A Modified Genetic Algorithm for Matching Building Sets

Andrew Buck, James Keller, Marjorie Skubic
Computational Intelligence Research Lab – University of Missouri

Abstract

We present an approach to the task of locating a group of buildings based solely on their relative spatial relationships. Of importance to us is the general text-to-sketch problem where a sketch of building locations must be matched to actual satellite imagery [1]. Information about the nature of these relative positions is captured by the histograms of forces (HoF) [2]. We consider a modified genetic algorithm that allows us to search for a specific group of buildings within a large geospatial database using the histograms of forces in the matching process. A novel mutation operator is introduced to adapt the standard GA to this environment.

Method

We initialize our search by creating n random chromosomes, each representing a possible building set. For each chromosome, we compute a set of force histograms that quantitatively capture the relative orientations and scales between each of its buildings. The target histogram set generated from an input sketch is used to evaluate the match quality of each chromosome. Mutation is performed on a chromosome by replacing a random building with the nearest neighbor that produces the highest match quality. Each new chromosome replaces it's parent with a probability proportional to its fitness. When a chromosome's match quality exceeds q_{min}, the top match results are returned.
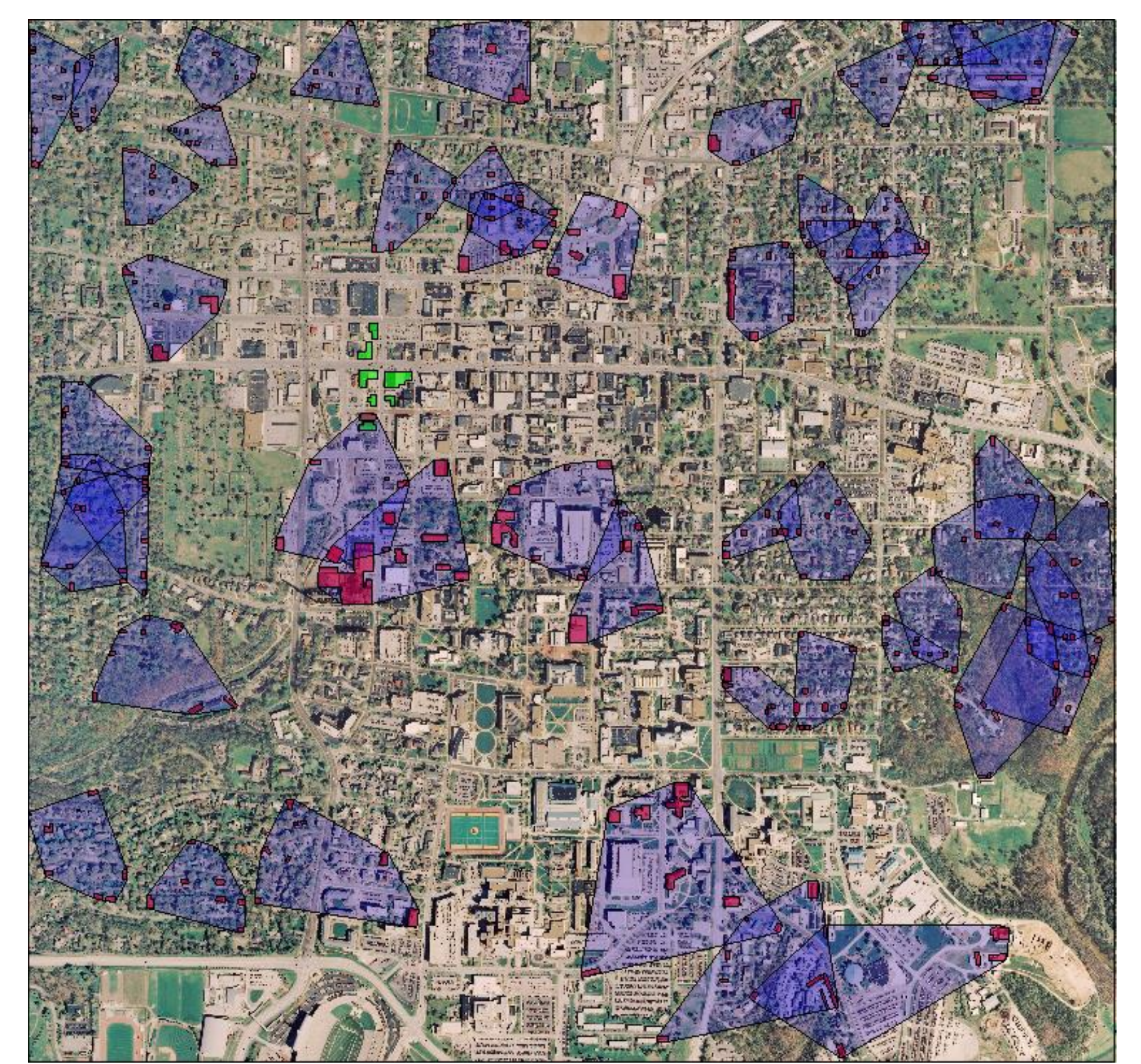
Input sketch
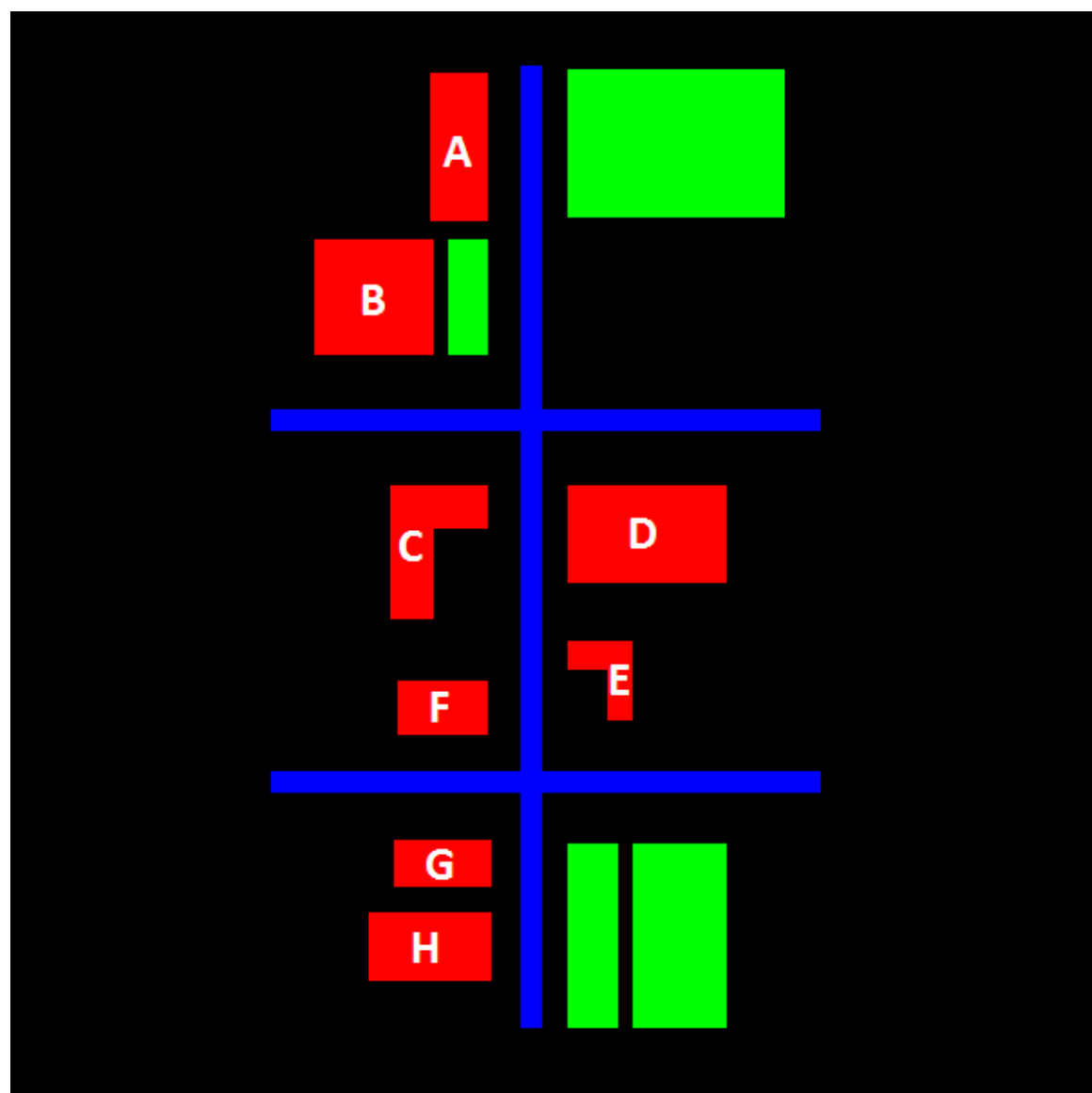
Evaluate target histogram set

