

Generation and Evaluation of Artworks

Penousal Machado

Instituto Superior de Engenharia de Coimbra
3030 Coimbra, Portugal
machado@alma.uc.pt

Amílcar Cardoso

Dep. Eng. Informática, Uni. Coimbra, Polo II
3030 Coimbra, Portugal
amilcar@dei.uc.pt

Abstract: This paper is dedicated to the development of *constructed artists*, i.e., computer programs capable of creating artworks with little or no human intervention. We make an analysis and critic of some of the most prominent work on this field. We give a description of the main characteristics that a system should have, in order to be considered a *constructed artist*. These characteristics include the capacity of making aesthetic judgments, which takes us to the origins of art and aesthetics, for which we present a brief theory. Finally we propose a model for the development of a *constructed artist*. This model has the capability of performing aesthetic evaluation, through the use of neural networks. The images are generated using a genetic algorithm, and represented using Fractal Image Encoding. This type of methodology allows the representation and, consequently, the generation of any type of image.

1. Introduction

In this paper we will talk about the development of computer programs capable of creating artworks. We would like to stress the importance of this type of systems, capable of performing creative tasks, as a way of showing AI's potential, and bringing AI closer to the layperson. The first applications of computers and AI to the field of the arts date from a long while. These applications had a stronger influence in music than in visual arts. This outcome is not surprising since the memory requirements necessary for image handling are substantially bigger than those necessary for sound. Additionally, music theory is more developed and quantitative than theory in visual arts (*Kurzweil 1990*).

In section two, we make an assessment of the current "state of the art" in this field. In doing so, we specify a set of features that current systems lack, and that should be present. The third section pertains to the origins of art and aesthetic judgement, we give a short biological explanation to the devotion of humans to art and to how natural evolution favoured the appearance of art. We also give evidence to the sharing of aesthetic values with other species. In section four we propose a model for a *constructed artist*. This model overcomes some of the flaws that the *current constructed artists* exhibit. Finally, in section five, we describe the current state of development of our system, draw some conclusions, and point towards unexplored aspects in this field.

2. State of the Art

The vast majority of AI applications in the field of the arts falls into two categories: Systems performing some sort of *art understanding* task, such as musical analysis, and systems that work as "intelligent" tools for human artists (Spector & Alpern 1994); and a new range of applications that is beginning to emerge, the *constructed artists* which "are supposed to be capable of creating aesthetically meritorious artworks on their own, with minimal human intervention." (Spector & Alpern 1994). Restricting ourselves to the field of visual arts, which is our main field of interest, we will describe two approaches that have gained a vast acceptance.

Harold Cohen can be considered as the precursor of the rule based approach (*Kurzweil 1990*). His system, Aaron, is probably the most acclaimed *constructed artist*. The development of Aaron took approximately one decade, which gives an idea of the amount of work involved in coding the knowledge necessary to do artworks into rules. This set of rules is extremely valuable since it provides an accurate description of the artwork's theory and structure (*Kurzweil 1990*).

The second approach is rooted in a computer program written by Richard Dawkins. This program evolves images of virtual creatures (biomorphs), through the use of a genetic algorithm (Dawkins 1987). From an initial random set of *biomorphs*, the user chooses the most aesthetically pleasing, the next generation is created through the mutation of the

genetic code of the selected *biomorph*; thus, the fitness function results are supplied by the user; this technique was named *interactive evolution*. This methodology provides a way of achieving flexibility and complexity with a minimum of user input and knowledge of details. This “simple” idea served as a base for a large number of applications in several fields, including the field of visual arts (some of the examples are (Sims 1991), (Graf & Banzhaf 1995), (Todd & Latham 1992)). The main difference between these applications is in the coding of the images. Karl Sims, for instance, uses mathematical functions, coded in the form of Lisp S-expressions. The program generates the images from the S-Expressions; mutation and crossover is also performed at the S-expression level. As far as we know, there is no evolutionary algorithm application that works directly with the images as bitmaps.

