MODEL PROPOSAL FOR A CONSTRUCTED ARTIST

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.

Keywords: Computer Art, Aesthetics, Genetic Algorithms, Neural Networks.

1. INTRODUCTION

The idea of using computers for generating artworks is present from the earliest dates of computer history. Their use as tools had a great impact in fields such as music, advertising and graphics - changing the way art is made. Some of this applications try to reproduce "natural" tools for artists while others provide completely different ones (e.g. virtual environments) enabling the appearance of new art forms.

In spite of their cleverness, sophistication and utility this applications are merely tools for the artist. The "dream" of creating a machine capable of generating art also has a long history. An illusionist from the 19th century created an automata named Zelda that was capable of drawing simple images. Zelda's rudimentary mechanism of dented wheels was quite an achievement for that time, and was able to fool people into believing that Zelda could really draw any image they asked.

Since then computers and AI have been applied to several fields of the arts, including poetry, music and image generation. The success of these applications has been bigger in music than in visual art. This can be explained by the higher quantity of information required by image handling and by the fact that music theory is more developed and quantitative than theory in visual arts [16].

The majority of AI applications to the arts falls into two categories: (1) Systems performing some sort of *art understanding* task, such as musical analysis, and systems that work as "intelligent" tools for human artists [9]; (2) 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." [9].

In this paper we talk about the development of computer programs capable of creating artworks. We focus in the field of visual arts. 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 judgment, we give a short biological explanation to the devotion of humans to art and to how natural evolution favored 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

We will start by describing two approaches that have gained a vast acceptance. The first approach is rule based while the second relies on the use of genetic algorithms.

Harold Cohen can be considered as the precursor of the rule based approach [16]. Cohen started his career as a painter and adquired a worthy reputation, he had a special interest in "the way in which symbols could evoke significance or meaning to the viewer" [4]. He became involved in computer programming as a hobby and joined its two interests when he started developing Aaron in 1972. Aaron, is probably the most acclaimed *constructed artist*, its paintings have been exhibited in several museums and art galleries. The first versions of Aaron generated monochromatic line drawings that were later colored by Cohen. The latest version of Aaron (1995) is already able to paint its own drawings.

Aaron's rules are based on Cohen's beliefs about his own image making process [4]. These rules concern to two types of knowledge: *Procedural Knowledge* that tells the system how to paint and D*eclarative Knowledge* that tells what to paint. The development of Aaron ranges for over two decades, 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 [16].

The second approach is rooted in a computer program written by R. Dawkins. This program evolves images of virtual creatures (*biomorphs*), using of a genetic algorithm [15].

The original program works in the following way: (1) Generates an initial random population of *biomorphs*, that becomes the active population. (2) The individuals of the active population are evaluated by the user according to some criteria. (3) A new population is generated by the mutation of the *genetic code* of the best individual. (4) This population becomes the active population and the process is repeated from step 2. This "simple" idea served as a base for a large number of applications in several fields including the field of visual arts.

By using as criteria for evaluating the individuals, the aesthetic value of the image that they represent, this method can be used to generate aesthetically pleasing images as was shown by [7][8][17]. Most of this applications are more complex than the original program specially in what concerns to the matting operators that generally include sexual reproduction. The main difference between these applications lie in the coding of the images. Karl Sims, for instance, uses mathematical functions, coded in the form of Lisp S-Expressions, ex:

> "(round (log (+ y (color-grad (round (+ (abs (round (log (+ (y (color-grad (round (+ y (log (invert y) 15.5)) x)3.1 1.86#(0.95 0.7 0.59) 1.35)) 0.19) x)) (log (invert y) 15.5)) x) 3.1 1.9 #(0.95 0.7 0.35) 1.35)) 0.19) x)" [8].

The program generates the images from the S-Expressions; mutation and crossover is also performed at the S-expression level[1].

The power of this method rests on the fact that the user doesn't need do have knowledge about the S-Expressions involved in the generation of the image, he/she only needs to select the most appealing ones [17]. The technique in which the user supplies the fitness function results, and thus guides de evolution process was named *interactive evolution*. This methodology has already proved to be extremely useful, some of its applications are: identification of the faces of criminal suspects by witnesses; generation images, animation, textures, music and 3D-Objects. Some of the potential applications are product design, e.g., cars, and architectural design [7].