Target Histogram Set
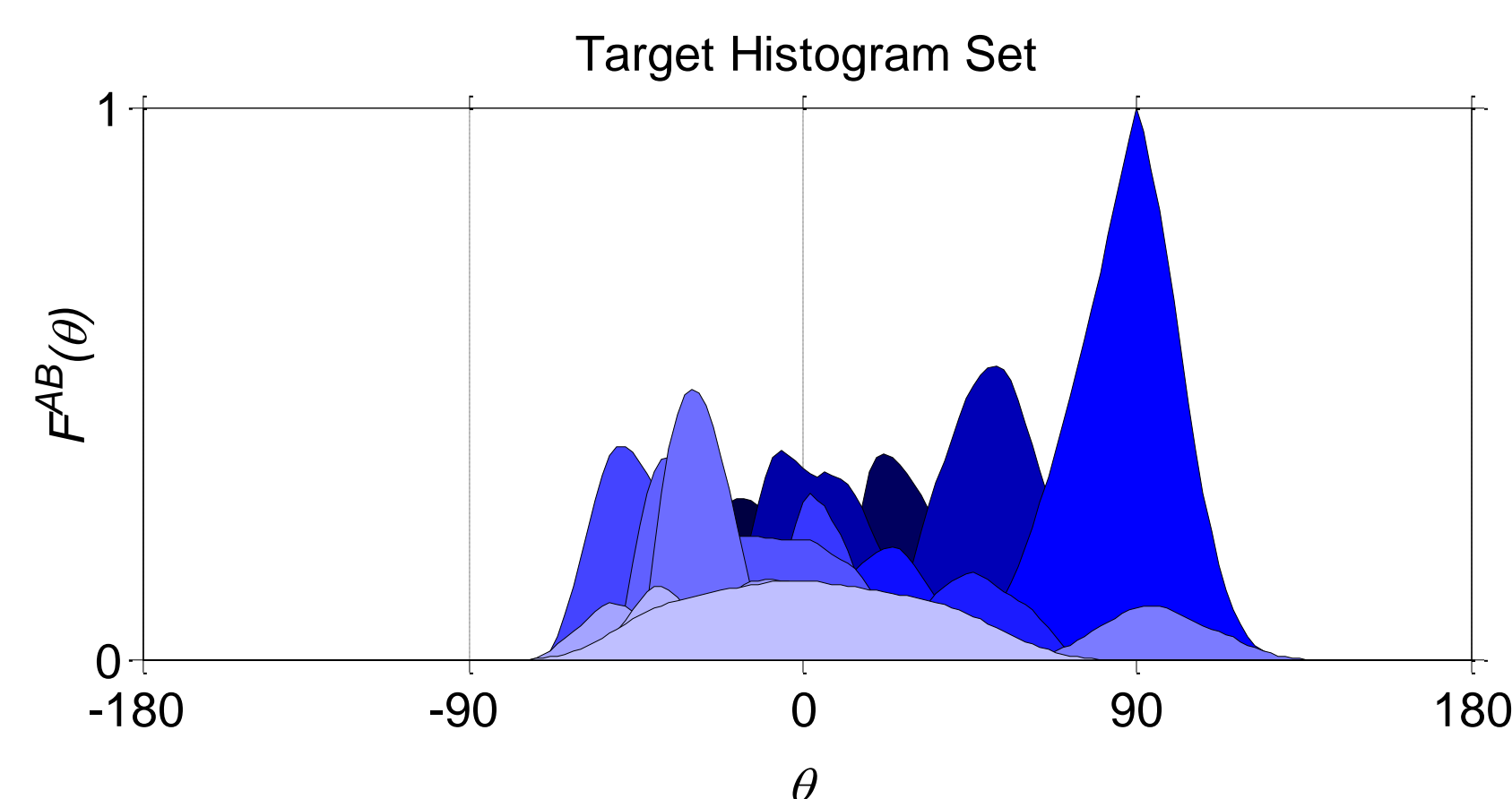
Add n random chromosomes to population

