Modelling human preference in evolutionary art

Anikó Ekárt[1], Divya Sharma[2], and Stayko Chalakov[1]

[1] Aston University
Aston Triangle, Birmingham B4 7ET, UK
Email: a.ekart@aston.ac.uk, stayko16@yahoo.com
[2] IIT Ropar, India
Email: divyas@iitrpr.ac.in

Abstract. Creative activities including arts are characteristic to humankind. Our understanding of creativity is limited, yet there is substantial research trying to mimic human creativity in artificial systems and in particular to produce systems that automatically evolve art appreciated by humans. We propose here to model human visual preference by a set of aesthetic measures identified through observation of human selection of images and then use these for automatic evolution of aesthetic images.

Keywords: aesthetic measure, human preference modelling, genetic programming, interactive vs automatic evolution

1 Introduction

Ever since the invention of the first computing device, humanity has been thinking about using them to perform creative activities. Producing aesthetically pleasing pieces of art is certainly one such creative activity. Beginning with the pioneering work of Dawkins [7] and Sims [19], over the past twenty years a lot of effort was spent on generating increasingly more effective evolutionary art systems that produce aesthetic artworks. Successful examples attracting substantial public attention include the Electric Sheep [10], the NEvAr system [14] and the Painting Fool [3].

The majority of evolutionary art systems are either interactive (for example [21]) or automatic (for example [1, 9]). Interactive systems tend to generate more aesthetic artworks, as their driving force is human selection, but at the same time need a lot of effort on the part of the human, may incur user fatigue and could be inconsistent over time. Automatic systems have the advantage of a built-in automatic fitness evaluation, so the human effort is reduced; however, the aesthetics of the resulting artworks may suffer as the automatic evaluation has not been perfected yet. To overcome the disadvantages and also combine the advantages of both approaches, Machado et al. propose partially interactive evolution [15], where the human user's contribution is much reduced compared to the fully interactive approach, but the human still guides the evolution.

Substantial efforts in evolutionary art research have been dedicated to studying and devising good aesthetic measures [8, 13, 17, 18]. It is generally agreed that formulating a universally valid and acceptable aesthetic criterion is not within our reach. Achieving automatic evolution that produces aesthetic images to the liking of the human user very strongly depends on the understanding of the particular user's aesthetic values. A recent study by Li and Hu [12] suggests using machine learning to learn the differences between aesthetic and non-aesthetic images, as indicated by image complexity and image order. Colton [4] produces new rules for forming fitness functions through the use of an inference engine. Greenfield proposes the technique of evolutionary refinement [11] to encourage aesthetic pattern formation through stages and concludes that "evolution in stages with radical changes in fitness criteria may be a profitable evolutionary exploration strategy".

Our contribution complements these previous approaches by considering four established aesthetic measures in interactive evolutionary art to model human preference. We monitored how these measures evolved over the generations when different users interacted with a simple evolutionary art system and fully drove the selection process. We found that a combination of aesthetic measures models user preference suitably well. We consequently employed this combination (MC and BZ) to automatically evolve further images starting from the result of interactive evolution.

2 Aesthetic measures

We study the evolution of four well-known aesthetic measures in an attempt to model human selection in interactive evolutionary art. Measure R is based on Ralph's work, measure MC is based on Machado and Cardoso's work, measure BZ on Birkhoff and Zurek's work and finally measure S on Shannon entropy.

2.1 Aesthetic measure R

This aesthetic measure is based on the mathematical model proposed by Ralph[18]. After analyzing hundreds of examples of fine art, it was found that many works consistently exhibit functions over colour gradients that conform to a normal or bell curve distribution.

The colour gradient for each pixel is computed as:

$$|\nabla r_{i,j}|^2 = \frac{(r_{i,j} - r_{i+1,j+1})^2 + (r_{i+1,j} - r_{i,j+1})^2}{d^2}$$

where $r_{i,j}$ is the value of the Red component for pixel (i, j) and d is a scaling factor which is taken to be 0.1% of the diagonal length, as suggested by Ralph's model [18] (leading to the value $d^2 = 3.277*10^{-2}$). $\nabla g_{i,j}$ and $\nabla b_{i,j}$ are computed similarly for the Green and the Blue colour components.