As we said before, these systems have been highly successful; yet, in our opinion, they have weaknesses that might prevent them from being considered *constructed artists*. Let us consider the characteristics that we think a *constructed artist* should display. The system should exhibit *generic representational capabilities*, thus, it should be possible to represent any kind of image. None of the systems described has this representational power, what hinders their generation ability. An human artist learns over time, a constructed artist should also be able to *learn*. Aaron is incapable of learning, Karl Sims’s system performs some kind of learning, since it gradually improves its performance, based on the user’s evaluation of the generated artworks; However, a human artist doesn’t start from zero, like Karl Sims’s system does¹, in fact a human artist has access to the artworks made by others. He/she is able of learning from these examples, and eventually using them as source of inspiration. *Integrating knowledge* in a constructed artist will certainly improve its performance, it is not by chance that the most successful constructed artist, Aaron, is deeply knowledge based. To be independent from humans, a constructed artist must be able to recognize an artwork when it sees it, this will enable it to evaluate its own artworks and guide the generation process. Thus, a *constructed artist* must be able to perform *aesthetic judgments*. This is probably the most important task to achieve, unfortunately it is also the most difficult one. The issue of aesthetic judgment was first addressed by Plato, and since then the debates go on (Quintás 1987). There is a large number of theories regarding aesthetic judgment, and their relative values are unknown. The lack of a strong theory of aesthetic value, poses several problems, including the issue of validation of the developed systems.

3. Aesthetic Judgment and The Origins of Art

If we ask someone why he/she likes a certain painting, people will usually talk about the emotions or feeling triggered by the artwork, the combination of colors, global composition of the painting, or, even more frequently, we will get the intriguing answer: “I just like it.”.

From our viewpoint, the assessment of the aesthetic value of an artwork is influenced by the “*content*” of the artwork (which can trigger emotions, etc) and also the “*visual aesthetic value*” of the artwork (color combination, composition, etc). The aesthetic value of an artwork rests on these variables and their interaction. It is possible to have an artwork that is visually pleasing, but whose *content* is displeasing, in fact many art styles rely on the mixed feelings caused by this discrepancy (e.g. many of Salvador Dali’s paintings). We also believe that these variables are independent. In other words, if an image has a high *visual aesthetic value*, it will have a high aesthetic value, independently from its *content*, and even if it is deprived of *content*². We don’t mean that *content* isn’t important, we just mean it is not indispensable.

It seems clear that the way how *content* influences the aesthetic value of a given artwork depends, mainly, on cultural issues. From our point of view, *visual aesthetic value*, is directly connected to visual image perception and processing and is, therefore, mainly: biological, hardwired, and thus universal. The remainder of this section is dedicated to the support of the previous statement.

¹ The initial population is generated randomly.

² The existence of images deprived of *content* or meaning is controversial.

3.1 Art's Origin

“Why did man devoted himself to art?” To give answer to this question we will give an explanation of how natural selection favored the appearance of art. Natural selection should favor the fittest individuals in a population, so why should a seemingly useless activity as art be favored by it?

To the vast majority of the animals, the struggle for survival takes all their time. Only in their infancy they have time to playing and games. The same happened to the primitive man. Only from a certain point in history man begun to have spare time.

The first manifestation of art, as we know, is the “stone of Makapansgat”. This stone was recollected 3 million years ago by our ancestors, the Australopithecus. It was not created by them, nevertheless the single act of recognition of the forms of an human face in a stone is significant by itself. The first act of artistic creation, dates from 300.000 years, and is the representation of a female figure.

The appearance of art may be explained by the necessity of using other forms of communication other than gesture or speech. There are, however, other explanations that can be considered complementary.

The coordination of hand movements had a great influence on human evolution. When a prehistoric man devoted himself to painting he was not doing a fruitless activity: by painting he was also developing, and making an exhibition of, its manual ability. The exhibition of ability is important since it would, eventually, give him an advantage by making him/her more desirable for mating, considering that animals tend to choose the most fit individual they can for mating partners.

Intriguingly, some animals develop characteristics that are disadvantageous for their immediate survival, e.g. the tail of the paradise bird. The explanation for this paradoxical event is that: to be able to bear this disadvantage an animal needs to be very well fit. Consequentially animals displaying simultaneously: these disadvantageous features, and successfulness, would be favored as mating partners (e.g. Dawkins 1987). By dedicating himself to art, an individual is showing that he is highly fit, since he has time to devote himself to activities other than those necessary for his/hers immediate survival.