Evaluate search histogram set and compare to target histogram set for each chromosome
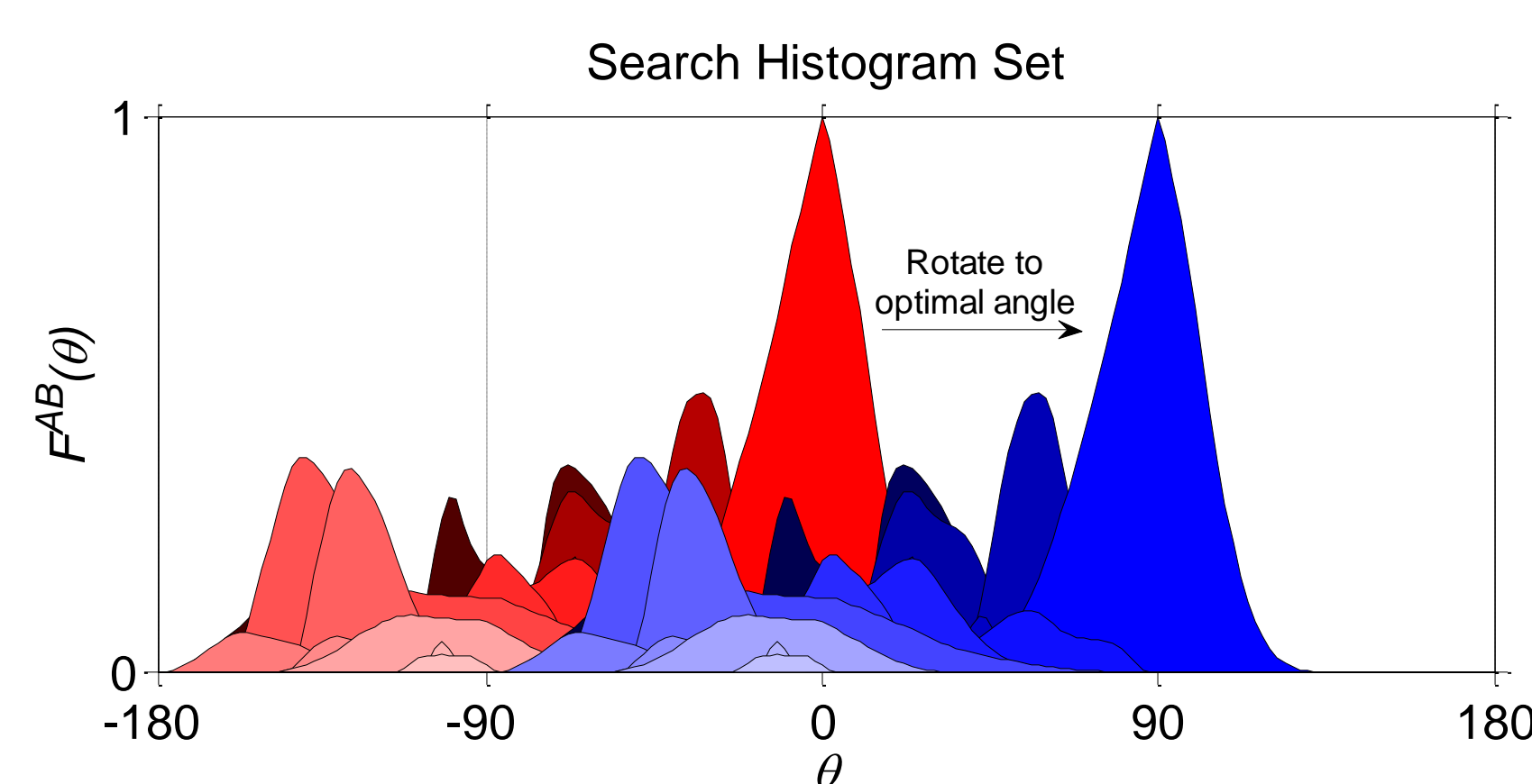
Search Histogram Set