The overall gradient, or the stimulus, of each pixel is calculated as follows:

$$S_{i,j} = \sqrt{|\nabla r_{i,j}|^2 + |\nabla g_{i,j}|^2 + |\nabla b_{i,j}|^2}.$$

Next, the viewer's response to each pixel is computed as

$$R_{i,j} = \log(S_{i,j}/S_0).$$

The range of values for $R_{i,j}$ is $[0, 1)$. R can never become negative. The minimum value of 0 corresponds to the case when there is no change in colour at a pixel at all; if there is no stimulus the response is 0. $S_{i,j}$ can never be less than S_0 due to the scaling factor. S_0 is the detection threshold taken to be 2, as suggested by Ralph's model of aesthetics. If $S_{i,j} = 0$ (no change in colour at a pixel), it is ignored.

The mean μ and standard deviation σ of the response values are calculated using the response values themselves as weights because the probability that a viewer pays attention to a detail of an image is considered proportional to the magnitude of the stimulus that resides at that detail. A histogram is built next to judge how close the distribution of response values is to the bell curve distribution. The "bins" will each represent an interval of size $\sigma/100$. Then the probability that a response value falls in a given bin is computed. This is repeated for all the bins by going through all the $R_{i,j}$ values where each $R_{i,j}$ updates its corresponding bin using a weight of $R_{i,j}$.

Then, the deviation (D) from the normal distribution is computed as follows:

$$D = \sum_i p_i \log\left(\frac{p_i}{q_i}\right)$$

where p_i is the observed probability in the i^{th} bin of the histogram and q_i is the expected probability assuming a normal distribution around μ with standard deviation σ. When $q_i = 0$, that bin is ignored. The value $e^{-|D|}$ will reported as the value of the aesthetic measure. With a value between 0 and 1, a low value will indicate a large deviation and hence a poor image, whereas a large value will correspond to a good image.

We justify this aesthetic measure as follows:

1. Aesthetic measure R discourages images which give rise to very high or very low response values. If a viewer gives very little response to something, it is too insignificant to be of interest. On the other hand, if a viewer gives a very large response to something, it is too disturbing or chaotic.
2. The response value increases as the gradient increases and decreases as the gradient falls. Very low gradients give rise to single coloured monotonous areas (which do not interest a viewer) whereas very large gradients give rise to sharp lines and boundaries separating areas with huge colour differences (which is undesirable). Aesthetic measure R discourages very high and very low gradients and encourages reasonable values of gradients.

2.2 Aesthetic measure MC

This measure is based on the aesthetic theory of Machado and Cardoso [13] asserting that the aesthetic value of an artwork is directly connected to Image

Complexity (IC) and inversely connected to Processing Complexity (PC). So, the value of the aesthetic measure is calculated as the ratio

$$\frac{IC}{PC}. \tag{1}$$

In order to compute IC, we first compress the image losslessly using JPEG compression and calculate the ratio (I) of the size of compressed image to the size of uncompressed image. We hypothesize that the IC is directly connected to the ratio I. The inherent unpredictability, or randomness can be measured by the extent to which it is possible to compress the data [6]. Low values of I indicate substantially compressible and low complexity image. High values of I indicate not very compressible and therefore more complex image. That is,

$$\text{more compressible} \equiv \text{less random} \equiv \text{more predictable} \equiv \text{less complex}$$

Hence, the less the value of ratio I (the less the size of the compressed file) is, the more compressible and hence, the less complex the image is. We substitute the ratio I for IC in Equation 1.

PC should reflect the complexity of the coding of the image. We encode each image by three expression trees, one for each of the R, G and B components, as detailed in Section 3. In order to compute PC, we compress the expression trees represented as strings in prefix notation and again find the ratio P of size after compression to size before compression. We argue that PC can be substituted by the ratio P. The aesthetic measure MC will be computed as

$$\frac{I}{P}.$$

In theory the value of this aesthetic measure could range from zero to infinity, where infinity corresponds to an image that cannot be compressed, but whose genetic expression tree can be compressed to the minimum. Zero corresponds to an image that can be compressed significantly, but with an expression that cannot be compressed. It is notable that the compression rate PC of the mathematical expressions could be replaced with the more exact rate computed for the minimum length of an equivalent mathematical expression. However, as arithmetic simplification is not applied on the mathematical expressions in our system, we consider that using the actual evolved expression is appropriate.

