the processes were generated, using a max tree depth of 5, providing the results as shown in Figure 4.27 with examples of the 3D content shown in Figure 4.28.

The AGP showed the greatest variations across the measures, followed by the particle generation process. For all five measures, the Axial process generally had the widest and most uniform spread of values. For the symmetry measures, this is expected due to the shapes being placed around a central axis, however, for the colour and form complexity, this result is more surprising, indicating that the AGP is capable of creating items that show a wide range of information. The voxel-based system produced the lowest range of values across most measures however, this can be explained by how the items were generated, where they were all cuboids composed of voxels. This means that the symmetry would be identical across all the items, and the number of objects visible from each view would also be constant, explaining the low range of values this process generates. One interesting

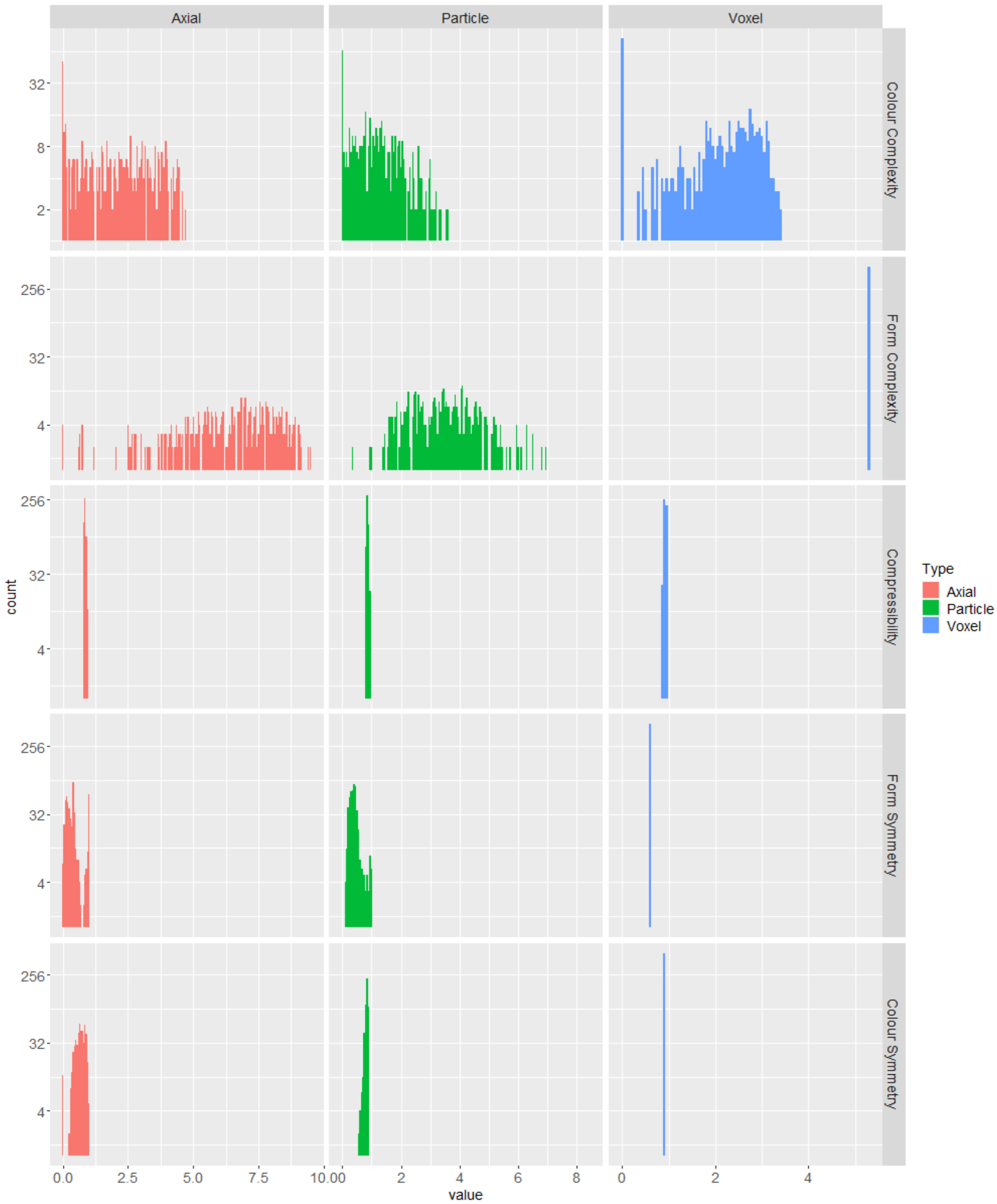
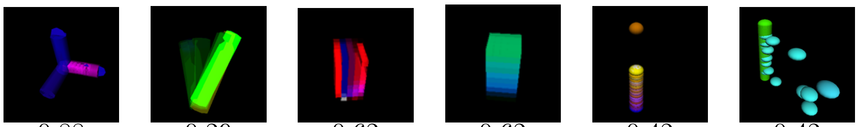


Figure 4.27: The distribution of the 3D randomly generated items across all generation types

aspect of all processes demonstrated is the colour complexity, where the highest number of items had the lowest value between 0 and 0.2. This is due to the initial random generation creating a high number of uninteresting and uncomplex items. This is an aspect where the AGP performed slightly better than the Voxel and Particle processes by creating a lower



Form Symmetry	0.88	0.29	0.62	0.62	0.42	0.42
Colour Symmetry	0.81	0.79	0.92	0.92	0.74	0.60
Form Complexity	1.21	4.71	5.31	5.31	5.08	4.58
Colour Complexity	0.37	1.75	2.17	1.98	3.38	1.63
Compressibility	0.86	0.83	0.91	0.91	0.89	0.66

Figure 4.28: Example 3D output across all processes

number of these types of items.

Finally, the compressibility encountered the same issues as with the base process evaluation and 2D evaluation, where only a small range of values are created across all three generation systems. Once again, due to the increased variability in how the Axial and Particle items are stored, it would lead them to have a higher range of values than the voxel-based system, whose definitions could only ever differentiate based on colour, leading to higher levels of redundancy within the files, leading to consistently high values being found.

4.5.2 Validating the AGP with human users

With this comparison in place to existing systems, in order to meet the goals outlined in Question 1, the sculptures that are generated also need to be able to seem visually diverse and potentially aesthetically appealing to humans, something which the above metrics are not guaranteed to be able to achieve. With this in mind, during the experiments (Chapters 5, 6 and 7) that were run as part of this project, a large amount of data was collected relating to a human's perception of the sculptures.

As part of the experiment, looking at determining how individuals perceive and use symmetry during an aesthetic judgement, more details are provided in Chapter 6. The aesthetic appeal of the sculptures was measured on a 0-10 Likert scale. The users were asked to provide a value for the best and worst sculpture they had seen, as well as the final sculpture they were presented. The collected data is shown in Figure 4.29.

In addition to this data, other user evaluations were collected as part of the aesthetic concept experiment, Chapter 5. Once again, the user was asked to rate the sculptures on a scale from 0 to 10 but were asked to perform this for each item, the highest and lowest rated sculptures are shown in Figure 4.30. However, as this experiment aimed to display a

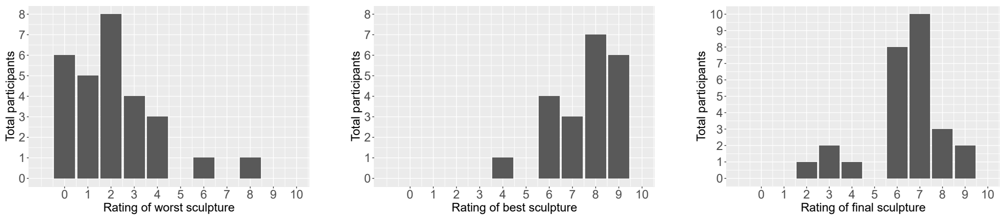
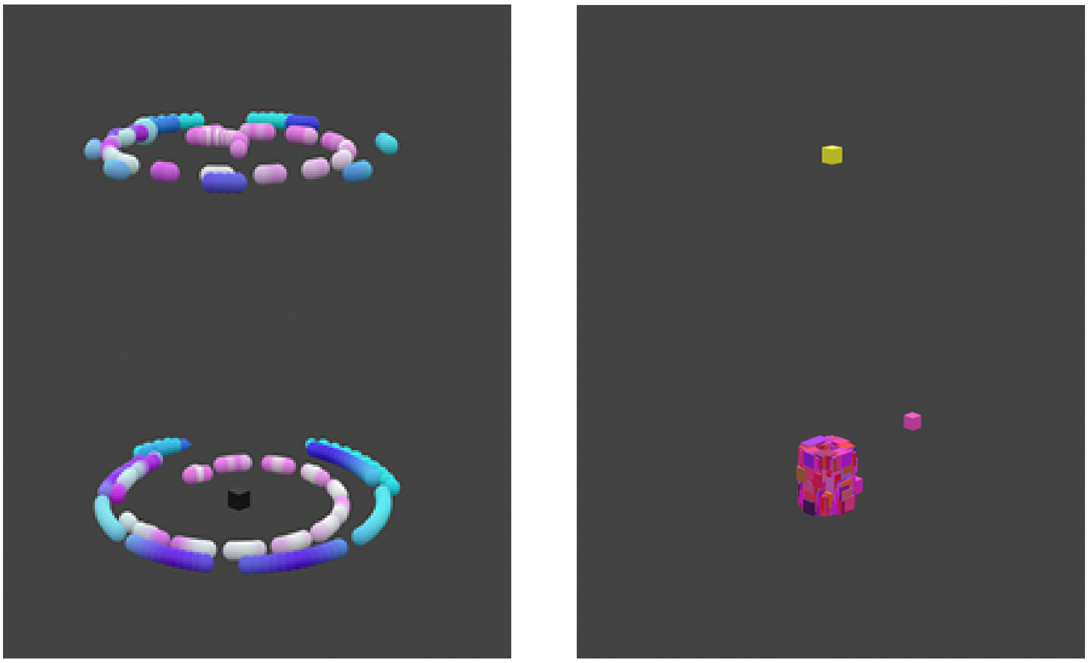


Figure 4.29: Overall aesthetic preference ratings provided for the sculptures in the system: left) Ratings for the worst sculpture, middle) Ratings for the best sculpture, right) Ratings for the final sculpture



(a) Highest rated sculpture shown to participants **(b)** Lowest rated sculpture shown to participants

Figure 4.30: Highest and lowest rated sculptures

wider range of sculptures to determine which aspects contribute to aesthetic judgement, it is expected and beneficial that these sculptures would display a wider range of ratings from the participants.

4.6 Conclusion

The extensions to the AGP presented all have a single focus: to improve the versatility of the Axial Generation Process, based on the initial base process (Easton et al., 2021). The process itself applies the Concretism style to constrain the search space whilst being versatile and capable of generating interesting, visually varied and novel content. The process can easily be applied to existing and future Evolutionary Art systems, and constraining the available search space has the potential to improve various aspects, such as user fatigue, by improving how quickly interesting content can be found. However, the adherence to the Concretism style limits the applicability of this system, as everyone may not consider the style aesthetically pleasing. Whilst the process may not be suitable for all scenarios, the benefits it does possess indicate that it may help improve the search for highly aesthetic items in both automatic and individual-led systems.

The AGP exposes multiple parameters, improving the variety of items that can be generated and giving another dimension for the output to be explored. The computational intensity of generating large or high-resolution items is reduced as the algorithm does not calculate a value for each pixel or voxel. This means that it is suitable for use in performance-intensive environments such as Augmented and Virtual Reality as well as online systems.

The content from the base process exhibited a wide range of values across symmetry and complexity measures; however, it had some issues, and it needed to be clarified whether this meant the AGP was limited or how these results compared to existing systems. In order to allow this comparison and increase the range of novel content generated by the AGP, a range of improvements were applied to the process. Comparing against the other two processes indicated two main themes; firstly, the compressibility measure generally fell within a small range of values for all three generation systems across both 2D and 3D items. In general, the AGP produced items in both 2D and 3D, which displayed a wide distribution of values across measures representing symmetry, complexity and compressibility, indicating that it is a suitable process to be used within an evolutionary environment as it has a high potential to create novel artefacts, as well as being capable of generating a wide range of visually varied, complex and interesting images as well as digital sculptures, comparing favourably to two well-established systems.

Comparing the output from the AGP to other notable 3D art generation systems such as Galapagos (Sims, 1997), Mutator (Latham et al., 2021), or xTNZ (Antunes and Leymarie, 2008) is a difficult premise. Whilst all of these systems sit within the same context of 3D art generation and being utilised as a tool in order to generate 3D artwork, each has its own context, which makes the comparison difficult. The Sculptures generated through the AGP have several benefits required for this project, which these other systems would not exhibit, for example, as this project has a high focus on utilising VR, as detailed in Chapter 3, due to its ability to allow complete control of the environment, preventing the need to create sculptures physically and allowing specific data to be collected from participants. The ease of use within a Virtual Reality Environment and the ability to showcase a controlled range of aesthetic values across different 3D artefacts, e.g. increasing from simple to complex, are specific considerations for this project, which the other systems do not need. Previous systems were created with a different purpose: to create artwork that the creators enjoyed rather than necessarily to study the judgement and aesthetic appeal of 3D artwork. In this sense, the artwork from the previous systems is often more visually complex than the output from the AGP and has a more distinct focus on natural and fantastical forms, and was designed to be more distinctive of the respective artist than the AGP (Lambert et al., 2013).

These results indicate that the AGP meets the constraints outlined for this project and can be used to help determine which aspects contribute to the assessment of 3D elements. This forms a complete answer to Question 2, allowing the creation of visually varied and diverse items. The AGP, combined with the VREs shown in Chapter 3, allows the extensive study of aesthetics at both the collective and individual levels, enabling answers for the remaining research questions.

Chapter 5

Aesthetics of 3D sculptures

The work presented in this chapter has been adapted from the following publication:
E. Easton, U. Bernardet, and A. Ekárt. Contributors to the aesthetic judgement of 3D virtual sculptures. In *2023 Third International Conference on Digital Creation in Arts, Media and Technology (ARTeFACTo)*. IEEE, 2023b.

The success of the AGP at generating a wide range of visually different sculptures provides an answer to Question 2 and allows Question 3 to be considered, determining the aspects and allowing their categorisation. The first task is to identify which aspects contribute to the aesthetic judgement of 3D sculptures, due to the time constraints on this project and the potential for the number of aspects to be almost infinite, the most popular and important aspects need to be identified in order to be focused on throughout the project.

A common method of investigating how artwork is judged starts with generating or analysing specific aspects of existing artwork. However, this approach may be limiting as it does not necessarily align with how humans judge artwork. Determining which terms are commonly used to describe artwork can help overcome this, these terms are the natural way to describe artwork and aesthetic concepts and provide an entry point for people who are less familiar with describing artwork to describe what contributes to their aesthetic preferences. Finding which terms are used by non-experts provides a good starting point for understanding the aesthetic judgement process, eventually enabling the creation of artwork applicable to a wider range of people.

As more aspects contributing to aesthetic judgement have been identified, it will be possible to map which attribute or combination of attributes, when applied to a piece of art, will contribute to the artwork exemplifying the selected term.

Some of the difficulty of formalising the aspects lies with the differences between the computational understanding of these terms compared to how they are understood by humans. Computational approaches generally look at attributes which are easily understood by computers, often visual aspects that are easy for a computer to process, such as symmetry (al Rifaie et al., 2017) and a variety of statistical analyses such as contrast (Matkovic et al., 2005) and naturalness (Acebo and Sbert, 2005). This is not to say more complex aspects have not been investigated, for example, aspects like ethical considerations (Brown and Ventura, 2022; Ventura and Gates, 2018), but the computational interpretation will not necessarily reflect the human understanding of the same term and generally does not allow for ambiguity in terms. As an example, the term complexity could relate to a piece of artwork with complex emotional content or subject matter and visual complexity such as multiple colours, lines and shapes (McCormack et al., 2021). This indicates why a disparity exists between human judgement and the analogue followed by auto-generation systems. Attempting to solve this disparity forms an integral part of the auto-generation of aesthetically pleasing artwork, and while this disconnect remains, it will be difficult for any generation system to reliably create aesthetically pleasing artwork.

Other approaches to the subject of aesthetic judgement concern themselves with the human understanding of the process, and some cross-over between these and computational approaches do exist (Gulhan et al., 2022), but are under-utilised, requiring further investigation to merge the two different understandings of aesthetic judgement. We present an initial step towards merging the two approaches.

5.1 Obtaining the relative importance of aesthetic attributes

The high variety of artwork reflects the complicated nature of aesthetic judgement. Out of all potential aspects, not all contribute to the same extent or even in the same way across different judgements in different contexts, suggesting that aesthetic judgement is not a static process (Tinio and Leder, 2009). The first step to addressing this is to identify which aspects form a fundamental part of the aesthetic judgement process while accounting for the subjectivity in the terms and their application. Within the context of the auto-generation of art, it is crucial to understand which aspects contribute significantly to the aesthetic judgement process.

The identified aesthetic terms would form the basis of generating artwork, however, to use them effectively, more details would be needed, such as the relative contributions of the terms, which ones provide the most helpful information about the items they are describing and finally, whether these terms contribute positively or negatively to the judgement. To collect this, item analysis was chosen as an effective approach that is often applied to designing exam questions (Bichi, 2016; Quaigrain and Arhin, 2017; Toksöz and Ertunç, 2017). With a few minor amendments, item analysis can also be applied to help understand the different aspects that contribute to the aesthetic judgment of artwork.

The tags that are the most commonly applied when judging artwork will represent the popularity of the terms being applied. The more popular the tag, the more important it will be to aesthetic judgement. As well as the popularity, the consistency of the application of the tags is also considered. This protects against ambiguous terms and ensures enough data will be collected to help formalise the aspect.

To calculate both consistency and popularity, the allocation of the tags to sculptures by participants needs to be defined, as shown in Equation 5.1:

$$allocation_{p\,t\,s} = \begin{cases} 1 & \text{participant } p \text{ assigns tag } t \\ & \text{to sculpture } s \\ 0 & \text{otherwise.} \end{cases} \tag{5.1}$$

Using the defined allocations, the mean assignment of each tag t to a sculpture A_t can be calculated by dividing the total number of times the tag has been assigned to each sculpture by the number of sculptures as shown in Equation 5.2:

$$A_t = \sum_{s=1}^{S} \sum_{p=1}^{P} \frac{allocation_{p\,t\,s}}{S}. \tag{5.2}$$

The consistency value for each tag and sculpture can then be calculated by dividing the mean assignment for a tag (A_t) by the total number of participants (P), as shown in Equation 5.3. The consistency indicates whether a specific tag has been reliably applied to sculptures which exhibit the same attributes and that the link between these attributes and the applied term is identifiable by the participants.

$$C_t = \frac{A_t}{P} \tag{5.3}$$

Two further details can also be obtained about each tag: how difficult the tag was to apply and how discriminating the tag is. A highly discriminating tag will be chosen less frequently against sculptures which do not represent the tag, and tags with a high difficulty will not be applied frequently. Both of these measures reveal important information: the difficulty of the application allows the discarding of tags which are too difficult or too easy to apply as they provide too little or too much information on which pieces of art exhibit the term. The discrimination allows the identification of tags which are representative of the sculptures they are applied to, items which do not discriminate enough would be an indication that the tag is ambiguous and can be mapped to multiple attributes. The tags with a medium level of application difficulty and a high level of discrimination would be suitable candidates for further analysis and implementation in art-generating systems. By setting these thresholds for both the difficulty and discrimination values obtained for each tag, a shortlist of aspects can be obtained where each item has sufficient data collected within this experiment in order to potentially learn what physical attributes map to the abstract aesthetic terms.

$$E_t = A_t \tag{5.4}$$

The endorsement E_t of a tag (Equation 5.4) is the mean number of times it has been applied to each item. It can be interpreted as how easily a tag can be applied to the piece of art, where the higher the number of average applications, the easier the tag is to apply.

We consider the difficulty of application (Di_t) to be the complement of ease of application, and as such, we calculate it as the involution function.

$$Di_t = 1/E_t \tag{5.5}$$

For the discrimination to be evaluated, first, a total score must be calculated for each piece of art, set as the sum of applications of all tags applied by all participants to the artwork:

$$Score_s = \sum_{t=1}^{T}\sum_{p=1}^{P} allocation_{p\,t\,s} \tag{5.6}$$

Calculating the discrimination and whether the term has a positive impact on aesthetic judgement requires calculating the correlation between the sculptures and the calculated data. The correlation[1] (Equation 5.9) is calculated using the standard deviation[2] (σ), shown in Equation 5.7, and the covariance[3] (s_{xy}), shown in Equation 5.8, representing how the two sets of data are linearly related.

$$\sigma = \sqrt{\frac{\sum (x_i - \mu)^2}{N}} \tag{5.7}$$

[1] https://www.r-tutor.com/elementary-statistics/numerical-measures/correlation-coefficient
[2] https://www.r-tutor.com/elementary-statistics/numerical-measures/standard-deviation
[3] https://www.r-tutor.com/elementary-statistics/numerical-measures/covariance

$$s_{xy} = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \hat{x})(y_i - \hat{y}) \tag{5.8}$$

$$r_{xy} = \frac{s_{xy}}{\sigma_x \sigma_y} \tag{5.9}$$

The discrimination value for each tag is calculated as the correlation between the total number of times the tag has been assigned to the artwork and the total number of tags assigned to the artwork. A sculpture representing a particular tag will have a higher correlation between these values.

$$D_t = r_{xy}(\sum_{p=1}^{P} allocation_{p\,t\,s}, Score_s). \tag{5.10}$$

The final measure, which provides valuable information about the terms, considers the positive or negative influence that each tag has on the aesthetic judgement of an artwork and requires the overall aesthetic rating for each item (Equation 5.11). Each participant was asked to rate the sculpture between 0 and 10 ($rating_s$), this is correlated to the total applications of each tag on that sculpture, which indicates whether the tag is considered as a positive or a negative aspect of the judgement, the more times a tag has been assigned to a sculpture with a higher rating, the higher the positive impact that that tag has.

$$Pos_{t\,s} = r_{xy}(rating_s, \sum_{p=1}^{P} allocation_{p\,t\,s}). \tag{5.11}$$

5.2 Methodology

It can be easy to determine whether someone likes a sculpture by collecting ratings for each item, an approach which has been used extensively to judge auto-generated artwork to test new generation methods or to investigate whether a particular feature contributes to aesthetic judgement. To collect the required data to determine contributing factors, several aspects need to be considered: what terms the participant can use to describe the artwork, what artwork to display to a participant and how the participant will be able to describe the artwork.

5.2.1 Concept Tags

Determining which terms to include is made more difficult due to several restrictions which need to be applied: (1) terms should not be too technical to ensure they can be understood

Quiet	0.17	Ordered	0.18
Boring	0.19	Complex	0.19
Dynamic	0.21	Simple	0.23
Stiff	0.25	Balanced	0.25
Disconnected	0.25	Unified	0.25
Irregular	0.27	Natural	0.27
Plain	0.28	Gentle	0.28
Neutral	0.29	Exciting	0.29
Interesting	0.29	Dull	0.32
Surprising	0.32	Alive	0.38
Freakish	0.38	Graceful	0.38
Lifeless	0.38	Messy	0.38
Practical	0.38	Predictable	0.38
Strange	0.38	Strong	0.38
Controlled	0.38	Ugly	0.38
Unemotional	0.38	Unfriendly	0.38
Unnatural	0.38	Unpleasant	0.38
Weak	0.38		

Table 5.1: Aesthetic terms

by all participants. (2) The conceptual aesthetic space needs to be covered to try and fully explore how individuals judge sculptures.

To compile the initial list of terms, an overview of aesthetic judgement was sought. Work undertaken by Sibley (1959), who investigated which terms were commonly used to describe artwork by art critics, resulted in a list of 134 terms founded the basis of our search. Due to their origin, these terms are inherently applicable to the judgement of artwork, and whilst they were compiled over 60 years ago, their meanings are still relevant in this context. To avoid overwhelming the participants with too many terms, the list was filtered further to include only 30 terms by checking the semantic similarity using the WordNet2 database[4]. This allowed the reduction of the list of terms whilst still maintaining the coverage of a wide range of the conceptual space. The similarity was calculated by measuring the distance to the lowest common linked word between each pair of terms, items which have a closer common ancestor are considered more similar. Once the value had been calculated between each pair of terms, an average was taken for each one to represent how generally different that term was from the others, examples of the terms and their average semantic similarity are shown in Table 5.1.

The terms were then manually sanitised to ensure the list contained no antonyms, synonyms or duplicates and that the term did not use outdated language, i.e. where the term's meaning may not be obvious to all participants, e.g. Gaudy. This allowed all terms to be presented in the most unambiguous form possible and avoided any issues with slight differences in meaning between the presented terms. Finally, any terms directly connected to emotional state were removed, for example, happy or sad, as their inclusion may unin-

[4]https://wordnet.princeton.edu/

Busy	Cold	Quiet
Angular	Unrefined	Ordered
Complex	Boring	Drab
Friendly	Static	Separate
Calm	Original	Natural
Curved	Warm	Loud
Simple	Sophisticated	Disordered
Unfriendly	Interesting	Bright
Unoriginal	Dynamic	Connected
	Unnatural	

Table 5.2: Final list of aesthetic terms

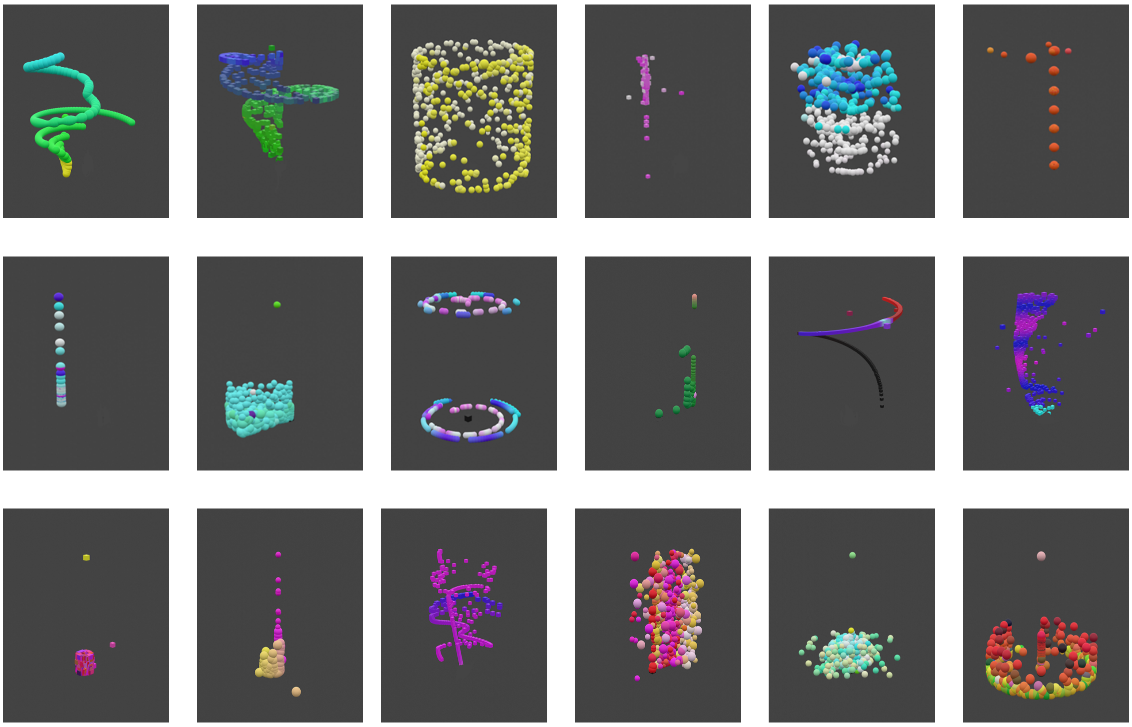


Figure 5.1: Sculptures available for participants to judge

tentionally affect the results due to differences in participants' emotional state before taking part in the experiment. The top 14 dissimilar terms were selected, and then to widen the area of semantic space being covered by this reduced list, their antonyms were added to the list to obtain the final list of aesthetic concepts, shown in Table 5.2. These terms represent the widest area of the aesthetic concept space available using the chosen inputs whilst removing as much subjectivity as possible. It should be noted that while these terms were removed here, they may still be suitable for investigation using a more suitable experimental paradigm.