Rotate to optimal angle

If match quality is not above threshold q_{min}, select a random building to change.
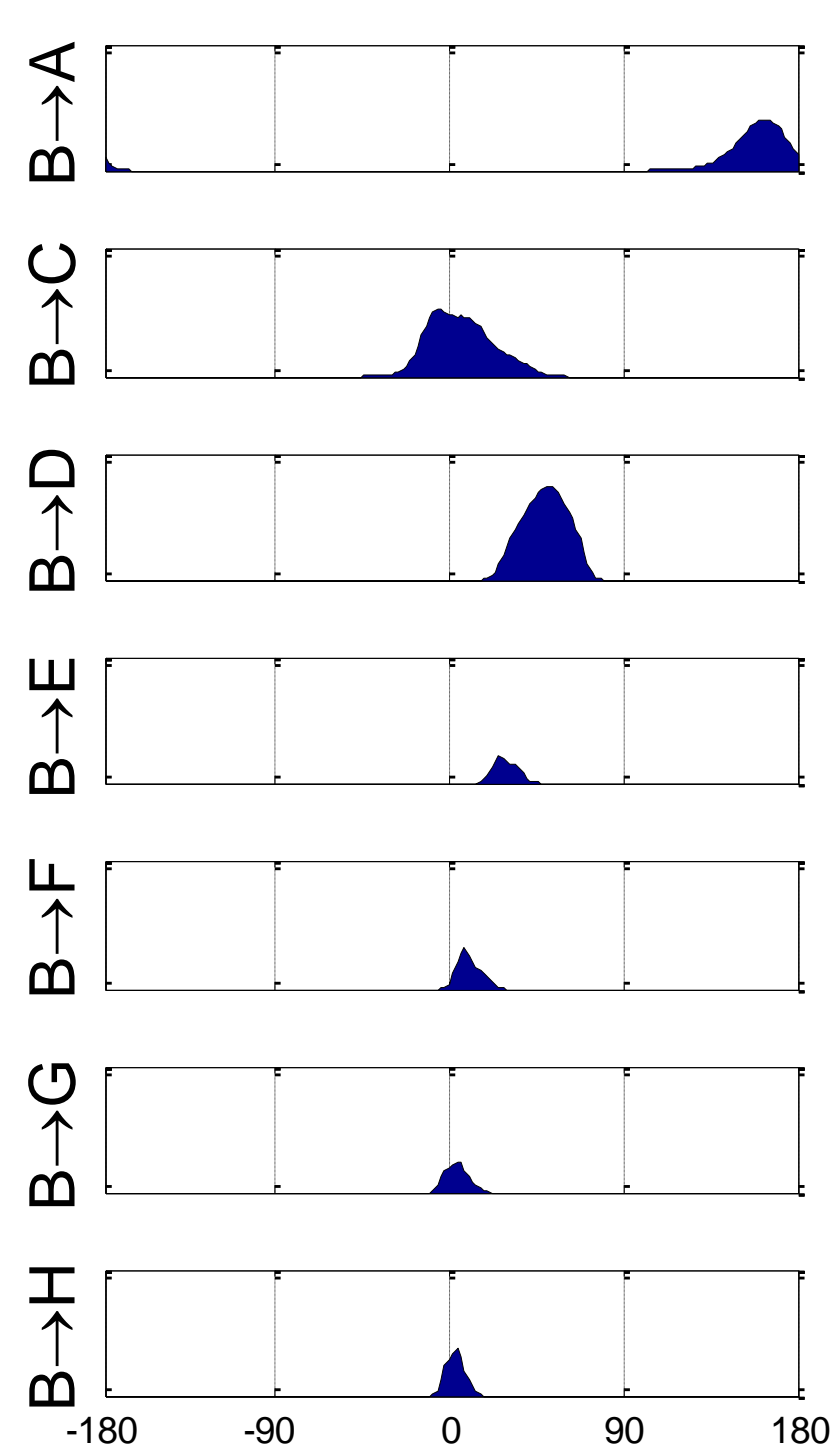
Building B chosen for replacement