2.3 Aesthetic measure BZ

This aesthetic measure is based on Birkhoff's measure [2] and Zurek's physical entropy [17]. We compute the value of Shannon's entropy as mentioned in [17] by creating a histogram of luminance values of pixels and computing Shannon's entropy H_p as follows:

$$H_p = -\sum_i p_i \log p_i$$

where p_i is the probability in the i^{th} bin of the histogram. The luminance value (L) for a pixel (i, j) is computed as follows:

$$L = (0.2126 * r_{i,j}) + (0.7152 * g_{i,j}) + (0.0722 * b_{i,j}).$$

Next, the Kolmogorov Complexity (K) [17] of the expression trees of the image is estimated by compressing the strings corresponding to expression trees and finding the length of the compressed string. The value of this aesthetic measure is given by

$$\frac{H_p}{K}$$

This aesthetic measure discourages very high and very low luminance values because it favours high values of H_p. Very high and very low luminance values lead to low values of H_p. Here, K is used as a measure of PC.

2.4 Aesthetic measure S

As stated in [17], to analyse an image's composition, the used measures must quantify the degree of correlation or similarity between image parts. We compute this degree of correlation by dividing the image into four equal squares and compute Shannon's entropy (H_{p_i}, $i = 1, \ldots, 4$) for each of these parts. We then compute the weighted mean of these values (the weight being the area of the part). Finally, we find the ratio of the weighted mean to the Shannon's entropy value of the image as a whole to obtain the value of the aesthetic measure. The value of the aesthetic measure is given by

$$\frac{H_{p_1} + H_{p_2} + H_{p_3} + H_{p_4}}{4H_p}.$$

3 The underlying evolutionary art system

A simple interactive evolutionary art system is used, where the user is presented with nine images and has to select two as parents for the next generation. Images are represented by triplets of functions corresponding to the three components R, G and B. For each pixel of the image, the values of these functions are calculated and produce the colour of the pixel as shown in Fig. 1. Genetic programming is employed for the evolution of the expression trees. Simple subtree crossover, subtree and point mutation are the allowed genetic operators. The user can set the operators to be used and their rates and can also introduce new random images at anytime during the interactive evolution. The system was implemented in Java and uses all mathematical functions provided in Java.Math. The terminals are Cartesian and polar coordinates of the points and also random constants.[3] Examples of images produced by the authors are shown in Fig. 2.

[3] The interactive system is available for download at http://www.evoartmedia.com.

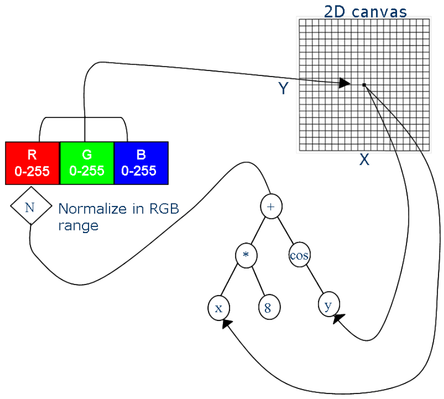


Fig. 1. The genetic representation used

4 Simple numerical analysis

To see whether there are any particular functions that are preferred more than others, we initially analysed their frequency in nice images. The number of occurrences of each function in 44 nice images manually selected from a set of images generated through interactive evolution by the authors in sessions of length varying between 15 and 30 minutes is shown in Fig. 3. The images were selected such that on visual inspection they looked substantially different. It can be seen that the preferred functions are SEC, CUBRT, EXP, LOG, SQRT +, -, MAX, AVG, *, as each of these occurs on average at least once in every aesthetic image. At 63.8% of all variables, polar coordinates were the preferred variables and variables were preferred over numeric constants, as 73.3% of terminals. The constant range $[0.5, 0.6)$ had the highest frequency of 20% of all constants, the next closest being $[0.6, 0.7)$ at 12.5%. Such a simple analysis does not really offer detailed understanding of human aesthetic selection, just indicates the more likely ingredients of aesthetic images within the given class of images.