This set of explanations gives a reasonable justification for the devotion of man to art, they don't justify, however, why do we find certain images beautiful, aesthetically pleasing, or artistic. As we said before, we believe that *visual aesthetic value* is directly connected to visual image perception, having biological roots. To justify this statement we must first show that there is a *visual aesthetic value* that is universal and independent from cultural issues, and thus hardwired. The analysis of children's paintings allows us to observe the development of aesthetic. Between the ages of one and three years, the infants have difficulties to accurately control their hand movements. Among the ages of two to three years, they are capable of making powerful lines, soon simple images begin to emerge from the chaos of lines. The next stage is the appearance of the first pictorial images. Usually, the first pictorial image is the human figure. This image is, also, always constructed in the same, rather puzzling way. It begins with an empty circle; in the next step bubbles are added to the inside of the circle; the bubbles are gradually transformed in eyes, mouth and nose; afterwards hair is included; Some of the hairs get longer, till they are transformed into arms and legs (Morris, 1994). Somewhere between the ages of six and twelve, these universal images begin to disappear due to educational influence.

It is safe to say that visual aesthetic judgment is not particular to humans, in fact we share this ability with other species of animals. Experiments with chimpanzees show that they follow the same steps of development of human infants. The first stages of development are similar, chimpanzees, however, aren't able to go beyond the phase of the circle into the phase of the filled circle. They never seem to be able to create a pictorial image. Nevertheless, their paintings show that the brain of a chimpanzee is capable of making simple aesthetic judgments, displaying control of the composition and thematic variation (Morris, 1994). It is important to notice that the paintings are created without any kind of training or conditioning. Experiments with other species of animals seem to indicate that there are other species (e.g. cats (Busch & Silver 1995)) also capable of making visual aesthetic judgments. Unfortunately many of the examples that we found, lack credibility.

Till the appearance of photography, visual arts bearded the burden of representation. This resulted in the development of a high technical competence, usually achieved by sacrificing the recreational aspects that are the roots of children's art (to an analysis on the connections of art and aesthetic with games and play see Quintás, 1987). With the appearance of photography, the artists became more experimental gradually returning to the recreational roots of art (the cubist's art, for instance, is very similar to some forms of tribal art).

To talk about visual aesthetic, it is important to consider how the visual perception system works. Today, we know the process of transformation of the ocular image into its digital representation. From this digital representation, the brain constructs internal representations, retaining only certain aspects of the image. The way how this process works is still source of much debate. The idea that there is a “pre-processing” of the digital image (shape and contour detection, color analysis, depth and movement analysis, etc), and that “recognition” and subsequent transformation to internal representations is made based on the results of this “pre-processing”, is usually accepted (to a brief description of the visual perception see Vignaux, 1991). From the previous statement, we can say that there is a difference between image complexity, and internal representation complexity; furthermore a complex image isn’t necessarily difficult to (pre)process. To clarify our previous statement, consider the following analogy: a fractal image is usually complex, and highly detailed, yet it can be compactly described by a simple mathematical formula.

In the book *The Society of Mind*, Minsky associates the concepts of fashion and style to the mental work necessary to process images: “...why do we tend to choose our furniture according to systematic styles or fashions? Because familiar styles make it easier for us to recognize and classify the things we see. If every object in a room were distracting by itself, our furniture might occupy our minds too much... It can save a lot of mental work...”³ (Minsky, 1986). If we accept this explanation we are lead to conclude that simpler, easier to process, images are more beautiful than complex ones. If we think about this theory we will conclude, rather paradoxically, that an empty white sheet is more beautiful than any artwork, since it is certainly more easy to process. Minsky’s examples are related to office furniture, when we are in an office we don’t want to be distracted, you want to work⁴; when we are admiring an artwork, we want to be distracted, that’s probably why we usually have, in our offices, a painting to look at when we want to distract ourselves.