	Building IDs							
Target Set	A	B	C	D	E	F	G	H
Search Set	1224	1255	1291	1292	1340	1338	1353	1362

Update population with new chromosomes

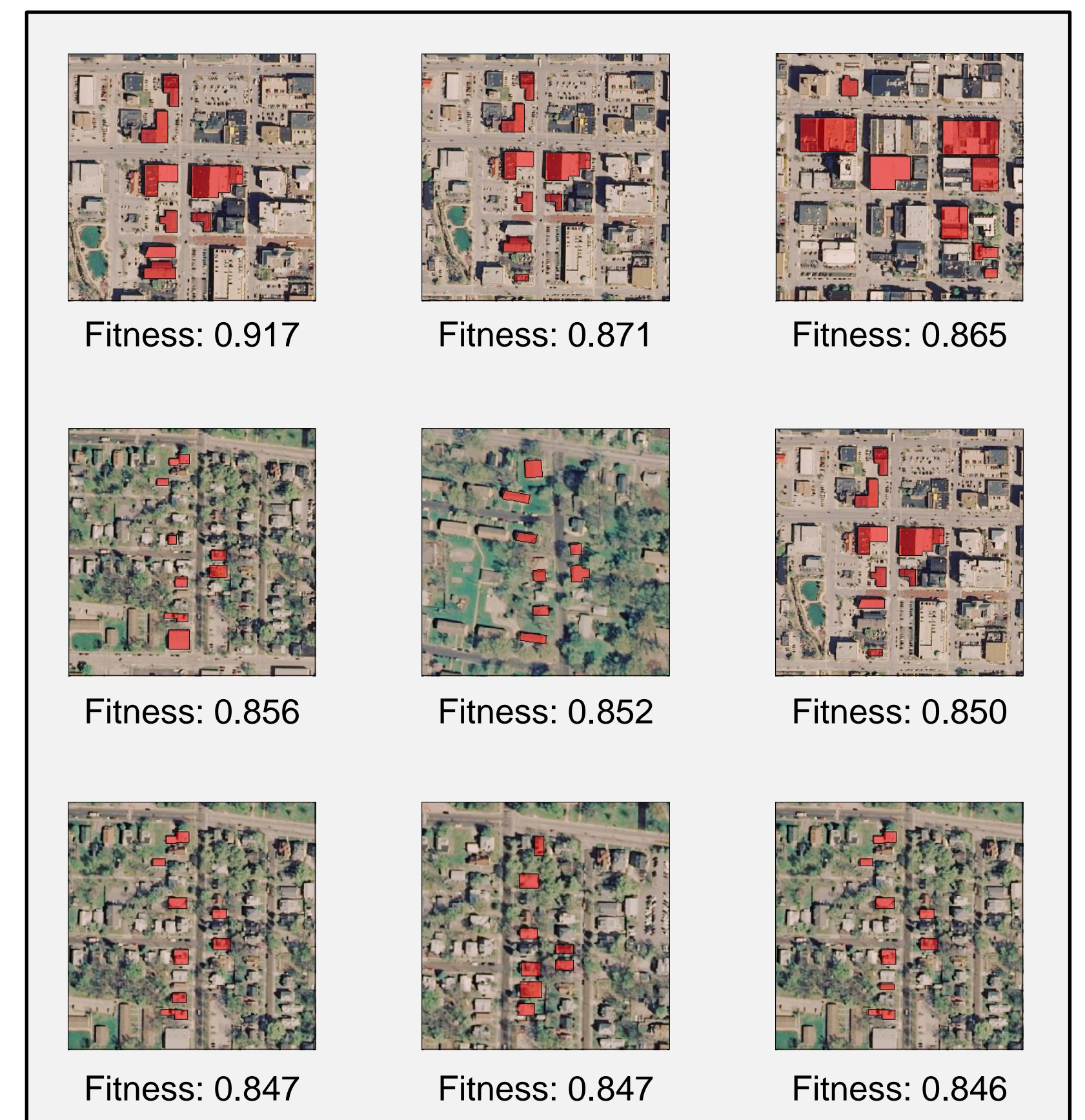Compute target histograms for the chosen building and compare against each nearest neighbor

