Genetic Programming for Edge Detection Based on Figure of Merit

Wenlong Fu
School of Mathematics, Statistics and Operations Research
Victoria University of Wellington
New Zealand
Wenlong.Fu@msor.vuw.ac.nz

Mark Johnston
School of Mathematics, Statistics and Operations Research
Victoria University of Wellington
New Zealand
Mark.Johnston@msor.vuw.ac.nz

Mengjie Zhang
School of Engineering and Computer Science
Victoria University of Wellington
New Zealand
Mengjie.Zhang@ecs.vuw.ac.nz

ABSTRACT

The figure of merit (FOM) is popular for testing an edge detector's performance, but there are very few reports using FOM as an evaluation method in the learning stage of supervised learning methods. In this study, FOM is investigated as a fitness function in Genetic Programming (GP) for edge detection. Since FOM has some drawbacks from type II errors, new fitness functions are developed based on FOM in order to address these weaknesses. Experimental results show that FOM can be used to evolve GP edge detectors that perform better than the Sobel detector, and the new fitness functions clearly improve the ability of GP edge detectors to find edge points and give a single response on edges, compared with the fitness function using FOM.

Categories and Subject Descriptors

I.2.m.c [**Artificial Intelligence**]: methodologies—*Evolutionary computing and genetic algorithms*; I.4.6.a [**Image Processing and Computer Vision**]: Edge and feature detection

General Terms

Algorithms

Keywords

Genetic Programming, Edge Detection, Figure of Merit

1. GOAL

The overall goals of this paper are to investigate the figure of merit (FOM) [6] as a fitness function in a GP system for evolving low-level edge detectors based on using an entire image as input, and to develop a new fitness function based on FOM for improving detection performance. FOM has weaknesses from type II errors [1], but whether the weaknesses make the GP system using FOM as the fitness function perform poorly when evolving low-level GP edge detectors for natural images is not clear. Specifically, we investigate the following objectives.

GECCO'12 Companion, July 7–11, 2012, Philadelphia, PA, USA.
ACM 978-1-4503-1178-6/12/07.

- Whether FOM as a fitness function in a GP system can be used to evolve good low-level edge detectors, compared with the Sobel detector.
- Whether fitness functions based on FOM variants can improve the detection performance, compared with standard FOM.
- What features exist in detected images from the evolved detectors using the fitness functions based on FOM?

2. FITNESS FUNCTIONS

We investigate fitness functions based on FOM and employ the GP system in [2]. The standard FOM (defined by Eq. (1)) is used as a fitness function. Here, $d_1(i)$ is the distance of predicted edge point i to the nearest true edge point, and α is a positive weight factor. In general, α is small and is set to $\frac{1}{9}$ based on the overlap of a 3×3 window. The value of FOM ranges from 0 to 1; higher values indicate predictions with better quality detection.

$$FOM = \frac{1}{\max\{N_T, N_P\}} \sum_{i=1}^{N_P} \frac{1}{1 + \alpha d_1^2(i)} \tag{1}$$

In order to improve the sensitivity to type II errors, modification of FOM focuses on the matching direction. An existing FOM variant only uses distances from ground truth to predicted edges, and includes a factor addressing false (unmatched) edge points [5]. F_{nn} is defined in Eq. (2), where N_{FM} is the number of false edge points, i is one of the true edge points, β is a factor for the response on the false edge points, and $d_2(i)$ is the nearest distance from true edge point i to a predicted edge point. β was suggested be set to 1 in [5].

$$F_{nn} = \left(\frac{1}{N_T} \sum_{i=1}^{N_T} \frac{1}{1 + \alpha d_2^2(i)} \right) \left(\frac{1}{1 + \frac{\beta N_{FM}}{N_T}} \right) \tag{2}$$

$$F_b = \frac{1}{N_{T \bigcup P}} \sum_{i=1}^{N_{T \bigcup P}} \frac{1}{(1 + \alpha d_1^2(i))(1 + \alpha d_2^2(i))} \tag{3}$$

In order to fairly *balance* the measurement between type I errors and type II errors, we propose a new FOM variant F_b based on $d_1(i)$ and $d_2(i)$ and inspired by the Hausdorff distance [3]. Function F_b is defined by Eq. (3), where, $N_{T \bigcup P}$ is the number of points that are either predicted as edge points or are true edge points, and i is one of these points.

(a) 207056 (b) 23080 (c) 62096 (d) 296059

Figure 1: Images from the BSD dataset.