In our opinion the *aesthetic visual value* of an artwork depends on two factors: (1) The amount of work necessary to process the image (the simpler, the better); (2) The image complexity of the image (complex images are more beautiful than simple ones). This seems contradictory, but as we said before a complex image isn’t necessarily difficult to process. Thus, images that are simultaneously visually complex and easy to process are the images that are more beautiful. Our state of mind influences how we value this factors, if we are tired, we will probably give more importance to processing simplicity. The importance of recognizability (in the sense of easiness of processing) is present in the works of many artists. M.C. Escher, for instance, devoted a lot of attention to how the coloring of images should be made, in order to increment the recognizability of its patterns (Schattschneider, 1990). Returning to the fractal example, fractal images are usually complex, having detail at all levels, the propriety of self-similarity makes these images easier to process, which gives an explanation to why we usually find fractal images beautiful. The characteristic of various levels of detail can also be found in many artworks, (e.g. Kandinsky’s works). When we look briefly at such an artwork we are automatically able to recognize its main shapes, if we give it more attention we will increasingly discover more detail. This is important since it makes the image easy to process, and thus less distracting if we don’t want to give it attention, simultaneously if we want to give it attention we will always find enough detail to “fill” our minds. If the image had only one level of detail, it would probably make it either difficult to process *rapidly* or with little complexity. This hinders the generality of the artwork in the sense that our willingness to look at it would largely depend on our state of mind.

One final remark goes to the fact that the above statements also could be made if we were referring to other fields of art like music.

4. Proposed Model

In the development of a model for a “*constructed artist*”, we took into consideration the previously referred characteristics that a constructed artist should have. In what concerns to representational issues we, have the problem of developing a generic system that, simultaneously, lessens the amount of information needed to represent the images. The amount of information needed to work with images in bitmap format makes it impossible to use this form of representation. Methods of image compression that rely on Huffman coding, or similar algorithms, aren’t a solution to this problem, since the amount of information remains the same, and we can’t work directly with the coded image. What is needed, is a method that relies on the transformation of the images, e.g. transformation of the bitmap image into a set of lines, and filling the shapes generated by these lines with an adequate colour. Considering these problems, we chose Fractal Image Encoding as a way of representing images. It was demonstrated by Barnesley that any image can be represented through a Partitioned Iterated Function System (PIFS), resulting, generally, in a significant reduction on the amount of information needed to represent the image (Fisher 1995). The possibility of representing any image is

³ This is in some way similar to the “principle of economy” found in evolution’s theory.

⁴ This statement is not based on personal experience, and remains to be proved.

important, since we want our *constructed artist* to have access to the works of human artists. The transformation of the an image into its PIFS representation, involves a high computational cost; fortunately this process can be easily be implemented by a parallel algorithm.

For the generation of the images, we rely on a Genetic Algorithm. This algorithm works with the images in their PIFS encoding. From an initial population of images, the system selects the ones with highest *visual aesthetic value* (the process of image evaluation will be described later). The next generation is created through mutation and recombination (through crossover) of the selected images. This type of approach relies on the assumption that the combination of two highly fit images results, at least generally, in a fit image. This assumption isn't always true, so, to improve the system's performance, we use a matting operator. This operator selects sets of "compatible" images, crossover is performed using these images.

In what concerns to the possibility of integrating background knowledge, we propose two different methods. The initial population does not need, necessarily, to be random: we can use any set of images, including famous artworks, as initial population. Additionally, we can maintain a knowledge base of images, which are not subject of selection. The genetic algorithm can select these images for matting with images of the current population or between themselves. Additionally, images of high aesthetic value can be added to the knowledge base, preventing the early lost of remarkable images.

As we said before, aesthetic evaluation is probably the most difficult problem to tackle. In our model, evaluation is performed through neural networks. The evaluation of the aesthetic value of an image is made by a set of neural networks; each of these networks is concerned with a different characteristic of the image. Thus, we decompose the task of aesthetic evaluation into several ones. There are two "layers" of neural networks. The neural networks in the first "layer" take, as input, the PIFS representation of the image; The results provided by this "layer" are used as input for the second layer (that has a single network). The network on the second layer gives the aesthetic value of the picture. The decomposition of the task of aesthetic evaluation is based on the aesthetic theory earlier described. Each of the networks in the first "layer" evaluates a basic principle of aesthetic order: unity, predominance, variety, balance, continuity, symmetry, proportion and rhythm.

The use of a set of networks, each assigned to a relatively simple task, enables easier training than would be possible otherwise. The training of the neural networks is made using examples taken from psychological tests that were developed to assess the aesthetic evaluation ability of humans, e.g. (Graves, 1977). The set of examples is completed with images of artworks. We test the evaluation system by comparing the results attained by the system in psychological tests (that were not used as examples) with the results of humans.

All the steps of the process work with the image in its PIFS representation. This saves a lot of computing time, since the transformation of bitmap images to PIFSs is a time consuming process. One of the advantages of the described model is its modularity; We can independently develop and test any of the modules - evaluation, generation of images, and fractal image encoding.