5.2.2 Sculptures

In addition to the tags needing to cover a wide area of the aesthetic concept space, the same criterion needed to be applied to the sculptures displayed to the participant. For this experiment, the AGP (see Chapter 4) was used with the sculptures being created by placing 475 spheres or cubes within the bounds of a 1x1x2m containing box.

To ensure that the sculptures were visually different and displayed a wide variety of attributes, a distance search was implemented using a Genetic Program, with a fitness function which determined how dissimilar one sculpture was from another. A standard approach to implementing this would involve measuring the geometric distance between each of the points within a sculpture and plotting these distances into a histogram, with the final measure being set as the Chi-Squared distance between each pair of histograms. The higher the resulting value, the more different the sculptures would be.

However, as noted by Biasotti et al. (2016), this process has limitations, and the signature generated by the histogram is not unique enough to determine true dissimilarity. In order to combat this, several histograms were created for the sculpture similar to the process used by Mahmoudi and Sapiro (2009). The entire process was not followed as some of the measures were not appropriate for the sculptures, for example, using the geodesic distance did not work well as there was no clear surface or path between the points in the sculpture. Instead, another approach was formulated, which created a histogram for the X, Y and Z distances individually. All histograms were combined and the Chi-Squared distance was taken on the resulting sculpture signature.

Once each run of the Genetic Program had been completed, the sculpture with the highest dissimilarity was selected and added to an archive. The archive was used in all future runs of the algorithm to compare all newly generated items, ensuring that all selected items were visually different from each other. Each run maximised the dissimilarity between the population and the archive sculptures for ten generations, and a total of 50 runs were completed, providing a selection of 50 sculptures. Similar to the aesthetic tags, 50 sculptures were deemed to be too many to be included in this experiment and so the 18 most visually different sculptures were selected by the authors, the resulting set of 18 sculptures is shown in Figure 5.1.

5.2.3 Environment and Process

The VR Experiment from Section 3.7 was utilised due to the main focus of collecting generalised details about the aesthetic judgement process. The experiment followed the process as outlined in Chapter 3, with participants offered £5 to take part based on the estimated 30 minutes to complete the entire process, with the list of tags available to assign to each group of sculptures shown in Table 5.2;

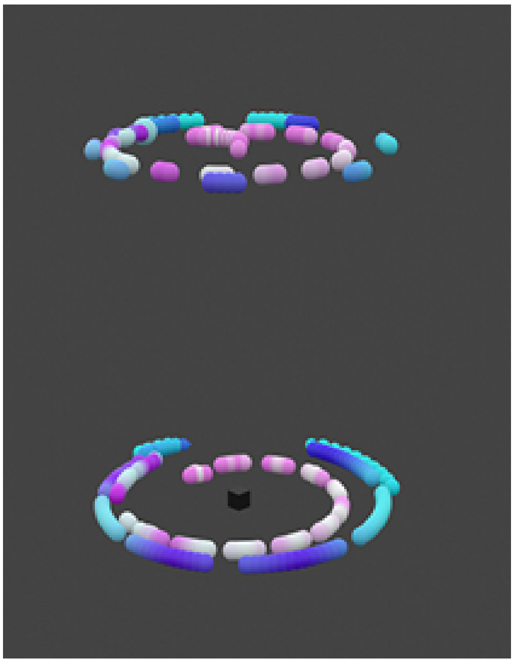
5.3 Results

To determine whether the chosen aesthetic terms formed any part of the aesthetic judgement of the sculptures and, if they did, the extent the terms contributed and whether these terms had a positive or a negative impact on how the sculptures were judged, the groups and tags the participant assigned as well as the time the participant spent in each stage was collected. The exit questionnaire collected details about the art expertise of the individual, a rating for each sculpture from zero to ten, and other general feedback on the process, such as whether the participant experienced any motion sickness and whether any other terms may be appropriate to describe the presented sculptures.

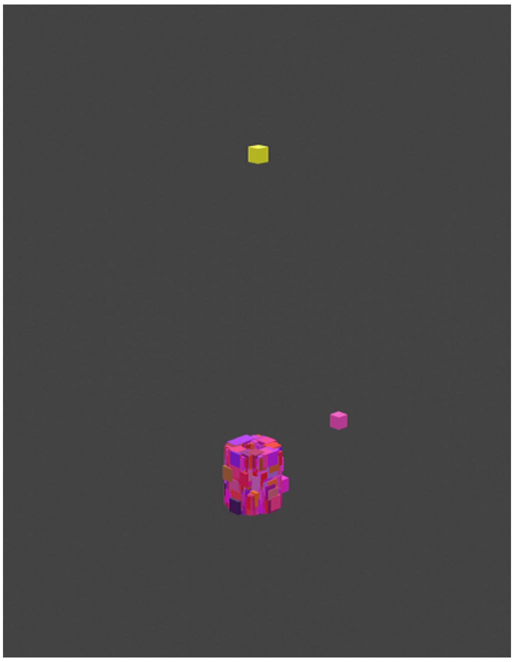
A total of 37 participants completed the experiment (10 female) with a mean age of 30, who spent an average of around 26 minutes within the VR environment. Across all participants, there was a low amount of art expertise as shown by the mode responses provided across all three collected metrics: how often the participant visited an art museum (Rarely), how often an art book was read (Never) and how often the participant practised any form of art (Never). This indicates that the data collected would not be influenced by a high level of expertise, allowing the data to be considered more generalised in terms of art expertise.

The presented sculptures represented a wide range of being aesthetically pleasing to the participants, with average ratings of the sculptures ranging from 1.8 (shown in Figure 5.2b) to 7.3 (shown in Figure 5.2a).

Table 5.4 shows the summarised data about each aesthetic tag, intriguingly, it can be seen that Dynamic, Curved, Interesting and Connected terms had the most assignments on average, an early indication of the application difficulty for the tags. One important result is that all available tags were applied to the sculptures, suggesting that the chosen tags



(a) Highest-rated sculpture shown to participants



(b) Lowest-rated sculpture shown to participants

Figure 5.2: Selection of sculptures with the highest and lowest average rating

were suitable for describing the presented sculptures. As shown in Table 5.3, dynamic was also the most consistently applied tag, indicating that the sculptures presented exhibited a high level of dynamism. Dynamic along with curved, interesting and connected are some of the main aspects contributing to the aesthetic judgement of 3D sculptures.

Table 5.3 also shows an indication of which tags had a suitable level of discrimination and application difficulty to be considered for further analysis by having a high level of discrimination and a medium level of application difficulty. The tags which fell within the acceptable criteria are highlighted in the table. Dynamic was the most popular tag, due to this, the application difficulty was too low to produce reliable information. Instead, the tags which fell within the criteria specified in Section 5.1 were connected, busy, interesting, ordered, complex, angular, friendly and calm. The final column shows the positivity rating for each of the tags, and all of the terms which met the outlined criteria had a positive influence on the judgement of the sculptures. None of the highlighted terms were the most positively associated terms, however, this accolade is applied to curved, with simple being the most negatively associated term.

	C	E	Diff	Disc	Positivity
Dynamic	0.32	5.78	0.68	0.48	0.22
Curved	0.26	4.67	0.74	0.77	0.31
Simple	0.23	4.17	0.77	-0.47	-0.27
Connected	0.23	4.06	0.77	0.81	0.18
Busy	0.18	3.22	0.82	0.42	0.07
Interesting	0.18	3.17	0.82	0.73	0.11
Ordered	0.17	3.00	0.83	0.74	0.11
Complex	0.16	2.94	0.84	0.67	0.13
Bright	0.16	2.83	0.84	0.71	0.08
Disordered	0.15	2.67	0.85	-0.36	-0.08
Unnatural	0.15	2.61	0.86	-0.07	0.00
Static	0.14	2.50	0.86	-0.31	-0.08
Angular	0.12	2.11	0.88	0.36	0.07
Friendly	0.12	2.17	0.88	0.59	0.10
Calm	0.12	2.22	0.88	0.61	0.04
Warm	0.12	2.17	0.88	0.00	0.08
Unfriendly	0.12	2.11	0.88	-0.45	-0.04
Cold	0.11	2.06	0.89	0.10	-0.03
Boring	0.11	2.00	0.89	-0.73	-0.17
Loud	0.11	2.00	0.89	-0.27	-0.01
Separate	0.10	1.89	0.90	-0.48	-0.05
Quiet	0.08	1.39	0.92	-0.07	-0.03
Original	0.08	1.39	0.92	0.43	0.05
Natural	0.08	1.50	0.92	0.49	0.09
Sophisticated	0.07	1.28	0.93	0.27	0.04
Drab	0.06	1.00	0.94	-0.81	-0.11
Unrefined	0.04	0.78	0.96	-0.52	-0.08
Unoriginal	0.03	0.61	0.97	-0.61	-0.08

Table 5.3: Consistency (C), Endorsement (E), Difficulty of application (Diff), Discrimination (Disc) and positivity rating for each tag, highlighted items represent tags which warrant further investigation

5.4 Discussion

The collected data raised some interesting patterns, all the available tags were used, with the lowest number of participants being drab, which was applied by 10 participants. This is a good indication that the selected terms were fit for purpose and successfully described the generated sculptures, the tags were also mainly rated as positive.

A critical aspect of this data is that the details were collected from participants who had a limited amount of art expertise, this potentially helps to explain some of the results that were obtained. As having a high level of art expertise can influence aesthetic judgement, the data collected in this experiment can be considered less biased in this respect. This potentially explains the difference in which terms were considered the most important to achieving a positive aesthetic rating over commonly used aspects such as ordered or complex.

The positivity or negativity ratings for the terms are mainly intuitive, for example, boring sculptures were rated less highly than interesting sculptures and generally, they corresponded to the antonym relationship between terms, e.g. Bright and Drab. However, a few anomalous aspects were identified, for example, whilst natural related to positive ratings, unnatural was neutral, which potentially questions the use of some measures which look at the naturalness of artwork, such as the fractal dimension or Benford's law, where having lower ratings on these scales may not be too detrimental. Similarly, the contribution of complexity is positive, however, the positive impact of items being complex is not as significant as the negative contribution if an item is considered simple. This indicates that while complexity is an important aspect of aesthetic judgement, it is more important, when trying to create aesthetically pleasing items, to ensure that items are not simple rather than specifically creating complex items. This potentially affects how items should be generated, especially within a Computational Creativity context, where instead of trying to maximise the positive influence of complexity, it should be more important for systems to minimise the negative aspect of simplicity.

The subjectivity of the process is exemplified by the contribution of the curved/angular antonym pair of terms, as both are shown to contribute positively to the judgement, albeit with angular being less positive than curved. Inspecting the sculptures the terms were applied, to shows that both terms were applied to the majority of the same sculptures, indicating that even though these terms are semantically opposite, both can be displayed within the same sculpture and have a positive effect on the overall rating.

The stalwart formal measures of complexity and order also feature as important aspects, backing up existing research by showcasing that they contribute positively to aesthetic judgement. However, despite their ubiquity, they are not the most positive items contributing to aesthetic judgement. Terms like curved, dynamic and connected had a higher positive impact on aesthetic judgement, indicating that in the process of auto-generating aesthetically pleasing items, these terms should be considered more often than complex and ordered.

One final aspect of the results is that negative tags seem more difficult to describe than positive items. The average difficulty for positive tags was 0.84. In contrast, for negative tags, it was 0.89, which may indicate that negative aspects are less easy to visually distinguish, at least within the sculptures used within this project.

5.5 Categorising the aesthetic aspects

With this data collected about each of the investigated tags, categorising them can begin. This forms a significant part of answering Question 4, where the categories will determine how the tags will be handled within the created user model, described in Chapter 8.

Two areas need to be considered when describing an aspect as constant, transient or subjective. The first relates to the term itself and its presence within the sculpture, e.g. all sculptures may have an element of dynamism, but only some may be considered complex or static. The second is to consider the values for the aspects, for example, symmetry is often considered constant. However, as investigated in Chapter 6, the preferred value can be subjective. To some extent, a tag may be present in all the sculptures, however, determining the preferred value for the current individual can be much more complicated. This illustrates the importance of considering both sides when generating a complete model, however, due to time constraints and being unable to fully formalise all of the investigated aspects, only the tags will be categorised.

Splitting the terms into the appropriate categories consists of a two-stage process: (1) the constant and subjective items need to be identified, due to the main goal of Question4 being to create a model of individual user preference, the assignment of tags per participant will be considered. If the tag has been applied by most or all of the participants, then it can be considered as constant, as the tag will most likely be present in all sculptures to a level which is identifiable by all individuals. If the tag has not been applied with regularity, it is considered subjective. Using the collected data, the tags and how many participants applied them are shown in Table 5.4. Due to that experiment focusing on displaying a wide range of sculptures and providing only a limited amount of aesthetic tags, the constant tag application threshold will be set at 75%, which for the collected data is 28 participants.

Using this threshold, the tags are split into the two initial categories, Table 5.5. There are significantly more items which fall within the subjective category than in the constant category, which suggests that the categorisation process is roughly correct, as it follows the general rule that aesthetic judgement is more subjective than objective.

There are some interesting points to note in the list, aspects such as ordered and complex are included within the subjective category, this suggests that these aspects do not need to be present when generating sculptures using this model. This reinforces the data collected, where aspects like simple and complex were treated differently, indicating that they might be more subjective. It is perhaps unintuitive that some antonyms appear within different categories, e.g. dynamic and static, it may be expected that a sculpture would sit on a spectrum between these two items so that either tag may be present at any time. However,

	Total Uses	Total Participants Used
Curved	172	31
Dynamic	182	30
Calm	111	29
Interesting	170	28
Simple	124	28
Bright	124	25
Disordered	97	25
Busy	103	24
Warm	86	24
Connected	132	23
Unnatural	102	23
Cold	103	22
Natural	77	22
Loud	97	21
Ordered	97	20
Complex	95	20
Boring	95	19
Angular	60	18
Friendly	69	18
Static	75	18
Quiet	51	17
Sophisticated	51	16
Original	62	14
Original	62	14
Separate	64	13
Unfriendly	52	12
Unrefined	42	12
Unoriginal	35	11
Drab	27	10

Table 5.4: How many times each tag was used and how many participants used the tags

this factor can be accounted for in three ways, firstly the display of each aspect within an art context is not a binary selection, i.e. a sculpture can display both friendly and unfriendly attributes. This is backed up by the positivity ratings obtained for each tag, obtained in Table 5.3, where aspects like curved and angular were both attributed to positive ratings for the sculptures. The second explanation is that the human understanding of these terms does not equate to exact opposites, especially when judging these sculptures. An alternative explanation would be that only limited amounts of information has been collected, and repeating this experiment with more participants across different art styles could produce different results.

5.6 Identifying Transient Aesthetics of 3D sculptures

The transient aspects represent a subset of the initial subjective aspects; each transient item will be subjective, but not all subjective items are transient.

It would be possible to identify which aspects are transient by running a longitudinal experiment that looks at how preference changes over time for the aspects. It would not be possible to run this experiment as part of this project due to time constraints, so instead, another approach needs to be considered. In order to do this, the semantic meaning of the

Constant	Subjective
Dynamic	Connected
Curved	Bright
Calm	Busy
Interesting	Cold
Simple	Unnatural
	Ordered
	Loud
	Disordered
	Complex
	Boring
	Warm
	Natural
	Static
	Friendly
	Separate
	Original
	Angular
	Unfriendly
	Quiet
	Sophisticated
	Unrefined
	Unoriginal
	Drab

Table 5.5: Initial categories of each tag

Original
Sophisticated
Unrefined
Unoriginal

Table 5.6: Subjective terms which change based on prior knowledge of the context

terms will be evaluated and combined with previous research in order to determine whether a subjective tag should be moved into the transient category.

This involves identifying any terms which relate to existing knowledge of the sculptures or the emotional states of the individual. The former suggests the term is transient as the longer someone spends with the sculptures, their understanding of the sculptures will change to include the new variations presented. A clear example of this is the "original" tag. The first time someone sees the sculptures presented, they may consider them as highly original due to the participant not having the required prior knowledge to compare the sculptures to. As more sculptures are presented, it would be likely that even dissimilar items would be marked as less original than the initial ones. These terms are context-dependant terms where the understanding of what is possible will potentially affect how they are viewed by the participant, the terms which fall under this aspect are listed in Table 5.6. Whilst this can be applied to all of the terms, other aspects such as "connected" can be assessed objectively, whereas "original" solely relies on knowing what other sculptures the participant has seen.

The second method of identifying the transient terms is to consider whether any of the terms are related to the emotional state of the individual. These items would be heavily influenced by a wide range of external aspects, such as how the participant is currently

Cold
Loud
Boring
Warm
Friendly
Unfriendly
Quiet
Drab

Table 5.7: Subjective terms which relate to the emotional state of the participant

Constant	Subjective	Transient
Dynamic	Connected	Original
Curved	Bright	Sophisticated
Calm	Busy	Unrefined
Interesting	Natural	Cold
Simple	Unnatural	Unoriginal
	Ordered	Loud
	Disordered	Boring
	Complex	Warm
	Natural	Friendly
	Static	Unfriendly
	Separate	Quiet
	Angular	Drab

Table 5.8: Final categories of each tag

feeling or whether the participant is in an excitable state (Emanuel and Eldar, 2022; Forgas, 1995; Lerner and Keltner, 2000). A participant's emotional state is likely to change regularly and is highly dependent on the context the participant is viewing the sculptures in and their current mindset when viewing the sculptures. Items which have been identified to fall within this are shown in Table 5.7.

Table 5.8 shows the final classification of the terms to be utilised within the model of individual preference. As this is the initial placement of the terms into categories, it is possible the categories may not be correct, however, through further investigation, these categories can be checked and improved.

5.7 Conclusion

The gallery VR environment, introduced in Chapter 3, was utilised in order to provide the initial answer to Question 3, allowing the extraction of terms that contribute to the aesthetic judgement of 3D sculptures providing a wealth of data about the judgement, along with this, terms which contribute both positively and negatively to the overall aesthetic rating of each sculpture were also identified.

The overall level of expertise of the participants helps the collected data be less influenced by artistic expertise, causing it to be less biased. Successfully collecting data provided by participants with limited expertise in artwork indicates the presented approach's success in determining which aspects contribute to aesthetic judgement. This approach can be utilised to help overcome the difficulty of describing which aspects contribute to aesthetic judgement, helping to remove this barrier with modelling aesthetic judgement.

By following set criteria, based on the results obtained through item analysis, potential tags were established which would make suitable candidates for further investigation. Other items would potentially be good criteria for investigation; however, they would need a wider set of sculptures to be presented before enough data could be collected to reliably determine the attributes causing their application. The established criteria considered each tag's application difficulty and discrimination, where a medium level of difficulty and a high level of discrimination was required to select a tag for further consideration. The collected data led to nine tags being identified: connected, busy, interesting, ordered, complex, bright, angular, friendly and calm, which can be used as a basis for formalisation. One interesting aspect of these tags is the inclusion of the complex and ordered aspects; this backs up existing research relating to the importance of complexity in the assessment of artwork. However, it must be noted that due to the lack of definition provided to the participants for the terms, how this was assessed by each individual may be distinct from the notion of complexity as introduced in Chapter 4 (McCormack et al., 2021). This potential disparity between the participants understanding and how this is measured is beyond the scope of this project.

However, there were limitations with this process, using a small set of geometric shapes to generate the sculptures, only spheres and cubes, may have influenced how easily terms such as curved or angular were applied. Whilst care was taken to try and present a wide range of terms for use, by only presenting a set of 18 terms, the process placed restrictions on how the presented items could be judged. The result suggests that the terms were suitable; however, it is possible that other terms could have a higher positive or negative impact than the ones identified. In addition, the restriction also means that the process cannot represent a full aesthetic judgement, only a small portion of it. The participants were given the opportunity to suggest other terms, which could help to rectify these limitations, however, there was no agreement, with each participant suggesting different terms, a further indication of how subjective the process of judging artwork is and how successfully this approach of extracting aspects was. The small amount of sculptures also represents a limitation of this system again, even though effort was taken to make the sculptures as different from each other as possible, only a small selection was presented, which may not have exhibited all of the provided terms in equal amounts.

With the collected data, it was possible to categorise each of the investigated tags into one of three categories: constant, subjective and transient. These categories will be utilised in order to influence how the terms should be handled when used as part of the individual aesthetic model as described in Chapter 8.

Overall, several important and relatively unconsidered high-level aesthetic aspects and their categories were identified through item analysis to determine how well the aesthetic tags described the presented sculptures. This provided a wealth of information on how the tags were applied and the positive and negative impact of the terms but also provided 3D models which exhibit the properties of each term, helping to identify some of the many aspects which contribute to the aesthetic judgement of 3D sculptures.

Chapter 6

Extracting individual aesthetic preferences

The work presented in this chapter has been adapted from the following publication:
E. Easton, U. Bernardet, and A. Ekárt. Is Beauty in the Age of the Beholder? In *Artificial Intelligence in Music, Sound, Art and Design: 12th International Conference, EvoMUSART 2023, Held as Part of EvoStar 2023, Brno, Czech Republic, April 12–14, 2023, Proceedings*, pages 84–99. Springer, 2023a.

As discussed in Chapter 1, two approaches of aesthetic appreciation need to be considered, individual and collective, with both scopes needed to understand aesthetic judgement. This chapter considers individual aesthetic preference, using the Mine Cart VRE introduced in Chapter 3. Chapter 5 enabled the identification of the most important terms used when judging 3D sculptures, and similar to prior research, one of the main aspects which contributes positively to aesthetic judgement is symmetry. The symmetry may not be the most important aspects contributing to the assessment when compared against the other tags, such as connected. These other tags need to be formalised before they can be applied to individual aesthetic preference (this formalisation process is completed in Chapter 7), however, well-established methods of calculating the order exist, meaning that the focus on this term can relate to how it fits in within aesthetic judgement.

Symmetry is considered one of the most prevalent aspects throughout the general population, with people having a high affinity for symmetrical features, especially in the faces of other humans (Rhodes, 2006). This has led to the speculation that a desire for symmetry is caused by an evolutionary process, for example, as the presence of symmetry indicates a higher presence of healthy levels of hormones within a human being. However, the investigation of which levels of symmetry are preferred within artwork has led to some varied results. This contrasts with the levels preferred within 3D, real-life objects such as architecture and faces, where the level is constant throughout all participants. This chapter considers symmetry with respect to sculptures, looking at whether it is possible to extract individual user preferences of symmetry, considering the role that symmetry plays in the appreciation of 3D virtual sculptures as well as any patterns of preferred values.

6.1 Identification of symmetry preference in 3D sculptures

The symmetry present in a piece of artwork has long been known as an essential aspect relating to aesthetic judgement, it was part of the earliest and more ubiquitous formal equations of beauty (Birkhoff, 1933), where the beauty of an artwork is the ratio between its order and complexity. Further analysis has been applied to the symmetry and why it is so important, with several theories suggesting that due to its prominent role in processing visual information, where the more ordered an item is, the easier it is for our brains to process (Bertamini et al., 2018). Extensive research has been performed, resulting in different methods being suggested for quantifying it, most attempts to define the order relate it to the symmetry of the artwork, for reasons summed up by Weyl (2015) "Symmetry is one of the ideas by which man through the ages has tried to comprehend and create order, beauty and perfection". Symmetry is something which has near universal importance across multiple research fields, such as architecture (Zeki, 2019; Salingaros, 2020; McDermott et al., 2012), inter-human attraction (Rhodes, 2006), user interface design (Bauerly and Liu, 2006) and the extensive work looking at it within 2D artwork (al Rifaie et al., 2017; Bergen and Ross, 2013; den Heijer and Eiben, 2010b; den Heijer, 2012; O'Reilly and Hemberg, 2007; Osborne, 1986; Vinhas et al., 2016). It is a safe assumption that symmetry plays a major role in the evaluation of artwork however, there are still a variety of factors which can influence an individual's symmetrical preference (Jones et al., 2007; Leder et al., 2019; Weichselbaum et al., 2018). Determining its relative importance is not straightforward, as it must be understood how symmetry fits within the wider context of art appreciation. Evidence needs to be collected which considers how people judge artwork when different levels of symmetry are

Total Generations	500
Mutation Probability	0.7
Initialisation Method	Full
Selection Method	Tournament (k=3)
Population Size	50
Fitness Measure	Average Symmetry
Tree size	4

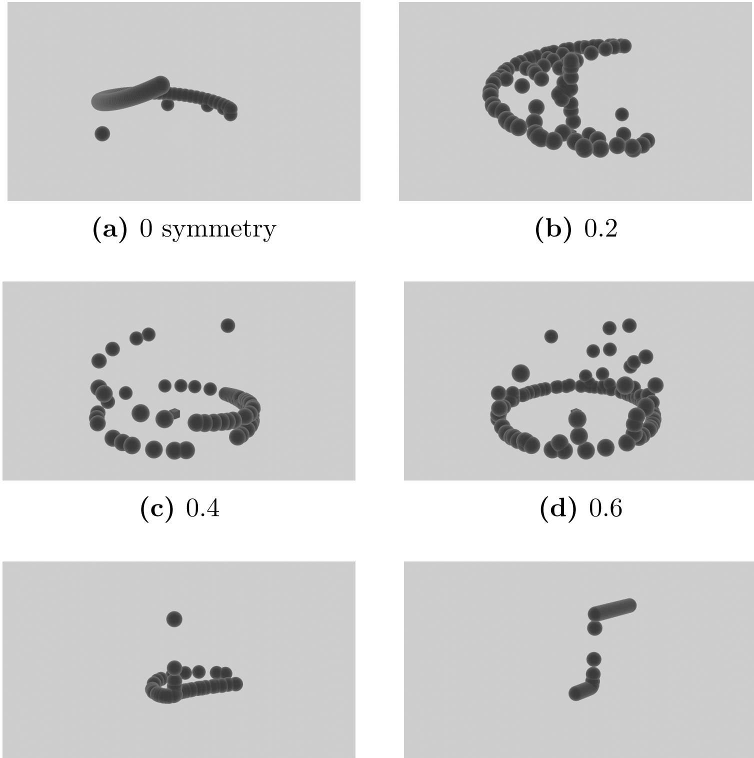
Table 6.1: Evolutionary Algorithm parameters

present, within the experiment presented here, this data is obtained through participants making multiple choices between varying levels of symmetry.