[1]As far as we know, there is no evolutionary algorithm application that works directly with the images as bitmaps.

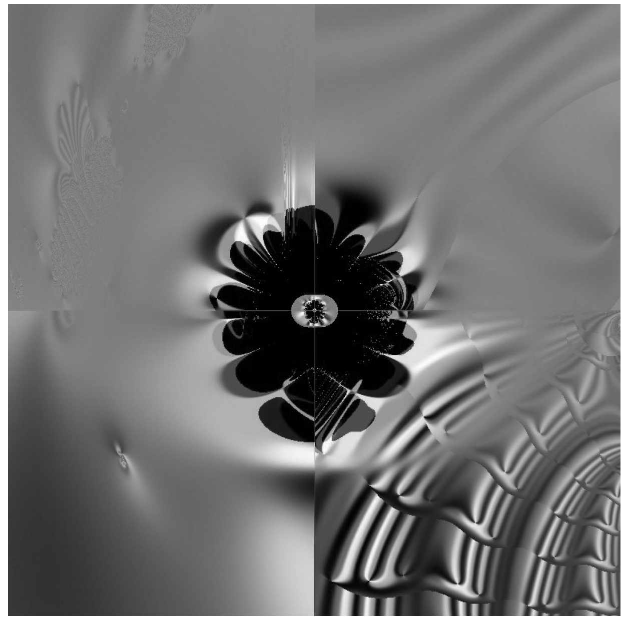

Fig.1 - "Blind Spot" from Interactive Genetic Art II, http://robocop.modmath.cs.cmu.edu:8001/, supplied by John Mount.

One of the major drawbacks of interactive evolution is that evaluating the individuals of a population is a time consuming process. Furthermore, in genetic algorithms applications it is common to have populations with a large number of individuals, population sizes of over an hundred are not uncommon, what only makes matters worst. Due to this problem the majority of *interactive evolution* applications use small population sizes, typically 10 to 20 individuals. According to [8] it is often needed at least 10 to 40 generations to evolve interesting images.

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. In Sims' system we are limited by the representational power of the used formulas, there is no procedure for converting a generic image to that type of representation. In Aaron we are limited by the procedural knowledge that the system has, this type of rules determine the *style* of Aaron paintings, and thus the generated images. We can make Aaron capable of drawing anything, by adding declarative knowledge, but it will always draw in its own style, unless we change the procedural knowledge. In other words, any change in the behavior of Aaron must be preceded by programming intervention. The work involved in changing or adding rules is extremely high. Aaron is incapable of learning, a *constructed artist* should be able to *learn*, like human artists do. We want a system that changes his behavior over time according to its experiences. Karl Sims' system gradually improves its performance, based on the user's evaluation of the generated images. Considering this type of development as learning might be considered as an overstatement, but at least we can change the system's behavior without programming. A human artist doesn't start from zero, like Karl Sims' system does (the initial population is random), humans have access to the artworks made by others; They are 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.

Finally, 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. As a result, 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 [9][11]. 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 [9].

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 filling 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.". The question "What is Art?" is an intriguing one, lets see a definition:

> "To evoke in oneself a feeling one has once experienced and having evoked it in oneself then by means of movements, lines colors, sounds, or forms expressed in words, so to transmit that feeling that another can experience the same feeling-this is the activity of art"[10]

This defines art as a form of communication, further, in the same book, art is defined specifically as an human activity. Another definition could be: the creation or expression of something beautiful, that gives pleasure to the senses or to the mind. Art and Aesthetics are different, yet highly related, subjects. We can say, to a certain point, that Aesthetics is a subset of the Arts, and we can define it as the study of the form in itself, striped from its meaning.

From our point of view, the assessment of an artwork is influenced by two factors:

- The "*content*" of the artwork, that relates to what is represented by the artwork, and which can trigger emotions and feelings. If we consider Art as a form of communication, then "*content*" is what is communicated.
- The "*visual aesthetic value*" of the artwork, that relates to color combination, composition, shapes, etc. We are talking about the form of the artwork, thus, how "*content*" is represented.

By assuming this point of view we aren't creating a false dichotomy between "form" and "content". We are aware of the fact that this factors aren't completely independent, consequently the value of an artwork rests on these factors and their interactions. It is important to notice that, 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 S. Dali's paintings).