5. Conclusions and Future Work

Our system is still under development, and no results are available yet. One of the advantages of our model is its great modularity. There are three main modules to implement: image generation, aesthetic evaluation and fractal image encoding. We can implement and test any of these modules individually. The generation and encoding modules are nearly finished. The first results of the system, with the user making the evaluation of the generated images and thus guiding the evolution process, should be available in the near future.

In this paper we specified a set of features that the current *constructed artists* lack. We justified the importance of this set of features. It is our conviction that the proposed model provides a feasible way of integrating these features. The design of the model was based on the theory of aesthetic that we presented. In this theory, we consider the biological roots of art and aesthetic knowledge. We state that the aesthetic value of an image depends on two features: image's *content* and image's *visual aesthetic value*, furthermore we claim that *visual aesthetic value* is directly connected to the image processing task of the brain.

Little as been said about music, we think that much of what we stated about visual arts could also be applied to music. One of the conjectures, we would like to test, is the sharing of aesthetic values between different fields such of visual art and music. Thus, can a neural network trained to make *visual aesthetic judgements* make, without changes, aesthetic judgements in the musical field?

6. Acknowledgments

We would like to thank Francisco Colunas, Paulo Gomes and Carlos Grilo for helping in the revision and criticism of the paper, giving different opinions and points of view on the subject.

7. References

- Baluja, S., Pomerlau, D., & Todd, J. (1994). *Towards Automated Artificial Evolution for Computer-Generated Images*. *Connection Science*, 6(2), 325-354.
- Barnsley, M. F. (1993). *Fractals Everywhere* (Second ed.). Academic Press Professional.
- Busch, H., & Silver, B. (1995). *Porque Pintam os Gatos - Uma Teoria da Estética Felina*. Centralivros, Lda.
- Dawkins, R. (1987). *The Blind Watchmaker*. W.W. Norton & Company, Inc.
- Fisher, Y. (Ed.). (1994). *Fractal Image Compression - Theory and Application*. Springer-Verlag.
- Graf, J., & Banzhaf, W. (1995). *Interactive Evolution Of Images*. In J. R. McDonnell, R. G. Reynolds, & D. B. Fogel (Ed.), *Evolutionary Programming IV*, MIT Press.
- Graves, M. (1977). *Test de Apreciacion de Dibujos*, The Psychological Corporation.
- Hurtgen, B., Mols, P., & Simon, S. F. (1994). *Fractal Transform Coding of Color Images*. In SPIE'94.
- Kurzweil, R. (1990). *The Age of the Intelligent Machines*.
- Minsky, M. (1986). *The Society of Mind*. Simon & Schuster.
- Moles, A. (1990). *Arte e Computador*. Afrontamento.
- Morris, D. (1994). *The Human Animal*. BBC Books.
- Pickover, C. A. (1991). *Computers and the Imagination - Visual Adventures Beyond the Edge*. New York: ST. Martin Press, Inc.
- Quintás, L. A. (1987). *Estética de la Creatividad*. Promociones e Publicaciones Universitarias, S. A.
- Schattschneider, D. (1990). *Escher, M. C. - Visions of symmetry*. New York: W. H. Freeman and Company.
- Sims, K. (1991). *Artificial Evolution for Computer Graphics*. In SIGGRAPH'91, in *Computer Graphics*, 25 (pp. 319-328). Las Vegas: ACM.
- Sims, K. (1994). *Evolving 3d Morphology and Behavior by Competition*. In R. Brooks & P. Maes (Ed.), *Artificial Life IV*, (pp. 28-39). MIT Press.
- Sims, K. (1994). *Evolving Virtual Creatures*. In *SigGraph '94*, *Computer Graphics*, (pp. 15-22).
- Spector, L., & Alpern, A. (1994). *Criticism, Culture, and the Automatic Generation of Artworks*. In AAAI-94. MIT Press.
- Todd, S., & Latham, W. (1992). *Evolutionary Art and Computers*. Academic Press.
- Vignaux, G. (1991). *Les Sciences Cognitives - Une Introduction*. Paris: Éditions La Découverte.
- World, L. (1996). *Aesthetic Selection: The Evolutionary Art of Steven Rooke*. *IEEE Computer Graphics and Applications*(16(1), January 1996).