SIMULATING ARTIST AND CRITIC DYNAMICS
An Agent-based Application of an Evolutionary Art System

Gary Greenfield
Department of Mathematics and Computer Science, University of Richmond, Richmond, Virginia, U.S.A.

Penousal Machado
Department of Informatics Engineering, University of Coimbra, 3030 Coimbra, Portugal

Keywords: Evolutionary art, Quantifying aesthetics, Agent-based simulation.

Abstract: We describe an agent based artist-critic simulation. Artist agents use a swarm based evolutionary art system to evolve images that try to match their preferences. Preferred images are submitted to critic agents who then decide, accordingly to their own criteria, which images should be displayed in a public gallery. The purpose of our model is to enable the implementation of a variety of behavioral policies which result in different dynamics. A reward system determines the impact of each critic and the success of each artist, which in turn leads to behavioral and preference changes. The experimental results indicate the emergence of novel styles and trends, artist-critic cooperation, and niche exploitation.

1 INTRODUCTION

We describe our implementation of an artists and critics simulation. The role of artists is to produce art. The role of critics is to decide what should be exhibited in a public gallery. Artists are software agents that have recourse to a generative art system to evolve images and decide whether or not to show them to critics. Critics are software agents that critique artwork with reference to examples of “masters” that they are partial to and previous images they have been asked to critique.

We wish to make clear at the outset that we expect every reader to disagree with our assumptions about how a [public] art gallery actually functions. This misses the point of our work. Our objective is to establish a framework for implementing a model supporting constraints and assumptions the reader might wish to impose, and to do so in such a way that the dynamics resulting from the simulation can be archived and analyzed.

2 BACKGROUND

The impetus for this investigation is three-fold. First, we reconsider the approach of (Saunders and Gero, 2002) which made use of agents sharing a common evolutionary art system in order to study artificial creativity via the dynamics of artists and critics. Second, we invoke a grayscale swarm based evolutionary art system (Aupetit et al., 2003). Third, we quantify aesthetics by using image comparison techniques that rely upon extracting information theoretic measurements from the images under consideration (see, for example, (Machado and Cardoso, 1998) (Machado et al., 2007)). Some of our previous work in evolutionary art involving co-evolving agent-critic systems (e.g. (Greenfield, 2007) (Machado et al., 2007)) provides further inspiration for this research.

Saunders and Gero. (Saunders and Gero, 2002) offer a conceptual framework for studying artificial creativity supported by an implementation where a community of software agents, each possessing a Sims’ style image generating system (Sims, 1991), attempts to generate a “cultural repository” of shared aesthetic imagery. Their agents make use of neural nets and self organizing maps. The alternative framework we present is informed in part by the following observations we made about their work: i) an agent is simultaneously an artist and a critic, ii) agents share genomes, iii) feature extraction and image evaluation use a neural net to evaluate a downsampled \(32 \times 32 = 1024\) bit filtered and thresholded image, iv) the policy for an image to be added to the cultural
repository is that one agent decides an image is interesting and better than its own work.

Swarm based Generative Art. Swarm art refers to a generative technique which uses large numbers of cooperative drawing agents to produce images in a style that is often reminiscent of Jackson Pollack. The term is perhaps attributable to Moura (Moura and Ramos, 2002) who, arguably, along with Jacob (Jacob et al., 2007), has done the most to popularize the art form. We are motivated by the formulation of swarm paintings as ant colony paintings in the manner of (Aupetit et al., 2003) (Greenfield, 2005b) (Greenfield, 2006) and (Urbano, 2005). Since we are interested in difficult questions pertaining to aesthetic merit and artistic style, to simplify matters, we work solely in grayscale. For a discussion of color considerations see (Greenfield, 2009).