If we restrict ourselves to the field of Aesthetics these factors gain independence, because, as stated before, Aesthetics focuses in the "*form*". 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*. Notice that the existence of images deprived of *content* or meaning is controversial, since reality is often in the viewer's mind. We don't mean that *content* isn't important, we just mean it is not indispensable form the Aesthetics point of view.

It is clear that "Every artwork exists within a rich cultural context, and many theorists argued that good art can be neither produced or assessed in ignorance of this context"[9]. It seems clear that the way how *content* influences the 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.

Art's Origin

To support the previous statement we will start by focusing on the origins of art and try to show how Natural Evolution could favor 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.

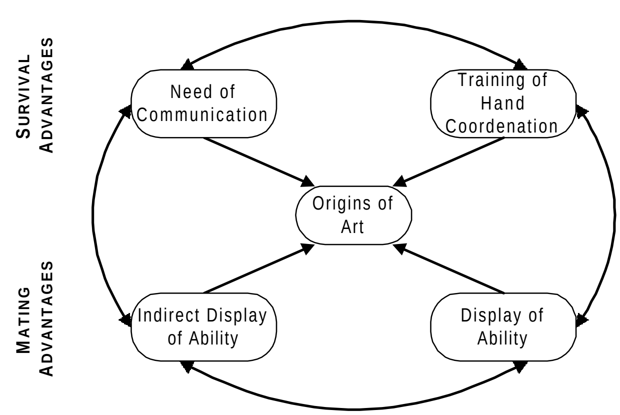

Fig.2- Set of explanations that justify the appearance of art.

The appearance of art may be explained by the necessity of using other forms of communication other than gesture or speech. This explication is usually accepted, 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 training and improving his motor coordination. This two factors would give him an immediate survival advantage

When it comes to natural selection, we also have to consider the matting advantages since that animals tend to choose the most fit individual they can for matting. By making an artwork he is making a *direct display of ability*, furthermore he is showing that he has time to spend on activities that aren't vital for his immediate survival, and thus making an *indirect display of ability*. In our opinion this set of explanations gives reasonable justification for the devotion of man to art, they don't justify, however, why we find certain images beautiful, aesthetically pleasing or artistic. As we said before, we believe that there is a visual aesthetic judgment that is universal, independent from cultural issues and thus hardwired, having biological roots. We aren't saying that it is coded in our genes, we just mean that what we are, and the way our visual image perception system works, makes us favor certain images to others.

The analysis of children's drawings 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, untill they are transformed into arms and legs [2]. Somewhere between the ages of six and twelve, these universal images begin to disappear due to educational influence.

It seems 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 also 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. Morris found six common principles between chimpanzee and human art: Self-Rewarding Activity, Compositional Control, Calligraphic Differentiation, Thematic Variation, Optimum Heterogeneity and Universal Imagery [1].

These leads us to conclude that the assessment of the visual aesthetic value of an image is directly connected to the visual image perception system. Let us consider how this system works.

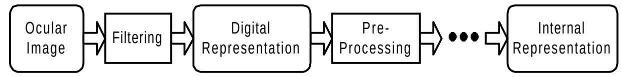

Fig.3 - Steps of visual image perception.

The process of transformation of the ocular image into its digital representation is well known. From this 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 [6].

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. Therefore, 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.

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…"[2] [13].

If we accept this explanation we are lead to conclude that simpler, easier to process, images have higher aesthetic value than complex ones. Or, in other words, low processing complexity implies high aesthetic value. If we follow this idea we come to the conclusion that, a completely blank image has higher aesthetic value 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 work; 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) Processing Complexity (the lower, the better); (2) Image Complexity (the higher, the better). 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 have higher aesthetic value. 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

[2] This is in some way similar to the "principle of economy" found in evolution's theory.

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 [3]. Returning to the fractal example, fractal images are usually complex, the propriety of self-similarity makes these images easier to process, which gives an explanation to why we usually find fractal images beautiful. Another important characteristic of fractal images is that they have several levels of detail, this characteristic can also be found in many artworks, (e.g. Kandinsky's works). This strike's us as very important specially if we notice that the act of "seeing" isn't instantaneous, it spans through a (sometimes long) period of time. 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 makes the image easy to process, and thus less distracting when we don't want to give it attention; simultaneously, when 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 would hinder the generality of the artwork in the sense that our willingness to look at it would largely depend on our state of mind. Thus, it is important to preserve an high *Image Complexity / Processing Complexity* ratio through all the period of seeing.

