Evolving Aesthetic Imagery using Multiobjective Optimization

Gary R. Greenfield
Mathematics & Computer Science
University of Richmond

December, 2003

Outline.

I. Introduction.

A. Computational Aesthetics — *algorithmic* aesthetic evaluation of (computer generated) images.

- Simple Generative Systems.
  e.g. Grid Colorings (see Staudek).
- Special Generative Systems.
  e.g. "Aaron" (see Cohen, Mohr, Knowlton).
- Mathematical Generative Systems.
  e.g. Fractal images (see Sprott)
- Evolutionary Generative Systems.
  e.g. Evolving Expressions (see Rooke, Greenfield).

B. Evolutionary systems using *user-guided* aesthetics.

Dawkins (1989). *Biomorphs* — Images from drawing programs.

Sims (1991). Images from expression trees.

Latham (1992). *Mutator* — Images from parameter lists.

Lund (1995). *Artificial Painter* — Images from neural nets.

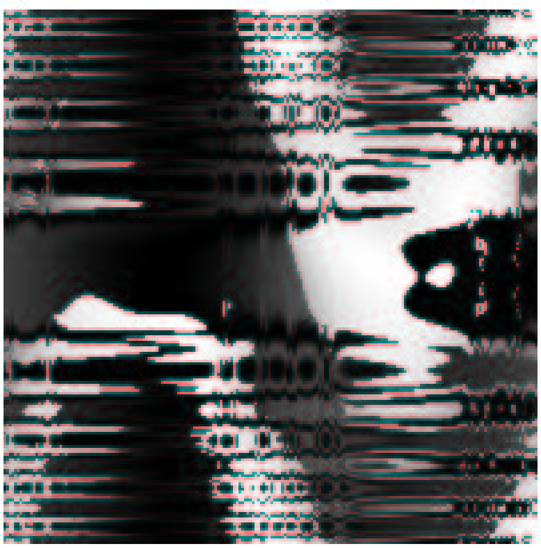
C. Evolutionary systems using *computational* aesthetics.

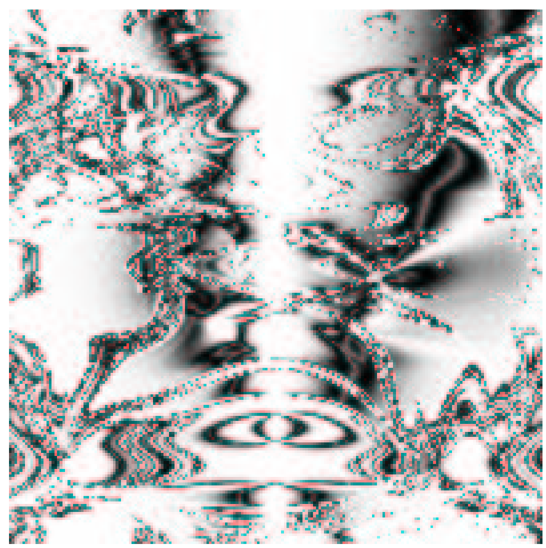
Baluja et al (1994). Neural net critic for expression tree images.

Rooke (1995). Expression tree critics for expression tree images.

