# Report of the Arc Segmentation Contest

Liu Wenyin

Dept of Computer Science, City University of Hong Kong, Hong Kong SAR, PR China
csliuwy@cityu.edu.hk

**Abstract.** The Arc Segmentation Contest, as the fifth in the series of graphics recognition contests organized by IAPR TC10, was held in association with the GREC'2003 workshop. In this paper we present the report of the contest: the contest rules, performance metrics, test images and their ground truths, and the outcomes.

## 1 Introduction

This contest on arc segmentation held at the fifth International Workshop on Graphics Recognition (GREC'2003), Barcelona, Spain, July 30-31, 2003 is the fifth in the series of graphics recognition contests organized by the International Association for Pattern Recognition's Technical Committee on Graphics Recognition (IAPR TC10). A brief history of the contest series is presented in [1]. The purpose of this series of contests is to encourage third-party independent and objective evaluation of the industrial and academic solutions to the graphics recognition problem and therefore push the research in this area.

This contest is a re-run of the fourth contest on arc segmentation [1], but with new test images. In this paper we briefly present the final report of the contest, including the contest rules, test images and their ground truths, the winners and their performance, and discussions.

## 2 General Rules

The rules are exactly the same to those of the fourth contest [1], except for the new test images. The contest rules are summarized below.

- Recognition accuracy was measured on only solid arcs.
- The tested systems were run as black boxes on-site, with their configurations fixed during the contest. No human intervention was allowed.
- 12 real life or synthesized drawing images were tested. See Section 3 for detail descriptions of these test images.
- The input file format was binary TIFF and the output file format was VEC.
- Zero score would be given to participants for images that cause system crash or non-stop.

- An overall average score based on each image's VRI [2] was used as the unique measure of performance of each participant's system. The performance evaluation software is also available at the contest website [4].

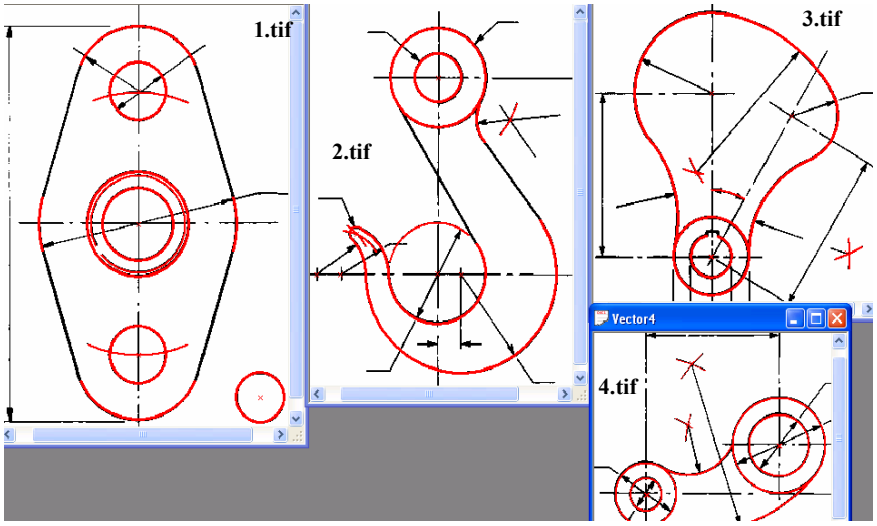


**Fig.1.** The ground truth vector drawing for the four synthesized test images.

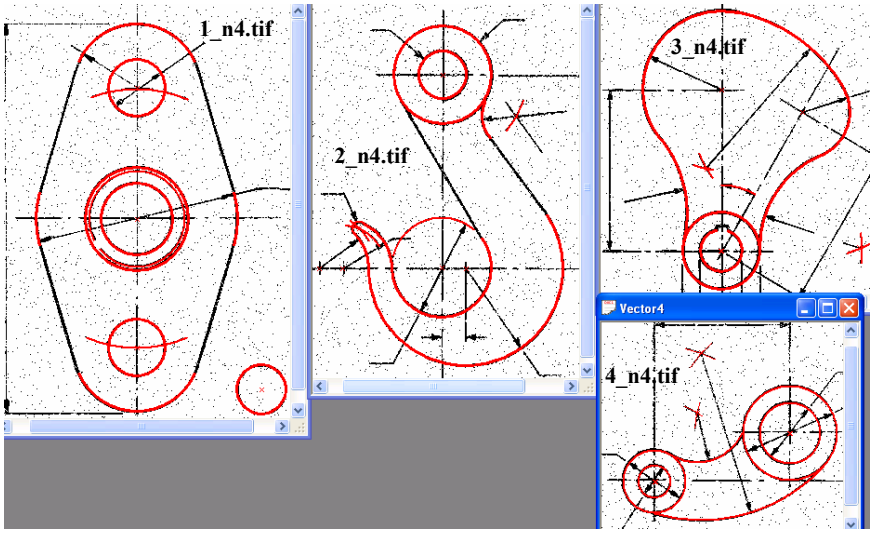


**Fig.2.** The ground truth vector drawing for the four synthesized test images.

## 3 Test Images and Their Ground Truths

In total we have used 12 test images. Four of them are generated by scanning four paper drawings in 256 grayscales and then binarizing with moderate thresholds. Their file names are 1.tif, 2.tif, 3.tif, and 4.tif, respectively, as shown in **Fig.1**. The grayscale files of 1.tif and 4.tif are binarized again, but with bigger threshold, resulting thicker line widths in 1_230.tif and 4_230.tif, respectively. The grayscale files of 2.tif and 3.tif are binarized again, but with smaller threshold, resulting thinner line widths, in 2_100.tif and 3_100.tif, respectively. Some synthesized noises defined in [3] are added to 1.tif, 2.tif, 3.tif, and 4.tif, respectively and the results are 1_n4.tif, 2_n4.tif, 3_n4.tif, and 4_n4.tif, respectively, as shown in **Fig. 2**. The ground truth arcs of these images are obtained by manually measuring their geometry parameters. They are displayed in **Fig.1** and **Fig. 2** in gray color over the test images. All ground truth vector files were stored in the VEC format. All these test images and their and ground truth files can be downloaded at the contest website [4].