Table 1: Means ± Standard Deviations of FOM for fitness functions FOM, F_b and F_{nn}

Image	FOM	F_{nn}	F_b	Sobel
3096	$0.588^{\downarrow} \pm 0.102$	$0.710^{\downarrow} \pm 0.069$	0.733 ± 0.068	0.740
37073	$0.693^{\uparrow} \pm 0.079$	$0.734^{\uparrow} \pm 0.051$	$0.761^{\uparrow} \pm 0.073$	0.436
42049	0.678 ± 0.094	$0.754^{\uparrow} \pm 0.052$	$0.797^{\uparrow} \pm 0.048$	0.646
62096	$0.369^{\downarrow} \pm 0.075$	$0.563^{\uparrow} \pm 0.127$	$0.453^{\uparrow} \pm 0.094$	0.402
101087	$0.648^{\uparrow} \pm 0.061$	$0.730^{\uparrow} \pm 0.035$	$0.765^{\uparrow} \pm 0.042$	0.444
106024	$0.416^{\uparrow} \pm 0.070$	$0.424^{\uparrow} \pm 0.048$	$0.471^{\uparrow} \pm 0.045$	0.307
253055	$0.488^{\uparrow} \pm 0.078$	$0.551^{\uparrow} \pm 0.048$	$0.589^{\uparrow} \pm 0.050$	0.408
296059	0.517 ± 0.114	$0.660^{\uparrow} \pm 0.054$	$0.636^{\uparrow} \pm 0.085$	0.536
299086	$0.535^{\uparrow} \pm 0.083$	$0.601^{\uparrow} \pm 0.049$	$0.649^{\uparrow} \pm 0.066$	0.426
361010	0.454 ± 0.057	$0.677^{\uparrow} \pm 0.049$	$0.631^{\uparrow} \pm .084$	0.437

3. EXPERIMENT

3.1 Image Dataset

We use training images from the Berkeley Segmentation Dataset (BSD) [4]. We only select six images 207056, 23080, 105019, 105053, 113044 and 216053 as the training dataset. Fig. 1 shows two training images, (a) and (b), two test images, (c) and (d), and the corresponding ground truth.

3.2 Results

Table 1 gives the means and standard deviations of the detection results (FOM) from FOM, F_b and F_{nn} for ten test images. The ten images can be found by their names in the BSD dataset. For comparison, the maximum FOM values from the Sobel detector based on 20 different threshold levels are given. Also, "$^{\uparrow}$" indicates the detection results based on FOM evaluation are significantly better than the Sobel detector using a t-test with significance level 0.05, and "$^{\downarrow}$" indicates the former is significantly worse than the latter. Compared with the Sobel detector, the three fitness functions all have better detection results for most images; only for image 3096, the three fitness functions do not have significantly better results than the Sobel detector. Also, FOM has worse detection results than the Sobel detector for image 62096, but results from F_{nn} and F_b are significantly better than the Sobel detector. Therefore, we conclude that the fitness functions based on FOM can be used to evolve good low-level edge detectors.

Table 2 shows the comparisons between the detection results from these fitness functions using t-tests with significance level 0.05. Here, "+" indicates the first fitness function is significantly better than the second fitness function, "−" indicates the former is significantly worse than the latter, and empty for non-significant difference. From the statistical tests, F_b has significantly better results for all test images, and F_{nn} for nine test images, compared with FOM. Also, F_b is slightly better than F_{nn} because F_b has significantly better results for five images and significantly worse results for only two images, compared with F_{nn}. It seems that F_b is the best fitness function among the three fitness functions for evolving edge detectors based on this experiment.

Table 2: Comparisons Among FOM, F_b and F_{nn}

Image	(F_b, FOM)	(F_{nn}, FOM)	(F_b, F_{nn})
3096	+	+	
37073	+	+	
42049	+	+	+
62096	+	+	−
101087	+	+	+
106024	+		+
253055	+	+	+
296059	+	+	
299086	+	+	+
361010	+	+	−

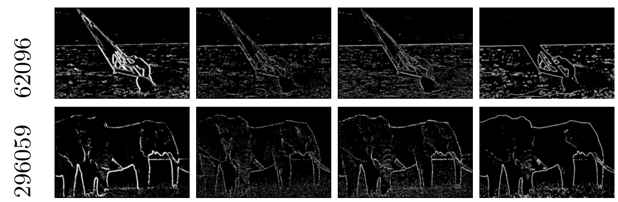

Figure 2: Detection from FOM, F_{nn}, F_b and Sobel.

3.3 Detected Images

Fig. 2 shows the detected images by the best detectors from the three fitness functions and the Sobel detector. From the detected images, the detectors evolved by the fitness functions based on FOM have high recall but bring low precision to the predicted texture edges, and the detectors from the fitness functions F_b and F_{nn} have single response for the boundaries.

4. REFERENCES

[1] J. A. Baddeley. An error metric for binary images. In *Proceedings of the International Workshop on Robust Computer Vision*, pages 59–78, 1992.

[2] W. Fu, M. Johnston, and M. Zhang. Genetic programming for edge detection: a global approach. In *Proceedings of the IEEE Congress on Evolutionary Computation*, pages 254–261, 2011.

[3] D. Huttenlocher, G. Klanderman, and W. Rucklidge. Comparing images using the Hausdorff distance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(9):850–863, 1993.

[4] D. Martin, C. Fowlkes, and J. Malik. Learning to detect natural image boundaries using local brightness, color, and texture cues. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(5):530–549, 2004.

[5] A. J. Pinho and L. B. Almeida. Edge detection filters based on artificial neural networks. In *Proceedings of the 8th International Conference on Image Analysis and Processing*, pages 159–164, 1995.

[6] W. K. Pratt. *Digital Image Processing: PIKS Inside, 3rd edition.* Wiley, 2001.