6.2 Generating the sculptures

Each choice the user makes is between two sculptures generated with the Axial Generation Process, as described in Chapter 4, with a couple of restrictions: using a reduced set of expressions in the genotype expression trees and only using a neutral stone grey colour to render them in the VR environment. To provide the ability for the user to choose between sculptures at differing levels of symmetry, six genetic algorithm runs were performed to create items which displayed symmetry between 0 and 1 in steps of 0.2, with the symmetry level calculated using a measure based on point clouds (Ecins et al., 2017). This measure works by finding multiple candidate symmetry planes and refining their positions by minimising the error between points reflected in the plane. The score calculated for each item is the average distance error across all detected planes for an item.

From each run of the GA, parameters shown in Table 6.1, the sculptures displaying symmetry values approximating the target values were automatically selected. The experimenters chose the final set of sculptures based on how interesting they looked, a selection of the sculptures at each symmetry level is shown in Figure 6.1. To avoid uninteresting sculptures being initially generated, the complexity, measured as the Global Normalised Kolmogorov Complexity (Rigau et al., 2007) which calculates the compressibility by encoding all points into text and then compressing the string using ZIP compression, similar to the process used by (Ekárt et al., 2011; Easton et al., 2022), was also maximised as part of the generation process.



(a) 0 symmetry **(b)** 0.2

(c) 0.4 **(d)** 0.6

(e) 0.8 **(f)** 1

Figure 6.1: Examples of the standard sculpture at each available symmetry level

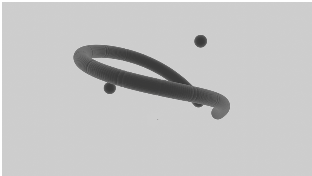
6.3 Creating the reward

As mentioned in Chapter 3, part of the incentive given to the user to compare the different sculptures is to locate the final sculpture, which has a symmetry level specific to their choices. The process to generate the final sculptures followed a similar method to the main sculptures but which created larger and more diverse items intended as a reward for navigating through the caves. The mean symmetry preference across the two sets of choices was calculated in 0.1 steps, the final sculptures also reflected this by being generated using 11 genetic algorithm runs, each focusing on a value from 0 to 1 in 0.1 steps, using the parameters shown in Table 6.2. Five sculptures were chosen at each level of symmetry to provide a wide range of options which could be displayed to the user in the final section of the experiment. Examples of the final sculptures are shown in figure 6.2. The pre-generation of the final sculptures ensured that the game experience was smooth, with no waiting period for live generation of the final sculpture. As it was solely the final symmetry value that would drive the generation of the final sculpture and there were only 11 possible values, the

Total Generations	500
Mutation Probability	0.7
Initialisation Method	Full
Selection Method	Tournament (k=3)
Population Size	50
Fitness Measure	Average Symmetry
Tree size	4

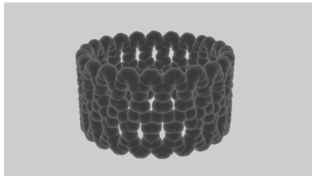
Table 6.2: Evolutionary Algorithm parameters

pre-generation of the set of possible final sculptures at each symmetry level did not detract from the quality of the observations.

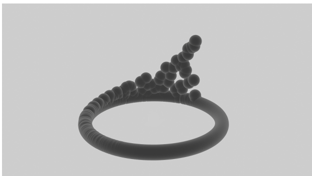


(a) 0 symmetry

(b) 0.1

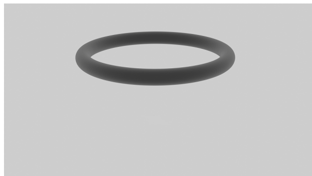


(c) 0.3



(d) 0.5

(e) 0.8



(f) 1

Figure 6.2: Examples of the final sculpture

6.4 Experimental setup

The experiment was run in two phases, the first phase was run entirely online, with participants sourced through the Prolific platform who had access to their own VR headsets. This phase specifically targeted individuals 18 years old and over with no specific age limit however, by requiring that the participant had access to their own VR headset, this potentially limited the available age range of participants, skewing the age range towards younger adults who are more likely to own a VR headset. Rather than developing multiple versions of the app for different headsets, the Meta Quest and Quest 2 headsets were focused on, with it being available to download via the Quest App Labs store.[1] The second phase was completed at the CityFest event, run by Aston University, showcasing technology projects to school children aged 11-13, and Meta Quest 2 headsets were used for the experiments.

[1]https://www.oculus.com/experiences/quest/3429965927112413

Group	Mean Symmetry	Symmetry difference	Min Age	Max Age	Count
1	0.50	0.28	13	17	5
2	0.30	0.23	18	30	15
3	0.37	0.31	31	60	13

Table 6.3: Details of the symmetry values for the three age groups

For the first phase, in order to run the experiment using the Prolific platform, the available headsets on the platform needed to be determined. This was achieved by running a short questionnaire which asked participants to list the VR headsets they had access to. From the results of this questionnaire, only participants with a Meta Quest headset could be targeted on the Prolific platform.

6.5 Results

A total of 33 participants completed the experiment (13 female, 20 male) between the ages of 13 to 52 (m=28.0). This data has been presented before in Chapter 3, where validating the process was considered. Two further aspects are considered here: looking at the overall level of symmetry preferred by the participants across different age groups and the participant's behaviour within the system.

6.5.1 Symmetry preference

The analysis of the process, Chapter 3 indicates that confidence can be had in the collected data and used for assessing the primary purpose of this experiment, extracting an individual's preference for symmetry.

Similarly to health-related age groupings in scientific studies, and also in line with the expectation that preference changes with experience and experience grows with age, the participants were grouped by age into three groups: under 18, 18-30 and 31+. Table 6.3 shows simple statistics of the age groups, including the mean of the chosen symmetry for each group, while the full distribution is shown in Figure 6.3a.

This simple analysis of the values of these groups show that an effect of age on the preferred level of symmetry is possible. Group 1 included the youngest participants, who consistently chose a higher level of symmetry compared to the more mature groups. This indicates that age potentially does have some impact on the level of symmetry preferred by a participant. However, the one-way ANOVA indicates there was no significant effect on the

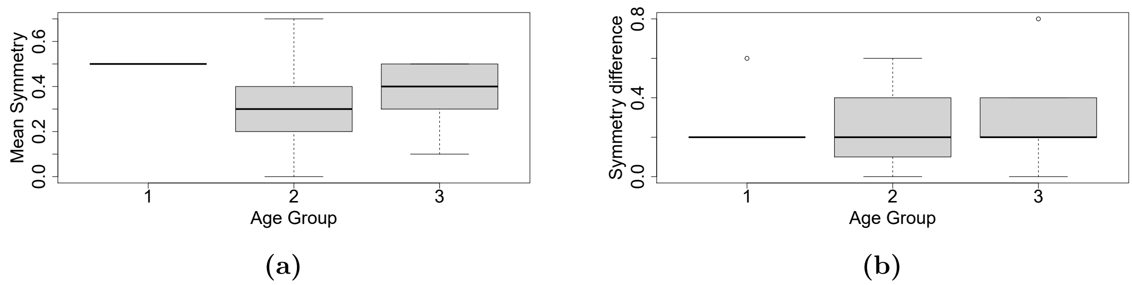


Figure 6.3: a) Distribution of mean symmetry levels per age group, b) difference in the chosen symmetry levels by age group

mean level of symmetry based on the age cluster [F(1,31)=0.777, p = 0.385]. This backs up existing research, which suggests that symmetry is a constant factor across humans. Another interesting aspect is the relatively low average across all participants (0.36), which indicates that less ordered sculptures may be preferable.

Some symmetry levels were impossible to reach in individual sets of choices (e.g. 0.5) therefore if the participant prefers a level of 0.5, this ideally results through the choice of sculptures with symmetry values of 0.4 and 0.6, respectively in the two stages, potentially explaining why a difference was present between the two sets of choices, Figure 6.3b. However, the similarity of the two paths taken by a participant also indicates whether symmetry was the main basis for choices being made in the system, as it could reasonably be expected that the participant would take the same or very similar route to reach the final value.

6.5.2 Participant behaviour

On top of the symmetry preference levels, more information can be obtained from the data collected, such as how an individual's preference can influence their confidence.

By considering each set of choices as a five-dimensional vector, where each value represents the level of symmetry selected from each pair of sculptures, the Euclidean distance can be calculated as a measure of how different the paths taken by the participant were, representing their confidence in the level of symmetry they chose.

Figure 6.4a shows the results split by age group: the distance between the paths was very low and consistent across all participants. If the distance is considered in terms of the preferred symmetry level, shown in Figure 6.4b, a clear pattern emerges, where the lower the preferred level of symmetry, the more consistent the participant was in their choices. Correlating these two values suggests that there is a significant effect of the level of symmetry preference on how confident the participants were with their choices. We found a moderate correlation, but which is highly significant, r(31) = .47, p = 0.005.

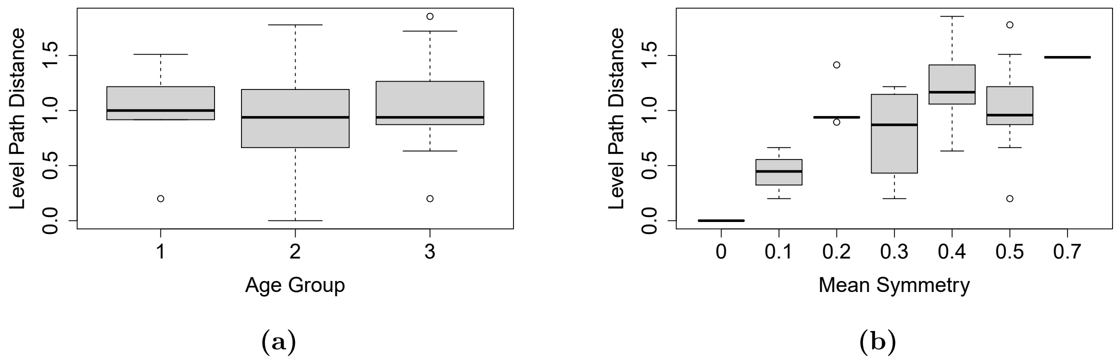


Figure 6.4: a) The difference between the two sets of choices per age, b) The difference between the two sets of choices per preferred symmetry level

This is further supported by inspecting the individual paths to determine how closely they followed an ideal route within the system based on the participant's preferred symmetry level. For example, if a participant prefers a level of 0.2, this can be achieved after making a single choice within the system. However, it would be more likely that the participant would choose the path which has the closest symmetry value sculpture for their preferred level; in this scenario, the expected path would be to choose the 0 symmetry value sculpture up until the last set of presented sculptures as this is the closest value to their preference. For users who chose a low level of symmetry, this often matched the path they chose exactly or with very minor discrepancies in the final step where both sculptures shown are of equal distance to the participant's optimal choice, as shown in Figure 6.5.

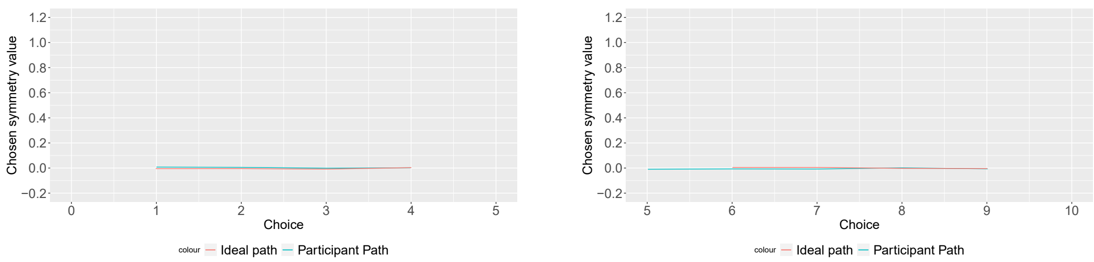


Figure 6.5: Distance between optimal and participant paths based on the selected level and the mean symmetry, left) First set, right) Second set

However, the adherence to the ideal paths is less consistent for higher levels of symmetry preference, as shown in Figure 6.6. It is visible that individuals are more confident about the symmetry level preference at the lower end of the spectrum. This, combined with mid-levels of symmetry being perhaps less easily perceived, suggests that participants preferring higher levels may not be as confident.

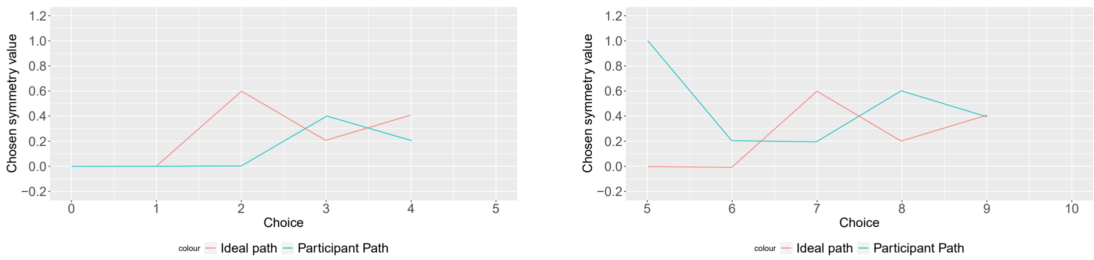


Figure 6.6: Distance between optimal and participant paths based on the selected level and mean symmetry, left) First set, right) Second set

6.6 Discussion

The consistency of the preferred symmetry level reveals two crucial aspects; firstly, the choices were being made, at least to some degree, based on the level of symmetry in the sculptures, which backs up existing research that indicates that symmetry is both an important part of the aesthetic judgement process and that it is relatively universal.

This data has two important effects when applied to auto-generating sculptures: rather than trying to display different levels of symmetry, aiming to keep the symmetry consistent may be a better approach for generating aesthetically pleasing artwork. Obtaining a baseline value for symmetry preference is vital as this can affect how the symmetry should be treated. Individuals did not always follow the ideal path for their chosen symmetry level, which means that variation in the symmetry level can be shown to a user while generating artwork. However, the amount of variation is highly dependent on the user's preferred level of symmetry: if an individual is found to prefer lower levels of symmetry, this consistently represents their choices and variation of this value should be presented, whereas, in the case of preference for higher levels of symmetry, a wider range of values should be displayed.

The average level of preferred symmetry is relatively low, around 0.3. This has implications for the equation proposed by Birkhoff, according to which in order to achieve a higher aesthetic rating, the complexity also needs to be a lower value. Intuitively, this does not seem to make sense as it would lead to boring pieces of art, suggesting that this measure of aesthetic value is not as relevant to sculptures and 3D art as it may be for 2D artwork.

6.7 Conclusion

Within this chapter, the Mine Cart VR environment was applied to extracting individual preferences for the presence of symmetry within virtual 3D sculptures. From the collected data, several interesting patterns were identified and the environment was validated as a suitable method for extracting individual user preferences. The data collected yielded information which can be applied to auto-generating sculptures. The information is also important for identifying other aspects which may be included in the assessment of sculptures and forms the basis for understanding how these other aspects may influence the preferred level of symmetry.

The closeness of the choices made in each stage indicates that the participants mainly made choices based on the symmetry, and the system reliably measured user preferences.

Whilst the presence of symmetry has been shown to be a subjective aspect, Chapter 5, its value may need to be handled differently. When considering which category symmetry value falls into, the similarity found across all participants suggests that symmetry fits nicely into the constant category and does not have to change or be re-measured to generate aesthetically pleasing sculptures. In addition, the consistency shown between the two sets of choices demonstrates that any interfering aspects do not change within small timescales, both of these reveal that symmetry can comfortably fit into the constant category. However, if a longer period of time is taken to describe someone's preferences, e.g. over months or years, then the same assumptions cannot be made.

The actual level of symmetry chosen by the participants had a surprisingly small variation across the age ranges. This is important when considering auto-generating 3D artwork as, from a general standpoint, a small range of values can be used across multiple contexts. The distance between the two paths the participant chose also impacts how symmetry can be used to auto-generate 3D artwork. The confidence the participant had in their choices was significantly affected by their preferred level of symmetry, where the lower the level of symmetry was, the more confident the participant was in their choices, indicating less variation in symmetry should be presented to people in this category.

Multiple attributes are used within aesthetic judgement, not just symmetry and the lack of a more complete understanding of aesthetic preference is a limitation of this system. This also prevents any conclusions from being drawn about how symmetry may be affected by the presence of other attributes within the sculpture. In addition, other forms of symmetry, such as colour symmetry were not investigated, so it is unknown whether the same principles would apply.

These results help to partially answer Question 4, allowing specified rules to be implemented, which help to model a particular individual's preferences for symmetry. For example, suppose someone is found to prefer low levels of symmetry. In that case, they may need to have symmetry handled in a different manner than others whose preferences fall closer to the overall average. These rules will not be included in the initial iteration of the aesthetic model presented in Chapter 8, as only having a single aspect handled in this manner would have a limited impact. However, as the knowledge around other aspects investigated in this project increases, these rules can be introduced to more completely model individual aesthetic preference.

Chapter 7

Formalising aspects of aesthetic appreciation

Following on from collecting the aesthetic judgement data and categorising the aesthetic aspects in Chapter 5 and before an aesthetic model of individual preference can be established in Chapter 8. The identified aspects need to be formalised into quantitative measures, which can be used to help identify an individual's preferences and generate matching sculptures. Three measures will be formalised within this chapter, representing one from each of the constant, subjective and transient categories: connected, calm and friendly. The measures chosen fell within the required range of values as discussed in Chapter 5 and cover all three categories, which enables a more detailed discussion of the aesthetic modelling process in Chapter 8, other aspects such as complex and order are not considered due to the wide body of research already available on these terms. Each aspect will have a corresponding algorithm designed to determine how much a sculpture displays the aspect. Validation for each formalised measure will be obtained, establishing whether the formalised measures represent human understanding of the aesthetic tags. Together, these complete the answer to Question 3 and will be utilised in Chapter 8.

7.1 Identifying suitable sculptures

Within Chapter 5, 18 sculptures were shown to participants, who were asked to assign aesthetic tags to the sculptures depending on their displayed aspects. These sculptures are shown in Figure 7.1.

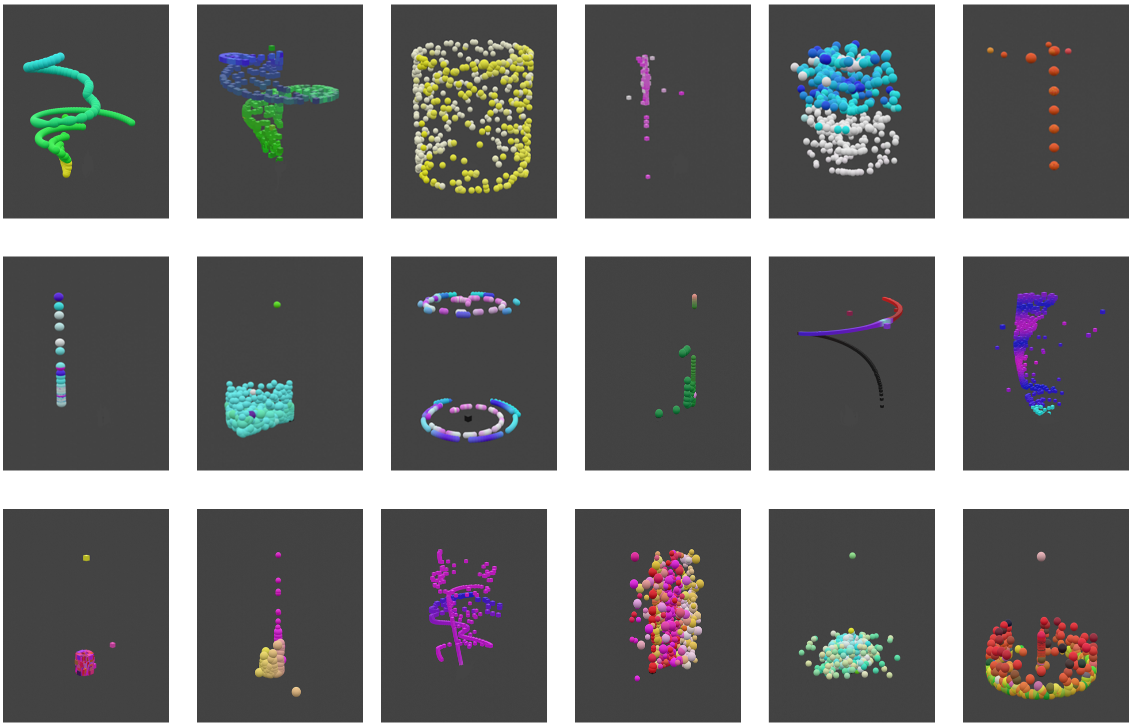


Figure 7.1: Sculptures available for participants to judge

The tag assignment details were collected from these sculptures, shown in Table 7.1. From these results, nine aesthetic concepts have been identified as potential items to be formalised: Connected, Busy, Interesting, Ordered, Complex, Bright, Angular, Friendly and Calm. These items met the threshold of being discriminating enough and did not have a high application difficulty, as discussed in Chapter 5. Ordered and complex will not be considered as significant research has been completed on formalising these aspects, and it would be unlikely that anything substantial would be added to this analysis.

One interesting aspect to note is that the same sculptures have the least amount of applications of the tags. This may be because these were the least aesthetically pleasing sculptures and, therefore, were harder to describe by the participants, making the formalisation process more difficult.

The accuracy of the measures does not need to be high due to their use within the aesthetic preference model. All that is needed is that the terms are described by the measures. The model allows a participant to determine their individual preference levels of values for these measures, i.e. it is not important to say 1 is a high level for connected and 0 is a low level, as an individual may prefer 0.3 as their maximum value, which represents

	C	E	Diff	Disc	Positivity
Dynamic	0.32	5.78	0.68	0.48	0.22
Curved	0.26	4.67	0.74	0.77	0.31
Simple	0.23	4.17	0.77	-0.47	-0.27
Connected	0.23	4.06	0.77	0.81	0.18
Busy	0.18	3.22	0.82	0.42	0.07
Interesting	0.18	3.17	0.82	0.73	0.11
Ordered	0.17	3.00	0.83	0.74	0.11
Complex	0.16	2.94	0.84	0.67	0.13
Bright	0.16	2.83	0.84	0.71	0.08
Disordered	0.15	2.67	0.85	-0.36	-0.08
Unnatural	0.15	2.61	0.86	-0.07	0.00
Static	0.14	2.50	0.86	-0.31	-0.08
Angular	0.12	2.11	0.88	0.36	0.07
Friendly	0.12	2.17	0.88	0.59	0.10
Calm	0.12	2.22	0.88	0.61	0.04
Warm	0.12	2.17	0.88	0.00	0.08
Unfriendly	0.12	2.11	0.88	-0.45	-0.04
Cold	0.11	2.06	0.89	0.10	-0.03
Boring	0.11	2.00	0.89	-0.73	-0.17
Loud	0.11	2.00	0.89	-0.27	-0.01
Separate	0.10	1.89	0.90	-0.48	-0.05
Quiet	0.08	1.39	0.92	-0.07	-0.03
Original	0.08	1.39	0.92	0.43	0.05
Natural	0.08	1.50	0.92	0.49	0.09
Sophisticated	0.07	1.28	0.93	0.27	0.04
Drab	0.06	1.00	0.94	-0.81	-0.11
Unrefined	0.04	0.78	0.96	-0.52	-0.08
Unoriginal	0.03	0.61	0.97	-0.61	-0.08

Table 7.1: Consistency (C), Endorsement (E), Difficult of application (Diff), Discrimination (Disc) and positivity rating for each tag, green-highlighted items represent terms which meet the criteria for further evaluation

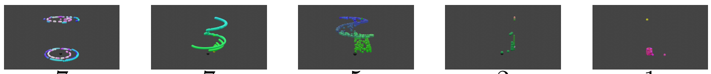
the right level for that individual. Ideally, the measures will correspond closely with the human understanding of the terms.

7.2 Connected

Connected is the first term to be considered, some examples of sculptures with high and low assignments are shown in Figure 7.2.

The role of connectedness in art and design can be linked to Gestalt psychology (Behrens, 1998), where it is the term which refers to the ability for humans to see connections in disjointed objects (Wertheimer, 1938). The Gestalt principles have widely been applied to art and design, and connectedness is a common guiding principle for many applications such as UI design. Over time, this theory has been expanded, eventually leading to the term Uniform Connectedness (Palmer and Rock, 1994), which considers regions to be connected if they share common properties such as colour, texture or motion (Wagemans et al., 2012). This theory seems particularly apt considering how the AGP generates the sculptures and can be used as the basis for formalising the measures in this project. Limited properties are available to choose as a candidate for this, with only colour and position being the two main variables amended within the AGP. The concept of colour connectedness is highly related to colour harmony, which has been described by Arnheim (2009, pp. 347-349) as "all colors of a composition must fit together in a unified whole if they are to be relatable to one another". Harmony is a concept which has long been considered an important part of art (Shelley, 2005; Margolis, 2005) and suggests that colours which are closely linked to each other may satisfy the requirements of being considered connected. Figure 7.2 shows some common properties of the sculptures, which could be used as the basis of the formal measure.

One complicating aspect is that the tag was applied at least once to all sculptures, which indicates that, at least to some extent, all of the sculptures appear connected and confirms the place connected has within the constant category, this suggests that all sculptures will display a value for the connected measure.



