Image Compression using Gene Expression Programming

John Zechlin

Background

- Need for compression of images.
- Not concerned about the time, only the size of the file.
- Lossless compression scheme

Basic Idea

- Use Gene Expression Programming to evolve a model that will solve most (if not all) of the pixel values for a given image. (alone, produces lossy compression)
- Using a combination of the evolved model, Huffman coding, and run-length encoding to achieve lossless compression.

Preprocessing

- Some methods tried...
  - Five Values: get an array of 5 columns and M rows, M being the number of pixels that we are trying to solve for. (Basic Method)
  - Four Values
  - Five Values w/Index
  - Time Series Traversal
  - Time Series Modified
  - Grid Approach
- The method currently being explored:
  - Five Values w/ Index and Zig-Zag Traversal

Other Preprocessing

- Discrete Wavelet Transform
- Discrete Cosign Transform
- Segmentation of the list of values
- Difference

GeneXproTools 4.0 - [zigzag]

File Edit View Reports Data Settings Functions Run Results Model Scoring Help

Runs

- Runs
- zigzag
  - Data
  - Settings
  - Functions
  - Run
  - History
  - Results
  - Model
  - Scoring

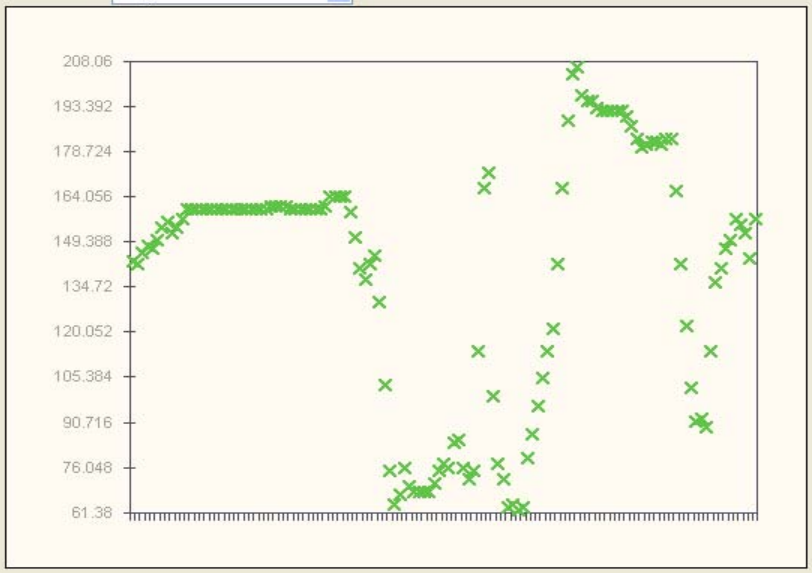
Home | zigzag

Training Set

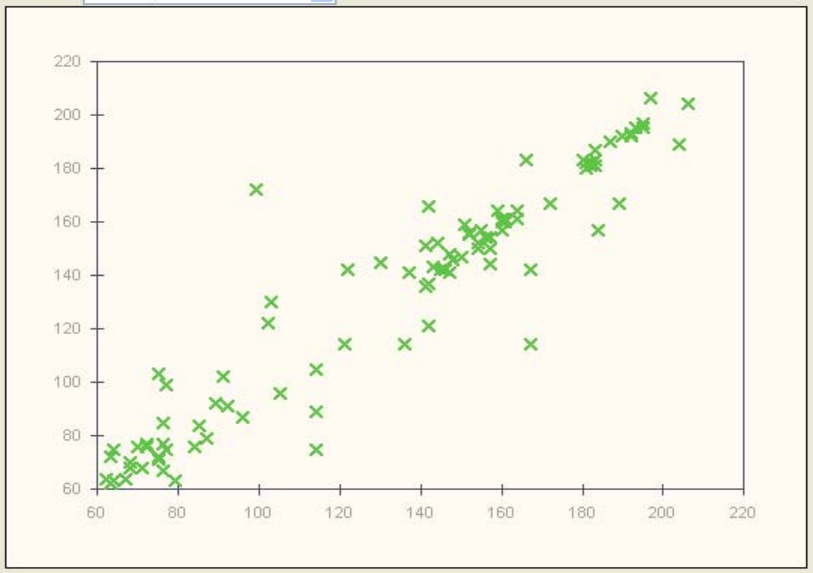
Samples: 127 Independent Variables: 5 Reset Refresh Change Data Set Remove Data Set