Image Evaluation and Aesthetics. Methods for automated image evaluation on the basis of aesthetics are eclectic and varied; the terminology itself is problematic (Greenfield, 2005c). In the domain of evolutionary art systems, the use of neural nets traces back to (Baluja et al., 1994) and is also prominent in (Machado and Cardoso, 1998) (Machado et al., 2007). Greenfield uses geometric measurements induced by image color organization (Greenfield, 2002). Noteworthy evaluation schemes based on theories of aesthetics include (Greenfield, 2005a) (Burns, 2006), (Ross et al., 2006), (Jacobsen, 2006), and (Schmidhuber, 2007). Most of these schemes make it difficult to account for factors such as culture, trends, or history which are commonly recognized as integral to aesthetics.

3 OVERVIEW OF OUR SYSTEM

To set the stage for the details to follow, it is first necessary to provide a brief overview of the agent based artist and critic simulation discussed in this paper. There are $N_a$ artist agents and $N_c$ critic agents prioritized by organizing them into a pyramid with $N_t$ tiers. The larger the number of artists and critics, the more difficult it becomes to analyze simulation results, therefore, in this paper we chose as a reasonable compromise $N_a = 4$ artists and $N_c = 7$ critics organized into $N_t = 3$ tiers where the top tier consists of one critic, the second tier consists of two critics, and the third tier consists of four critics. At all times, each artist agent has one target image from a database of images painted by “masters” while each critic agent has two images from this database. All artist agents make use of the same swarm based evolutionary art system. Artist agents access this system to evolve small populations of art works from which to select images to submit to the critics. Simulation proceeds by rounds. At the start of each round every artist evolves its image population for a fixed number of generations and then is invited to submit one image to the critics. Image populations are evolved and evaluated by artist agents according to differing criteria. If an artist agent decides to submit an image, that image is evaluated by $N_c$ critics selected in such a way that one critic from each tier is represented. If at least one critic accepts the image, then that image replaces the oldest image in a public gallery. Artist agents are provided with the evaluation results of each image that they submit to the critics. Critics receive impact points for accepting images. At the conclusion of each round, both artist agents and critic agents are given the opportunity to modify their strategies and behaviors based on the results, and the critic tier structure is updated on the basis of impact points. The policies that govern the various behaviors of artists and critics are presented as they arise during the more comprehensive description of our system that follows.

4 THE IMAGES

In this section we discuss how we manage images in our artist and critic simulation. All our images are $256 \times 256$ pixel grayscale images. A distinguishing feature of our system is that it uses a database of 25 images consisting of five images selected from each of five “masters” — Dali, Mondrian, Monet, Picasso, and van Gogh. This database was culled from online archives and subsequently resized and converted to grayscale. Examples of four representative images from this database are shown in Figure 1. For system testing we also make use of a database of 13 grayscale “noise” images. All other images in our system are generated from genomes that are fed to a swarm-based generative algorithm. Genomes and pixel maps of images generated from genomes are the property of artist agents. They are never shared. Feature vectors extracted from those images are shared.

4.1 Feature Vectors

An image is viewed as organized into six regions — the entire image, the upper left quadrant, the upper right quadrant, the lower right quadrant, the lower left quadrant, and the central “quadrant” whose area is one-fourth the size of the entire image. For each region, along with the mean and standard deviation of
the luminance values of its pixels, its complexity is estimated. This yields $6 \times 3 = 18$ components which make up the feature vector.

### 4.2 Complexity Measure of a Region

Our approach to measuring the complexity of a region is founded upon two assumptions: i) region complexity is an aesthetically relevant characteristic, and ii) perceived complexity can be estimated using fractal image compression. To estimate the perceived complexity of a region, we use quad-tree fractal image compression (Fisher, 1995). Our rationale is that complex images are harder to compress, resulting in larger files than simple images, therefore we assume that the compression ratio is negatively correlated with image complexity.

The fractal image compression scheme we use for each region is lossy so there will be compression error, i.e., the compressed image won’t exactly match the original. All other factors being equal, complex images will tend toward higher compression errors and simple images will tend toward lower compression errors. Thus, the compression error is positively correlated with image complexity.