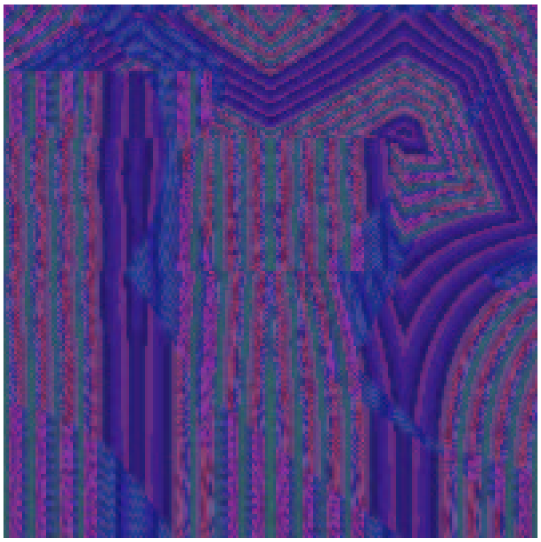
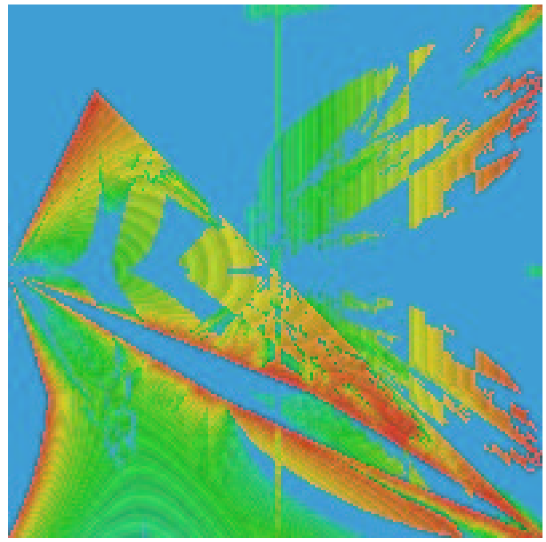
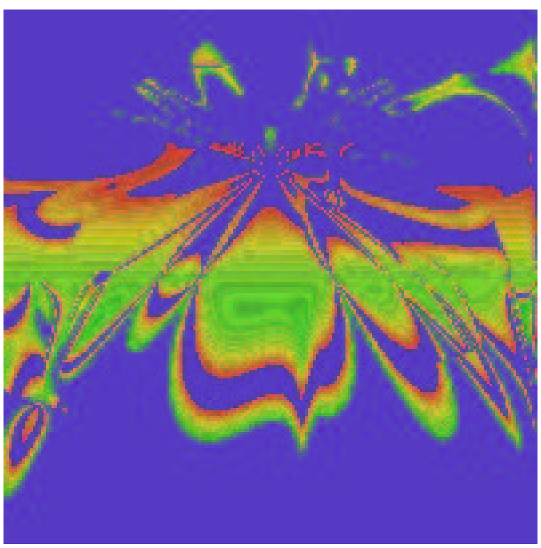
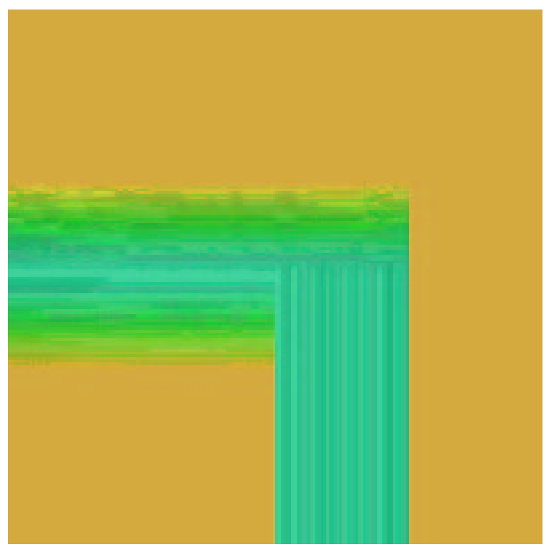
Machado & Cardoso (2000). Neural net critic (?) for neural net images.

Greenfield (2000). Digital filter critics for expression tree images.





Greenfield (2002). Single objective optimization based on geometric analysis of color segmentations for expression tree images.

II. Image Generation.

A. Consider prefix representation of a function

$$F : I \times I \longrightarrow I.$$

e.g., $F(V_0, V_1) = \max(\mathrm{mul}(0.758, V_0), \mathrm{sqrt}(V_1))$

B. Construct postfix tree using symbolic equivalents.
e.g., $F(V_0, V_1) = V_1\ U_2\ V_0\ C_{758}\ B_0\ B_6$

C. Form pixel image, $I = (p_{i,j})$, $0 \leq i, j < N$.

D. Define pixel density, $d_{i,j} = F(i/N, j/N)$.

E. Assign pixel color from HSV look up table:

$$\begin{array}{lcl}
d_{i,j} \in [0/450, 1/450) & \longrightarrow & (0.00, 0.70, 0.70) \\
d_{i,j} \in [1/450, 2/450) & \longrightarrow & (0.00, 0.70, 0.80) \\
d_{i,j} \in [2/450, 3/450) & \longrightarrow & (0.00, 0.70, 0.90) \\
d_{i,j} \in [3/450, 4/450) & \longrightarrow & (0.00, 0.80, 0.70) \\
 & . & \\
 & . & \\
 & . & \\
d_{i,j} \in [8/450, 9/450) & \longrightarrow & (0.00, 0.90, 0.90) \\
d_{i,j} \in [9/450, 10/450) & \longrightarrow & (0.12, 0.70, 0.70) \\
 & . & \\
 & . & \\
 & . &
\end{array}$$