Assignments	7	7	5	2	1
Average Max. item distance	15.8	18.95	20.9	10.98	14.41
Avg. item distance	10	5.44	14.43	5.78	8.3
Avg. colour distance	0.09	0.06	0.05	0.04	0.08
Total colours	20	10	13	13	52
Connected Item ratio	0.56	0.25	0.21	0.36	0.33

Figure 7.2: Connected sculptures shown in descending order of total assignments of the connected tag (highest at the left)

From the data in Figure 7.2, there seems to be a clear indication of which attributes of the sculpture contribute to the application of this tag. Firstly, the form of the sculpture, where the distance between the items needs to be low, making the sculpture appear to be a single solid object, matching the first two sculptures in the example above. However, this does not hold true for all of the sculptures, the third item has a higher level than the first. As discussed, it is expected that colour may also form part of the judgement. This matches the third sculpture, which has a lower colour distance than the first two (how far apart these items are in RGB space). The same is true of the fourth sculpture, which has a higher colour distance but a lower physical distance, making it clear that both factors must be considered when creating a quantitative measure. Calculating the ratio between how many elements within the sculpture have another element within close spatial or colour proximity together compared to the total number of elements in the sculpture leads to additional measurement, termed Connected Item Ratio. This has a positive correlation to the number of assignments across all sculptures, as shown in Figure 7.3 and was chosen as the basis for the connected measure.

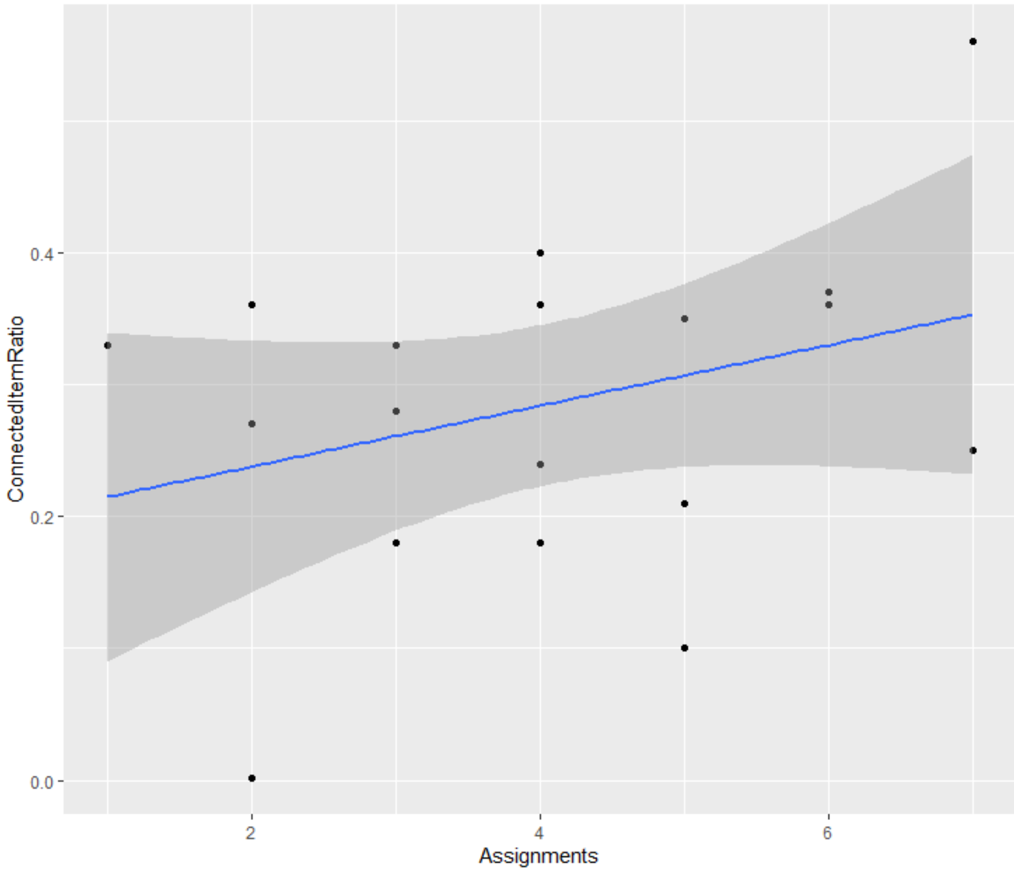


Figure 7.3: Scatter plot comparing the total assignments to the ratio of spatially and colour-connected items for each sculpture, with confidence level indicated in the shaded region.

One aspect amended when calculating the connected value of a sculpture was to consider the sculptures as a connected set of items. Building trees based on how physically close the items are and how close the colours were in LUV space. LUV space is used as it should more accurately represent the human interpretation of the colours. These trees represent a collection of items which are connected; the higher the number of trees identified for a sculpture, the less connected the item will appear. This process is outlined in Algorithm 2. The algorithm has two parts, both following the same outline. While there are still points that have not been added to a tree, a random starting point is selected, from there, children are recursively added to the tree depending on whether the point has any neighbours within a specified threshold within physical or LUV space. If no items are close enough, the node is left as a leaf, and once all leaf nodes have been identified for a tree, the process repeats, and a new tree is created.

Algorithm 2 Algorithm for calculating how connected a sculpture appears

```
Points[] ← [totalItems]
totalTrees ← 0
matchedCount ← 0
while matchedCount < totalItems do
    PositionTree ← Tree
    startingPoint ← Rand(0, totalItems)
    posNode ← PositionTree.AddNode(startingPoint)
    matchedCount + +
    posNode.Children ← GetClosestChildrenRecursive(startingPoint, posNode, "Position")
    totalTrees + +
end while
PositionVal ← 1 − (totalTrees/count(totalItems)
totalTrees ← 0
matchedCount ← 0
while matchedCount < totalItems do
    ColourTree ← Tree
    startingPoint ← Rand(0, totalItems)
    colourNode ← ColourTree.AddNode(startingPoint)
    matchedCount + +
    colourNode.Children ← GetClosestChildrenRecursive(startingPoint, colourNode, "Colour")
    totalTrees + +
end while
ColourVal ← 1 − (totalTrees/count(totalItems)
return (ColourVal + PositionVal)/2
```

Algorithm 3 Recursive method for building the colour and position trees

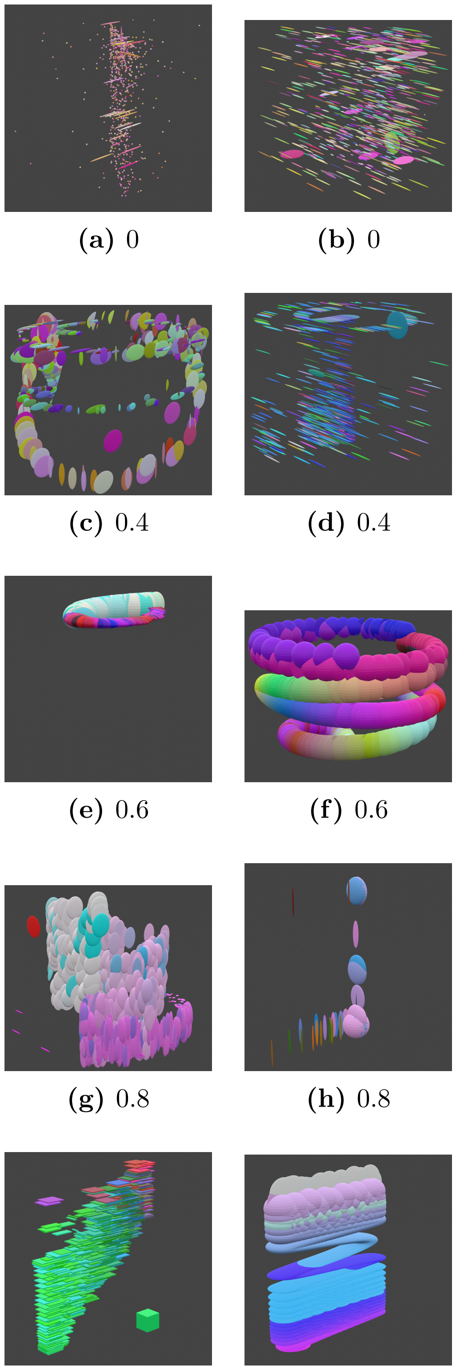
```
procedure GetClosestChildrenRecursive(startingPoint, node, type)
    neighbours ← GetNeighbours(startingPoint, type)
    if neighbours.Any() then
        return null
    else
        childNodes ← List()
        for each neighbour ∈ neighbours do
            matchedCount + +
            childNode ← node.addChild(neighbour)
            childNodes.Add(childNode)
            childNode.children ← GetClosestChildrenRecursive(neighbour, childNode, type)
        end for
        return List(childNodes)
    end if
end procedure
```

The value for each part is then calculated as the inverse of the total number of trees divided by the total number of items. This means that if all points fall within a single tree for physical and colour space, the sculpture is completely connected, and if the number of trees is the same as the number of items, the sculpture is not connected. The first part of the Algorithm looks at the form connections, and the second part looks at the colour connections. Once both parts have been calculated, the score for the sculpture is the average score across the form and colour connectivity values. Sculptures which have been generated at different levels of connectedness are shown in Figure 7.4.

7.3 Calm

The term calm is very subjective and can mean different things to different people; however, calm is often considered an antonym for stress. Figure 7.5 shows the sculptures with the most and least applications of the calm tag, alongside some common attributes calculated for the sculptures.

There have been many examples of utilising art to help relieve stress; in many studies, the presence of fractals has appeared to have this effect, often being cited as the reason why walking in nature has a relaxing effect (Taylor, 2006; Taylor et al., 2011; Tyler, 1996). However, fractal dimension has been studied in various contexts, suggesting that it may not be suitable for this project to determine the level of calm a sculpture exhibits. For example, it can be an indication of symmetry (Friedenberg et al., 2022), which, due to the rotational nature of sculptures generated through the AGP, may lead to any fractal-based measure being misleading. The link between fractals and their appearance in nature may also not be suitable for this context; naturalness was a term presented in Chapter 5 and was found to be very difficult to apply to the sculptures, making any link between the two largely speculative. In addition, calculating fractal dimension has also been used as a measure of complexity (Johnson et al., 2019; McCormack et al., 2021; Spehar et al., 2003), indicating again that they may not be suitable for use to determine the level of calmness displayed by the sculpture, as instead, the measure may be more related to one of these other uses, potentially making the calm measure less effective or even meaning any proposed measure was not computing how calm a sculpture appeared. For these reasons, other aspects have been considered when attempting to formalise a measure of the calmness



(a) 0 **(b)** 0

(c) 0.4 **(d)** 0.4

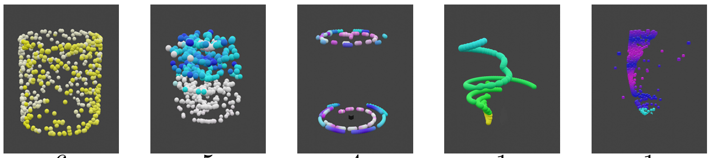
(e) 0.6 **(f)** 0.6

(g) 0.8 **(h)** 0.8

(i) 1 **(j)** 1

Figure 7.4: Created connected sculptures in ascending value order

in sculptures. However, in addition to the fractal dimension, other aspects also influence how calm something appears, for example, the colours used, with neutral colours often being preferred over bright colours (Stone, 2003), along with a limited variance of the colours being shown. Too much contrast between the items is also less likely to have a calming effect. This can be seen with the sculptures shown in Figure 7.5. The highest-rated sculptures have limited colour variation and have brighter colours present, contrasted with the items with the lowest application, which use less bright colours or have multiple variations in colour.



Assignments	6	5	4	1	1
Total colours	4	9	21	10	12
Avg. brightness	0.55	0.76	0.66	0.22	0.34
Max. brightness	1	1	0.99	0.89	0.7
Avg. saturation	0.99	0.69	1	0.99	0.96

Figure 7.5: Calm sculptures and associated values

From these stats, the assignments correlate to the brightness. The average brightness roughly decreases as the number of assignments does, and all sculptures have a high average saturation. The brightness and saturation have been considered as an important aspect which can influence the visual appearance of a colour, in terms of being warm or cold (Arnheim, 2009, pp. 348), and a similar pattern is found within the collected data. The average brightness compared to the assignments for all sculptures is shown in Figure 7.6. The average saturation should also be included as part of the measure to keep the items in line with the data collected and prevent any unintended changes caused by saturation differences. Within the measure Algorithm, 4, a threshold of 0.8 has been added for the brightness to prevent the created items from being entirely white, something not present in the sculptures used for this analysis. This is introduced because all of the sculptures display colours, similarly to the saturation, the same values should be replicated. This leads to highly calm sculptures having a high saturation and a brightness level of 0.8, non-calm sculptures would have a low saturation and a low brightness level. Examples of the sculptures created using these equations are shown in Figure 7.7.

7.4 Friendly

Similar to the other aspects being formulated, the sculptures with the most and least applications of friendly are shown in Figure 7.8. The concept of friendliness is much less well defined than the previous two concepts, with the meaning behind a piece of artwork often

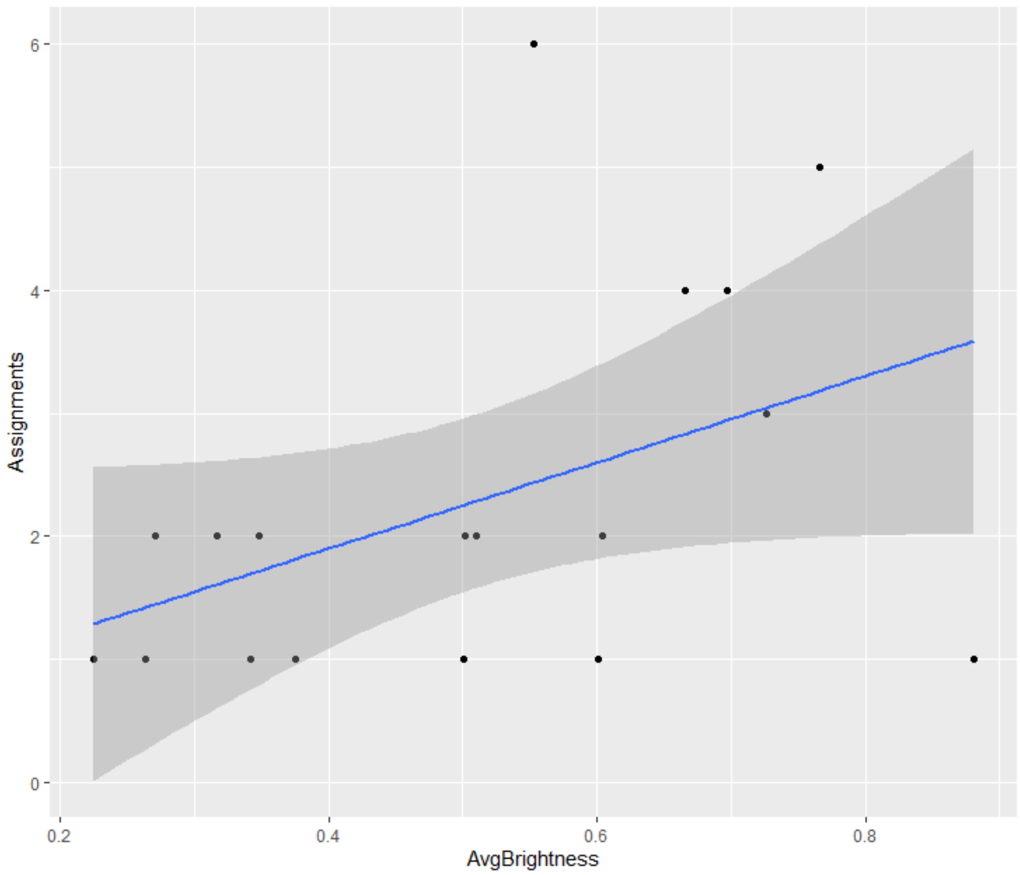


Figure 7.6: Scatter plot comparing the total assignments to the average brightness for each sculpture, with confidence level indicated in the shaded region.

Algorithm 4 Algorithm for calculating how calm a sculpture appears

$Points[] \leftarrow [totalItems]$
$saturationSum \leftarrow 0$
$brightnessSum \leftarrow 0$
for each $point \in Points$ **do**
    $saturationSum += GetSaturation(point)$
    $brightnessSum += GetBrightness(point)$
end for
$avgSaturation \leftarrow saturationSum/count(totalItems)$
$avgBrightness \leftarrow brightnessSum/count(totalItems)$
$brightnessError \leftarrow calcRelativeError(avgBrightness, 0.8)$
return $(avgSaturation + brightnessError)/2$

influencing how it may be judged, making it difficult to discern any reliable rules. Due to this, a similar approach to the definition of Calm will be utilised as a starting point. Two noticeable differences exist between the sculptures with the highest and lowest assignments. The first is that the items with the most assignments have a similar colour range to those

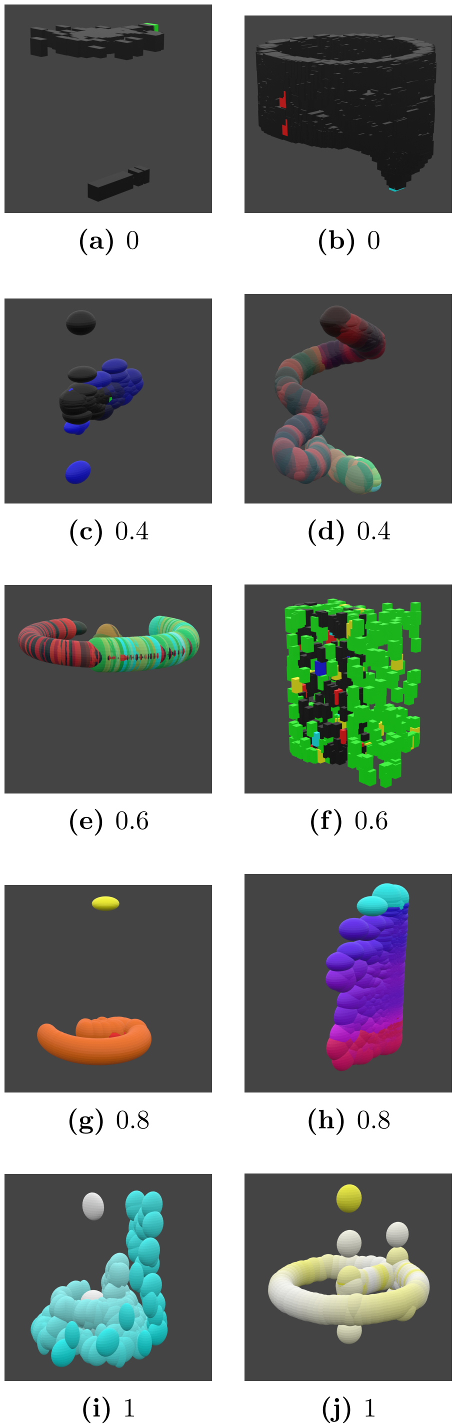
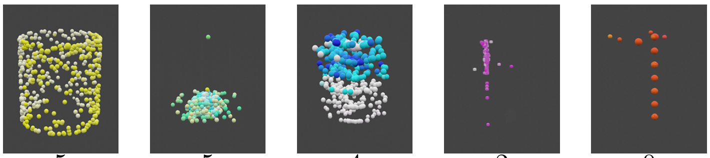


Figure 7.7: Created calm sculptures in ascending value order

identified for the Calm aspect; however, the lowest assigned items are very simple, suggesting that friendly may be related to the colour combinations of the sculptures. This aligns with aspects that contribute to something appearing as warm or cold, which, in a similar sense to friendly, is an expressive quality a sculpture may have (Arnheim, 2009, pp. 370). It can be considered as a combination of the dominant hue and other colours it is paired with, providing the starting point of how friendly may be classified, comparing how many times the tag was assigned to the sculpture to attributes of the dominant colour and to all the colours present in the sculpture.



Assignments	5	5	4	2	0
Total detected colour modes	4	16	8	3	5
Max brightness	1	1	1	1	0.89
Avg. brightness	0.55	0.72	0.77	0.6	0.5
Dominant colour brightness	0.89	0.91	0.68	0.42	0.43
Colour ratio	1	1	0.63	1	1

Figure 7.8: Friendly sculptures and associated values

The dominant colours can be obtained by running a mean shift algorithm (Derpanis, 2005) on the sculptures in the RGB colour space. The most popular mode containing the highest number of points is deemed the dominant colour as it is the colour which will be the most visible in the sculpture. The relationship between the dominant colour brightness and the assignments is shown in Figure 7.9. This positive correlation indicates that there is a relationship between the friendliness of a sculpture and the brightness of the dominant colour, where the higher the brightness of the dominant colour, the more friendly the sculpture appears. Once again, the colours feature heavily within the assessed sculptures, making preserving them in all the measures a vital consideration. In an alternative approach to the calm measure, for the friendly sculptures, the dominant colour is calculated, excluding the white and black elements, and in addition to ensuring that the sculptures display a high level of colour, the ratio of non-black and non-white items to the black and white items is maximised, this is termed the colour ratio within Figure 7.8.

To calculate a value for the friendliness of a sculpture, Algorithm 5 is used. Within this process, the Mean Shift algorithm determines how many distinctly different colours are present within a sculpture. It then obtains the dominant colour in terms of how many points in the sculpture have the same or similar colour and gets the average brightness of these dominant colours. The value for friendly is then taken as the average of the dominant colour brightness and the ratio of how many items do not have a brightness value of 0 or 1. This process looks to maximise both the number of items which are not black or white and

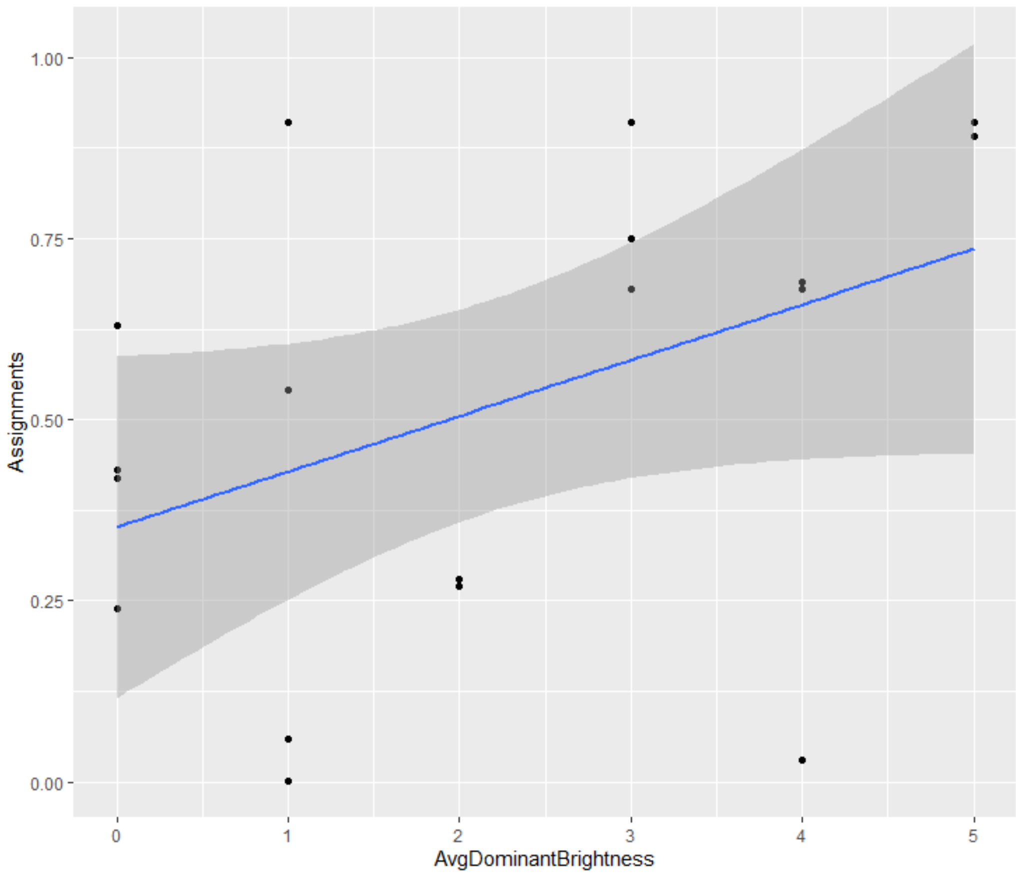


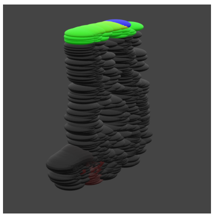
Figure 7.9: Scatter plot comparing the total assignments to the average brightness of the dominant colour for each sculpture, with confidence level indicated in the shaded region.

Algorithm 5 Algorithm for calculating how friendly a sculpture appears

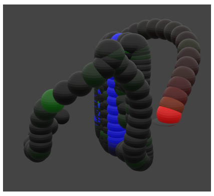
$allModes \leftarrow MeanShift(items)$
$dominantColours \leftarrow getMaxItemsInMode(allModes)$
$avgBrightness \leftarrow Average(dominantColours.Brightness)$
$totalColourItems \leftarrow 0$
for each $point \in Points$ **do**
  if $!point.IsBlack$ **then**
    if $!point.isWhite$ **then**
      $totalColourItems++$
    end if
  end if
end for
$colourRatio \leftarrow (totalColourItems/count(items))$
return $(avgBrightness + colourRatio)/2$

---

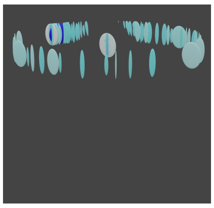
the brightness of the dominant colour shown within the sculpture. Examples of friendly sculptures calculated using this method are shown in Figure 7.10.



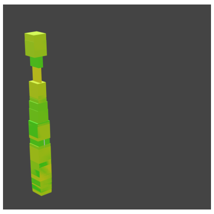
(a) 0



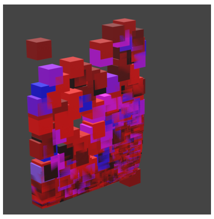
(b) 0



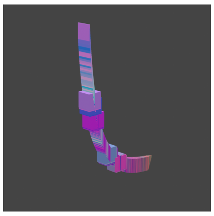
(c) 0.4



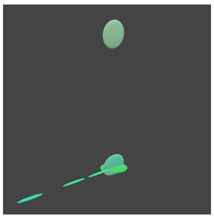
(d) 0.4



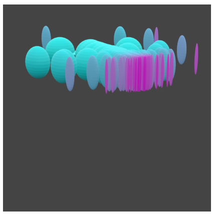
(e) 0.6



(f) 0.6

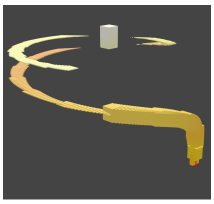


(g) 0.8



(h) 0.8

(i) 1



(j) 1

Figure 7.10: Created friendly sculptures in ascending value order

7.5 Validating the measures

While the measures may fit the sculptures in Figure 7.1. Two outstanding areas need to be considered before the measures may be suitable for use within the aesthetic model. The

Rating	Connected	Calm	Friendly
1	Not connected	Not calm	Not friendly
2	A little connected	A little calm	A little friendly
3	Somewhat connected	Somewhat calm	Somewhat friendly
4	Mostly connected	Mostly calm	Mostly friendly
5	Entirely connected	Very calm	Very friendly

Table 7.2: Available terms participants could apply to each group of sculptures

first is whether the measure can be applied to a broader range of sculptures and still hold the same relationship. More importantly, whether the measures are representative of the human understanding of the terms needs to be known. In order to validate the measures, a human-subject experiment was run, utilising the gallery VR environment introduced in Chapter 3, to collect data about whether humans understand the terms in a similar way to the proposed measures.