	A	B	C	D	Index	Target
1	143	146	146	146	31242	143
2	146	143	143	142	31243	142
3	148	144	143	143	31244	146
4	147	147	144	143	31245	148
5	150	148	147	144	31246	147
6	154	149	148	147	31247	150
7	156	149	149	148	31248	154
8	152	152	149	149	31249	156
9	154	152	152	149	31250	152
10	157	153	152	152	31251	154
11	160	153	153	152	31252	157
12	160	157	153	153	31253	160
13	160	158	157	153	31254	160
14	160	158	158	157	31255	160
15	160	158	158	158	31256	160
16	160	161	158	158	31257	160
17	160	161	161	158	31258	160
18	160	161	161	161	31259	160
19	160	161	161	161	31260	160

Distribution: Target



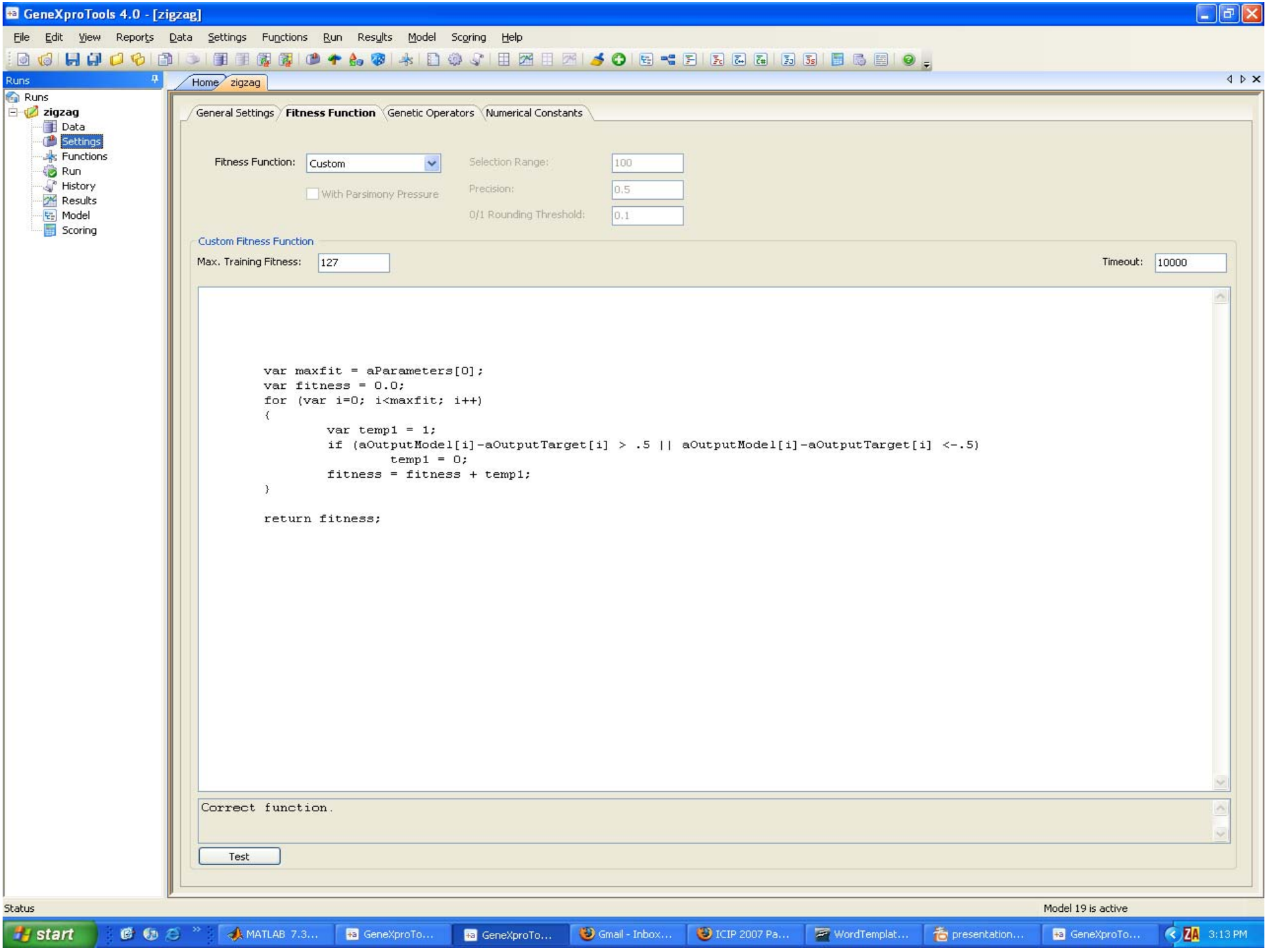