III. Color Segmentation.

A. Consider 32×32 pixel "thumbnail" image.

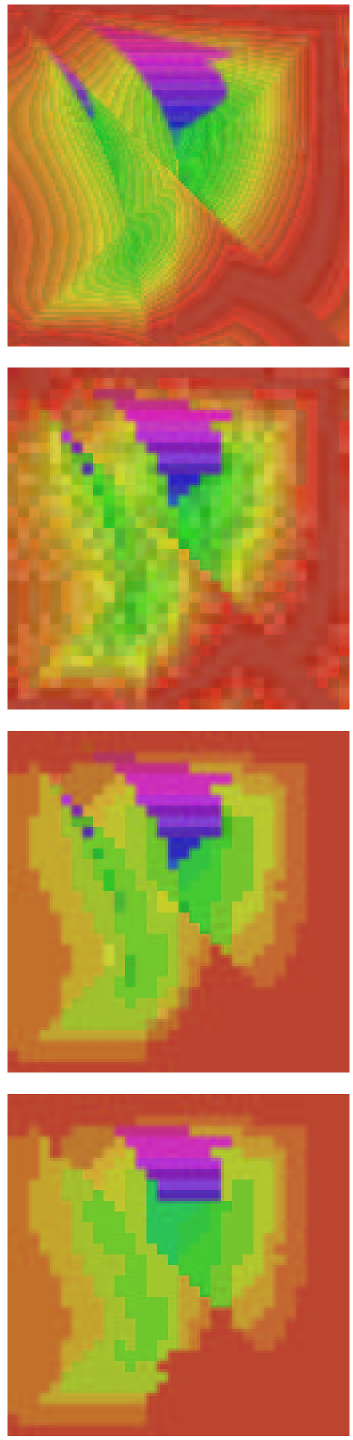
B. Aggregate pixels using region-merging algorithm requiring a segmentation vector

$$(k_{vv}, k_{vs}, k_{ss}, k_v, k_s, k_h).$$

C. Merges are triggered by selecting the edge of minimal priority. The priority $p(e)$ of edge e depends on HSV color differentials as follows:

$$p(e) = (k_h + k_{h,v}\Delta_v + k_{h,s}\Delta_s + k_{h,s,v}\Delta_s\Delta_v)\Delta_h + k_s\Delta_s + k_v\Delta_v.$$

D. There is an image "clean-up" phase.



IV. Aesthetic Metrics.

A. Sort merged regions by area and define aesthetic primitives:

$$\begin{aligned} a(X_i) &= \text{ area of region } X_i \\ b(X_i) &= \text{ boundary length of region } X_i \\ j(X_i) &= \text{ adjacency tally of region } X_i \end{aligned}$$

B. For $1 \leq s < t \leq n$, define aesthetic components:

$$\begin{aligned} A_{s,t} &= \sum_{k=s}^{t} (k+1)a(X_k), \\ B_{s,t} &= \sum_{k=s}^{t} b(X_k), \\ J_{s,t} &= \sum_{k=s}^{t} j(X_k). \end{aligned}$$

V. Multiobjective Framework.

A. Essentially the NGSA-II algorithm.

B. Half the population is replaced after each generation by crossing *clones* of tournament selection winners.

C. Fitness is integer valued so ties in rank are to be expected and thus crowding distances are important.

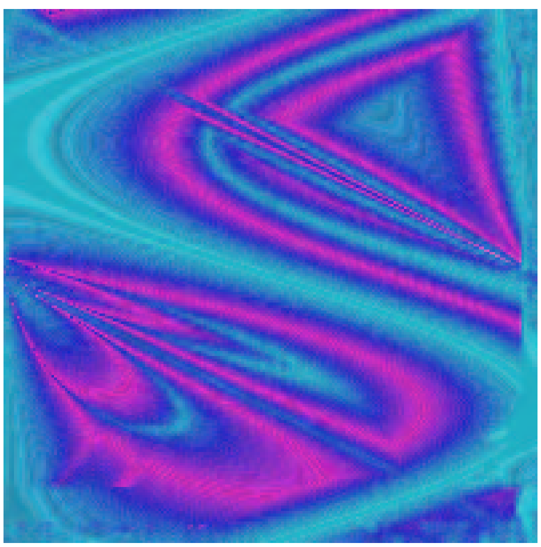
D. This leads to technical issues concerning fitness equivalence classes *e.g.* ordering and selection criteria.