We estimate image complexity of region $i$ from image $I$ using the following formula (Machado and Cardoso, 1998):

$$\text{Complexity}(i) = \text{RMSE}(i, FC(i)) \times \frac{s(FC(i))}{s(i)}$$

where $\text{RMSE}$ stands for the root mean square error, $FC$ is the fractal compression transformation, and $s$ is the file size function.

Fractal image compression can provide a compact encoding to images with high apparent visual complexity, in fact this characteristic of is the basis for the aesthetic judgement scheme described in (Machado and Cardoso, 1998). Although the estimate used has its shortcomings, previous work (Machado and Cardoso, 2002) indicates that images with similar complexity estimates tend to have similar visual complexity. In the experiments described here we use the set of parameters for our compression scheme given in Table 1. Note that letting the minimum partition level be 3 implies that the selected region is always first partitioned into 64 blocks. Subsequently, at each step, for each block, if one finds a transformation that gives a good enough pixel by pixel match, then that transformation is stored and the image block isn’t further partitioned. (Here, pixel by pixel match is with respect to the usual 0 to 255 grayscale interval encoding.) If the pixel by pixel match error is more than 8 for at least one of the pixels of the block in the partition, that image block is further partitioned into 4 sub-blocks, the level increases, and the process is repeated. Since the maximum partition level is 6, when that level is reached the best transformation found is stored even if the pixel by pixel match error for the block exceeds 8.

### 4.3 Comparing Two Images

Given an image $I$ with feature vector $X$, and a region $i$ of $I$, we let $c_X(i)$ be the complexity component and $m_X(i)$ be the mean luminance component from $X$ of region $i$. We define an image comparison metric $C(A, B)$ for comparing two feature vectors $A$ and $B$ by adding a complexity term together with a mean luminance term as follows. Our complexity term is a weighted, relative absolute difference of region complexity terms obtained by summing over regions:

$$\Sigma_i 100\frac{|c_A(i) - c_B(i)|}{\max(c_A(i), c_B(i))}$$.  

For our luminance term, we take into account both the ordering of the mean luminances of regions within the two vectors and the mean luminances between corresponding regions of the two vectors. More precisely,
we let \( r_A(i) \) denote the rank of the mean luminance of region \( i \) within its feature vector \( X \), and obtain by summing over regions:

\[
\sum_i 10|r_A(i) - r_B(i)| + \sum_i |m_A(i) - m_B(i)| / 5. \tag{3}
\]

Note that we have a perfect match between \( A \) and \( B \) when the comparison value \( C(A, B) \) is zero.

5 THE GENERATIVE SYSTEM

Our generative system is a swarm system modeled after a colony of ants. It most closely follows (Greenfield, 2006). Each virtual ant, or organism, is tracked by maintaining a location and a compass heading for it. At each time step, each organism senses the three cells immediately in front of it in order to ascertain their color luminances and their concentrations of each of two organism produced pheromones. Each organism then deposits color in the cell it currently occupies and advances between 1 and 4 cells in a direction determined by one of these three sensed cells.

5.1 Swarm Genomes

Each swarm painting is generated from a swarm genome. A swarm genome has four global attributes: a pseudorandom number generator seed value, the number of organisms, the time to execute the painting (viz. number of execution cycles), and the organism footprint size. Since organism behavior is stochastic, to make swarm paintings reproducible each painting must be associated with an integer seed to initialize the pseudorandom number generator. The most significant attribute is footprint size. Consistent with an organism’s current heading, when it is depositing color, footprint size determines the stroke width.

For each swarm painting, all its organisms are assumed to be of the same \textit{species}. Thus a swarm genome must also include the following species attributes: the number of color scents to recognize or deposit (here six), the various scent detection thresholds, the probability of following a color scent if one is detected, the probability of following (or avoiding) a virtual scent if one is detected, and the probabilities for which direction (forward, left, or right) to follow when either no scent is detected or scent is being ignored. In support of color scents, a swarm genome also includes a color scheme for the painting; here a palette consisting of six shades of gray.