One final remark goes to the fact that until 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 [11]). 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).

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. A constructed artist encompasses two main modules: Generation and Evaluation. For the generation of images we rely on a Genetic Algorithm (GA). This choice is based on the robustness of this method and its performance in complex and irregular search spaces, furthermore [7][8][17] show that this method can be successfully used for the generation of interesting images. For the evaluation of images we use Neural Networks (NNs). The main reason for choosing NNs over other methods was that NNs can be trained to perform some task by the presentation of examples. This allows us to achieve our goal without having to create rules to evaluate the images[17].

Of outmost importance is the issue of how the images should be represented. The amount of memory needed to store images in bitmap format is extremely high and this type of storage doesn't seem appropriate for the use with genetic algorithms. Using mathematical formulas to represent the images solves this two problems at the expense of the generality. We have the problem of developing a system that is, simultaneously, generic, appropriate for the GA approach and that lessens the amount of information needed to represent the images.

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. What we need, 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 color. 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 [18]. The possibility of representing any image is important, since we want our *constructed artist* to have access to the works of human artists.

Fractal Image Compression

A detailed description on how fractal image compression works, is beyond the scope of this article and can be found in [18], nevertheless we will make a short introduction to the subject since it is essential for the full understanding of our approach. We will leave much of the mathematical details out.

Consider a special type of photocopying machine, that reduces the image to half and reproduces it three times [18]. If we feed back the output this will converge to the image known as the Sierpinski Triangle, the *attractor* of these transformations.

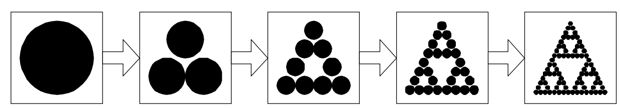

Fig.4 - Successive images generated by the copying machine, converging to the Sierpinski triangle. The initial image (a circle) is reduced to half its size, by repeating this process to infinity the original image is reduced to a "dot".

Since that the *transformations* are *contractive*, when the process is repeated an infinite number of times the initial image becomes reduced to a single point (see Fig.4). Therefore these transformations will always converge to the Sierpinski triangle, independently from the initial image. In a more formal way, defining $W(x) = x_1$ as the above mentioned transformations, x and x_1 as the images fed and generated by the copying machine (respectively), and A as the Sierpinski triangle, we have that:

$$\lim_{n \to \infty} W^{on}(x) = A \text{ }, \forall x.$$

and that:

$$W(A)=A$$

In mathematical terms W is an Iterated Function System (IFS), i.e., a collection of contractive transformations $\{w_i: \Re^2 \to \Re^2 \mid i=0,...,n\}$ which map the plane $\Re^2$ to itself [18]

How can this be used for image compression? Well, the Sierpinski triangle is a complex image yet it can be represented by a simple IFS consisting in three transformations, so all we have to do is to store the transformations. We can generate the image by applying the transformations to a random initial image, since as we stated before, this process will always converge to the *attractor*. Now consider that we want to compress an arbitrary image f, then we must find the set of transformations W that has that image as an attractor, $W(f)=f$. We have just described the basic idea behind fractal image compression. We will now generalize this idea so we can use it to compress gray-scale images.

The global self-similarity found in the Sierpinski triangle isn't usually present. To cope with this we will use Partitioned Iterated Function Systems (PIFS) instead of IFS. In an IFS each transformation is applied to the whole image, e.g. for the Sierpinski triangle we have three transformations, that scale the whole initial image to half and copy it to a specific location. In a PIFS each transformation is applied to a part of the image, i.e. instead of copying the whole image it copies a part of the image.

In the IFSs that we have discussed each transformation consisted in: scaling, rotation, and position of the copy. To enable the coding of gray-scale images, we must add, a contrast and brightness adjustment. Thus, each transformation of our PIFS will consist on the following factors: part of the original image to be copied, position of the copy, scaling, rotation, contrast adjustment and brightness adjustment.