7.5.1 Environment and Process

The same VR environment and process described in Section 3.7 was used. The differences between this iteration and previous uses of the VRE came in the form of how many sets of sculptures the participant would be grouping and the instructions and tags available to assign to the groups. In this experiment, the participant repeated the main process of grouping sculptures together three times, being presented with a different set of sculptures each time and asked to consider a different measure. The first looked at identifying connected sculptures, the second calm sculptures and the final considered friendly sculptures. Instead of a variety of aesthetic description terms, the participants were asked to assign one of five terms to a group depending on how well the sculptures within the group exemplified the current measure. These terms corresponded to a rating, shown in Table 7.2.

7.5.2 Sculpture generation

The sculptures utilised in the validation experiment were all generated using the corresponding formalised measures: for connected, Algorithm 2 was used to calculate how connected the sculptures were; for calm, Algorithm 4 was used and for friendly, Algorithm 5 was used. To obtain examples of sculptures at intervals of 0.2 across the spectrum of each measure, six genetic algorithm runs were performed to create items which displayed symmetry between 0 and 1 in steps of 0.2, similar to the process detailed in Chapter 6.

From each run of the GA, using parameters shown in Table 7.3, the sculptures displaying

Total Generations	250
Mutation Probability	0.7
Initialisation Method	Ramped half & half
Selection Method	Tournament (k=3)
Population Size	25
Fitness Measure	Connected—Calm—Friendly
Tree size	4

Table 7.3: Evolutionary Algorithm parameters

values which approximated the target values were automatically chosen. This generated a high number of sculptures at each target value, this list was reduced manually by the authors to only include two examples of each level of the target value, giving a total of 12 sculptures for each measure.

7.5.3 Results

A total of 21 participants completed the experiment (3 female, 18 male) between the ages of 18 to 40 (m=26.0). To determine whether the formalised measures match up with the human perception of the terms, the ratings provided by the participants were compared to the calculated value for the specific sculpture.

For each measure, an independent t-test was performed to compare the average rating applied by the participants and the calculated ratings for each sculpture. Each measure found a significant difference between these two groups of ratings, connected: $t(21.99) = 17.87$, $p < 0.001$, calm: $t(11) = 19.87$, $p < 0.001$ and friendly: $t(11.81) = 16.08$, $p < 0.001$; indicating, that all the correlations were not obtained by chance.

Connected

The entire collected ratings for the connected measure are shown in Table C.1. Performing a Pearson's Correlation between the average participant rating and the calculated rating returned a significant strong positive correlation between these two sets of values, $r(10) = 0.57$, $p = 0.05$, further indicated by the scatter plot shown in Figure 7.11.

Calm

All collected ratings for the calm measure are shown in Table C.2, and performing a Pearson's Correlation between the average participant rating and the calculated rating returned a significant strong positive correlation between these two sets of values, $r(10) = 0.86$, $p <$

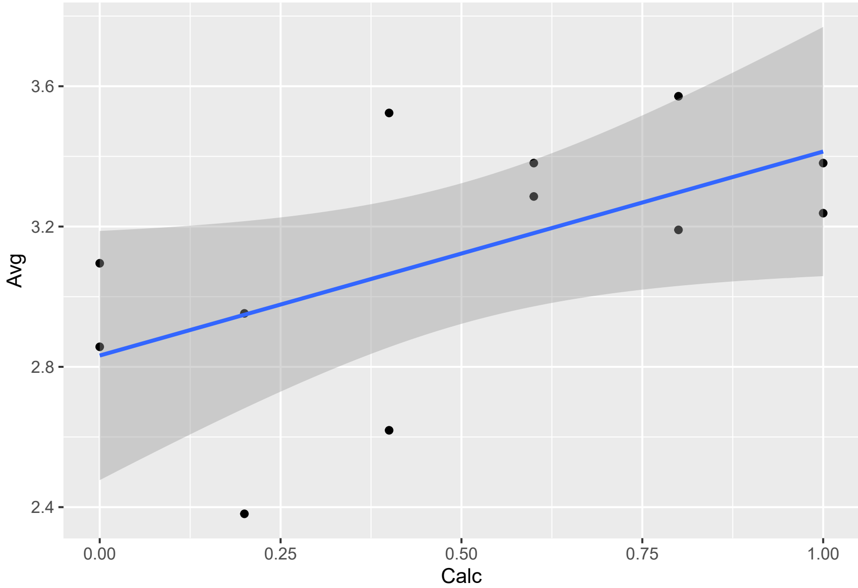


Figure 7.11: Scatter plot comparing the average participant ratings (Avg) to the calculated (Calc) ratings for each connected sculpture, with confidence level indicated in the shaded region.

0.01, further indicated by the scatter plot shown in Figure 7.12.

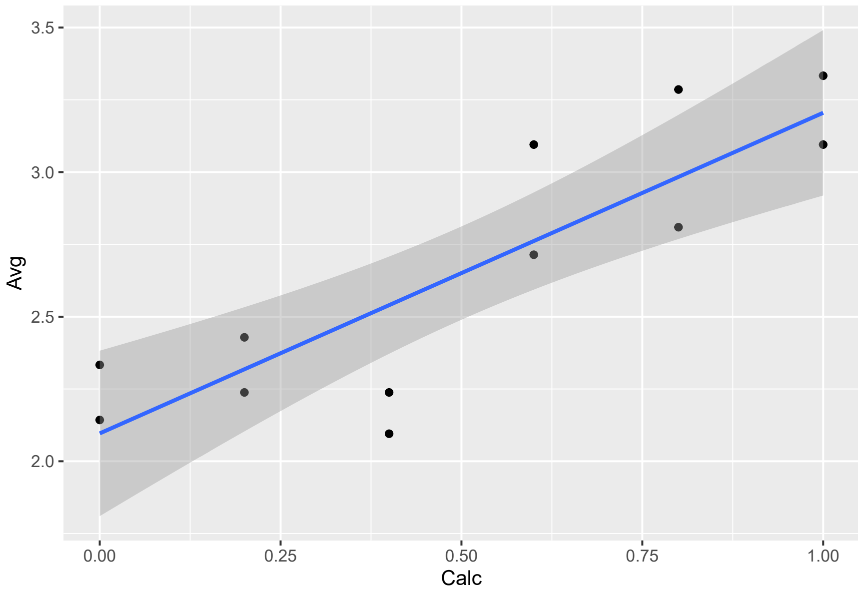


Figure 7.12: Scatter plot comparing the average participant ratings (Avg) to the calculated (Calc) ratings for each calm sculpture, with confidence level indicated in the shaded region.

Friendly

The collected ratings for the friendly measure are shown in Table C.3. Through performing a Pearson's Correlation between the average participant rating and the calculated rating returned the highest positive correlation across all three measures, $r(10) = 0.86$, $p < 0.01$, further indicated by the scatter plot shown in Figure 7.12.

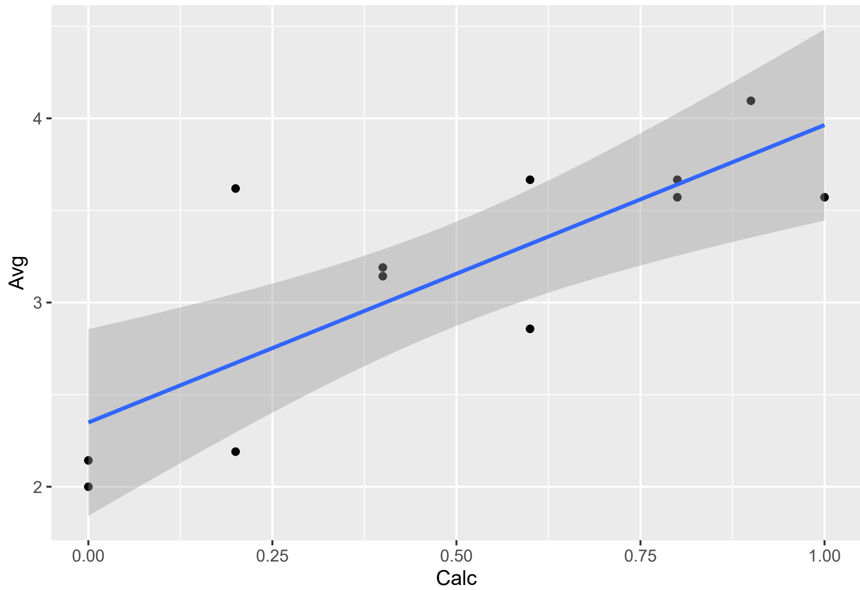


Figure 7.13: Scatter plot comparing the average participant ratings (Avg) to the calculated (Calc) ratings for each friendly sculpture, with confidence level indicated in the shaded region.

7.5.4 Discussion

The results obtained from the experiment all showed a clear pattern across all of the formalised measures. Each one showed a significant strong positive correlation between the rating obtained from the participants to the calculated ratings for each sculpture, indicating that the measures represented the human understanding of each of the terms when being applied to virtual sculptures and in line with the rules specified in Chapter 1, can be considered successful. The success also helps to show that the item analysis approach used within Chapter 5 to identify which terms were most suitable for formalising was appropriate as the sculptures allowed the features which contributed to each term to be identified.

Whilst the correlation was not exact, this is to be expected when accounting for several factors, such as the participants' subjectivity on the definition of the term. For example, the calm measure prefers lighter colours, which may not match every person's understanding of the term, who instead may prefer other colours. In addition, it may also be due to the experimental setup, due to the lack of restrictions on art expertise, it may be possible that not all participants were able to identify the features which contributed to each measure. This may have been exacerbated by the unintentional presence of other non-tracked aspects, which may have masked the presence of the main aspect. Another indication of this is that generally while the participant ratings increased alongside the calculated ratings, the extremes of each measure were not reflected. This indicates that it may be challenging to identify sculptures lacking these aspects, e.g. not being connected. One final potential limitation is that the chosen attributes included in each measure may not be the best items to choose, and it remains possible that other attributes would receive a higher correlation.

In the case of the connected measure, the identification of the non-connected items may have been especially difficult as the sculptures were presented on a pedestal which would indicate that all the items on the pedestal were included as part of the same sculpture, and therefore to some degree could be considered connected. This may explain why the connected measure received the lowest correlation out of the three investigated measures.

7.6 Conclusion

Within this chapter, three aspects: connected, calm and friendly, were chosen from the identified terms in Chapter 5, and a formal measure was presented for each term, which allowed a value between 0 and 1 to be obtained for a sculpture indicating how well the sculpture exhibits the particular value. These measures were used to generate a range of sculptures which displayed all possible values for each of the formalised measures. In order to validate the measures, these sculptures were shown to human participants within the Gallery VRE, who were asked to group and rate the sculptures to determine how well the human understanding of the term matched the calculated value for the sculpture. The data analysis from this experiment indicates that all three measures match the human perception of the term with all having a significant high positive correlation between the human ratings and the calculated value for the sculptures. This indicates that all three measures can be considered successful and used within the theoretical model presented in Chapter 8.

Analysing the data from Chapter 5 led to specific methods to calculate the level of Connected, Calm and Friendly within the sculptures generated by the AGP. The chosen attributes all displayed a positive correlation to the number of assignments. However, each did have a high amount of variation, indicated by the confidence in the correlation. This highlights the level of subjectivity of all of the descriptive terms; each of these terms would have slightly different meanings to each individual, and it is this subjectivity that is being demonstrated with these results and potentially suggests that these initial measures are not wholly suitable to reflect these nuances in every scenario, and further data collection would be required in order to achieve this. However, the results obtained from presenting sculptures created using the final measures show the validity of this approach.

There was a high level of variation observed in the ratings observed for each measure, often the same sculpture received the highest and lowest ratings from different participants. This is similar to the results obtained in Chapter 6, where whilst generally, different age groups preferred similar levels of symmetry, there is still a high level of subjectivity of what value of the measures are preferred by individual people, and it is assumed that the same is happening with these measures. However, this needs to be investigated further, which can be done using the sculptures in the mine cart VRE, described in Chapter 3. This would allow a much deeper analysis of individual preferences, this information could then be used to extend the model presented in Chapter 8.

Using these measures within a Genetic Program to create sculptures representing the full spectrum of values also provides additional validation on the AGP, presented in Chapter 4. It was shown within that chapter that the AGP could generate sculptures which displayed a wide range of values across several existing measures, however, this assessment was only performed with a limited number of measures. With these newly formalised measures, it has also been shown that the process is capable of generating sculptures representing a wide range of values across other measures, further strengthening the conclusion that the AGP is versatile and suitable for investigating aesthetic preferences.

The three measures and the collected data complete the answer to Question 3. Not only were the aspects formalised, but these measures were also successful in closely matching the human understanding of the term when aesthetically judging a sculpture. It was also shown that the measures can be used within a Genetic Program in order to generate sculptures which displayed different levels of each measure. This means that the formalised measures are suitable for use within Chapter 8 to help model individual aesthetic preference and answer Chapter 4.

Chapter 8

Modelling individual aesthetic preferences of 3D sculptures

The work presented in this chapter has been adapted from the following publication:
E. Easton, U. Bernardet, and A. Ekárt. Modelling individual aesthetic preferences of 3D sculptures. In *Artificial Intelligence in Music, Sound, Art and Design: 13th International Conference, EvoMUSART 2024, Held as Part of EvoStar 2024, Aberystwyth, United Kingdom, April 3–5, 2024, Proceedings*. Springer, 2024.

Having formalised three aspects, it is now possible to apply these terms to help represent a user's preferences. In addition, by considering the data from a different perspective, helping to represent a particular art style. This chapter represents a complete answer to Question 4 and presents a method of representing individual aesthetic preferences as well as how this model can be utilised within a Genetic Program to semi-automatically generate sculptures unique to an individual. Modelling individual aesthetic preference is not a new approach to be used within this field, other examples exist, such as (Ekárt et al., 2012; Li et al., 2012; Machado et al., 2021; Romero et al., 2003), all of which are different to the approach presented here.

The starting point of building this model is the three categories of aesthetic aspects: constant, subjective and transient. The category determines how an aspect should be updated within the aesthetic model based on user feedback. The categories each aspect falls within from a user perspective have been identified in Chapter 5. Through using this information, a theoretical overview of how the categories can be used to model an individual's aesthetic preferences will be presented, along with examples that showcase how quickly individual preferences can be identified and subsequently influence the generation of unique sculptures. Finally, creating new sculptures in the style of a specified set of sculptures will be considered, with the categorisation and modelling process completed.

Both types of model update in line with changes in the preferences and generate unique sculptures for that user or within the specified art style. Several experimental designs will be suggested, which would enable the further investigation of the concepts introduced.

Only a theoretical outline of the model and some worked examples will be presented due to the lack of formalised measures. With the current set of measures, it would not be possible to closely represent aesthetic judgement, which would limit the effectiveness of experimental results with the model.

8.1 Creating an aesthetic model individual preference of 3D sculptures

Due to the eclectic nature of aesthetic judgement, using the positivity ratings collected in Chapter 7 as part of this model would not be appropriate. It should be possible for a user to generate both positive and negative items, keeping the process as flexible as possible. Instead, the category alone indicates how an aspect is treated within the model. It is important to note that the model works at the tag level, allowing different aspects to be tracked at different times. The values of each aspect will be handled in two ways; for constant and subjective aspects, an average value will be taken and used across user choices. The reason for using this process is to accommodate the information learned in Chapter 6, where the overall preference of a particular aspect is influenced by the user's choices. As multiple sculptures may be selected in each round of choices, the value calculated for each set of choices is the average of the measure across all chosen sculptures. If the aspect is not selected in a set of choices, the value is not updated to avoid influencing the model without enough information to accurately make the change. Due to how the transient aspects have no persistence in the model, only the values from the most recent set of choices will be used. The whole process is detailed in Algorithm 6.

The constant aspects are the most straightforward items to handle within the model; they are constantly present for the entirety of the time the user is generating new sculptures.

Algorithm 6 Algorithm for creating and updating the model of aesthetic appreciation

```
procedure UpdateModel(sculptures: Sculpture[], currentIteration: int, currentModel: Dictionary)
    choiceValues ← Dictionary
    for each sculpture ∈ sculptures do
        for each measure ∈ sculpture.measures do
            if notExists(choiceValues, measure.name) then
                choiceValues.Add(measure.name, measure.value)
            else
                choiceValues[measure.name]+ = measure.value
            end if
        end for
    end for
    if currentIteration == 0 then
        for each measure ∈ choiceValues do
            currentModel.Add(measure.name, measure.value/sculptures.count())
        end for
    else
        for each measure ∈ currentModel do
            if measure.isTransient then
                currentModel.remove(measure)
            else
                if choiceValues.has(measure.name) then
                    measure.lastUpdatedIteration ← currentIteration
                    average ← choiceValues[measure.name]/sculptures.count()
                    measure.value ← (measure.value + average)/2
                else
                    if measure.lastUpdatedIteration <= (currentIteration − 2) then
                        if measure.isSubjective then
                            currentModel.remove(measure)
                        end if
                    end if
                end if
            end if
        end for
        for each measure ∈ choiceValues do
            if measure.isTransient or measure.isSubjective then
                if not currentModel.has(measure.name) then
                    currentModel.Add(measure.name, measure.value/sculptures.count())
                end if
            end if
        end for
    end if
end procedure
```

The only changing aspect will be the value, following the rules outlined above.

The subjective aspects need to be handled in a manner which accounts for their changeability, with different combinations utilised at each stage of the generation process. However, as subjective aspects should be similar throughout an individual's choices, the tags will only be amended based on multiple choices. The model starts by adding all subjective aspects present in the initial choices of sculptures. Then, in the second choice, any additional subjective tags present in the choices will be added. In the third set of choices, any new subjective aspects will once again be added, however, if any of the aspects from the first set of choices have not been present in the current set or second set of choices, they will be removed from the mode. This last step is repeated for subsequent choices, keeping the subjective aspects up to date with the user preferences.

Finally, the transient items represent the most volatile aspects contributing to the judgement. These items are assumed to be constantly changing, and their presence in a previous selection, either the aspect or the value of the measure, is not considered enough evidence that the user wants to view these aspects within all of the sculptures they are presented. Due to this, the transient tags used within the model will just be taken from the most recent choices, this is also true for the values for the transient aspects, which will also have no persistence within the model. This approach solves three potential problems: situations where the presence of a tag is masked by others and was unintentionally selected by the user, the potential volatility of user choices and finally, the potential for the miscategorisation of the subjective and transient aspects.

The methods of determining which tags should be included in the model and at which values allows the model to converge separately on the user's preferred tags and values.

Applying this model to the Genetic Program requires an objective function to be determined that identifies the sculptures which most closely match the user model, guiding the evolutionary search. To achieve this, the Euclidean Distance will be used, Equation 8.1, which determines how well the sculpture meets the current user model. The objective function will be minimised during the running of the GP, which finds sculptures which are closer matches.

$$d(model_{user}, model_{sculpt}) = \sqrt{\sum_{i=1}^{n} (model^{i}_{sculpt} - model^{i}_{user})^2} \tag{8.1}$$

Total Generations	10
Mutation Probability	0.7
Max tree depth	5
Initialisation Method	Ramped Half and Half
Selection Method	Tournament (k=3)
Population Size	14
Fitness Measure	User model replication

Table 8.1: Evolutionary Algorithm parameters

Choices	Aspect	Category	Initial value
8.1a	Connected	Constant	0.78
8.1b	Order	Subjective	0.7
8.1c	Complexity	Subjective	0.8
8.1d	Calm	Subjective	0.72
8.1e	Friendly	Transient	0.66
8.1f			
8.1g			

Table 8.2: Initial model for example one

8.1.1 Example individual models

In order to demonstrate the model, three examples are presented here. Calculating the initial user profile, simulating user choices and presenting the final generated sculptures after two choices from each user, all starting from the same set of sculptures. The three profiles will make different choices from the initial set of sculptures; two will make the opposite choices to one another, and the third will select half of the items from the other two. This allows the final set of sculptures to be visually compared, showcasing how the process is capable of quickly generating unique sculptures based on individual choices. The worked examples will consider the aspects formalised in Chapter 7, along with the form symmetry and Normalised Kolmogorov Complexity calculations used in Chapter 4.

The initial sculptures have been generated through the AGP (Chapter 4), using the distance search approach used to generate the sculptures in Chapter 5, to ensure the range of the presented values is as wide as possible, these are shown in, Figure 8.1. As indicated by the results presented in Chapter 4, the range of values across all measures and sculptures is wide. Figure 8.2 shows the range of all sculptures across all formalised measures, with only friendly not being fully represented.

User one selects the first seven sculptures, and from this, the initial model can be calculated, represented in Table 8.2. User two selects the other seven sculptures with their entire model progression shown in Table 8.4, and user three selects four sculptures from the first user and three from the second user, with their full model shown in Table 8.5.

Using each of the initial models as the initial population, ten generations are run in

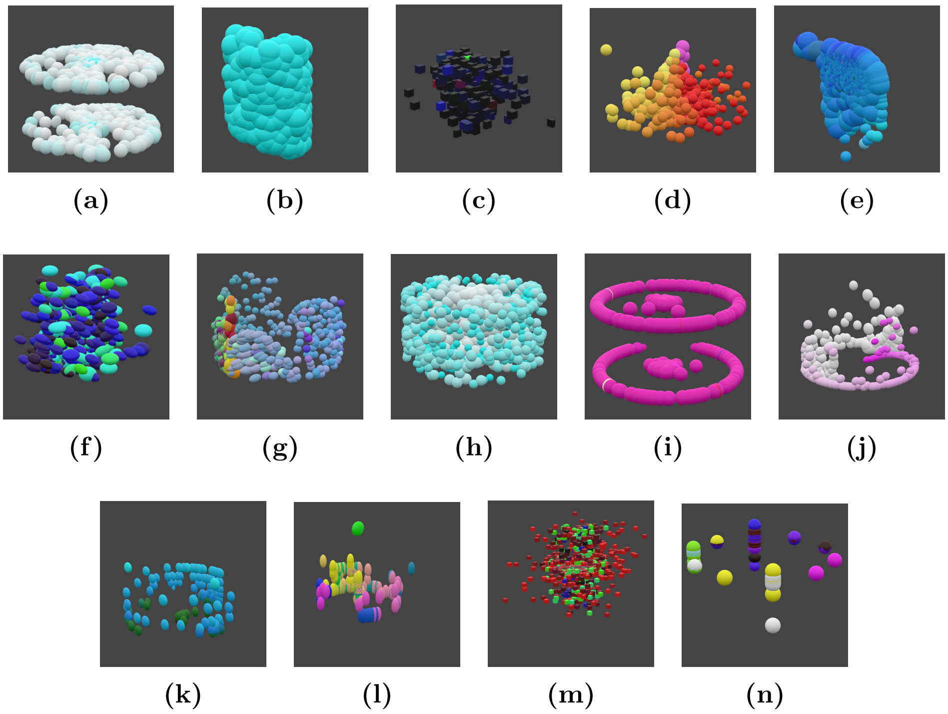


Figure 8.1: Initial sculptures generated on which the user model emulation is based

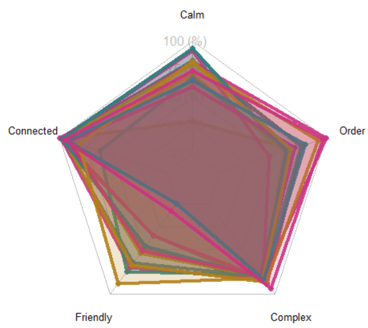


Figure 8.2: Range of values exhibited by the initial set of sculptures

the GP, taking the initial set of sculptures as the population, crossing over and mutating (as described in Chapter 4). The fitness is then calculated as the Euclidean distance, Equation 8.1, between the 5-vector generated for the sculpture and the user model. The lower the distance, the better the generated solution. Several examples of the sculptures are shown in Figures 8.3, 8.8 and 8.9, for Users one, two and three respectively.

This second set of sculptures is then presented to each user, who once again selects their favourites. Instead of specified choices, each user selects the seven sculptures that most closely match their current aesthetic model. With the new choices complete, the model is

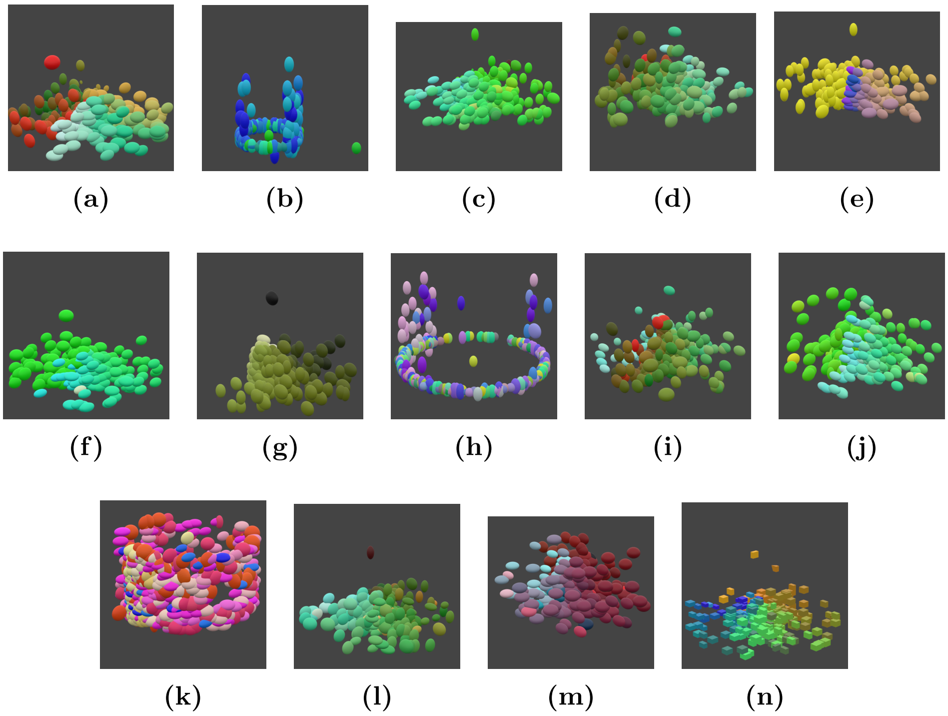


Figure 8.3: First set of sculptures generated for Example One

Choices	Aspect	Category	Model value
8.3n	Connected	Constant	0.66
8.3a	Order	Subjective	0.68
8.3b	Complexity	Subjective	0.77
8.3l	Calm	Subjective	0.66
8.3d	Friendly	Transient	0.75
8.3i			
8.3h			

Table 8.3: Second model for the first example

then updated and refined. The constant and subjective aspects have had different values selected, and the transient values have been replaced, leading to the model being updated to the definitions shown in Table 8.3, along with the user's choices.