Which species attributes will actually come in to play is determined by the \textit{caste} an individual organism belongs to. There are three castes representing three different behaviors. The \textit{explorer} caste tries to find unvisited cells. The \textit{color} caste is sensitive to detection of either the color it tries to deposit or the color it is trying to find i.e., if the deposited color is sensed then the ant tries to avoid it, but if the color it seeks is sensed then (almost always) the ant follows it. Similarly, the \textit{pheromone} caste is sensitive to the presence of virtual scent: one of the scents is an attracting scent, the other is a repelling scent.

Finally, a swarm genome also includes attributes for each of the individual organisms. Organisms are differentiated by caste, color to deposit, color to follow, and number of cells traversed at each time step. Crossover and mutation operators needed to breed swarm painting genomes are consistent with those described in (Greenfield, 2006).

5.2 Generating a Swarm Painting

The number of organisms per swarm painting is restricted to be between 20 and 80, in multiples of five. Because we wish to evolve swarm paintings, it is necessary to be consistent in the initial placement of the organisms on the canvas. Rather than initially placing all of the organisms at pseudorandom locations on the grid, we cluster them into five equally sized groups. One group is placed at the center of our canvas, while the other four groups are placed at the centers of each of the four quadrants. Further, we use identical initial heading orientations for organisms within clusters.

5.3 Behavior Statistics

When the sense-decide-color-move cycle is performed for each organism according to the number of times determined by the time limit parameter in the swarm genome, five statistics \( B_1, \ldots, B_5 \) are collected for future use. Respectively, they are the number of instances where an organism: i) visited a previously unvisited cell, ii) was in “wandering” mode, iii) pursued (or avoided) color, iv) pursued (or avoided) scent, v) blended the color it deposited. Regarding this last statistic, there is some duplication of effort. If color is being deposited in a previously unvisited cell, then it replaces the background color (black), otherwise it blends with the existing color.

5.4 System Capabilities

From our cursory description of the generative system, in light of master images such as those in Figure 1 that the artists and critics are assigned, it should be clear that the generative system will never be able to approximate such images closely! This is our intention. We treat images in the database done by “mas-
ters” either as if they were created by a process that has been lost or forgotten, or as if they were retrieved from an alien extraterrestrial site. Either way, they have now become the inspiration for artists and critics alike.

6 THE ARTISTS

Each of our four artists has a population of swarm genomes. Genome population size is 10, with 6 genomes culled after each generation and the remaining 4 used to repopulate via standard evolutionary art methods. In all the experiments run here, when an artist agent invokes the generative system, its genome population is evolved for 6 generations according to the criteria it provides before examining the results. At all times, each artist has a target image from the database of masters, and a binary coefficient vector \((c_1, \ldots, c_5)\) for use with the behavior statistics \(B_1, \ldots, B_5\) that are gathered during the image generation procedure.

**Image Fitness Calculation.** When an artist calculates image fitness it compares its current target image feature vector \(T\) with the feature vector \(X\) extracted from the image under consideration \(I\) and obtains fitness value \(F\) by letting

\[
F = C(T, X) + \sum_j c_j w_j B_j,
\]

where the \(w_j\)'s are scaling weights that are constant for the simulation.

**Submitting an Image to the Critics.** Our policy is: When queried, an artist submits an image to the critics by submitting its feature vector if and only if its generative system has produced a more fit image than the last time it was queried. In response, the simulation provides the ranks (between 0 and 9) that were assigned by the three critics — the tier 1 critic, a tier 2 critic, and a tier 3 critic — that were selected to critique it, and the decision about whether the image was accepted for the gallery or not.

**Changing an Artist’s Target.** The policy we use to change an artist’s target is: If the number of consecutive submissions rejected reaches the threshold value \(ATAR\) then it is time to change one’s target!