Let us see how a copying machine based on this scheme would work. For each transformation w_i belonging to the PIFS W, a portion of the original image D_i, is copied to a part R_i of the generated copy. We call the D_i*s domains* and R_i*s ranges*. As it is copied the D_i suffers a brightness, contrast and rotation (*BCR*) transformation. Given an image f, a single copying step with this special copying machine can be written as $W(f)=w_1(f) \cup w_2(f) \cup \ldots \cup w_n(f)$, n=Nº of transformations of the PIFS. To encode an image f we must find a PIFS W such that, $W(f)=f$, generally we can't find an exact match and we have to settle for $W(f) \cong f$.

Fig.5 - The original Lenna image (top). And the same image compressed 16 (left) and 32 (right) times using the Fisher's Quadtree-Partition method.

There are several ways to partition the image, a description and comparison of these methods can be found in [18] here we describe the Quadtree Partitioning algorithm presented in [18]. Assuming that the image, f, size is 256*256, we start by creating a domain pool D consisting in all the squares of the image of size 8, 16, 32 and 64. Then we partition the image in non-overlapping squares of size 32, this will be our R_i*s*. For each R_i we try to find a D_i and corresponding *BCR* transformation such thus $w_i(f) \cong R_i$, we only consider D_i*s* that are twice the size of the R_i. If the error is less than a predetermined value e, we mark the R_i and store the transformation. If not, we subdivide the square in four and repeat. The algorithm ends when all the R_i*s* are marked. This algorithm works surprisingly well, as can be proved by the above images.

Using this method for encoding image involves a high computational cost; fortunately this process can be easily be implemented by a parallel algorithm. The decoding process is relatively fast.

Generation

As we said before, we the generation of images is made through a GA. This algorithm is close to the one presented in [8]. The differences are that we use fractal image encoding to represent images, and added a Knowledge Base of Images.

In its simplest form the algorithm works in the following way: 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.

The choice of fractal image encoding for the representation of images allows us to integrate background knowledge. This can be achieved in two different ways:

1. The initial population does not need, to be random: we can use any set of images, including famous artworks, as initial population.
2. We can maintain a knowledge base of images, which are not subject of selection. The images in this knowledge base can be selected for matting with images of the current population or between themselves. The knowledge base can be initialized with any set of images we choose. Additionally, images of high aesthetic value generated by the system can be added to the knowledge base.

The use of a knowledge base of images has additional benefits. It is a reasonable way of "increasing" population diversity, if population diversity is low we can use images from the knowledge base as matting partners. It also prevents the early lost of remarkable images. Imagine that a image, f, of high aesthetic value is generated in one of the first populations, this image is selected for matting and a new population is created, but since the other matting partners had low aesthetic value the resulting images are probably worst f. Within a few generations the f can be "irreparably" lost. The use of a knowledge base of images prevents this from happening.

In most genetic algorithm application the individuals are represented by bit strings. In this case a mutation can be defined as a alteration of the value of a randomly chosen bits. Crossover can be defined in a simple way: swap a randomly chosen sub-string between the two parents, the resulting individuals are the offspring.

Our individuals are represented by quad-trees, although this quad-trees can easily be represented in a bit string format (it is the format used for file storage of the images); Using this format in conjunction with the "standard" mutation and crossover operators isn't appropriate. Our *muttator* operator can be defined as follows: 1) randomly choose one of the nodes of the tree; 2) If the chosen node is an internal node, i.e. a square that is further subdivided, it becomes a leaf, and the values of the *BCR* transformations and D_i are randomly set. If the chosen node is a leaf we randomly choose one of its fields (i.e. one of the *BCR* transformations or D_i) for mutation, and mutate it. The *crossover* operator consists in randomly selecting a sub-tree from each of the parents, and swapping this sub-trees. If the sub-trees are located at different levels, there is a difference in the size of the D_i*s*, this can result in D_i*s* that go beyond the edges of the image; When this happens the x,y position of the D_i*s* is changed to the nearest admissible position.

Evaluation

As we said before, aesthetic evaluation is probably the most difficult problem to tackle. Automating the evaluation process serves two purposes. First, we consider that a *constructed artist* must be able to perform aesthetic evaluation. Second, and from a more pragmatic point of view, the automation of the evaluation process enables greater population sizes, improving the performance of the system.