The values for the constant and subjective aspects are calculated using the averaging function (Equation 8.2), which calculates the average between the existing value for the measure and the newly selected value. The same aspects are persisted in these examples due to the limited number of formalised aspects available. As this list of aspects increases, different terms are expected to form the model at each stage. Once again, the GP runs another ten generations on the population of sculptures, optimising the distance between the newly created items and the latest user model. This leads to new items being generated and presented to the user, Figure 8.4. From these items, several closely match the initial

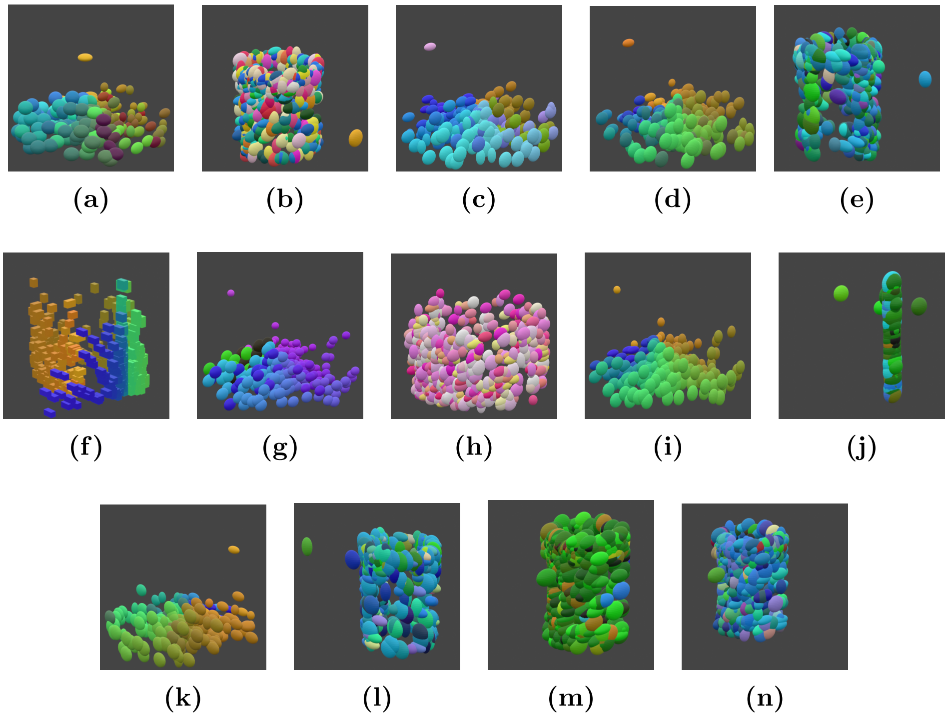


Figure 8.4: Final set of sculptures generated for Example One

user model and would be considered the best-fit for the user, shown in Figure 8.5. The same process is repeated for the remaining two examples with the model progression for user two shown in Table 8.4 and their generated best fit sculpture shown in Figure 8.6. The model for user three is shown in Table 8.5 and their final sculpture in Figure 8.7.

$$val_{new} = (val_{exist} + val_{select})/2 \tag{8.2}$$

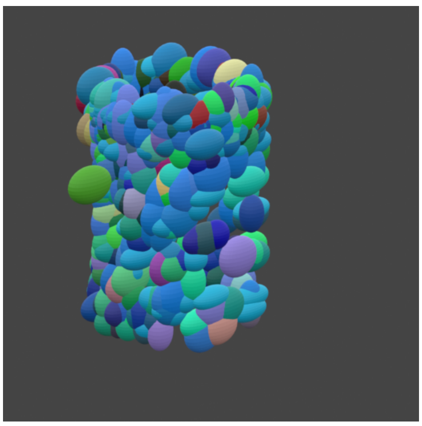


Figure 8.5: Best sculpture generated for Example one

Visually, all three best-fit sculptures generated are very different, even after two choices

Initial choice	Aspect	Initial value	Second value
8.1h	Connected	0.93	0.97
8.1i	Order	0.87	0.91
8.1j	Complexity	0.9	0.88
8.1k	Calm	0.72	0.8
8.1l	Friendly	0.76	0.69
8.1m			
8.1n			

Table 8.4: Model values for Example Two

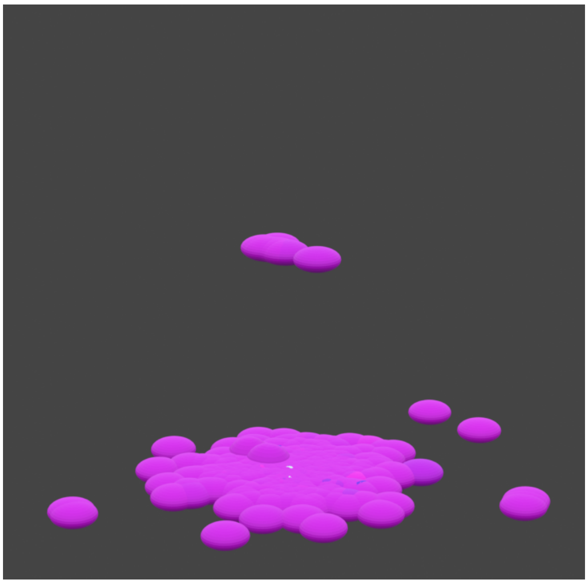


Figure 8.6: Best sculpture generated for Example Two

and minimal algorithmic input, indicating the effectiveness of this approach for generating unique sculptures specific to an individual's preferences. Interestingly, the influence of the initial choices can be seen in the final sculptures presented to each user. Users one and three have similar forms, and Users two and three have similar colours, with limited similarity between users one and two. With the additional subjectivity introduced by using real humans, these similarities would be much less defined, indicating how capable the model is at representing different preferences, a very important aspect considering the subjectivity of art preference. The results from the model also indicate how well the formalised measures are used in combination. For example, in Figure 8.5, the user model asks for a relatively

Initial choice	Aspect	Initial value	Second value
8.1d	Connected	0.95	0.96
8.1e	Order	0.7	0.73
8.1f	Complexity	0.83	0.8
8.1g	Calm	0.87	0.81
8.1h	Friendly	0.67	0.67
8.1i			
8.1j			

Table 8.5: Model values for Example Three

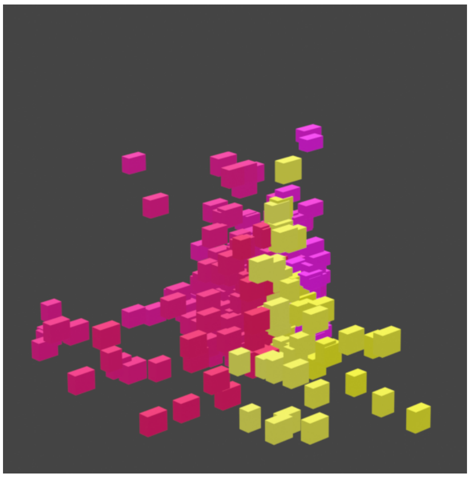


Figure 8.7: Best sculpture generated for Example Three

low level of each of the aspects, which the final sculpture reflects by having elements of each but no dominant factor. This is in contrast to the examples for users two and three, who both have a high level of connectivity, an aspect which is more clearly visible in their respective best sculptures.

8.2 Modelling an art style

The approach of considering tags is not only limited to modelling a user; by considering the data collected in Chapter 5 from a different perspective, generating new sculptures in a specified style is possible. An art style can be defined in terms of the aesthetic aspects the artwork produced in the style has, this fits in with the constant, subjective and transient categories model, allowing new items representing the specified style to be generated.

Similar to modelling an individual, modelling a style can be done by inspecting the number of times each tag was applied, but instead of looking at how often the tag was applied by users, how often the tag was applied to the range of sculptures within the style is considered. The same categorisation process can be followed, where if a tag has been applied to all or most sculptures, it is likely to fall within the constant category. If not, likely to fall within the subjective/transient categories, the assignment data collected for the sculptures used in the tag application experiment (Chapter 5) is shown in Table 8.6.
The same process from Chapter 5 can be used to identify the constant, subjective and transient aspects of the art style. However, having more data for the sculpture assignments makes it possible to be more strict on the applications of the tags, where constant tags are

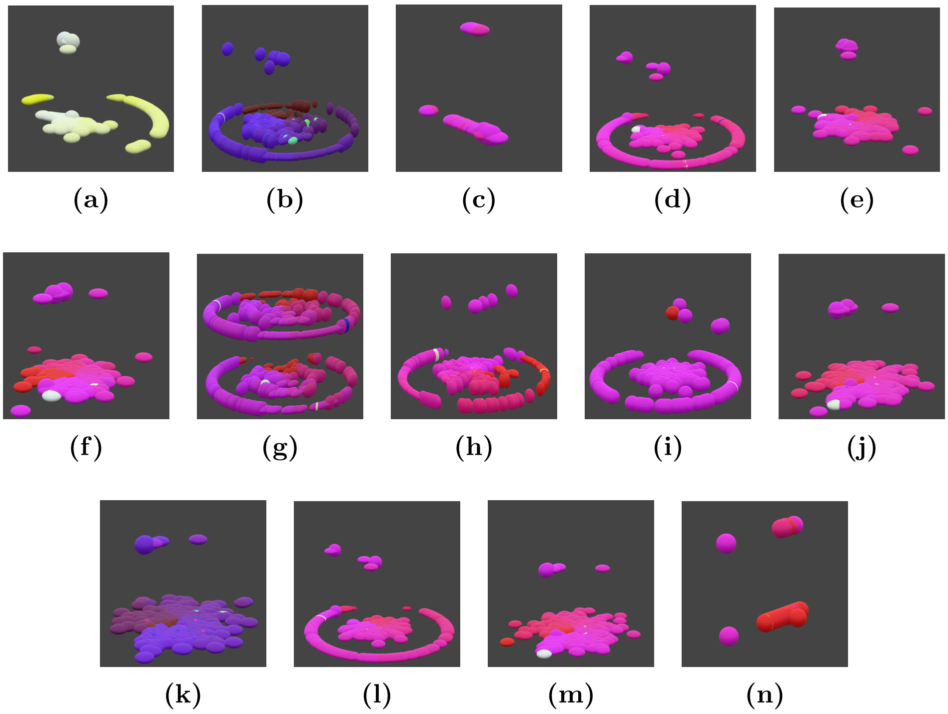


Figure 8.8: Final set of sculptures presented for Example Two

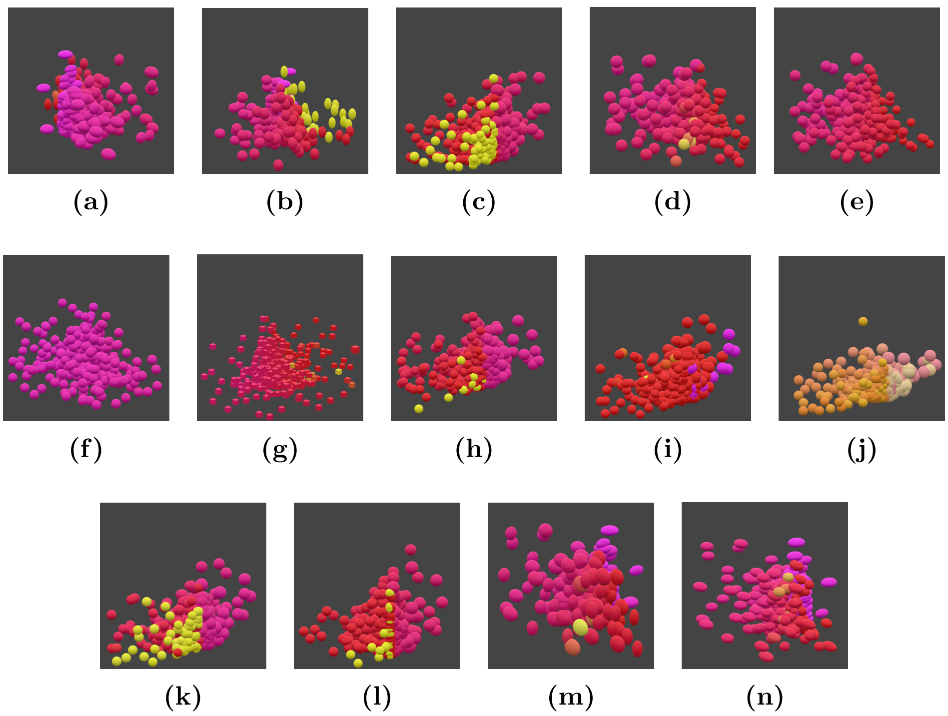


Figure 8.9: Final set of sculptures presented for Example Three

Tag	Total assignments
Busy	18
Warm	18
Disordered	18
Dynamic	18
Connected	18
Unnatural	18
Ordered	17
Static	17
Separate	17
Calm	17
Loud	17
Simple	17
Cold	16
Complex	16
Interesting	16
Bright	16
Unfriendly	16
Boring	15
Original	15
Curved	15
Friendly	14
Natural	14
Quiet	13
Angular	13
Sophisticated	13
Drab	9
Unrefined	8
Unoriginal	6

Table 8.6: Number of sculptures each tag was applied to

Constant	**Subjective**	**Transient**
Busy	Ordered	Original
Warm	Static	Sophisticated
Disordered	Separate	Unrefined
Dynamic	Calm	Cold
Connected	Simple	Unoriginal
Unnatural	Complex	Loud
	Interesting	Boring
	Bright	Warm
	Natural	Friendly
	Curved	Unfriendly
	Angular	Quiet
		Drab

Table 8.7: Final categories of each tag to recreate the tag application experiment sculptures

only those applied to all sculptures, e.g. dynamic, which leads to the categories shown in Table 8.7.

From this data, the set of tags in this model has different items being considered subjective and constant than for individual users. This is to be expected due to the differing context, these categories would also be different if a different style were being replicated. With the classification of the tags complete, it is now possible to try and replicate the style of the sculptures used in this experiment.

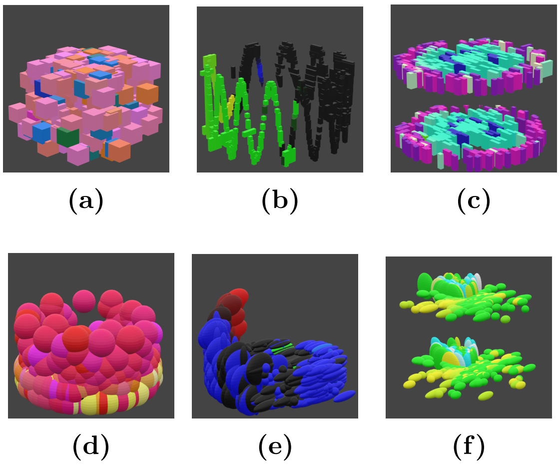


Figure 8.10: Starting sculptures for recreating the style of the aesthetic tag sculptures

8.2.1 Replicating the Chapter 5 art style

As discussed in Chapter 4, the sculptures generated within this project have a high affinity to the Concretism art style. However, the following model should not be considered analogous to Concretism, instead, it represents the style of the sculptures used in the tag application experiment. As the sculptures from Chapter 5 are the target style, a new set of 50 sculptures has been generated to act as the starting population for the replication.

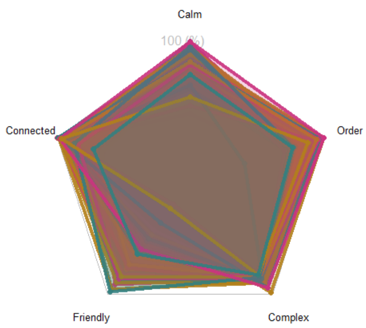


Figure 8.11: Range of values exhibited by the initial set of sculptures for replicating the style

A selection of the initial sculptures generated is shown in Figure 8.10, with the range of values these sculptures represent shown in Figure 8.11. As there is no user input to rely on to generate the model, the calculated average values from the original set of sculptures will be used as the target values for each tag. As with the user model, only the formalised measures from Chapter 7, along with symmetry and complexity, will be used within the

Total Generations	50
Mutation Probability	0.7
Max tree depth	5
Initialisation Method	Ramped Half and Half
Selection Method	Tournament (k=3)
Population Size	50
Fitness Measure	Style replication

Table 8.8: Evolutionary Algorithm parameters

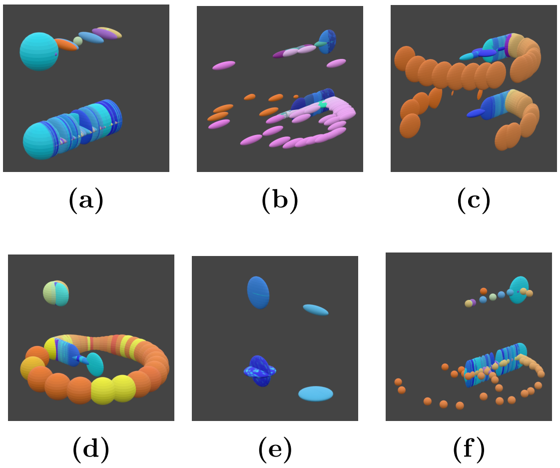


Figure 8.12: Example sculptures generated in the style of the aesthetic tag sculptures

model. Unlike the user model process, the art style process can be subjected to high numbers of generations to produce as accurate results as possible. The sculptures were evolved over 50 generations, using the parameters shown in Table 8.8, with a selection of the resulting sculptures shown in Figure 8.12.

The resulting sculptures are similar to each other, however, this could be mitigated by applying a novelty search process within the GP. Visually, the sculptures have similarities to the target set, especially with respect to the colours, but are still visually distinct enough to be identified as new sculptures, achieving the stated goals for this approach. This shows that the process can replicate a specific style of art as defined within a set of sculptures, and instead of exact replicas, novel sculptures are generated.

8.3 Conclusion

This chapter presented a novel approach to modelling individual aesthetic appreciation, which is suitable for generating sculptures unique to individuals or creating new examples of a specified art style.

The model and its application represent a lightweight approach to modelling individual user preferences across a wide range of aesthetic aspects. This model provides several benefits: it allows easy integration into a Genetic Program, which allows sculptures to be generated using the model, updating in real time based on user choices and changing preferences. In addition, the feedback loop allows the model to become more refined based on more choices being made by a user, generating more personalised sculptures. The model considers each user as an individual, allowing the generated sculptures to be unique to that individual and their preferences, even if multiple users overlap in their choices. Very quickly, the process is capable of generating unique sculptures, this means the user can spend less time making choices in order to generate sculptures which appeal to them, helping to keep the user engaged in the process.

As mentioned in Chapter 1, one of the current concerns with the current proliferation of AI-based art generation systems is their ethical impact. This is another of the benefits of the modelling approach presented, as it can generate artwork specific to a person without requiring large amounts of existing artwork to be used to train the system. It also allows these items to be generated without the user having to be able to explicitly identify their preferences, keeping the entire process for the user at a very high level, making it more usable, especially by people who have less expertise in art.

There are some limitations with this approach, even though individual preferences can be identified quickly within the model, accurately creating the model of the user's preferences may require the user to spend more time refining the model in order to generate the artwork; this runs the risk of hitting issues relating to user fatigue, and it is possible that the user would get bored with making choices before their model became refined enough to generate sculptures that accurately represent their preferences. Another difficulty is not knowing how the aspects interact with one another; the presence of one aspect may prevent another from being strongly present, either due to limitations of the sculpture generation method or a conflict between how the values are calculated. This could lead to the model trying to introduce a specific aspect to the sculptures that is unnecessary and may impact the presence of other aspects. The AGP was a generation process designed to generate a wide range of visually different sculptures, however, it is not capable of generating all possible sculptures. This means that the process in its current form may not be applicable to all users, however, due to the modular design of the process, the generation method can easily be amended, allowing the process to be applied to other art styles and keeping it flexible.

In its current state, the approach cannot represent a full aesthetic judgement model. However, with additional experimentation, the scope of the system detailed here can be improved in both impact and reliability. One approach to determine the success of this modelling approach would be to use the art gallery VR system described in Chapter 3, where it would be possible to implement the modelling and generation process using the grouping mechanism to identify a user's preferred sculptures. User feedback and choices about preferred sculptures can determine whether the system markedly improved the quality of the sculptures generated in terms of user assessment. However, without increasing the number of formalised measures available, the ability of the process to improve the sculptures for an individual user will be limited.

Overall, the presented theoretical model meets all of the stated aims by quickly generating sculptures unique to an individual's choices or representative of a specified style of art, depending on how the categorisation process has been completed. The model represents a full answer to Question 4, which, whilst being a theoretical approach, has shown high potential for future use from the emulating of different user choices. This indicates that once some of the mitigations have been applied to overcome the limitations mentioned above, this process is a viable candidate for becoming a powerful tool for automatically and semi-automatically generating unique sculptures based on a specified aesthetic judgement model.

Chapter 9

Conclusion

This thesis focused on the topic of aesthetics, presenting multiple novel avenues of research collectively aimed at further understanding the aesthetic judgement of sculptures. Helping to bridge the gap between aesthetic theory and computational aesthetics. The project has presented two VR-based systems which consider the topic of aesthetics from two levels, individual and collective; a novel artwork generation system allowing the creation of both 2D and 3D pieces of art, three experiments utilising the VREs investigating individual preferences of symmetry, which aspects contribute to the aesthetic judgement of sculptures and how closely human users perceive attributes of sculptures compared to computational measures. In addition to this, three aspects contributing to aesthetic judgement have been formalised, allowing the level of connectivity, calm, and friendliness to be quantified for a sculpture. Finally, a theoretical model was presented that aims to overcome some of the limitations of the existing computational models, allowing an individual's aesthetic preferences to be tracked in line with user choices, handling constant, subjective and transient aspects, meaning sculptures, which meet the individual's preferences to be generated.

The project approached the problem of aesthetic judgement by considering high-level concepts within aesthetic judgement identified through philosophers such as Plato and Frank Sibley, including aspects such as ethics, friendly, dynamic and boring were introduced representing a novel approach to computational aesthetics, with a high majority of existing research considering aesthetics in terms of low-level aspects such as symmetry. These approaches introduced a gap between the research and the application, limiting the effectiveness of these systems. The challenge of this project was split into four different research questions, detailed in Chapter 1, each looking at a specific area allowing the completion of the goal.

The first question considered whether a reusable framework could be developed which allowed the effective study of aesthetic judgement of 3D sculptures at both a collective and individual level. This was achieved by creating two novel flexible VR-based systems, each designed to look at aesthetics at one of the defined levels. The mine-cart based VR system allowed the extraction of individual preferences by asking participants to make pairwise decisions between two sculptures displaying differing levels of symmetry. This allowed individual preferences and selection behaviour for a particular aspect to be extracted without adding any restrictions on expertise level. This unique system was used within the project by running an experiment to understand individual preferences of symmetry described in Chapter 6, with two notable conclusions being obtained. The experiment consisted of two stages of five choices, narrowing the preference for symmetry between 0 and 1 to within an error value of 0.2. The first was that the approach used within the experiment was effective at obtaining a specific value for user preference with exceptionally consistent data collected. The second conclusion from this experiment was that users who preferred lower levels of symmetry were significantly more consistent in their choices than for other preference levels. This information signifies an important piece of information when considering how symmetry should be used when displaying generated art to users within an Evo-Art context.

In order to understand aesthetic judgement at an individual level, collective-level details need to be collected. This was the focus of the second VRE introduced in this project. It was based around a warehouse gallery setting and was designed to collect general aesthetic information. The environment allowed the display of a variety of different sculptures and allowed the participants to group similar sculptures, assigning terms to give meaning to each group, such as judgement details, ratings, and sculpture preferences to be recorded. This environment was used to run two different experiments within this project. The first collected data about which descriptive terms contributed to the aesthetic judgement of sculptures, this experiment provided a set of potential descriptive terms such as curved, angular, static and dynamic. The main result of this experiment was obtaining a list of common terms used to describe the presented sculptures alongside how positively or negatively these terms affect aesthetic judgement. Aside from the collection of the data, which was utilised throughout the remainder of the project, this experiment also showcased a vital ability of this VRE, allowing non-experts to succinctly describe their aesthetic judgement, something which has not often been considered before and ultimately helps to reduce the bias of the results presented in this project.

The second experiment using the gallery VRE considered the outcome of the measure formalisation process, where participants were shown sculptures which displayed a different value of one of the specified measures and asked to group the items based on how strong the presence of the aspect was within the sculpture. The participant rated 12 sculptures for each of the three measures. The data indicated that the measures met all the expectations, finding that all three measures had a significant high or very high correlation to the ratings. There were some noticeable differences in the selections made by the participants, however, due to the variability and subjectivity of the high-level terms presented, especially when using them in a potentially unfamiliar context such as judging sculptures, this was expected, having an exact correlation would be extremely unlikely.

These VREs combined answer to the first question, whether it was possible to develop a reusable and flexible framework to effectively study the aesthetic judgement of 3D sculptures at both a collective and individual level. Utilising the VREs presented within this thesis allows both levels to be investigated, with many potential avenues on how they can be used to to understand aesthetic judgement further.

With the ability to investigate and collect data around the aesthetic judgement of 3D sculptures completed, the second question concerned whether it was possible to develop a 3D sculpture generation process capable of exhibiting multiple aspects of aesthetic judgement. Several potential generation methods were considered, however, the first four were disregarded due to potential issues with generating diverse sculptures. The fifth process, the Axial Generation Process (AGP), used a simple algorithm to place multiple geometric objects in 3D space. After the initial prototype indicated that it may be a suitable method for generating sculptures, it was subsequently extended to cover a variety of features such as colour, dynamic sizing and the use of different geometric shapes. The AGP was subject to analysis comparing its output to two other established generation processes, particle-based generation and pixel/voxel-based generation, comparing inline or favourably with these two other generation methods. Another major factor of the AGP is that it is capable of generating 2D as well as 3D content. This novel feature allowed specific functionality to be implemented, for example, being able to render 3D content from a variety of different angles to convert the 3D item to 2D. Whilst this was not an approach that was considered within this project, it would be possible to use this feature to compare 2D artwork to 3D artwork to determine what differences exist. How the AGP compared to other generation systems and the specific benefits that were part of it means that Question 2 was also fully answered and that the method which allowed the creation of diverse 3D artwork is the AGP.

The remaining two questions were investigated using the VR environments detailed in Chapter 3 and the generation system from Chapter 4. The data collected from the gallery VRE established the terms which were part of the aesthetic judgement of 3D sculptures. It also revealed sculptures which exhibited the terms to varying degrees; analysing this information allowed the formalisation of three aspects, quantifying how connected, calm or friendly a sculpture looks. These measures were formalised using the tag assignment data and different low-level attributes displayed by the sculptures. Running the experiment detailed above successfully validated these measures, answering the third research question and providing a method of formalising and validating other aspects, along with the three presented algorithms within this project.