**Changing an Artist’s Behavior.** Because of the way artist’s calculate image fitness, artists can choose not to be slaves to the pursuit of trying to match their target image’s characteristics. This means they can pursue their own unique style, or further develop successful styles, or change the evolutionary pressure currently being exerted by a target that is not producing gallery acceptances. Artist behavior is changed by flipping one or more of the bits in the binary coefficient vector \((c_1, \ldots, c_5)\). This is implemented by making the probability of each bit being flipped 0.2. The policy we use to change an artist’s behavior is: Within the last \(ALAG\) rounds, if either no submission has been made, or every submission has been rejected, then it is time to change one’s behavior!

7 THE CRITICS

As previously stated, we use seven critics. At all times, each critic has two target images from the database of masters. In addition it has a rank-ordered list of its preferences, or favorites, resulting from the top ten evaluations it has given to all the images it has critiqued so far.

**Critiquing an Image.** This favorites list makes the actual act of critiquing an image simple. With target feature vectors \(T_1\) and \(T_2\), given the feature vector \(X\) from an image \(I\) to critique, our policy is: Take the smaller of \(C(T_1, X)\) and \(C(T_2, X)\), insert \(I\) in the favorites list by replacing the value of the highest ranked image in the list of favorites with that value, sort the favorites list in increasing order, and returns the rank that \(I\) receives as a result.

**Managing the Preferences List.** During initialization we instantiate a “random artist” and generate an initial population of images. We use that set of feature vectors to initialize all the critic preference lists. Because critics have different targets, the way they rank these images is different. Thus critic diversity is present right from the start. In our model, critic preferences are not fixed. Their tastes change over time. The way we chose to implement this notion was to age each preference list by artificially inflating the ranking of the oldest critiqued image after every 3 rounds. This causes the next image the critic is asked to critique to force this image to be dropped from the list.

**Changing a Critic’s Target.** If a change in target is triggered, one of the two target images the critic currently has is replaced by an image chosen at random from the masters database. Note that as a side effect, some of the images in the preferences list must be re-evaluated — in particular, those that were present because their comparison values were based on an im-
age that is now gone — and the preferences list must be re-ranked. The policy we use for triggering a critic target change is: If a critic was unsuccessful in finding an image to recommend for the gallery in the last CTAR rounds it is time to change one of its targets!

8 THE SIMULATION

The overarching simulator for our artists and critics is straightforward. After initialization, for each of $N$ rounds the artists are queried and invited to make submissions. If a submission is made, the simulator selects and asks three critics to evaluate it. If one or more decide to accept it to the gallery, then the simulator handles all the record keeping as well as enforcing the various policies that apply to artists and critics that we have previously described.

Managing the Gallery. The gallery of eight images is initialized by selecting image from the masters database. As artist generated swarm images are accepted the oldest image in the gallery is replaced.

Selecting the Critics. Each critic has an impact rating. This rating increases when images are accepted to the gallery by the critic. At the start of each round all critics are sorted on the basis of this rating to determine the tier 1 critic, the two tier 2 critics, and the remaining four tier 3 critics. When a submission is received, one critic from each tier is selected and asked to evaluate the image. Our policy is: An image is considered to be accepted by the tier 1 critic if it places in the top three of its preferences list, by the tier 2 critic if it places in the top two of its preferences list; and by the tier 3 critic if it places at the top of its preferences list. For breaking acceptance ties our policy is: In case of a tie the lower ranking critic gets credit for having accepted the image. This last policy has profound implications with respect to the dynamics of impact ratings.

Assessing Artists and Critics. For artists, we keep a running tally of the number of submissions and the number of acceptances. For critics, in addition to keeping a running tally of the number of critiques and the number of acceptances, as previously mentioned, we maintain an impact rating. We add a further twist to the calculation of the impact rating by rewarding critics for introducing novelty into the gallery as follows. The critic’s impact rating increases by 1 if it accepts a submission to the gallery, but by 2 if, in addition, the submission is sufficiently different from the images currently in the gallery. Here “different” is determined by considering the region that consists of the entire image and checking to see if either its complexity or average luminance is more than one standard deviation away from either of those quantities averaged across the entire gallery.