Replace chosen building with the nearest neighbor that most closely matches the target histograms

Top match results

Nearest Neighbors for Building 1255
1256
1248
1223
1291
1224
1225
1261
1303
1289
1139
1259
1162
1232
1292
1266
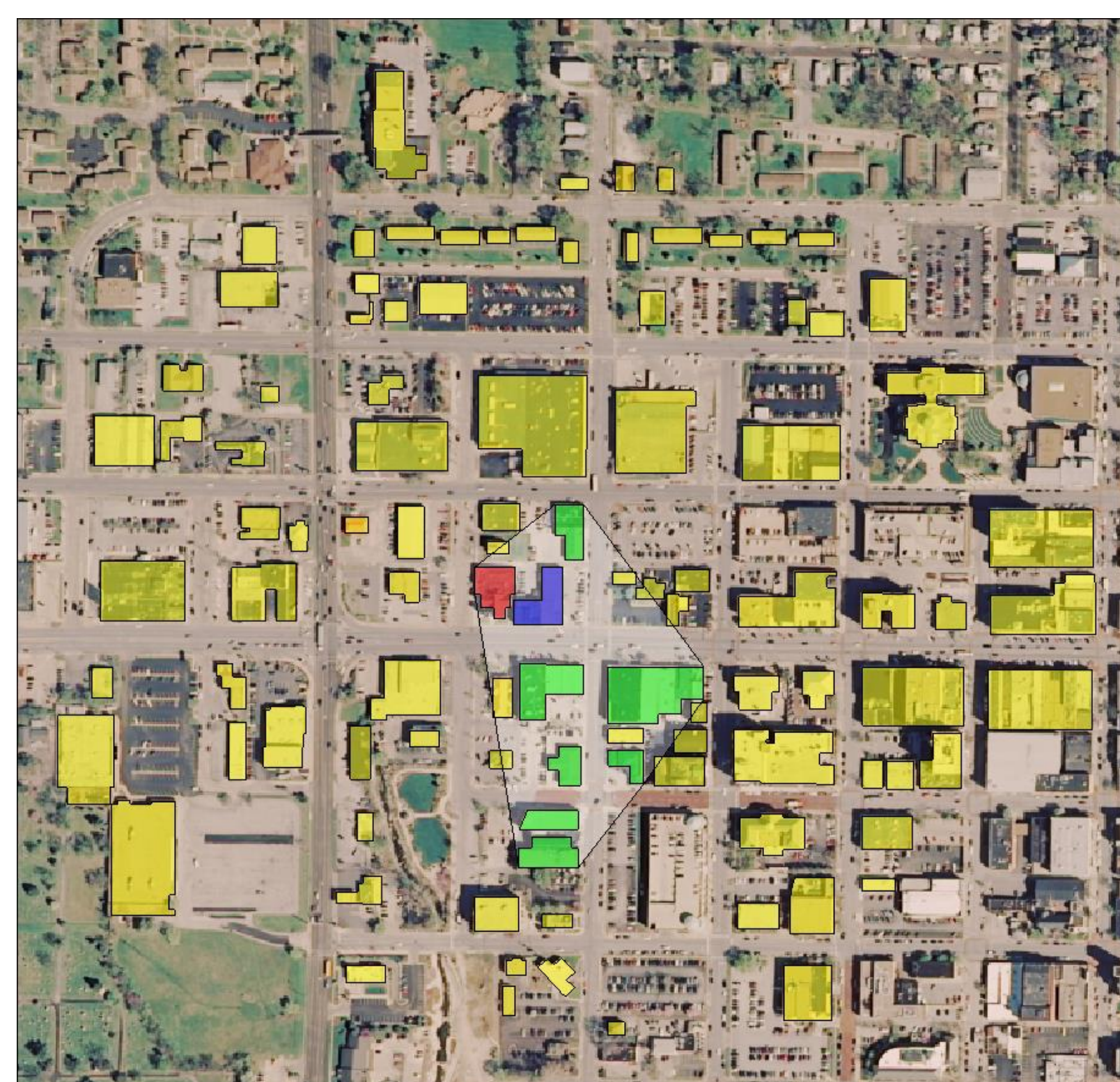
When a chromosome's match quality exceeds q_{min}, output results

Fitness: 0.917 Fitness: 0.871 Fitness: 0.865
Fitness: 0.856 Fitness: 0.852 Fitness: 0.850
Fitness: 0.847 Fitness: 0.847 Fitness: 0.846

Experimental Results

Our tests were conducted on a dataset of 2467 buildings from Columbia, MO. We created 100 test sets with five buildings each, with the restriction that each building must be one of the 30 nearest neighbors of each other building in the test set. In the first experiment, we evaluate the performance of the HoF GA using direct resubstitution, whereas we apply a random global transformation to each test set for the second experiment. In both cases, the average fitness value rapidly increases during the first 50 generations and then levels off. Larger populations have higher fitness values and tend to take fewer generations to converge. The global transformations also required more generations to converge, with several failing to converge at all.