Scatter: A vs Target

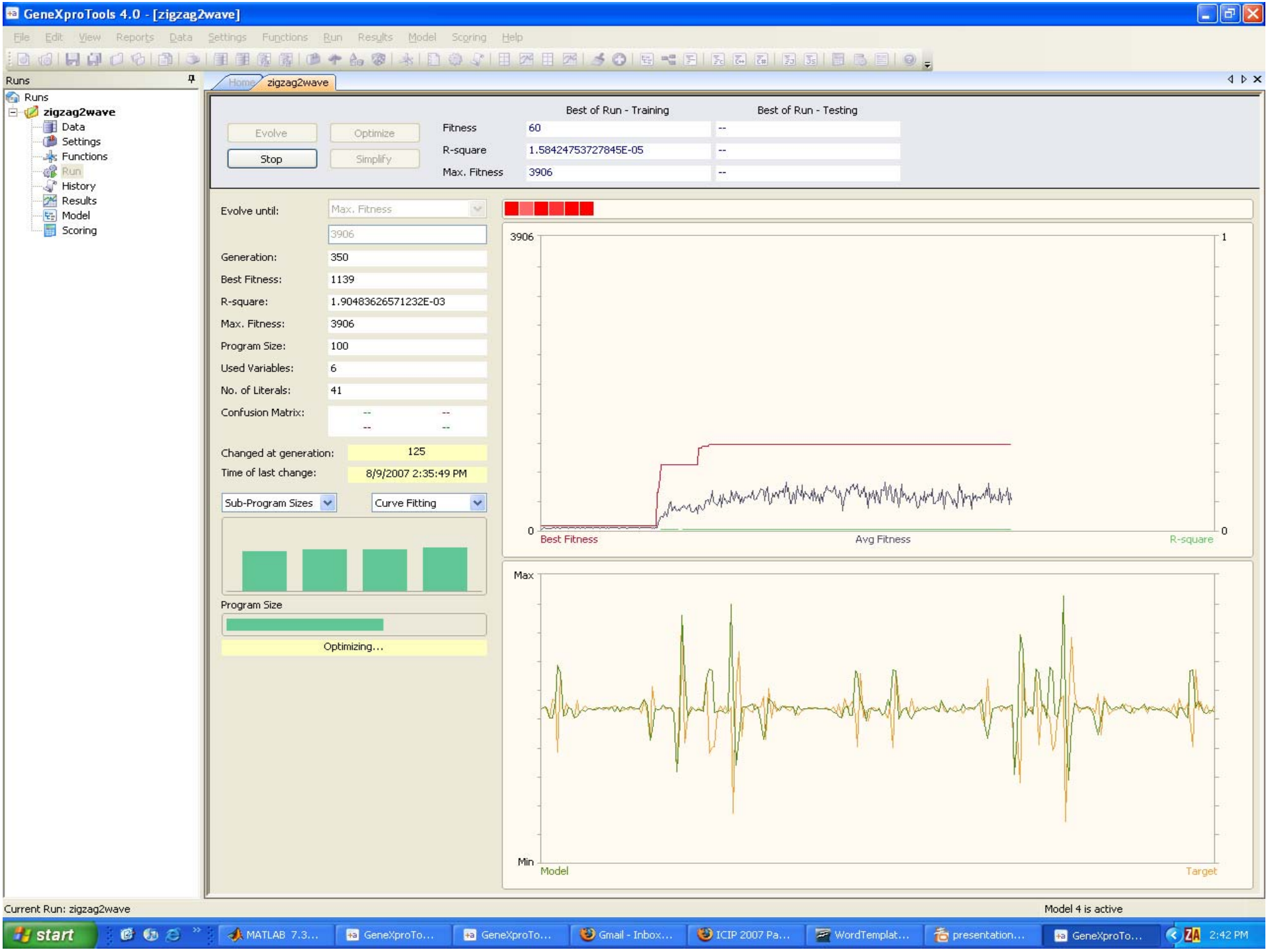


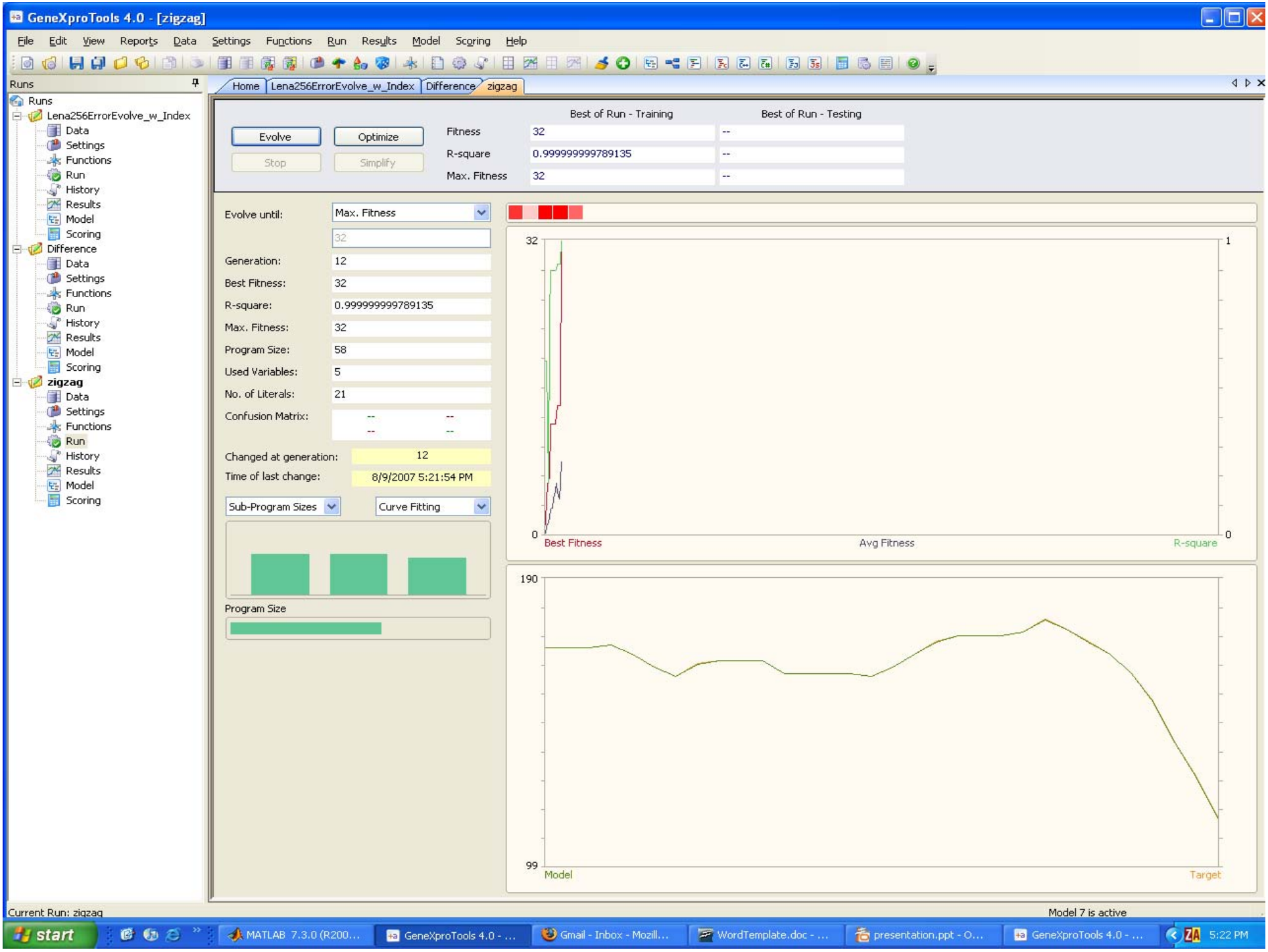
Current Run: zigzag

Model 19 is active

start MATLAB 7.3.0 (... GeneXproTools ... GeneXproTools ... ICIP 2007 Paper... WordTemplate.d... presentation.ppt... GeneXproTools 4... 5:06 PM