9 SIMULATION RESULTS

To examine the dynamics resulting from our model, we made several runs of our simulation lasting $N = 50$ submission-evaluation rounds using as default settings $ATAR = 5$, $ALAG = 3$, and $CTAR = 10$. Figure 2 shows image diversity from one of these runs. Figure 3 shows the style development of an artist that had 8 of 29 submissions accepted during this same run. The graphs in Figure 4 show the rank ordering of artists on the basis of number of acceptances, and of critics on the basis of impact rating, as well as the cumulative totals of these quantities by rounds during the course of the run. Note that by round #25 critic rankings have stabilized. This is consistent with most of the other experiments we performed. Under our policies it is difficult for lower ranked critics to displace higher ranked ones. Consequently, as a run progresses there is tendency for critic ranking to stabilize and tier changes to become less frequent. Although tier changes are uncommon, they still do sometimes occur, and in such cases may even trigger dramatic changes in both critic and artist rankings as well as in the type of imagery accepted for the gallery.

10 FURTHER EXPERIMENTS

Our artist-critic simulation has several parameters and many features. To further examine its capabilities and help ascertain its limitations we performed several tests. These tests included runs with:

- $N = 30$, $ATAR = 10$, and $ALAG = 1, 3, 5$, and 7;
- $N = 30$, and $ATAR = 5, 10$, and 15;
- $N = 30$, and $CTAR = 5, 10$, and 15;
- $N = 30$, where one master image is selected and assigned to be the same target for all artists;
- $N = 30$, where one master is first selected and then each critic is assigned two (not necessarily distinct) target images from the five that are available in the master’s database by that master;
- $N = 20$, where one artist functions as a “random artist” by submitting its highest ranked image.
Figure 2: Artist diversity as evidenced by an accepted submissions of each of the artists during rounds #7 through #10 of an $N = 50$ round run.

Figure 3: Artist style development is observed by considering an artist’s accepted images during the course of a run. For the virtual artist chosen here, its acceptances are shown (clockwise from top left) for rounds #3, #19, #23 and #46 of an $N = 50$ round run. Its accepted image from round #7 is shown at the top left in the previous figure.

from a random population each round (Reassuringly, no image from this artist was ever accepted by the critics beyond round 5);

• $N = 20$, where the master’s database is replaced by “noise” images, the number of critics is reduced to 3, one master image is chosen, and both targets for all critics are assigned to be this image (This tests how successful the generative system is at “solving” the image duplication problem).

Due to space limitations we are unable to discuss all of these experiments. We focus on two runs where a single master was first identified; the two assigned targets for each critic were selected from that master; and critic targets did not change during the course of an $N = 30$ round run. Because there were only five targets available for critics, most critics shared at least one target with another critic. However, if a critic had an unshared target this created a niche opportunity such that one artist could develop an exploitative relationship supporting a sustained period of submissions and acceptances that catapulted both artist and critic to higher rankings. These runs produced the greatest turmoil within the tier system we used for critics (see Figure 5). Interestingly, for both these runs the critic that came to dominate had only one target image i.e., its two targets were the same. Evidently, for artists, the existence of a critic with an unambiguous preference was useful.

11 CONCLUSIONS

We have described a flexible artist-critic agent based simulation. It is based on a model whereby different policies can be implemented in order to test the consequences of what one might expect to occur. The policies we chose supported phenomena such as the development of artistic styles; artist-critic cooperation;
critic diversity (because critics have different preferences and these preferences may change over time); and artistic freedom (because artists can exhibit “freedom of expression” by deciding not to blindly follow the critics). Also noteworthy is that our model establishes a critic hierarchy such that over time some critics become more equal than others. A distinguishing feature of our model is that it values the discovery of novel imagery by requiring that in addition to being new, images must also be “fit”. More importantly, it appears our policies allow for artistic trends to arise via a strategy whereby a lower ranked critic locates a niche to exploit and teams with an artist to furnish images so that the resulting flurry of acceptances cause images in this artist’s style to populate the public gallery.

REFERENCES