(a) Spider (b) Flower (c) Fish

Fig. 2. Example images generated by interactive evolution.

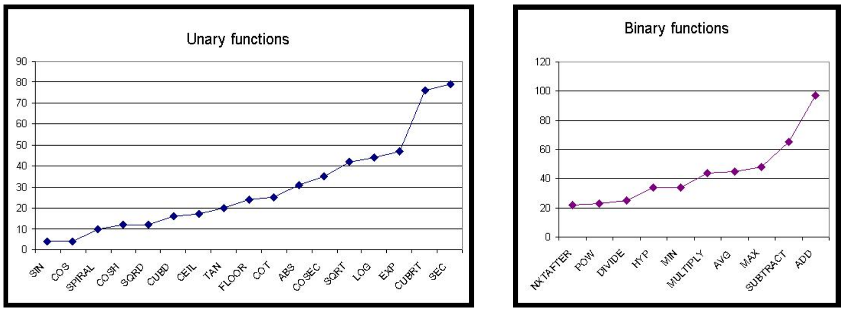
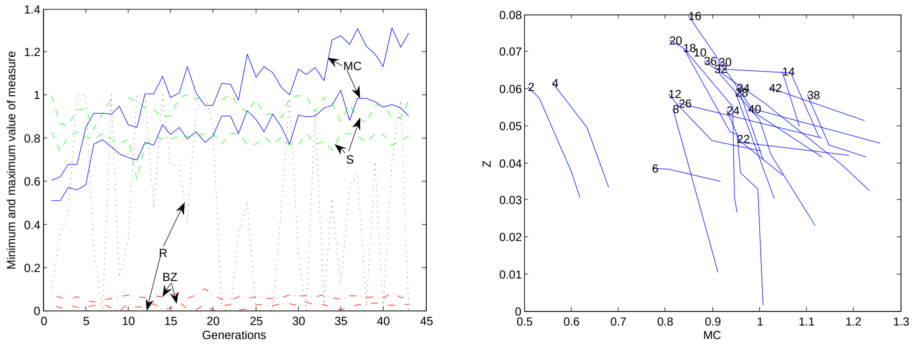


Fig. 3. Occurrences of various functions in 44 different aesthetic images obtained in different runs of the system.

5 Aesthetic measures to model human selection

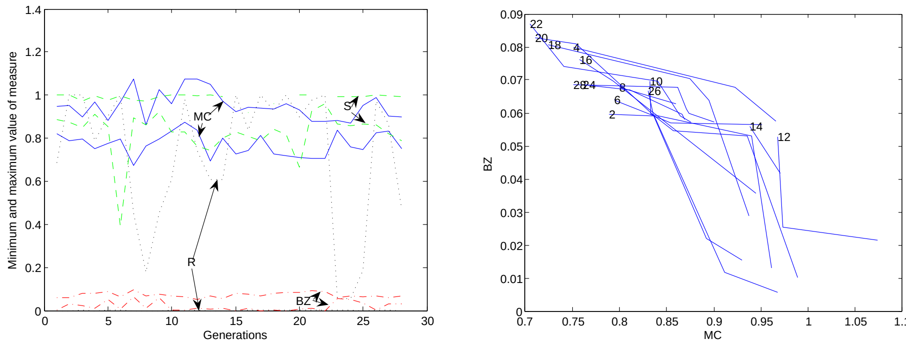
Individual interactive evolution experiments were analysed to better understand their driving force: human selection. Then automatic evolution was applied to the resulting image, using the selection criteria revealed by the analysis. The analysis involved monitoring the evolution of the four aesthetic measures described in Section 2 during interactive experiments performed by four different people. We found that although there are similarities for all users, the occupation of the user substantially influences their use of the system and their image selection preference.[4] There is no universally applicable simple model. Computer scientists tend to use the system for longer than graphic designers. The value of measure MC shows a clear growth over generations in the case of computer scientists, as in Fig. 4(a), both for the minimum and the maximum value taken in each generation. At the same time, there is no clear tendency in the evolution

[4] The users were computer scientists and graphic designers.



(a) Evolution of measure values (b) Pareto front (MC,BZ)

Fig. 4. Evolution of image through interactive evolution by computer scientist.