## 4 Winners and Their Scores

This time, only two participants submitted their systems for the contest: Dave Elliman [5] and Song JiQiang [6]. The scores (*VRI*s) of their systems are listed in Table 1. Song's system obtained an overall score of 0.609 and while Elliman's obtained an overall score of 0.487. So the winner was Song JiQiang.

**Table 1.** The scores of the participants.

| Image (*.tif) | Song JiQiang's scores | | | Dave Elliman's scores | | |
|---|---|---|---|---|---|---|
| | Dv | Fv | VRI | Dv | Fv | VRI |
| 1 | 0.553 | 0.272 | 0.641 | 0.462 | 0.329 | 0.567 |
| 2 | 0.789 | 0.283 | 0.753 | 0.359 | 0.482 | 0.439 |
| 3 | 0.482 | 0.417 | 0.532 | 0.197 | 0.653 | 0.272 |
| 4 | 0.742 | 0.272 | 0.735 | 0.000 | 0.000 | 0.500 |
| 1_230 | 0.553 | 0.273 | 0.640 | 0.481 | 0.302 | 0.589 |
| 2_100 | 0.780 | 0.208 | 0.786 | 0.426 | 0.400 | 0.513 |
| 3_100 | 0.290 | 0.688 | 0.301 | 0.417 | 0.379 | 0.519 |
| 4_230 | 0.803 | 0.223 | 0.790 | 0.188 | 0.543 | 0.323 |
| 1_n4 | 0.446 | 0.428 | 0.509 | 0.488 | 0.159 | 0.664 |
| 2_n4 | 0.700 | 0.294 | 0.703 | 0.506 | 0.282 | 0.612 |
| 3_n4 | 0.199 | 0.750 | 0.224 | 0.327 | 0.425 | 0.451 |
| 4_n4 | 0.680 | 0.305 | 0.688 | 0.283 | 0.485 | 0.399 |
| Average | 0.585 | 0.368 | 0.609 | 0.345 | 0.37 | 0.487 |

## 5 Summary and Discussion

Compared with the previous contest, the test drawings this time are more difficult since there are many arcs tangent and connected to one another. The tangent points are hard to locate precisely, which also have strong impact on detection accuracy of the arc centers and radius. Hence, the overall scores of both participants are lower than the last time. The purpose of using all real life images with some extreme difficulties is to test the algorithms' robustness since almost all algorithms are quite good at segmenting simple isolated arcs/circles, as we saw in the last contest [1].

Again, we find from the contest results that although different approaches may work better on some images but worse on others, the overall performance on these difficult drawings is still not so satisfactory. This can be seen from their scores. After a rough analysis, we may draw a preliminary conclusion that a score of over 0.8 (VRI) may be relatively satisfactory/acceptable.

The impacts of binarization threshold and noise are no so clear since the impacts were not on the same directions. Sometimes, thick line width resulted in higher performance but sometimes, thinner width outperformed. The same cases happened for noises. Especially for Elliman's algorithm, which sometimes generated wrong arc directions, the impacts of noises and binarization threshold were even more difficult to obtain since the results were not trustable.

There is a special case for Elliman's algorithm on image 4.tif, which generated no arcs. Hence both detection rate and false alarm rate are zero. However, the VRI is 0.5, which is higher than many other images. Hence, we think we may have to re-define VRI the geometric average instead of arithmetic average of the two rates. We will try in the future contests.

Anyway, the contest had successfully attracted new participants and re-tested previously tested algorithms. We hope we can attract more participants in future contests and accumulate more and more data for a more comprehensive understanding of arc segmentation algorithms.

## 6 References


1. Liu W., Zhai J., and Dori D.: Extended Summary of the Arc Segmentation Contest. In: Graphics Recognition: Algorithms and Applications, Lecture Notes in Computer Science, volume 2390, Springer (2002), pp. 343-349.
2. Liu W. and Dori D.: A Protocol for Performance Evaluation of Line Detection Algorithms. Machine Vision and Applications 9 (1997) 240-250.
3. Zhai J., Liu W., Dori D., Li Q: A Line Drawings Degradation Model for Performance Characterization. In: Proc. ICDAR (2003), pp. 1020-1024.
4. http://www.cs.cityu.edu.hk/~liuwy/ArcContest/ArcSegContest.htm
5. Elliman D.: The Nottingham Symbol Recognition Algorithm. In: Post-Proc. of GREC2003, Lecture Notes in Computer Science (this volume), Springer (2004).
6. Song J. and Lyu M.R.: A Multi-Resolution Arc Segmentation Method. In: Post-Proc. of GREC2003, Lecture Notes in Computer Science (This volume), Springer (2004).