Resubstitution

Average Fitness Value per Generation for Resubstitution

Average Generations to Convergence per Test for Resubstitution

Global Transformation

Average Fitness Value per Generation for Global Transformation

Average Generations to Convergence per Test for Global Transformation
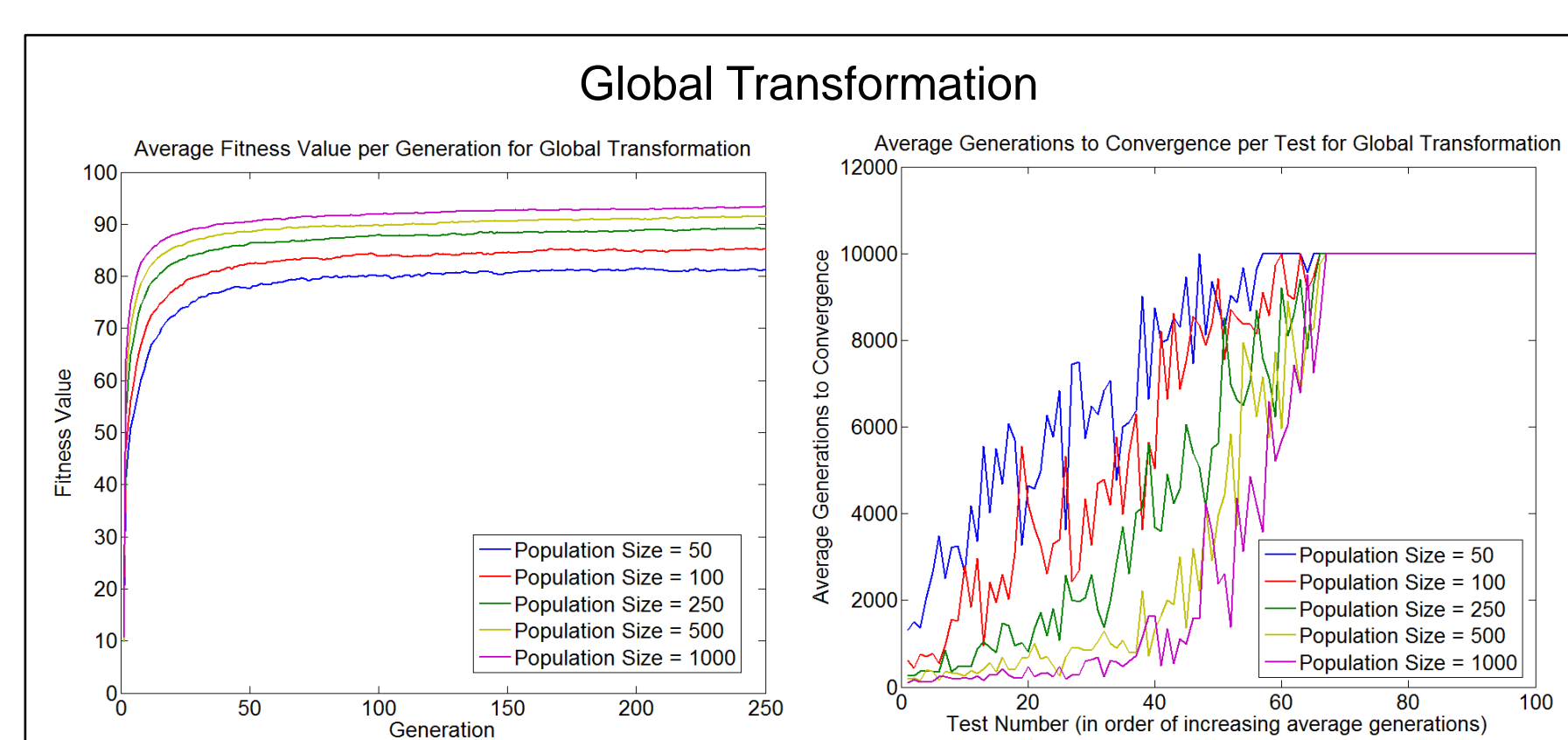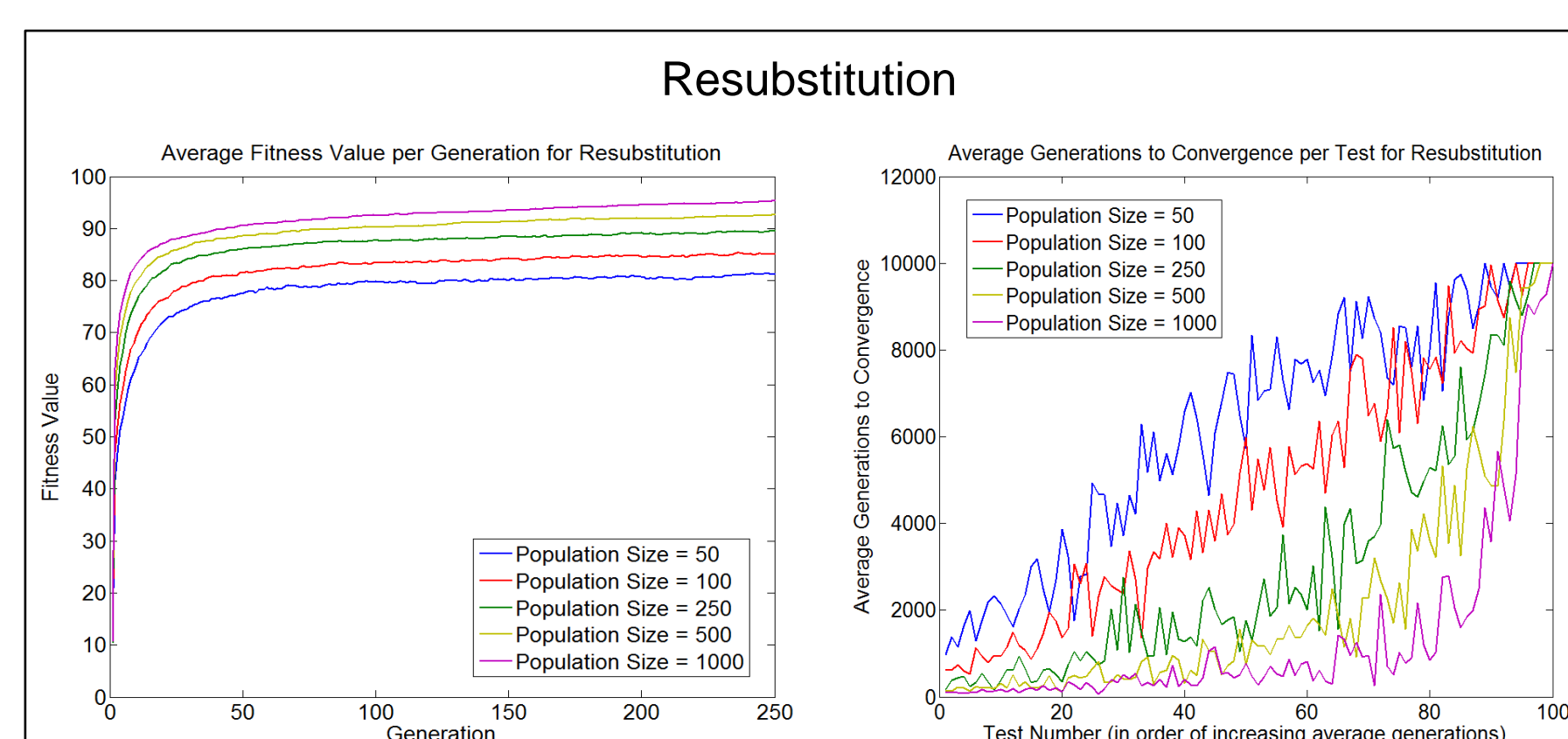
PERCENT OF TESTS WHICH CONVERGE WITHIN 10,000 GENERATIONS