The final research question considered whether an individual's aesthetic preferences could be modelled in a suitable manner to allow the semi-automatic generation of sculptures unique to that individual. This question required combining all of the previous work within this thesis, and the model was presented in Chapter 8. A theoretical approach to modelling individual aesthetic preference, utilising a novel approach of categorising aesthetic aspects into constant, subjective and transient categories, using this category to inform how the aspects are handled. Through emulating multiple users, this modelling approach was shown to be effective at quickly being able to generate novel sculptures for individual preferences, albeit with more investigation required. This modelling process was also applied to generating new sculptures within a specified art style, and to demonstrate the efficacy of this approach, several sculptures were generated in the style of the sculptures used within the experiment in Chapter 5, with some positive results.

9.1 Re-examining the contributions

Answering all four of the proposed research questions led to multiple contributions being obtained, they are as follows:

1. The two VR environments allow the investigation of individual and collective level aesthetic judgement. Each of these environments had a singular focus on either individual or collective level aesthetics and were found to be successful in their goals across multiple experiments. In their current state, they are suitable to be used to investigate different aspects of the aesthetic judgement process. These environments were detailed in Chapter 3, and utilised in Chapters 5, 6 and 7.

2. The AGP art generation system allowed a wide range of visually diverse 2D and 3D artwork to be generated. The process was found to compare well against existing generation methods, creating artwork which exhibited a wide range of values over several established aesthetic measures. The AGP can be used with or without Evolutionary Computation, however, using it in conjunction with EC techniques extends its ability to generate visually varied artwork.

3. Three formalised measures, quantifying levels of connected, calm and friendly within sculptures. The measures were formulated based on the data collected in Chapter 5 and subsequently validated by comparing the calculated value against ratings provided by human participants. All three of these measures were found to be representative of human judgement and are therefore suitable for use with or without the other contributions mentioned from this project being applicable to other 3D generation systems. However, utilising other contributions, such as the AGP, would allow these measures to be used to generate 2D artwork.

4. The model of individual aesthetic preference using the categorisation of all aspects completed in Chapter 5 to model a user's preferences. Taking into account any subsequent choices a user makes, the model is updated, efficiently changing the search for sculptures unique to their individual tastes. This modelling process can be used with various art generation systems, providing an effective solution to model individual aesthetic preferences.

9.2 Limitations

Whilst the information presented in this thesis had some significant benefits, some limitations must be considered.

The most prominent is that this project focused on sculptures within a 3D virtual environment, which limits the extent to which any conclusions can be applied. For example, it would not be possible to directly utilise the measures on 2D items without the items being generated through the AGP and expect the same results as in this project.

In addition to the focus on sculptures, the art style introduced for the generation of sculptures in this project also represents a limitation of this project. Using the Concretism style has many benefits, as detailed in Chapter 4; however, the style of sculpture generated by the AGP may not be considered aesthetically pleasing to everyone, which could have impacted the data collected. Fortunately, the sculptures were often highly rated by participants of the experiments, but this does not make the details generalised to everyone, and more extensive studies with different styles of sculptures would need to be conducted in order to determine these additional details.

The final major limitation was the amount of data collected, especially during the formalisation process. Some of the measures only had a limited amount of data available about the sculptures, and the number of sculptures available to assess was limited. This meant it was more difficult to identify clear patterns; the conclusions drawn about which attributes contributed to a particular aspect may be different if more sculptures were available.

Overall, even with the limitations, this project has successfully been able to meet all of the goals, providing comprehensive answers to each of the proposed research questions alongside multiple contributions which can be used as standalone items or in conjunction with one another, extensively expanding the knowledge surrounding aesthetics and aesthetic judgement, allowing individualised artwork to be generated. Each aspect of the project provided insight into the judgement process, including the multiple presented systems, measures and models, illuminating the transient nature of aesthetic judgement and how different aspects affect aesthetic perception in a positive or negative way.

9.3 Further Work

All of the contributions from this project have many different potential paths for future research. Details on some of the potential identified paths are below.

9.3.1 Minecart VR environment

There are two main paths of work that can be continued; the first is to further the understanding of symmetry by considering how it interacts with other aspects that contribute to the aesthetic judgement of sculptures. The second is that the system will be used to measure other aspects of aesthetic appreciation, e.g. complexity, enabling their values to be categorised as constant, transient or subjective. This will include the formalised descriptive terms from this project: connected, calm and friendly, extracting preference levels for these high-level concepts, providing a much greater understanding of how aesthetic judgement is performed.

Another aspect which would be beneficial to investigate would be to utilise this VRE to help elicit the impact each aesthetic tag had on each other. For example, friendly sculptures could be more highly rated when combined with calm aspects but are judged to be less positive when combined with connected aspects. By keeping one aspect the same across all sculptures presented to the participant and amending the other aspects present, it would be possible to determine a rough outline of which aspects work well together, creating sculptures which have a more positive impact on the viewer.

9.3.2 Gallery VR environment

The gallery VRE also has significant potential to be used for further research. Using the configurations introduced in this project, this system could be used to investigate further tags which contribute to the aesthetic judgement of 3D and 2D artwork, allowing additional aspects to be formalised and added to the model described in this thesis. It can also be used to provide aesthetic ratings for a wider range of sculptures, allowing the AGP or other art generation methods to be evaluated. Finally, the environment would be suitable to investigate the model presented in Chapter 8, re-using the grouping structure to allow the participant to pick favourites, informing the next stages of the model.

9.3.3 Aesthetic tags

The data collected in Chapter 5 opens up multiple avenues for further research, such as looking at the difference between the positive or negative interpretation of the aesthetic terms. The presented approach can also be amended to extend the amount of information collected; for example, it could be applied to investigating 2D images, which can then be compared to 3D items to determine the differences between the judgement of the two types of items. A potentially wider array of terms and sculptures could also be presented, further extending the aspects that can be considered part of aesthetic judgement and further understanding the interaction between the terms and whether this affects the positive or negative connotations.

9.3.4 Aesthetic model

This model can be applied to a wide range of areas which would benefit from the real-time individual generation process. For example, if the process could determine user preference using eye gaze rather than explicit choices or ratings, the process could function entirely in the background. This would be useful for many virtual applications, where the scenery or objects within the environment could be updated to reflect the user's preferences, only requiring the user to spend time within the application. This could be applied to industry applications such as interior design.

Considering the approach where specific art styles could be replicated, this could also be applied to showing specific themes, allowing the environment around the player to be updated in line with the specific part of the story the user is experiencing. Not only giving the user a more engaging experience but also a unique one still reflecting the feel of the storyline, or if the application was relating more serious endeavours, such as counselling systems, similar to the one described by Osimo et al. (2015), the environment could be updated to make the user more comfortable and engaged in the system, increasing the immersion, improving their effectiveness at helping with mental health issues.

The modelling approach is not limited to virtual applications; when combined with other technologies such as 3D printing, the system could be used to help users design unique and individualised physical products, which, considering the expansive nature of the customisation market, could be a very exciting future direction for this project. A product could be made available as a set of 3D printing plans, however, through using a process similar to the one described could be followed to amend the visual design of the product without affecting the functionality. This could then be printed by the user, meaning that a wide range of products would no longer need to be mass-produced and easily individualised.

List of References

S. Aaronson, S. M. Carroll, and L. Ouellette. Quantifying the rise and fall of complexity in closed systems: the coffee automaton. *arXiv preprint arXiv:1405.6903*, 2014.

E. Acebo and M. Sbert. Benford's law for natural and synthetic images. In *First Eurographics conference on Computational Aesthetics in Graphics, Visualization and Imaging*, 2005.

M. al Rifaie, A. Ursyn, R. Zimmer, and M. Javid. On symmetry, aesthetics and quantifying symmetrical complexity. In *International Conference on Evolutionary and Biologically Inspired Music and Art*, pages 17–32. Springer, 2017.

R. C. Alamino. Measuring complexity through average symmetry. *Journal of Physics A: Mathematical and Theoretical*, 48(27):275101, 2015.

J. S. Allen. Plato: The morality and immorality of art. *Arts Education Policy Review*, 104 (2):19–24, 2002.

L. Anderson and H.-C. Huang. Chalkroom, 2017. [VR Animation].

R. Antunes and F. Leymarie. Generative choreography: animating in real-time dancing avatars. In *International Conference on Evolutionary and Biologically Inspired Music and Art*, pages 1–10. Springer, 2012.

R. Antunes and F. F. Leymarie. xTNZ: an evolutionary three-dimensional ecosystem. In *Proceedings of the 4th International Conference on Digital Arts, Artech2008*, pages 201–204, 2008.

R. Arnheim. *Art and visual perception: A psychology of the Creative Eye;*. University of California Press, 2009.

S. Bak, N. Rask, and S. Risi. Towards adaptive evolutionary architecture. In *International Conference on Computational Intelligence in Music, Sound, Art and Design*, pages 47–62. Springer, 2016.

M. Bauerly and Y. Liu. Computational modeling and experimental investigation of effects of compositional elements on interface and design aesthetics. *International journal of human-computer studies*, 64(8):670–682, 2006.

R. R. Behrens. Art, design and gestalt theory. *Leonardo*, 31(4):299–303, 1998.

C. Bell. Significant form. In *Philosophy: Basic readings*, pages 349–343. Routledge, 2005.

V. Bell. A history of media technology scares, from the printing press to facebook., Feb 2010. URL `https://slate.com/technology/2010/02/a-history-of-media-technology-scares-from-the-printing-press-to-facebook.html`.

S. Bergen and B. Ross. Aesthetic 3D model evolution. *Genetic Programming and Evolvable Machines*, 14(3):339–367, 2013.

M. Bertamini, J. Silvanto, A. M. Norcia, A. D. Makin, and J. Wagemans. The neural basis of visual symmetry and its role in mid-and high-level visual processing. *Annals of the New York Academy of Sciences*, 1426(1):111–126, 2018.

S. Biasotti, A. Cerri, A. Bronstein, and M. Bronstein. Recent trends, applications, and perspectives in 3d shape similarity assessment. In *Computer graphics forum*, volume 35, pages 87–119. Wiley Online Library, 2016.

A. A. Bichi. Classical Test Theory: an introduction to linear modeling approach to test and item analysis. *International Journal for Social Studies*, 2(9):27–33, 2016.

G. Birkhoff. *Aesthetic Measure*. Harvard University Press, Cambridge, MA, USA, 1933.

M. A. Boden et al. *The creative mind: Myths and mechanisms*. Psychology Press, 2004.

D. Bountouridis, D. Brown, H. Koops, F. Wiering, and R. Veltkamp. Melody retrieval and classification using biologically inspired techniques. In *International Conference on Evolutionary and Biologically Inspired Music and Art*, pages 49–64. Springer, 2017.

T. Braun Janzen, B. de Oliveira, G. Ventorim Ferreira, J. R. Sato, C. Feitosa-Santana, and P. Vanzella. The effect of background music on the aesthetic experience of a visual artwork in a naturalistic environment. *Psychology of Music*, 51(1):16–32, 2023.

H. Bredin. Medieval Art Theory. *A companion to art theory*, pages 29–39, 2002.

D. Brieber, M. Nadal, and H. Leder. In the white cube: Museum context enhances the valuation and memory of art. *Acta psychologica*, 154:36–42, 2015.

D. G. Brown and D. Ventura. Ethics, Aesthetics and Computational Creativity. In *ICCC*, 2022.

J. Byrne, E. Hemberg, M. O'Neill, and A. Brabazon. A methodology for user directed search in evolutionary design. *Genetic Programming and Evolvable Machines*, 14(3):287–314, 2013.

R. Canaan, S. Menzel, J. Togelius, and A. Nealen. Towards game-based metrics for computational co-creativity. In *2018 IEEE Conference on Computational Intelligence and Games (CIG)*, pages 1–8. IEEE, 2018.

S. Carliner. Modeling information for three-dimensional space: Lessons learned from museum exhibit design. *Technical communication*, 50(4):554–570, 2003.

C. Carter. Sculpture. In *The Routledge Companion to Aesthetics*, pages 503–517. Routledge, 2005.

P. Castelló, M. Sbert, M. Chover, and M. Feixas. Techniques for computing viewpoint entropy of a 3d scene. In *International Conference on Computational Science*, pages 263–270. Springer, 2006.

A. Chatterjee, P. Widick, R. Sternschein, W. B. Smith, and B. Bromberger. The assessment of art attributes. *Empirical Studies of the Arts*, 28(2):207–222, 2010.

E. Chu. Evolving evocative 2d views of generated 3d objects. *arXiv preprint arXiv:2111.04839*, 2021.

M. Claypool, K. Claypool, and F. Damaa. The effects of frame rate and resolution on users playing first person shooter games. In *Multimedia Computing and Networking 2006*, volume 6071, page 607101. International Society for Optics and Photonics, 2006.

M. Cohen, L. Cherchiglia, and R. Costa. Evolving Mondrian-Style artworks. In *International Conference on Evolutionary and Biologically Inspired Music and Art*, pages 338–353. Springer, 2017.

S. Colton. Automatic invention of fitness functions with application to scene generation. In *Workshops on Applications of Evolutionary Computation*, pages 381–391. Springer, 2008a.

S. Colton. Creativity Versus the Perception of Creativity in Computational Systems. In *AAAI spring symposium: creative intelligent systems*, volume 8, page 7. Palo Alto, CA, 2008b.

S. Colton. Evolving a library of artistic scene descriptors. In *International Conference on Evolutionary and Biologically Inspired Music and Art*, pages 35–47. Springer, 2012.

S. Colton, M. Cook, and A. Raad. Ludic considerations of tablet-based evo-art. In *European Conference on the Applications of Evolutionary Computation*, pages 223–233. Springer, 2011.

M. Cook and S. Colton. Redesigning computationally creative systems for continuous creation. 2018.

J. Cowles, P. Chan, T. Voight, and I. Parkinson. A Show of Kindness. In *ACM SIGGRAPH 2018 Virtual, Augmented, and Mixed Reality*, SIGGRAPH '18, New York, NY, USA, 2018. Association for Computing Machinery. ISBN 9781450358217. doi: 10.1145/3226552.3226572. URL `https://doi.org/10.1145/3226552.3226572`.

D. W. Crawford. Kant. In *The Routledge companion to aesthetics*, pages 51–64. Routledge, 2001.

C. Davies. Osmose: Notes on Being in Immersive Virtual Space. In *ISEA '95*, 1995.

E. Davies, P. Tew, D. Glowacki, J. Smith, and T. Mitchell. Evolving atomic aesthetics and dynamics. In *International Conference on Computational Intelligence in Music, Sound, Art and Design*, pages 17–30. Springer, 2016.

N. Daykin, E. Byrne, T. Soteriou, and S. O'Connor. The impact of art, design and environment in mental healthcare: a systematic review of the literature. *Journal of the Royal Society for the Promotion of Health*, 128(2):85–94, 2008.

E. den Heijer. Evolving art using measures for symmetry, compositional balance and liveliness. In *International Conference on Evolutionary Computation Theory and Applications*, volume 2, pages 52–61. SciTePress, 2012.

E. den Heijer and A. Eiben. Comparing aesthetic measures for evolutionary art. In *European Conference on the Applications of Evolutionary Computation*, pages 311–320. Springer, 2010a.

E. den Heijer and A. Eiben. Evolving pop art using scalable vector graphics. In *International Conference on Evolutionary and Biologically Inspired Music and Art*, pages 48–59. Springer, 2012.

E. den Heijer and A. E. Eiben. Comparing aesthetic measures for evolutionary art. In *European Conference on the Applications of Evolutionary Computation*, pages 311–320. Springer, 2010b.

K. G. Derpanis. Mean shift clustering. *Lecture Notes*, 32:1–4, 2005.

K. Dev, N. Villar, and M. Lau. Polygons, points, or voxels? stimuli selection for crowdsourcing aesthetics preferences of 3D shape pairs. In *Proceedings of the symposium on Computational Aesthetics*, pages 1–7, 2017.

J. Dewey. Art as experience. In *The richness of art education*, pages 33–48. Brill, 2008.

E. Easton. Investigating User Fatigue in Evolutionary Art. Master's thesis, Aston University, 2018.

E. Easton, U. Bernardet, and A. Ekart. Tired of Choosing? Just Add Structure and Virtual Reality. In *International Conference on Computational Intelligence in Music, Sound, Art and Design (Part of EvoStar)*, pages 142–155. Springer, 2019.

E. Easton, A. Ekárt, and U. Bernardet. Axial Generation: A Concretism-Inspired Method for Synthesizing Highly Varied Artworks. In *10th International Conference on Artificial Intelligence in Music, Sound, Art and Design, EvoMUSART 2021 held as Part of EvoStar 2021*, pages 115–130. Springer, 2021.

E. Easton, A. Ekárt, and U. Bernardet. Axial Generation: Mixing Colour and Shapes to Automatically Form Diverse Digital Sculptures. *SN Computer Science*, 3(6):505, 2022.

E. Easton, U. Bernardet, and A. Ekárt. Is Beauty in the Age of the Beholder? In *Artificial Intelligence in Music, Sound, Art and Design: 12th International Conference, EvoMUSART 2023, Held as Part of EvoStar 2023, Brno, Czech Republic, April 12–14, 2023, Proceedings*, pages 84–99. Springer, 2023a.

E. Easton, U. Bernardet, and A. Ekárt. Contributors to the aesthetic judgement of 3D virtual sculptures. In *2023 Third International Conference on Digital Creation in Arts, Media and Technology (ARTeFACTo)*. IEEE, 2023b.

E. Easton, U. Bernardet, and A. Ekárt. Modelling individual aesthetic preferences of 3D sculptures. In *Artificial Intelligence in Music, Sound, Art and Design: 13th International Conference, EvoMUSART 2024, Held as Part of EvoStar 2024, Aberystwyth, United Kingdom, April 3–5, 2024, Proceedings*. Springer, 2024.

A. Ecins, C. Fermuller, and Y. Aloimonos. Detecting reflectional symmetries in 3d data through symmetrical fitting. In *Proceedings of the IEEE International Conference on Computer Vision Workshops*, pages 1779–1783, 2017.

A. Eiben and J. Smith. *Introduction to evolutionary computing*, volume 53. Berlin: springer, 2003.

A. E. Eiben and J. E. Smith. *Introduction to evolutionary computing*. Springer, 2015.

A. Ekárt, D. Sharma, and S. Chalakov. Modelling human preference in evolutionary art. In *European Conference on the Applications of Evolutionary Computation*, pages 303–312. Springer, 2011.

A. Ekárt, A. Joó, D. Sharma, and S. Chalakov. Modelling the underlying principles of human aesthetic preference in evolutionary art. *Journal of Mathematics and the Arts*, 6 (2-3):107–124, 2012.

A. Emanuel and E. Eldar. Emotions as computations. *Neuroscience & Biobehavioral Reviews*, page 104977, 2022.

M. Farrer. Who was Edward Colston and why was his Bristol statue toppled?, Jun 2020. URL `https://www.theguardian.com/uk-news/2020/jun/08/who-was-edward-colston-and-why-was-his-bristol-statue-toppled-slave-trader-black-lives-`

L. J. Fogel, A. J. Owens, and M. J. Walsh. Intelligent decision making through a simulation of evolution. *Behavioral science*, 11(4):253–272, 1966.

J. P. Forgas. Mood and judgment: the affect infusion model (AIM). *Psychological bulletin*, 117(1):39, 1995.

J. Friedenberg, P. Martin, N. Uy, and M. Kvapil. Judged beauty of fractal symmetries. *Empirical Studies of the Arts*, 40(1):100–120, 2022.

J. Gaiger. The aesthetics of Kant and Hegel. *A companion to art theory*, pages 125–138, 2002.

P. Galanter. Computational aesthetic evaluation: past and future. *Computers and creativity*, pages 255–293, 2012.

A. Gartus, N. Klemer, and H. Leder. The effects of visual context and individual differences on perception and evaluation of modern art and graffiti art. *Acta psychologica*, 156:64–76, 2015.

B. Gaut. Art and ethics. In *The Routledge Companion to Aesthetics*, pages 394–403. Routledge, 2013.

V. Giannouli. Visual symmetry perception. *Encephalos*, 50:31–42, 2013.

M. Gircys and B. Ross. Image Evolution Using 2D Power Spectra. *Complexity*, 2019a.

M. Gircys and B. J. Ross. Image Evolution Using 2D Power Spectra. *Complexity*, 2019, 2019b.

J. Grefenstette, R. Gopal, B. Rosmaita, and D. Van Gucht. Genetic algorithms for the traveling salesman problem. In *Proceedings of the first International Conference on Genetic Algorithms and their Applications*, volume 160, pages 160–168. Lawrence Erlbaum, 1985.

G. M. Grube. Plato's theory of beauty. *The Monist*, pages 269–288, 1927.

J. Guetzkow. *How the arts impact communities.* Princetion University, 2002.

D. Gulhan, S. Durant, and J. M. Zanker. Aesthetic judgments of 3D arts in virtual reality and online settings. *Virtual Reality*, pages 1–17, 2022.

G. Harris. New report recommends Bristol's controversial slaver statuetorn down by BLM protestors-be permanently installed in museum, Feb 2022. URL `https://www.theartnewspaper.com/2022/02/04/new-report-recommends-bristols-controversial-slaver-statue-be-permanently-placed-in-mus`

C. Haskins. Kant and the Autonomy of Art. *The Journal of Aesthetics and Art Criticism*, 47(1):43–54, 1989.

G. U. Hayn-Leichsenring and A. Chatterjee. Colliding terminological systems—Immanuel Kant and contemporary empirical aesthetics. *Empirical Studies of the Arts*, 37(2):197–219, 2019.

J. Holland. Adaptation in natural and artificial systems. *Ann Arbor*, 1975.

B. Hollingsworth and J. Schrum. Infinite art gallery: a game world of interactively evolved artwork. In *2019 IEEE Congress on Evolutionary Computation (CEC)*, pages 474–481. IEEE, 2019.

G. S. Hornby and J. Bongard. Learning comparative user models for accelerating human-computer collaborative search. In *International Conference on Evolutionary and Biologically Inspired Music and Art*, pages 117–128. Springer, 2012.

D. Hume. Of the standard of taste. In *Aesthetics*, pages 483–488. Routledge, 2017.

M. Hungerford. *Molly Bawn: In Two Volumes*, volume 1788. Bernhard Tauchnitz, 1878.

A. I. Isik and E. A. Vessel. Continuous ratings of movie watching reveal idiosyncratic dynamics of aesthetic enjoyment. *PloS one*, 14(10):e0223896, 2019.

T. Jacobsen and L. Höfel. Aesthetic judgments of novel graphic patterns: Analyses of individual judgments. *Perceptual and motor skills*, 95(3):755–766, 2002.

D. Jacquette. Idealism: Schopenhauer, Schiller and Schelling. In *The Routledge Companion to Aesthetics*, pages 66–76. Routledge, 2013.

C. Janaway. Plato. *The Routledge companion to aesthetics*, pages 3–13, 2013.

C. G. Johnson, J. McCormack, I. Santos, and J. Romero. Understanding aesthetics and fitness measures in evolutionary art systems. *Complexity*, 2019, 2019.

B. C. Jones, L. M. DeBruine, and A. C. Little. The role of symmetry in attraction to average faces. *Perception & psychophysics*, 69(8):1273–1277, 2007.

K. Kabassi and E. Maravelakis. Walkthrough evaluation of a VR museum for the physical environment. In *2015 6th International Conference on Information, Intelligence, Systems and Applications (IISA)*, pages 1–5. IEEE, 2015.

P. Karimi, K. Grace, M. L. Maher, and N. Davis. Evaluating creativity in computational co-creative systems. *arXiv preprint arXiv:1807.09886*, 2018.

B. Katz. What makes a polygon pleasing? *Empirical studies of the Arts*, 20(1):1–19, 2002.

M. Kieran. Value of art. In *The Routledge companion to aesthetics*, pages 313–326. Routledge, 2005.

K. Koidan. Legal & ethical aspects of using dall-E, Midjourney, & Stable Diffusion, Mar 2023. URL `https://medium.com/@katekoidan/legal-ethical-aspects-of-using-dall-e-midjourney-stable-diffusion-cc5606a76d8e`.

C. Korsmeyer. Taste. In *The Routledge companion to aesthetics*, pages 257–266. Routledge, 2013.

J. R. Koza. *Genetic programming: A paradigm for genetically breeding populations of computer programs to solve problems*, volume 34. Stanford University, Department of Computer Science Stanford, CA, 1990.

J. R. Koza. Genetic programming as a means for programming computers by natural selection. *Statistics and computing*, 4:87–112, 1994.

S. Lakhal, A. Darmon, J.-P. Bouchaud, and M. Benzaquen. Beauty and structural complexity. *Physical Review Research*, 2(2):022058, 2020.

C. Lamb, D. G. Brown, and C. L. Clarke. Evaluating computational creativity: An interdisciplinary tutorial. *ACM Computing Surveys (CSUR)*, 51(2):1–34, 2018.

N. Lambert, W. Latham, and F. Fol Leymarie. The emergence and growth of evolutionary art: 1980–1993. In *ACM SIGGRAPH 2013 Art Gallery*, pages 367–375. 2013.

W. Latham and S. Todd. *Evolutionary art and computers.* Academic Press Inc, 1992.

W. Latham, S. Todd, P. Todd, and L. Putnam. Exhibiting mutator VR: procedural art evolves to virtual reality. *Leonardo*, 54(3):274–281, 2021.

K. Layng, K. Perlin, S. Herscher, C. Brenner, and T. Meduri. CAVE: Making Collective Virtual Narrative: Best Paper Award. *Leonardo*, 52(4):349–356, 2019.

H. Leder and M. Nadal. Ten years of a model of aesthetic appreciation and aesthetic judgments: The aesthetic episode–developments and challenges in empirical aesthetics. *British journal of psychology*, 105(4):443–464, 2014.

H. Leder, J. Goller, T. Rigotti, and M. Forster. Private and shared taste in art and face appreciation. *Frontiers in human neuroscience*, 10:155, 2016.

H. Leder, P. P. Tinio, D. Brieber, T. Kröner, T. Jacobsen, and R. Rosenberg. Symmetry is not a universal law of beauty. *Empirical Studies of the Arts*, 37(1):104–114, 2019.

J. Lehman and K. O. Stanley. Exploiting open-endedness to solve problems through the search for novelty. In *ALIFE*, pages 329–336, 2008.

J. S. Lerner and D. Keltner. Beyond valence: Toward a model of emotion-specific influences on judgement and choice. *Cognition & emotion*, 14(4):473–493, 2000.