To automate the evaluation process we chose to use NNs, this choice is also a pragmatic one. We can train a NN to perform some task by giving examples to it. If we chose a rule based approach we would have to construct rules to perform the task. The task of aesthetic evaluation appears to be extremely complex, furthermore, there is no strong theory in this domain, as result of this the rule based approach seems inadequate and would certainly be extremely time costly and complex. We have considered two other approaches: Case Based Reasoning (CBR) and Genetic Programming (GP). CBR has been used, with success, to create *constructed artists* in the field of music [5][9], this systems, however, don't perform aesthetic evaluation. It doesn't seem clear to us, how such a system could be created using this approach, at least without the expense of generality. GP looks like a good choice, yielding the same advantages (possibly more) of NNs, unfortunately, GP doesn't appear to be able to cope with problems as complex as aesthetic evaluation (at least presently).

To train a NN we must provide it a set of examples that is representative of the whole population. In our case we have to construct a set that covers the population of all possible images of a given size [17]. To deal with this problem we make several independent runs of our system, with an user making the evaluation of the image. From the set of images generated by this process, we randomly chose a subset for training purposes [17]. This choice is made taking in account that we must have enough examples of images of all "different" aesthetic values. Generally, most of the images are uninteresting, if we chose completely at random we would end up many uninteresting images and only a few interesting ones.

It is important to notice that, while in Sims' system the user only needed to evaluate the images in comparison with the others of the same population, in our case the user must provide a global value for each image.

An interesting question is what shall serve as input for the NN. We have two possibilities, either we use the quad-tree encoding of the image to generate the image and use the bitmap representation as input, or, we use the quad-tree representation as input, directly. This second approach seams more promising by several reasons. Considering that we are working with 256*256 gray-scale images, the amount of memory necessary to store this images in bitmap format is 64K. We can reduce the image to 64*64 and use the resulting images as input for the NN, although the error involved in this operation is usually negligible, this still means that we must feed the neural network with 4K, if we use the quad-tree representation, directly, 2K are enough[3]. Furthermore, by using the quad-tree representation as input we are also giving some information about the structure of the image, i.e. about the self-similarities present in the image.

For the validation of the results of the NN we propose three methods: Using the images generated by the system that weren't used for training, compare the values supplied by the neural network to those supplied by the user [17]; Use a "random" set of images, taken from magazines, books, etc., and compare the assessment made by the NN to the assessment made by a human; Use psychological tests that were developed to assess the aesthetic evaluation ability of humans, e.g. [12], and see how the NN scores in comparison with humans. These three methods have different difficulty levels and allow us to determine the quality and generality of the results supplied by the neural network.

The Whole Picture

Let's review how our system works: we start with a working set of images built by the knowledge base of images and the current population. The images in this set are already in its quad-tree representation and evaluated.

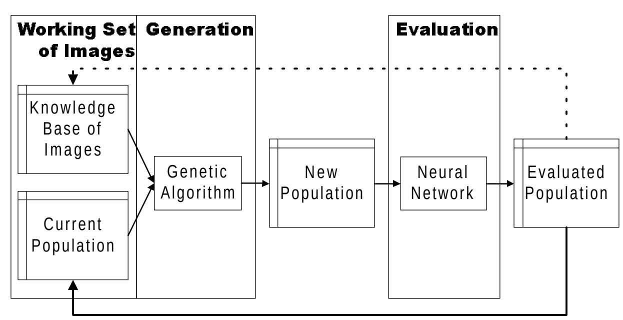

Fig 6- The proposed model.

[3] Considering that the artworks integrated in the knowledge base were compressed 32 times.

From the working set, the genetic algorithm creates a new population of images through mutation and crossover. Images of high aesthetic value have more chances of being selected for matting. The images of this new population are, independently, evaluated by a neural network, resulting in a evaluated population. This population becomes the current population and, additionally, images that: have a aesthetic value that is higher than *c*, or superior to the average of the population by a value *d*, are added to the knowledge base (*c* and *d* are empirically specified constants).

The use of the knowledge base can be turned on/off, and we can change the probabilities of using images from the knowledge base for matting.