VI. Results.

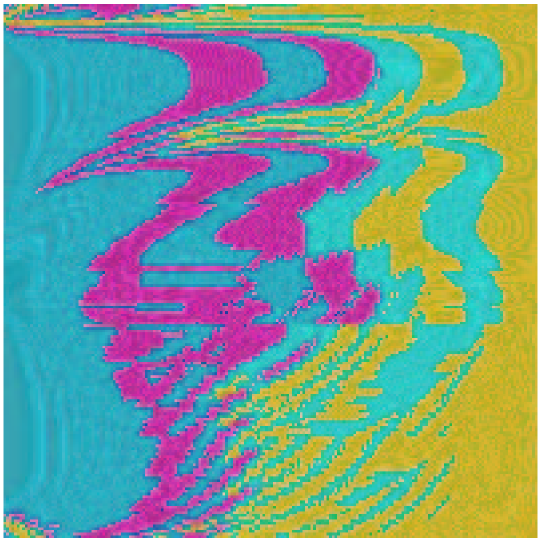
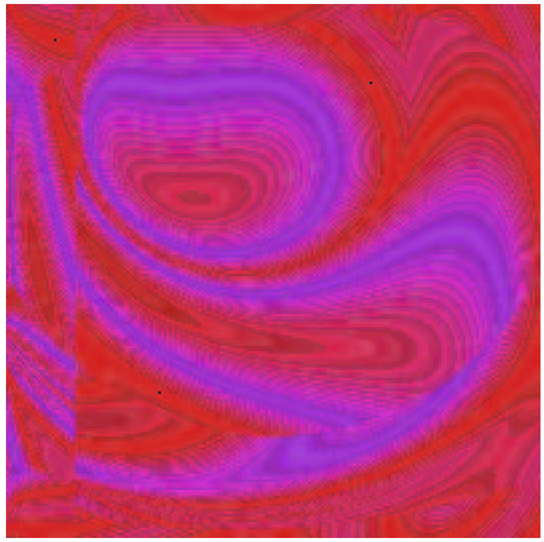
A. Using a *round robin* of fitness functions whose component weights were determined by using a preliminary run to estimate component ranges.

$$\begin{aligned} f_1 &= 10J_{1,25} + B_{1,4} \\ f_2 &= B_{1,4} + A_{1,4}/5 \\ f_3 &= A_{1,4}/5 + 10J_{1,25} \end{aligned}$$

- From generations #300 and #500 of a sample run.



• From generations #100 and #400 of a sample run.

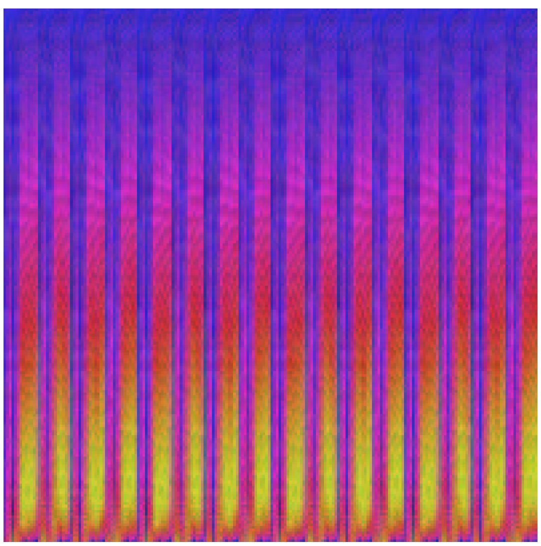
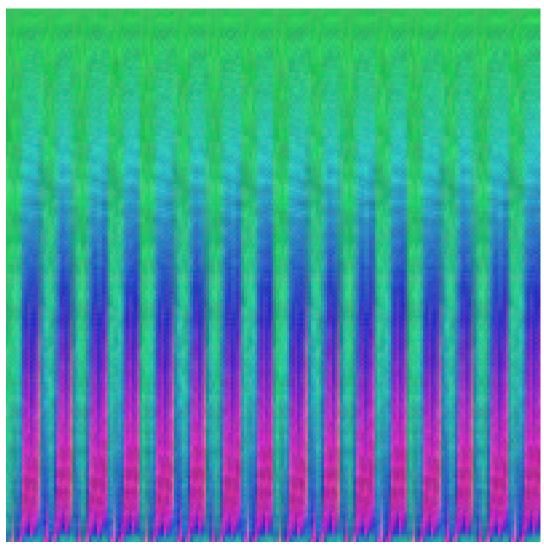
 