Population Size	Resubstitution	Global Transformation
50	55.0%	34.4%
100	70.7%	43.2%
250	84.5%	52.7%
500	90.1%	58.1%
1000	94.8%	60.5%

Convergence is defined as having at least one chromosome in the population with a fitness value of 0.99 or greater. Tests using strict resubstitution have a higher convergence rate than those using global transformations. Larger population sizes are also shown to improve the convergence rate.

Conclusion & Future Work

The histograms of forces provide a useful framework for matching building sets by emphasizing orientation and scale rather than specific shapes. The HoF GA uses a novel mutation operator to produce a set of matches that reflect the spatial relations of an input sketch. Although the ideal match is often returned at the top of the results, other matches may be rotated or have other flaws which we would like to filter out. Future work will impose additional search criteria such as road networks and specific object types (parking lots, restaurants, etc.).

Acknowledgements & References

[1] I. Sledge and J. Keller, "Mapping Natural Language to Imagery: Placing Objects Intelligently," IEEE Proceedings, International Conference on Fuzzy Systems (FUZZ-IEEE), Jeju Island, Korea, August, 2009, pp. 518-524.

[2] P. Matsakis, J. M. Keller, O. Sjahputera, and J. Marjamaa, "The Use of Force Histograms for Affine-Invariant Relative Position Description," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 26, pp. 1-18, Jan. 2004.