5. CONCLUSIONS AND FURTHER WORK

In this paper we specified a set of features that the current *constructed artists* lack and justified the importance of this set of features. It is our conviction that the proposed model provides a feasible way of integrating these features. We also presented a brief theory of aesthetics. In this theory, we consider the biological roots of Art and Aesthetics. We stated that the assessment of an artwork depends on two factors: *content* and *form*, furthermore we claim that *aesthetic value* is directly connected to the image processing task of the brain.

Our system is still under development, no results are available yet. There are three main modules: fractal image encoding image generation and aesthetic evaluation. The generation and encoding modules are nearly operational. 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.

We use genetic algorithms for image generation, 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 could develop a matting operator. This operator would select sets of "compatible" images and crossover would performed between this images.

It seams possible to attack the problem of developing a *constructed artist* in a radical different way. We stated that *aesthetic value* is connected to the image processing tasks, furthermore we claimed that images of high aesthetic value are the ones that have a high *image complexity/processing complexity ratio*. Therefore if we develop an "complete" image processing system, and assuming that our claims are correct, we can use the mentioned ratio as a measure of the aesthetic value. Testing this idea would be interesting on to accounts: It might provide a way of proving our claims; It can result in the development of a system capable of making aesthetic judgments that aren't based on those made by humans. This would probably be the first *non-anthropocentric constructed artist*, resulting in the creation of truly alien artworks.

Little as been said about music, we think that much of what we stated about visual arts could also be applied to music. We are currently involved in the development of a system for recognizing musical styles and the emotions generated by musical pieces.

One of the conjectures, we would like to test, is the sharing of aesthetic values between different fields such as visual arts and music. Thus, can a neural network trained to make *visual aesthetic judgments* make, aesthetic judgments in the musical field?

6. ACKNOWLEDGMENTS

We would like to thank Prof. Pedro Pita for sharing his views on Aesthetics. His ideas have been of key importance in the development of the presented theory of aesthetics. We would also like to thank Francisco Colunas and Paulo Gomes for helping in the revision and development of the paper. Finally we would like to thank APPIA (*Portuguese Association For Artificial Intelligence*) and ISEC (*Coimbra's Superior Institute of Engineering*) for financial support.

7. REFERENCES

[1] D. Morris, *The Biology of Art*.
[2] D. Morris, *The Human Animal*, BBC Books, 1994.
[3] D. Schattschneider, *Escher, M. C. - Visions of symmetry*, W. H. Freeman and Company, 1990.
[4] E. Burton, "Representing Representation: AI and Drawing", *Computers & Art*, Intellect Books 1997, pp.55-75.
[5] F. Pereira, L. Macedo, C. Grilo, A. Cardoso, "Towards a Computational Case-Based Model for Creative Planning", *First European Workshop on Cognitive Modeling*, 1996.
[6] G. Vignaux, *Les Sciences Cognitives - Une Introduction*, Éditions La Décoverte, 1992.
[7] J. Graf, W. Banzhaf, "Interactive Evolution Of Images", *Evolutionary Programming IV*, MIT Press, 1995.
[8] K. Sims, "Artificial Evolution for Computer Graphics", SIGGRAPH'91, *Computer Graphics*, Vol. 25, No. 4, pp. 319-328, ACM.
[9] L. Spector, A. Alpern, "Criticism, Culture, and the Automatic Generation of Artwork"*s, AAAI-94*, MIT Press.
[10]L. Tolstoy, *What is Art*,.
[11]L.A. Quintás, *Estética de la Creatividad*. Promociones e Publicaciones Universitarias S. A.,1987.
[12]M. Graves, *Test de Apreciacion de Dibujos*, The Psychological Corporation, 1977.
[13]M. Minsky, *The Society of Mind*. Simon & Schuster, 1986.
[14]M.F. Barnsley, *Fractals Everywhere (Second ed.)*, Academic Press Professional, 1993.
[15]R. Dawkins, *The Blind Watchmaker*, W.W. Norton & Company Inc., 1987.
[16]R. Kurzweil, *The Age of the Intelligent Machines*, 1990.
[17]S. Baluja, D. Pomerlau, J. Todd, "Towards Automated Artificial Evolution for Computer-Generated Images". *Connection Science*, 6(2), 1994, pp. 325-354.
[18]Y. Fisher (Ed.), *Fractal Image Compression - Theory and Application*, Springer-Verlag, 1994.