B. Using a round robin with only one component

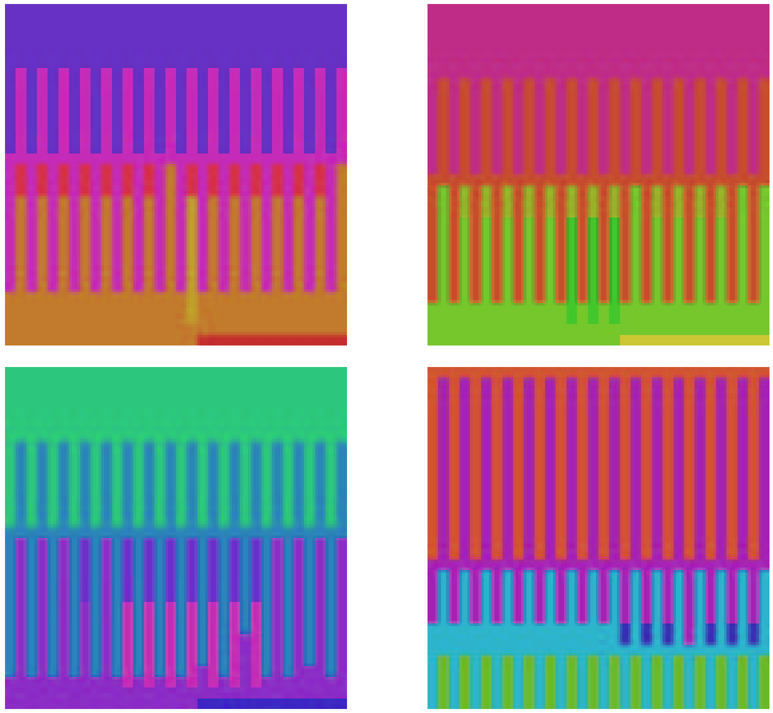
$$f_1 = B_{1,25} + A_{1,3}$$
$$f_2 = 10J_{1,25} + B_{1,25},$$

the images tuned "degenerate", the leading front exploded, but the segmentations were interesting.

• Degenerate images from generation #400.

- Segmented images from the same run.