F. F. Leymarie. Art. Machines. Intelligence. *Enjeux Numeriques*, (13):82–102, 2021.

Y. Li, C. Hu, M. Chen, and J. Hu. Investigating aesthetic features to model human preference in evolutionary art. In *International Conference on Evolutionary and Biologically Inspired Music and Art*, pages 153–164. Springer, 2012.

R. O. Lopes. Museum curation in the digital age. *The future of creative work*, pages 123–139, 2020.

N. Lourenço, F. Assunção, C. Maçãs, and P. Machado. EvoFashion: customising fashion through evolution. In *International Conference on Evolutionary and Biologically Inspired Music and Art*, pages 176–189. Springer, 2017.

P. Machado and A. Cardoso. Computing aesthetics. In *Brazilian Symposium on Artificial Intelligence*, pages 219–228. Springer, 1998.

P. Machado and A. Cardoso. All the truth about NEvAr. *Applied Intelligence*, 16(2): 101–118, 2002.

P. Machado, J. Romero, and B. Manaris. Experiments in computational aesthetics: An iterative approach to stylistic change in evolutionary art. In *The art of artificial evolution: a handbook on evolutionary art and music*, pages 381–415. Springer, 2008.

P. Machado, A. Vinhas, J. Correia, and A. Ekárt. Evolving ambiguous images. In *Twenty-fourth international joint conference on artificial intelligence*, 2015.

P. Machado, J. Romero, and G. Greenfield. Aesthetics, Artificial Intelligence, and Search-Based Art. *Artificial Intelligence and the Arts: Computational Creativity, Artistic Behavior, and Tools for Creatives*, pages 27–60, 2021.

M. Mahmoudi and G. Sapiro. Three-dimensional point cloud recognition via distributions of geometric distances. *Graphical Models*, 71(1):22–31, 2009.

J. Margolis. Medieval aesthetics. In *The Routledge Companion to Aesthetics*, pages 45–54. Routledge, 2005.

R. Marriner. Derrida and the Parergon. *A Companion to Art Theory*, pages 349–359, 2002.

K. Matkovic, L. Neumann, A. Neumann, T. Psik, and W. Purgathofer. Global contrast factor-a new approach to image contrast. *Computational Aesthetics*, 2005:159–168, 2005.

P. M. Matthews. Hutcheson on the Idea of Beauty. *Journal of the History of Philosophy*, 36(2):233–259, 1998.

J. McCormack and A. Lomas. Understanding Aesthetic Evaluation using Deep Learning. In *International Conference on Computational Intelligence in Music, Sound, Art and Design (Part of EvoStar)*, pages 118–133. Springer, 2020.

J. McCormack, B. Porter, and J. Wetter. A scriptable, generative modelling system for dynamic 3D meshes. *arXiv preprint arXiv:2004.10354*, 2020.

J. McCormack, C. Cruz Gambardella, and A. Lomas. The enigma of complexity. In *Artificial Intelligence in Music, Sound, Art and Design: 10th International Conference, EvoMUSART 2021, Held as Part of EvoStar 2021, Virtual Event, April 7–9, 2021, Proceedings 10*, pages 203–217. Springer, 2021.

J. McDermott. Graph grammars as a representation for interactive evolutionary 3d design. In *International Conference on Evolutionary and Biologically Inspired Music and Art*, pages 199–210. Springer, 2012.

J. McDermott. Graph grammars for evolutionary 3D design. *Genetic Programming and Evolvable Machines*, 14(3):369–393, 2013.

J. McDermott, J. M. Swafford, M. Hemberg, J. Byrne, E. Hemberg, M. Fenton, C. McNally, E. Shotton, and M. O'Neill. String-rewriting grammars for evolutionary architectural design. *Environment and Planning B: Planning and Design*, 39(4):713–731, 2012.

A. Mills. Animating typescript using aesthetically evolved images. In *International Conference on Computational Intelligence in Music, Sound, Art and Design*, pages 126–134. Springer, 2016.

L. C. P. Monteiro, V. E. F. d. Nascimento, A. Carvalho da Silva, A. C. Miranda, G. S. Souza, and R. C. Ripardo. The role of art expertise and symmetry on facial aesthetic preferences. *Symmetry*, 14(2):423, 2022.

M. Muehlbauer, J. Burry, and A. Song. Automated shape design by grammatical evolution. In *International Conference on Evolutionary and Biologically Inspired Music and Art*, pages 217–229. Springer, 2017.

M. Nadal and A. Chatterjee. Neuroaesthetics and art's diversity and universality. *Wiley Interdisciplinary Reviews: Cognitive Science*, 10(3):e1487, 2019.

A. M. Nguyen, J. Yosinski, and J. Clune. Innovation engines: Automated creativity and improved stochastic optimization via deep learning. In *Proceedings of the 2015 Annual Conference on Genetic and Evolutionary Computation*, pages 959–966, 2015.

M. Nicolau and D. Costelloe. Using grammatical evolution to parameterise interactive 3D image generation. In *European Conference on the Applications of Evolutionary Computation*, pages 374–383. Springer, 2011.

S. Niedenthal. What we talk about when we talk about game aesthetics. 2009.

F. Nietzsche. The birth of tragedy. In *Aesthetics*, pages 330–334. Routledge, 2017.

M. O'Neill, J. McDermott, J. M. Swafford, J. Byrne, E. Hemberg, A. Brabazon, E. Shotton, C. McNally, and M. Hemberg. Evolutionary design using grammatical evolution and shape grammars: Designing a shelter. *International Journal of Design Engineering*, 3(1): 4–24, 2010.

H. Osborne. Symmetry as an aesthetic factor. *Computers & Mathematics with Applications*, 12(1-2):77–82, 1986.

J. Osborne and F. Farley. The relationship between aesthetic preference and visual complexity in absract art. *Psychonomic Science*, 19(2):69–70, 1970.

S. A. Osimo, R. Pizarro, B. Spanlang, and M. Slater. Conversations between self and self as Sigmund Freud—A virtual body ownership paradigm for self counselling. *Scientific reports*, 5(1):13899, 2015.

U.-M. O'Reilly and M. Hemberg. Integrating generative growth and evolutionary computation for form exploration. *Genetic Programming and Evolvable Machines*, 8(2):163–186, 2007.

S. Palmer and I. Rock. Rethinking perceptual organization: The role of uniform connectedness. *Psychonomic bulletin & review*, 1(1):29–55, 1994.

M. Pelowski. Tears and transformation: Feeling like crying as an indicator of insightful or "aesthetic" experience with art. *Frontiers in Psychology*, 6:1006, 2015.

M. Pelowski, P. S. Markey, J. O. Lauring, and H. Leder. Visualizing the impact of art: An update and comparison of current psychological models of art experience. *Frontiers in human neuroscience*, 10:160, 2016.

M. Pelowski, M. Forster, P. P. Tinio, M. Scholl, and H. Leder. Beyond the lab: an examination of key factors influencing interaction with 'real'and museum-based art. *Psychology of Aesthetics, Creativity, and the Arts*, 11(3):245, 2017.

L. Plunkett. AI creating "art" is an ethical and copyright nightmare, Aug 2022. URL `https://kotaku.com/ai-art-dall-e-midjourney-stable-diffusion-copyright-1849388060`.

C. Post. The art of digital curation: Co-operative Stewardship of net-based art. *Archivaria*, (92):6–47, 2021.

C. Pugach, H. Leder, and D. J. Graham. How stable are human aesthetic preferences across the lifespan? *Frontiers in human neuroscience*, 11:289, 2017.

K. Quaigrain and A. K. Arhin. Using reliability and item analysis to evaluate a teacher-developed test in educational measurement and evaluation. *Cogent Education*, 4(1): 1301013, 2017.

A. Radford, J. W. Kim, C. Hallacy, A. Ramesh, G. Goh, S. Agarwal, G. Sastry, A. Askell, P. Mishkin, J. Clark, et al. Learning transferable visual models from natural language supervision. In *International conference on machine learning*, pages 8748–8763. PMLR, 2021.

C. Reynolds. Evolving textures from high level descriptions. *International Journal of Arts and Technology*, 9(1):26–38, 2016.

G. Rhodes. The evolution of facial attractiveness. *Annual Review of Psychology*, 57:199–226, 2006.

J. Rigau, M. Feixas, and M. Sbert. Conceptualizing Birkhoff's Aesthetic Measure Using Shannon Entropy and Kolmogorov Complexity. In *Computational Aesthetics*, pages 105–112, 2007.

A. Rodrigues, A. Cardoso, and P. Machado. Generation of Aesthetic Emotions guided by Perceptual Features. 2018.

R. Rombach, A. Blattmann, D. Lorenz, P. Esser, and B. Ommer. High-resolution image synthesis with latent diffusion models. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 10684–10695, 2022.

J. Romero, P. Machado, A. Santos, and A. Cardoso. On the development of critics in evolutionary computation artists. In *Applications of Evolutionary Computing: EvoWorkshops 2003: EvoBIO, EvoCOP, EvoIASP, EvoMUSART, EvoROB, and EvoSTIM Essex, UK, April 14–16, 2003 Proceedings*, pages 559–569. Springer, 2003.

B. Ross, W. Ralph, and H. Zong. Evolutionary image synthesis using a model of aesthetics. In *IEEE International Conference on Evolutionary Computation*, pages 1087–1094. IEEE, 2006.

N. A. Salingaros. Symmetry gives meaning to architecture. *Symmetry Cult. Sci*, 31:231–260, 2020.

L. Salkind and N. J. Salkind. Gender and age differences in preference for works of art. *Studies in Art Education*, 38(4):246–256, 1997.

J.-P. Sartre. *The imaginary: A phenomenological psychology of the imagination*. Routledge, 2010.

J. Secretan, N. Beato, D. D Ambrosio, A. Rodriguez, A. Campbell, and K. Stanley. Picbreeder: evolving pictures collaboratively online. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1759–1768. ACM, 2008.

J. Shelley. Empiricism: Hutcheson and Hume. In *The Routledge companion to aesthetics*, pages 55–68. Routledge, 2005.

R. Shusterman. Pragmatism. In *The Routledge Companion to Aesthetics*, pages 96–105. Routledge, 2013.

F. Sibley. Aesthetic concepts. *The philosophical review*, 68(4):421–450, 1959.

K. Sims. Artificial evolution for computer graphics. In *Proceedings of the 18th annual conference on Computer graphics and interactive techniques*, pages 319–328. ACM, 1991.

K. Sims. Galapagos interactive exhibit, 1997. URL `http://www.karlsims.com/galapagos/`.

M. Skov. Aesthetic appreciation: The view from neuroimaging. *Empirical Studies of the Arts*, 37(2):220–248, 2019.

M. Slater. A note on presence terminology. *Presence connect*, 3(3):1–5, 2003.

E. Specker, P. P. Tinio, and M. van Elk. Do you see what I see? An investigation of the aesthetic experience in the laboratory and museum. *Psychology of Aesthetics, Creativity, and the Arts*, 11(3):265, 2017.

B. Spehar, C. Clifford, B. Newell, and R. Taylor. Universal aesthetic of fractals. *Computers & Graphics*, 27(5):813–820, 2003.

N. J. Stone. Environmental view and color for a simulated telemarketing task. *Journal of Environmental Psychology*, 23(1):63–78, 2003.

H. Takagi. Interactive evolutionary computation: Fusion of the capabilities of EC optimization and human evaluation. *Proceedings of the IEEE*, 89(9):1275–1296, 2001.

Tate. Concrete art, 2017. URL `https://www.tate.org.uk/art/art-terms/c/concrete-art/`. [Online; Accessed: 2020-11-20].

J. Taylor. From Trump Nevermind babies to deep fakes: Dall-E and the ethics of AI Art, Jun 2022. URL `https://www.theguardian.com/technology/2022/jun/19/from-trump-nevermind-babies-to-deep-fakes-dall-e-and-the-ethics-of-ai-art`.

R. P. Taylor. Reduction of physiological stress using fractal art and architecture. *leonardo*, 39(3):245–251, 2006.

R. P. Taylor, B. Spehar, P. Van Donkelaar, and C. M. Hagerhall. Perceptual and physiological responses to Jackson Pollock's fractals. *Frontiers in human neuroscience*, 5:60, 2011.

R. Thornhill and S. W. Gangestad. Facial attractiveness. *Trends in cognitive sciences*, 3 (12):452–460, 1999.

Y. Tian. Evolving Three Dimension (3D) Abstract Art: Fitting Concepts by Language. *arXiv preprint arXiv:2304.12932*, 2023.

P. P. Tinio and H. Leder. Just how stable are stable aesthetic features? Symmetry, complexity, and the jaws of massive familiarization. *Acta Psychologica*, 130(3):241–250, 2009.

S. Toksöz and A. Ertunç. Item analysis of a multiple-choice exam. *Advances in Language and Literary Studies*, 8(6):141–146, 2017.

I. Tweraser, L. E. Gillespie, and J. Schrum. Querying across time to interactively evolve animations. In *Proceedings of the Genetic and Evolutionary Computation Conference*, pages 213–220, 2018.

C. W. Tyler. Human symmetry perception. *Human symmetry perception and its computational analysis*, pages 3–22, 1996.

I. Valori, P. E. McKenna-Plumley, R. Bayramova, C. Zandonella Callegher, G. Altoè, and T. Farroni. Proprioceptive accuracy in Immersive Virtual Reality: A developmental perspective. *PloS one*, 15(1):e0222253, 2020.

T. Van Doesburgh. ase de la peinture concret. *Art Concret*, 1930.

P.-P. Vázquez, M. Feixas, M. Sbert, and W. Heidrich. Viewpoint selection using viewpoint entropy. In *VMV*, volume 1, pages 273–280, 2001.

D. Ventura and D. Gates. Ethics as Aesthetic: A Computational Creativity Approach to Ethical Behavior. In *ICCC*, pages 185–191, 2018.

A. Vinhas, F. Assunção, J. Correia, A. Ekárt, and P. Machado. Fitness and novelty in evolutionary art. In *International Conference on Computational Intelligence in Music, Sound, Art and Design*, pages 225–240. Springer, 2016.

J. Wagemans, J. H. Elder, M. Kubovy, S. E. Palmer, M. A. Peterson, M. Singh, and R. Von der Heydt. A century of Gestalt psychology in visual perception: I. perceptual grouping and figure–ground organization. *Psychological bulletin*, 138(6):1172, 2012.

R. Wages, S. M. Grünvogel, and B. Grützmacher. How realistic is realism? Considerations on the aesthetics of computer games. In *International Conference on Entertainment Computing*, pages 216–225. Springer, 2004.

T. E. Wartenberg. Heidegger. In *The Routledge Companion to Aesthetics*, pages 116–125. Routledge, 2013.

S. Weale. Cambridge college seeks to remove memorial to patron with links to slave trade, Nov 2021. URL https://www.theguardian.com/education/2021/nov/15/cambridge-college-seeks-to-remove-memorial-to-patron-with-links-to-slave-trade.

H. Weichselbaum, H. Leder, and U. Ansorge. Implicit and explicit evaluation of visual symmetry as a function of art expertise. *i-Perception*, 9(2):2041669518761464, 2018.

M. Wertheimer. Gestalt psychology. *Source Book of Gestalt Psychology. New York: Harcourt, Brace and Co*, 1938.

H. Weyl. *Symmetry*, volume 47. Princeton University Press, 2015.

WHO. Mental health, 2023. URL https://www.who.int/health-topics/mental-health#tab=tab_2.

G. A. Wiggins. A preliminary framework for description, analysis and comparison of creative systems. *Knowledge-Based Systems*, 19(7):449–458, 2006.

R. Winegar. Good Sense, Art, and Morality in Hume's 'Of the Standard of Taste'. *Journal of Scottish Philosophy*, 9(1):17–35, 2011.

R. Wollheim. Nelson Goodman's languages of art. *The journal of philosophy*, 67(16):531–539, 1970.

J. Yang, Y. Sun, J. Liang, B. Ren, and S.-H. Lai. Image captioning by incorporating affective concepts learned from both visual and textual components. *Neurocomputing*, 328:56–68, 2019.

A. Zayed. Statue of slave trade investor may be removed from Cambridge uni, Nov 2020. URL https://www.cambridge-news.co.uk/news/cambridge-news/slavetrade-investor-statue-removed-university-19217850.

S. Zeki. *Inner Vision: An Exploration of Art and the Brain.* Wiley-Blackwell, 2002.

S. Zeki. Beauty in Architecture: Not a Luxury-Only a Necessity. *Architectural Design*, 89 (5):14–19, 2019.

Appendix A

Function definitions included in the Genetic program

Function Name	Description	Formula	Arity
Random	A random number between 0 and 1	Rand	0
Index	The index of the shape within the sculpture	index	0
Sine	Calculates the quotient between the opposite and hypotenuse	sin(x)	1
Cosine	Calculates the quotient between the adjacent and hypotenuse	cos(x)	1
Power	Returns the provided value raised to the power of 1.5	x1̂.5	1
Absolute	Returns the absolute value of the provided number	Abs(x)	1
Floor	Returns the largest integral value lower than the specified value	Floor(x)	1
Ceiling	Returns the smallest integral value higher than the specified number	Ceiling	1
Square Root	Returns the Square root of a provided number	Sqrt(x)	1
Log2	Returns the base 2 logarithm for a provided number	Log2(x)	1
Log10	Returns the base 10 logarithm for a provided number	Log10(x)	1
Tangent	Returns the tan value for the provided number	tan(x)	1
Exponential	Returns the exponential factor for the provided number	exp(x)	1
Power3	Returns the cube of the provided number	x^3	1
Cube root	Returns the cube root of the provided number	$\sqrt[3]{x}$	1
Add	Returns the sum of 2 numbers	x + y	2
Subtract	Returns the difference between two numbers	x - y	2
Multiply	Returns the product of 2 numbers	x * y	2
Divide	Returns the quotient between 2 numbers	x / y	2
modulo	Returns the remainder of the division of 2 numbers	x % y	2
Max	Returns the higher number out of the provided options	(x ¿= y) ? x : y	2
Min	Returns the smallest number out of the provided options	(x ¡= y) ? x : y	2
Average	Returns the average between two numbers	(x + y) / 2	2
Egg Crate	Formula defines an egg carton shape when plotted on a 3D graph	$X^2 + y^2 + 25(sin^2(x) + sin^2(y))$	2
Matyas	Formula which defines a single minima and two maxima	$0.26(x^2 + y^2) - 0.48xy$	2
Circle	Formula calculating the radius of a circle	$x^2 + y^2$	2
If Then Else	Compare the first value to a random number	(x ¿ rand) ? y : z	3

Table A.1: Available Expressions for the Expression Trees

Appendix B

Sculpture tag choices

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Busy	5	4	4	1	4	2	6	7	4	1	4	2	4	1	4	2	1	2
Cold	3	2	2	2	0	4	1	3	3	3	3	0	2	2	3	1	1	2
Quiet	2	3	1	2	1	2	0	0	0	3	2	4	1	1	0	2	0	1
Angular	1	1	3	3	0	0	2	2	5	0	2	0	5	4	5	3	2	0
Unrefined	0	0	0	2	0	3	1	0	0	3	1	1	0	1	0	0	0	2
Ordered	3	6	5	2	2	1	5	4	3	0	4	2	3	3	2	2	5	2
Complex	4	3	5	0	2	1	3	4	6	1	3	1	7	2	5	2	4	0
Boring	0	0	2	4	1	5	1	1	1	5	0	2	1	2	1	5	1	4
Drab	1	0	0	3	1	3	0	0	0	3	0	1	0	1	0	2	0	3
Friendly	4	5	3	0	3	0	2	5	1	0	4	3	1	1	1	2	4	0
Static	4	4	3	3	1	3	2	2	2	4	1	3	0	2	2	4	1	4
Separate	1	1	3	3	1	3	3	2	1	3	2	2	1	1	0	2	1	4
Original	0	2	2	1	1	0	1	1	2	1	2	4	2	1	2	2	1	0
Natural	2	2	1	1	0	1	0	2	4	0	1	1	2	3	2	0	4	1
Calm	5	6	2	2	1	1	2	3	2	1	4	4	1	2	1	2	1	0
Warm	2	4	2	1	5	1	5	3	1	1	1	2	2	1	1	2	2	3
Loud	2	1	3	2	4	2	4	3	1	1	1	2	2	2	0	1	2	3
Curved	4	7	3	1	0	2	7	2	10	2	10	3	6	8	9	0	10	0
Sophisticated	1	3	1	1	2	0	2	0	0	1	1	3	1	3	0	2	2	0
Disordered	4	1	4	2	5	4	1	4	2	4	1	2	4	2	2	2	2	2
Simple	3	4	4	7	0	9	1	2	2	10	3	5	1	6	1	5	4	8
Interesting	4	6	4	1	3	0	5	2	2	1	6	6	5	2	2	4	4	0
Bright	3	5	4	1	3	1	4	2	2	1	5	5	5	0	3	3	4	0
Unfriendly	1	0	0	2	1	5	2	1	4	4	1	1	2	3	3	2	2	4
Dynamic	7	5	4	5	8	3	6	6	8	2	7	5	9	7	9	3	8	2
Connected	3	6	4	3	3	2	6	4	5	2	7	4	5	4	5	1	7	2
Unoriginal	0	0	0	2	0	2	0	0	0	2	0	0	0	2	0	2	0	1
Unnatural	3	1	3	1	4	2	2	4	4	2	2	2	3	1	5	4	2	2

Table B.1: Sculpture tag matrix - number of times tags applied to each sculpture

Appendix C

Collected data for the sculpture measures

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	Avg	Calc	Measure
c3d5c	4	4	3	3	4	5	4	3	3	4	4	3	3	3	4	3	4	4	4	3	3	3.571429	0.8	Connected
8eeb4	2	3	2	3	3	3	3	3	1	2	2	1	3	3	3	3	2	3	4	3	3	2.619048	0.4	Connected
c2946	2	3	2	2	3	4	3	1	1	1	2	1	3	3	3	3	3	3	4	1	2	2.380952	0.2	Connected
9d500	5	3	5	2	3	3	3	3	3	2	5	1	3	4	4	3	3	3	4	1	2	3.095238	0.0	Connected
ae4f6	4	3	4	3	3	4	3	3	1	5	4	3	3	5	4	3	2	2	4	4	4	3.380952	1.0	Connected
d39fb	4	4	3	3	4	5	4	3	5	5	1	2	3	4	1	1	1	2	4	5	5	3.285714	0.6	Connected
0f083	4	4	3	3	4	5	4	3	5	4	4	3	3	3	4	1	4	4	4	2	3	3.523810	0.4	Connected
3a195	4	3	3	3	3	5	3	3	5	4	1	2	3	4	1	3	2	2	4	5	4	3.190476	0.8	Connected
f0248	4	3	4	3	3	5	3	3	3	5	4	3	3	5	1	3	2	2	4	4	4	3.380952	0.6	Connected
3138c	4	3	4	3	3	4	3	3	3	4	4	1	3	5	4	3	2	2	4	3	3	3.238095	1.0	Connected
0649f	5	1	5	2	3	4	3	1	1	2	5	2	3	4	4	3	3	3	4	2	2	2.952381	0.2	Connected
dfd65	5	3	5	3	3	3	3	1	1	2	2	2	3	3	4	3	3	3	4	2	2	2.857143	0.0	Connected

Table C.1: Connected Ratings

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	Avg	Calc	Measure
cd866	2	3	5	3	3	1	1	2	5	2	3	1	3	3	1	4	2	3	1	1	2	2.428571	0.2	Calm
12328	1	3	5	3	3	4	4	3	5	3	4	2	2	4	2	4	3	3	5	3	3	3.285714	0.8	Calm
ae0fe	1	3	3	1	3	2	3	3	5	3	3	2	1	4	2	3	3	2	5	2	3	2.714286	0.6	Calm
d8a7f	3	3	5	4	4	1	1	2	2	1	1	1	3	1	1	4	2	2	1	1	2	2.142857	0.0	Calm
68907	1	3	3	4	4	1	1	2	2	2	1	1	3	4	1	4	1	2	1	1	5	2.238095	0.4	Calm
129dd	2	3	2	1	3	1	3	3	4	2	3	1	1	2	1	3	1	2	1	2	3	2.095238	0.4	Calm
c48ae	2	3	1	3	1	5	3	1	4	4	5	3	5	5	3	1	5	3	1	4	3	3.095238	1.0	Calm
501c3	2	3	5	3	3	1	1	2	2	2	1	2	2	3	1	4	2	3	1	2	2	2.238095	0.2	Calm
e4d7b	1	3	2	1	1	3	3	1	4	4	5	1	5	5	2	1	5	2	5	2	3	2.809524	0.8	Calm
c1423	2	3	2	1	3	5	4	1	4	4	4	3	5	5	3	1	5	3	5	5	2	3.333333	1.0	Calm
65cf0	3	3	5	4	3	3	3	3	5	2	3	2	3	4	1	4	3	2	5	2	2	3.095238	0.6	Calm
e6ebe	1	3	5	3	4	1	1	2	5	1	1	1	3	4	1	4	1	1	1	1	5	2.333333	0.0	Calm

Table C.2: Calm Ratings

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	Avg	Calc	Measure
bc447	3	4	3	3	3	5	4	2	3	3	5	4	4	5	5	2	5	2	4	3	3	3.571429	0.8	Friendly
955a1	3	4	4	3	5	5	3	3	3	3	5	4	3	5	5	2	5	2	3	4	3	3.666667	0.8	Friendly
37526	1	2	3	1	3	1	1	2	5	2	4	1	2	1	2	3	1	4	4	2	1	2.190476	0.2	Friendly
724c6	4	4	4	5	5	5	3	4	3	4	5	3	3	4	5	2	4	1	4	3	1	3.619048	0.2	Friendly
c7087	4	4	4	5	5	5	3	1	3	3	5	3	4	5	5	3	5	1	3	3	3	3.666667	0.6	Friendly
ca390	3	2	4	1	3	4	2	4	5	4	4	2	5	3	5	2	1	1	3	4	4	3.142857	0.4	Friendly
7102d	3	2	2	1	3	1	2	3	5	1	2	1	1	1	1	3	1	4	3	1	1	2.000000	0.0	Friendly
906ff	5	4	3	5	2	4	4	4	5	5	4	4	5	4	5	3	5	4	3	3	5	4.095238	0.9	Friendly
bf03d	5	2	2	5	3	1	1	1	1	1	2	1	1	2	1	3	3	4	3	1	2	2.142857	0.0	Friendly
6defc	5	2	3	1	3	4	4	1	1	3	3	4	5	5	5	3	4	2	4	4	1	3.190476	0.4	Friendly
d8dab	3	2	3	1	3	4	2	3	3	4	2	1	2	3	5	2	3	4	4	2	4	2.857143	0.6	Friendly
00c8b	4	3	4	5	5	5	1	3	5	5	4	2	3	4	5	2	4	1	3	3	4	3.571429	1.0	Friendly

Table C.3: Friendly Ratings