(a) Evolution of measure values

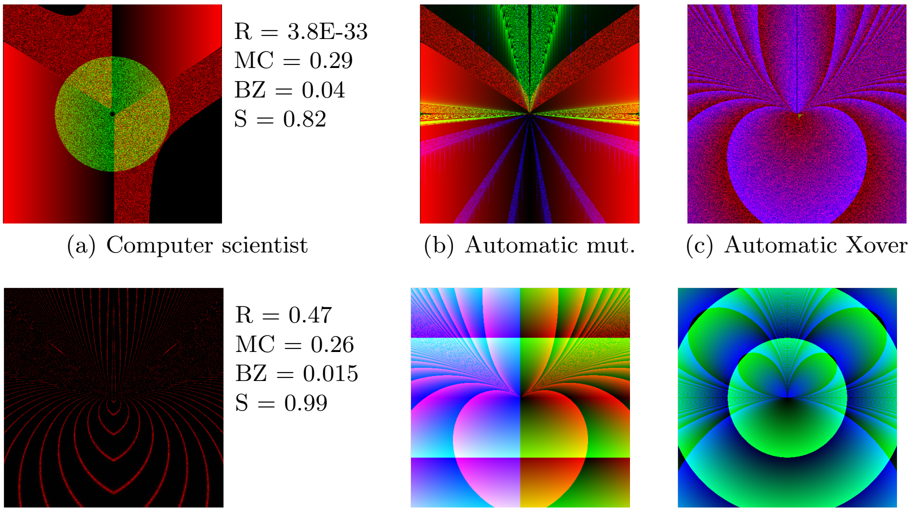
(b) Pareto front (MC,BZ)

Fig. 5. Evolution of image through interactive evolution by graphic designer.

of values for measure MC in the case of graphic designers (see Fig. 5(a)). The evolution of measures BZ and S are similar for both types of users: variation within similar ranges is observed, approximately $[0, 0.05]$ for BZ and $[0.8, 1]$ for S, respectively. The R measure has a lot of variation across its full range for computer scientists and somewhat less variation for graphic designers, but follows no clear pattern. Interestingly, if we consider the evolution of two measures together and draw the Pareto front of non-dominated solutions[5] in each generation, we notice some trends. Figures 4(b) and 5(b) show the evolution of the Pareto front for measures MC and BZ. In both shown examples, with a few exceptions over the full experiment, the front is moving toward better (i.e. higher values of aesthetic measures) non-dominated solutions. We interpret this as an indication that human users may not select the images during evolution in a way that consistently follows a single aesthetic measure, but more likely a set of aesthetic measures. In fact if we compare the two images in Fig. 6(a) and 6(d), we notice that the first image scores better over measures MC and BZ, while the second image scores better over measures R and S.

When attempting partial automatic evolution [15, 20] we propose that the human's previous selections are analysed and modelled by the best fitting set of measures and then the automatic evolution subsequently uses these measures. It is then more likely that images to the particular human user's liking are produced by automatic evolution. We therefore applied automatic evolution with the combination of the MC and BZ fitness measures to create images starting from the preferred images of the human users. We experimented with mutation only or both crossover and mutation, various settings of population sizes (15-40) and generation numbers (30-100) allowing the computer to spend different amounts of time on creating new images. Evolved images are shown in Figures 6(b) 6(c) and 6(e), 6(f), respectively.

[5] A point (x_1, y_1) is part of the Pareto front of a set if there is no other point (x_2, y_2) in the same set such that $x_2 > x_1$ and $y_2 > y_1$.



(a) Computer scientist (b) Automatic mut. (c) Automatic Xover

(d) Graphic designer (e) Automatic mut. (f) Automatic Xover

Fig. 6. Images created by computer scientist and graphic designer. Subsequent images evolved from these by automatic evolution using mutation or crossover.

6 Conclusion

We proposed modelling human user preference by a set of aesthetic measures monitored through observation of human selection in an interactive evolutionary art system. Although our evolutionary art system is very simple and is only capable of generating images within a limited set, it provides a suitable environment for studying human aesthetic judgment. The same principles could be applied using an extended set of aesthetic measures on more sophisticated evolutionary art systems and then different combinations of aesthetic measures may be found to model individual users best.