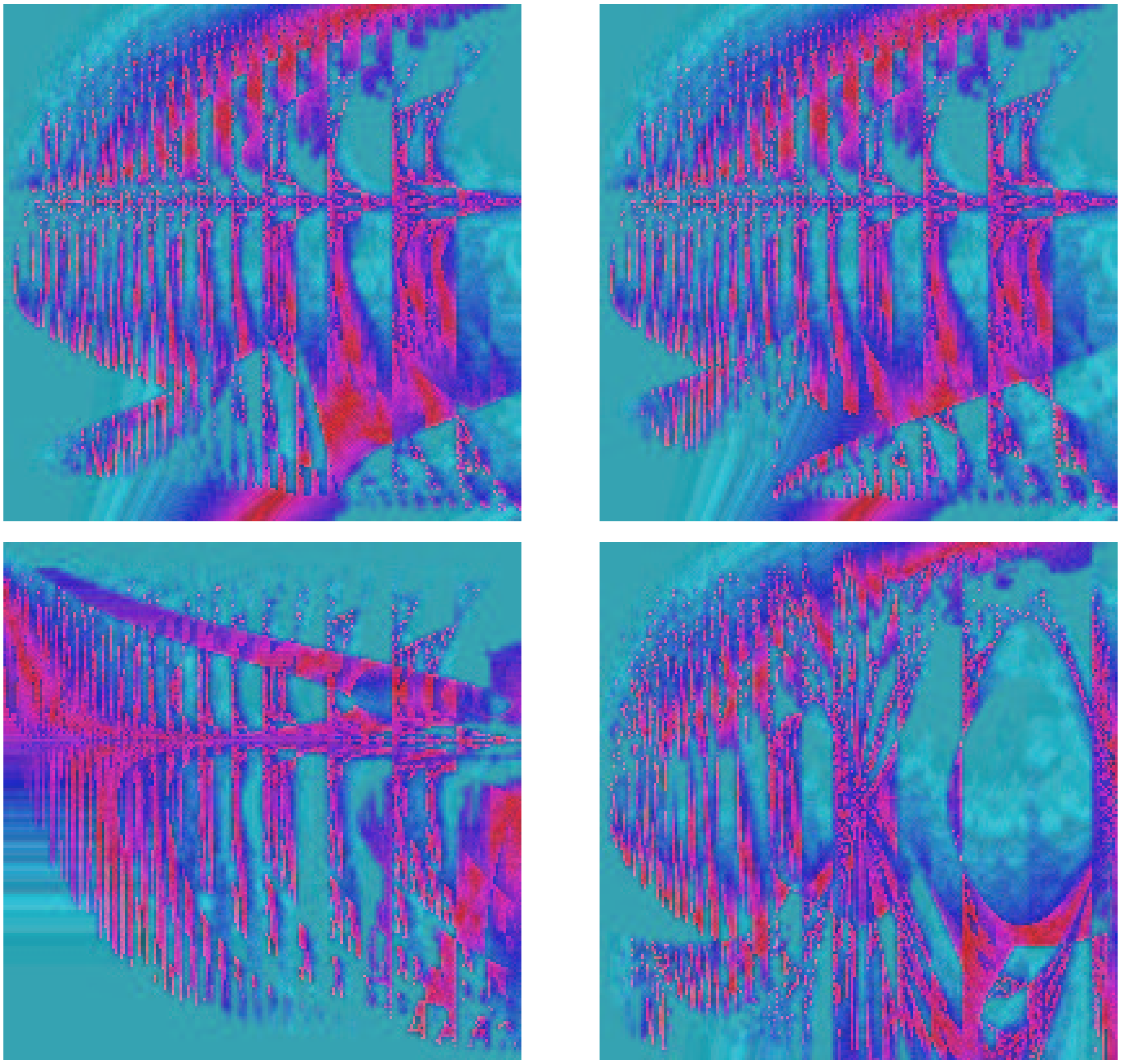


C. Using a “fix” to the previous example,

$$f_1 = B_{1,25} + A_{1,3}$$
$$f_2 = J_{12,25} + B_{1,4}$$

yielded *lineages*.

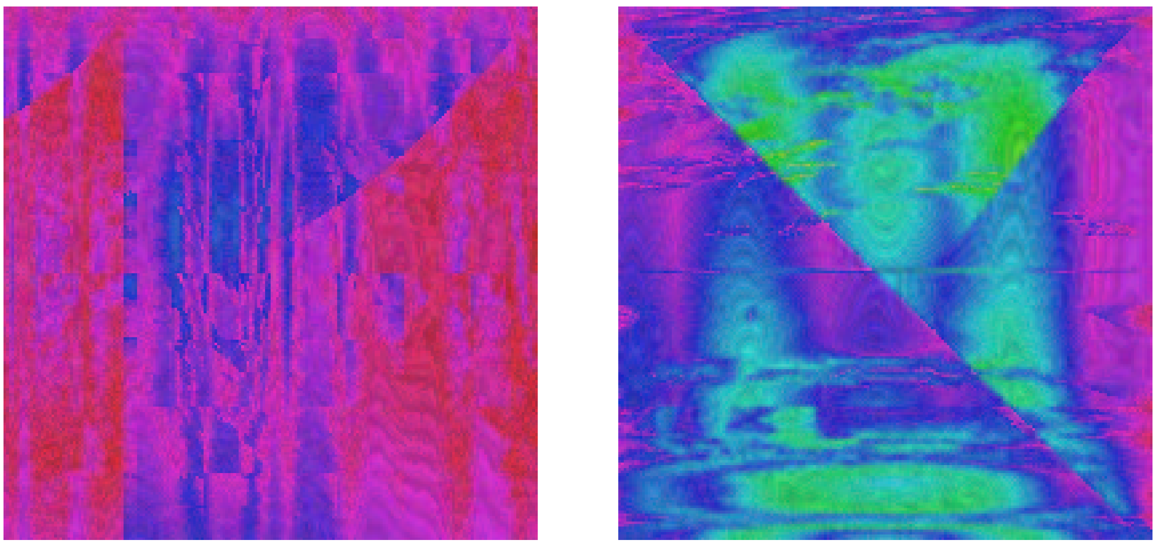
• Generations #100, #200, #300, #400.