GeneXproTools 4.0 - [zigzag]
File Edit View Reports Data Settings Functions Run Results Model Scoring Help
Runs
Runs
zigzag
Data
Settings
Functions
Run
History
Results
Model
Scoring
Home zigzag
General Settings Fitness Function Genetic Operators Numerical Constants
Fitness Function: Custom
With Parsimony Pressure
Selection Range: 100
Precision: 0.5
0/1 Rounding Threshold: 0.1
Custom Fitness Function
Max. Training Fitness: 127
Timeout: 10000
var maxfit = aParameters[0];
var fitness = 0.0;
for (var i=0; i<maxfit; i++)
{
var temp1 = 1;
if (aOutputModel[i]-aOutputTarget[i] > .5 || aOutputModel[i]-aOutputTarget[i] <-.5)
temp1 = 0;
fitness = fitness + temp1;
}
return fitness;
Correct function.
Test
Status
Model 19 is active
start
MATLAB 7.3...
GeneXproTo...
GeneXproTo...
Gmail - Inbox...
ICIP 2007 Pa...
WordTemplat...
presentation...
GeneXproTo...
3:13 PM


GeneXproTools 4.0 - [zigzag2wave]
File Edit View Reports Data Settings Functions Run Results Model Scoring Help
Runs
Runs
zigzag2wave
Data
Settings
Functions
Run
History
Results
Model
Scoring
Home
zigzag2wave
Best of Run - Training
Best of Run - Testing
Evolve
Optimize
Stop
Simplify
Fitness 60 --
R-square 1.58424753727845E-05 --
Max. Fitness 3906 --
Evolve until: Max. Fitness
3906
Generation: 350
Best Fitness: 1139
R-square: 1.90483626571232E-03
Max. Fitness: 3906
Program Size: 100
Used Variables: 6
No. of Literals: 41
Confusion Matrix: -- -- -- --
Changed at generation: 125
Time of last change: 8/9/2007 2:35:49 PM
Sub-Program Sizes
Curve Fitting
Program Size
Optimizing...
3906
0
1
0
Best Fitness
Avg Fitness
R-square
Max
Min
Model
Target
Current Run: zigzag2wave
Model 4 is active
start
MATLAB 7.3...
GeneXproTo...
GeneXproTo...
Gmail - Inbox...
ICIP 2007 Pa...
WordTemplat...
presentation...
GeneXproTo...
2:42 PM


GeneXproTools 4.0 - [zigzag]
File Edit View Reports Data Settings Functions Run Results Model Scoring Help
Runs
Runs
Lena256ErrorEvolve_w_Index
Data
Settings
Functions
Run
History
Results
Model
Scoring
Difference
Data
Settings
Functions
Run
History
Results
Model
Scoring
zigzag
Data
Settings
Functions
Run
History
Results
Model
Scoring
Home
Lena256ErrorEvolve_w_Index
Difference
zigzag
Evolve
Optimize
Stop
Simplify
Best of Run - Training
Best of Run - Testing
Fitness 32 --
R-square 0.999999999789135 --
Max. Fitness 32 --
Evolve until: Max. Fitness
32
Generation: 12
Best Fitness: 32
R-square: 0.999999999789135
Max. Fitness: 32
Program Size: 58
Used Variables: 5
No. of Literals: 21
Confusion Matrix: -- -- -- --
Changed at generation: 12
Time of last change: 8/9/2007 5:21:54 PM
Sub-Program Sizes
Curve Fitting
Program Size
32
0
1
0
Best Fitness
Avg Fitness
R-square
190
99
Model
Target
Current Run: zigzag
Model 7 is active
start
MATLAB 7.3.0 (R200...
GeneXproTools 4.0 - ...
Gmail - Inbox - Mozill...
WordTemplate.doc - ...
presentation.ppt - O...
GeneXproTools 4.0 - ...
5:22 PM

Encoding

- The presence of a residual array is the concern here.
- Hopefully only have several values that need to be expressed.
- Use Huffman coding to represent those values.
- Use run-length encoding to further compress those values.

Major Challenges

- The existence of duplicate lines of input with different output.
- Files with numerous zeros.
- Large amounts of variance between neighboring pixels.
- Some parts of the file are radically different than others.

Current Method

- Take DWT of Image
- Use zigzag method of preprocessing on “A” array.
- Evolve a model using the resulting file
- Then, use a similar method to evolve models for the “B”, “C”, and “D.” Also use Huffman and run-length encoding.

Questions

- Answer
- Answer
- Answer
- Answer
- Answer
- Answer