McCormack [16] criticises aesthetic selection itself and proposes an open problem "to devise formalized fitness functions that are capable of measuring human aesthetic properties of phenotypes". The key is to *model and measure* human aesthetic properties by the available means.

We argue that once a combination of measures that models increasing human preference during interactive evolution is identified, automatic evolution is provided with a suitable fitness evaluation method. We are planning to conduct more experiments using an approach similar to that of Colton et al. [5]. Also we are planning to employ machine learning techniques to find potentially more accurate functions driving human aesthetic judgment and to subsequently apply these functions for evaluation and selection in automatic evolution.

Acknowledgments

The authors acknowledge the support of Aston University's summer internship scheme which made it possible for the second author to take part in the reported work. The authors are grateful to Philip Welch, Ken Murphy, Adrian Carolan and Daniel Baxendale for taking part in the user experiments.

References

1. S. Baluja, D. Pomerleau, and T. Jochem. Towards Automated Artificial Evolution for Computer-generated Images. *Connection Science*, 6:325–354, 1994.
2. G. D. Birkhoff. *Aesthetic Measure*. Harvard Univ. Press, 1933.
3. S. Colton. The Painting Fool. www.thepaintingfool.com, 2007.
4. S. Colton. Automatic Invention of Fitness Functions with Applications to Scene Generation. In *EvoApplications 2008*, LNCS 4974, pages 381–391, 2008.
5. S. Colton, J. Gow, P. Torres, and P. Cairns. Experiments in Objet Trouvé Browsing. In *Conference on Computational Creativity*, pages 238–247, 2010.
6. S. Dasgupta, C. Papadimitriou, and U. Vazirani. *Algorithms*. Tata McGraw-Hill Publishing Company Limited, 2008.
7. R. Dawkins. *The Blind Watchmaker*. 1986.
8. E. den Heier and A. E. Eiben. Comparing Aesthetic Measures for Evolutionary Art. In *EvoApplications 2010, Part II*, LNCS 6025, pages 311–320, 2010.
9. S. DiPaola and L. Gabora. Incorporating Characteristics of Human Creativity into an Evolutionary Art System. *Genetic Programming and Evolvable Machines*, 2(1), 2009.
10. S. Draves. Electric Sheep. electricsheep.org, 1999.
11. G. Greenfield. Generative Art and Evolutionary Refinement. In *EvoApplications 2010, Part II*, LNCS 6025, pages 291–300, 2010.
12. Y. Li and C.-J. Hu. Aesthetic Learning in an Interctive Evolutionary Art System. In *EvoApplications 2010, Part II*, LNCS 6025, pages 301–310, 2010.
13. P. Machado and A. Cardoso. Computing Aesthetics. In *SBIA'98*, LNCS 1515, pages 219–228, 1998.
14. P. Machado and A. Cardoso. All the truth about NEvAr. *Applied Intelligence, Special issue on Creative Systems*, 16(2):101–119, 2002.
15. P. Machado, J. Romero, A. Cardoso, and A. Santos. Partially interactive evolutionary artists. *New Generation computing*, 23:143–155, 2005.
16. J. McCormack. Open Problems in Evolutionary Music and Art. In *EvoWorkshops 2005*, LNCS 3449, pages 428–436, 2005.
17. J. Rigau, M. Feixas, and M. Sbert. Informational Aesthetics Measures. *IEEE Computer Graphics and Applications*, pages 24–34, 2008.
18. B. J. Ross, W. Ralph, and H. Zong. Evolutionary Image Synthesis Using a Model of Aesthetics. In *IEEE Congress on Evolutionary Computation*, pages 1087–1094, 2006.
19. K. Sims. Artificial Evolution for Computer Graphics. In *SIGGRAPH 1991*, volume 25, pages 318–328, 1991.
20. S. Todd and W. Latham. *Evolutionary Art and Computers*. Academic Press, 1992.
21. J. J. Ventrella. Evolving the mandelbrot set to imitate figurative art. In P. F. Hingston, C. L. Barone, and Z. Michalewicz, editors, *Design by Evolution*, pages 145–167. 2008.