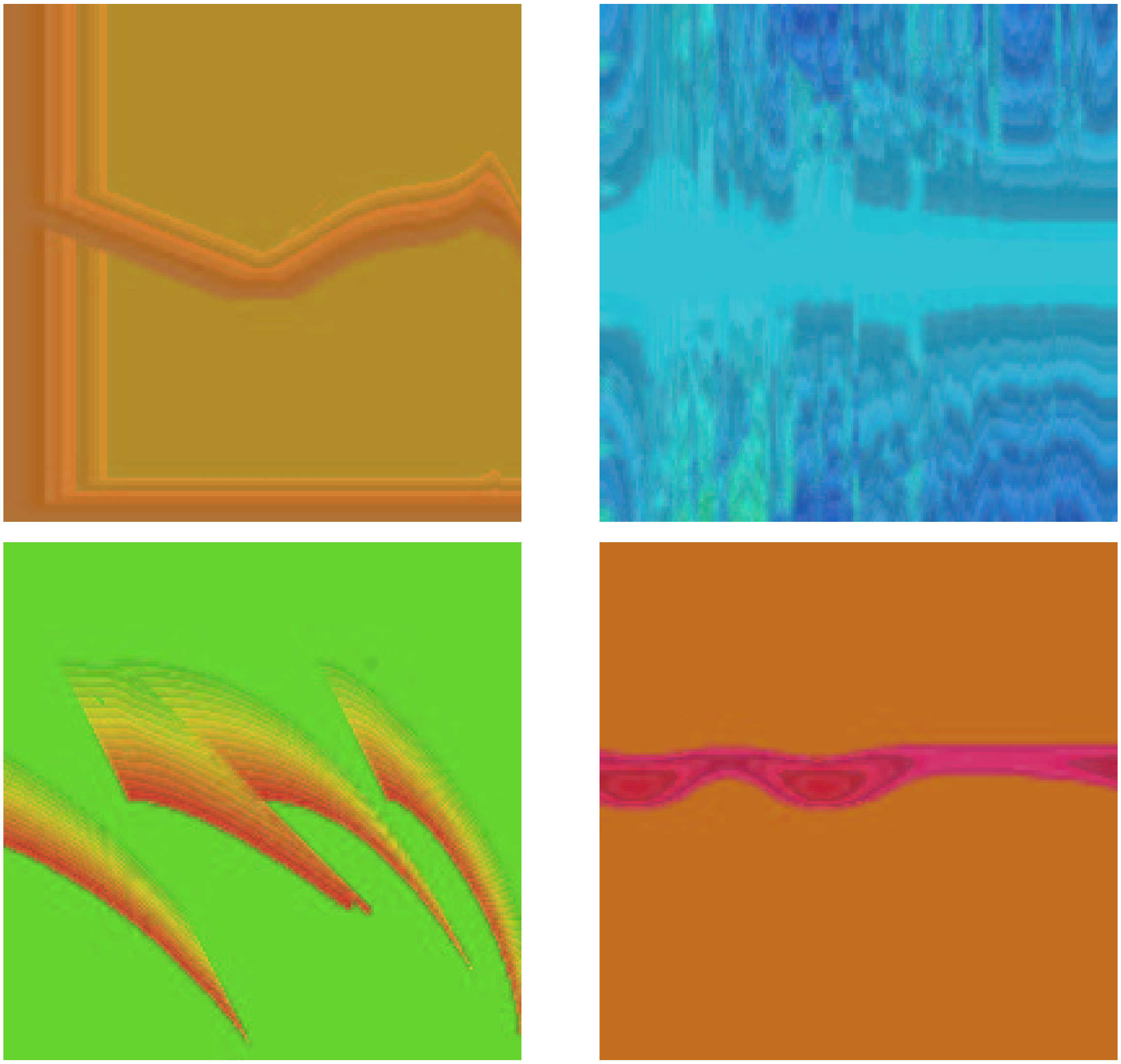
D. An experiment using the aesthetic components in the nonlinear scheme:

$$\begin{aligned} f_1 &= (B_{1,4} + A_{1,4})J_{12,25} \\ f_2 &= (J_{12,25} + B_{1,4})A_{1,4}, \end{aligned}$$

produced after only 25 generations:



E. Some miscellaneous examples.



VII. Conclusions.

A. Principal Advantage — Better novelty and variation.

B. Principal Disadvantage — Still too hard to craft fitness functions.