RANVEC and the Arc Segmentation Contest: Second Presentation

Xavier Hilaire

RANVEC and the Arc Segmentation Contest: Second Presentation

Xavier HILAIRE

LORIA – Université Henri Poincaré
B.P. 239, 54506 Vandœuvre-lès-Nancy, France
hilaire@loria.fr

Abstract. This paper provides some information regarding the winning system at the GREC'2005 contest on arc segmentation. Important facts are first recalled, then the changes made on the system since its first presentation at GREC'2001 are detailed. The obtained results are briefly commented, and the paper finally provides some clues about possible, future improvements regarding the system.

1 Some historical notes

The system mentioned in this paper was first presented during the GREC'2001 workshop, and took part of the arc recognition contest held at the same time in Kingston, Canada. At that time, results from the competition were not really enthousiastic [6]: the average VRI value did not exceed 0.63, many arcs were misdetected, and the presented prototype even crashed on an image.

Although the reader may find it suprising, *the method has almost not changed since that time*; but *its implementation definitely has.*

Indeed, to the exception of what is presented in the next section, the system described in this paper strictly follows what is detailed in [3] and [2]. It has been reimplemented only recently as a 64-bit PowerPC application, and actually runs on any Apple computer fit out with a G5 processor. The main goal of this short paper is therefore to inform the reader that the average results obtained in 2001 are mostly explained by a poor implementation of the method, not by an intrinsic default in the method itself.

2 The changes

Strictly speaking, there has been only two changes made on the vectorization method since its first presentation in [3]: the former concerns the thickness evaluation, while the latter is a revision of the reconstruction procedure, and has almost no incidence on the topic of arc extraction. We briefly describe them in this section.

2.1 Thickness estimation

The first change made concerns the estimation of the thickness. To explain it, let us first recall that a discrete circular ring $\mathcal{R}(x_0, y_0, \rho, w)$ with center (x_0, y_0), radius ρ, and thickness w (all possibly real) is the set of *integer* points (x, y) satisfying

$$(\rho - \frac{w}{2})^2 \leq (x - x_0)^2 + (y - y_0)^2 < (\rho + \frac{w}{2})^2$$

If (x_0, y_0) is known, and the ring is drawn without noise, then finding ρ and w is straightforward. However, in real life we have to cope with noise, which slightly complicates the problem. In [2], it has been proven, using Kanungo's document degradation model [4], that an elementary increment of the thickness of any primitive due to noise was very unlikely. On the other hand, we also know that a labeled skeleton, obtained with the $(3, 4)$-distance transform also gives us a lower estimate of the thickness at any skeletal point, and that the corresponding relative error decreases as the ground truth thickness increases [1]... This rapidly suggests us what to do:

1. Build a set E from the labeled skeleton as follows: for each skeletal point p with $(3, 4)$-DT value v, if p has less than 3 neighbours with value v, then add $v/3$, else add $1/2 + v/3$ to E;
2. Robustly estimate the thickness from E: $\hat{w} = \lfloor 2LMS(E) \rfloor$, where LMS stands for *least median of squares*;
3. Let $\mathcal{I}$ be the source image, $|.|$ denote cardinality, and put

   $$\Delta(\mathcal{X}, \mathcal{Y}) = |\mathcal{X} \cap \mathcal{Y}| - |\mathcal{X} \cap \mathcal{Y}^C|$$

   for any discrete sets $\mathcal{X}$ and $\mathcal{Y}$. If $\Delta(\mathcal{R}(x_0, y_0, \rho, \hat{w}+1), \mathcal{I}) > \Delta(\mathcal{R}(x_0, y_0, \rho, \hat{w}), \mathcal{I})$ then retain $\hat{w} + 1$ as the thickness, else retain $\hat{w}$.

In other words, the above procedure determines a lower bound $\hat{w}$ ot the thickness, and then checks wether it is more interesting to reconstruct the shape using a ring with thickness $\hat{w}$ or with thickness $\hat{w} + 1$.

2.2 Junction reconstruction

Another change made concerns the reconstruction of the junctions. The reconstruction method presented in [3] assumes that the junctions can be reconstructed by first determining the largest possible subsets of primitives with non-empty intersection, and then by connecting these subsets one to each other.

Although this approach is valid at a pixel point of view, it does not always lead to realistic results in practice, and has been replaced by a simplified procedure – which, nevertheless, still lets the problem of the reconstruction open.

This change, however, only deals with junctions, and may have only *very limited* influence on the obtained results when it comes to arc extraction. We therefore do not think it is worth detailing it in the scope of this paper.

3 Parameter setup

An important aspect, often kept silent in the literature, is how to parametrize a given recognition method in order to obtain acceptable results. Although the method commented here uses a reduced number of parameters, we still have to provide values for all of them. Keeping the notations of [3, 2], these parameters are: the thickness f, the noise tolerance m, a lower bound τ for the probability to achieve a correct extraction, and, most important, validity bounds for circular patterns ρ_{min}, θ_{min}, and ρ_{max}.

All parameters were set more or less empirically. For τ, the arbitrary value of 0.9999 was used. On the opposite, setting m was driven by a clue observed in Liu and Dori's evaluation protocol [5]. To summarize this clue, let us simply recall some equations from [5]: on the one side, we have

$$Q_v(c) = (Q_{pt}(c).Q_{od}(c).Q_w(c).Q_{sh}(c).Q_{st}(c))^{\frac{1}{5}} \tag{1}$$

and

$$Q_{pt}(c) = \exp -\frac{d_1(c) + d_2(c)}{W(g)} \tag{2}$$

$$Q_{od}(c) = \exp \frac{-2d_{overlap}(c)}{W(g)} \tag{3}$$

which define the basic quality of a candidate vector against its ground truth g, given their overlapping vector c. On the other side

$$Q_{fr}(k) = \frac{\sqrt{\sum_{g \in G(k)} l(k \cap g)^2}}{\sum_{g \in G(k)} l(k \cap g)} \tag{4}$$

characterizes the fragmentation rate of a given candidate k. Now, consider the two following situations:

(1) We detect a given arc without fragmentation, but with poor accuracy ($d_1(c)+ d_2(c) + 2d_{overlap}(c) \neq 0$);
(2) We detect a given arc with fragmentation $1 : n$, but with good accuracy ($d_1(c) = d_2(c) = 2d_{overlap}(c) = 0$).

Assuming that $Q_w(c) = Q_{sh}(c) = Q_{st}(c) = 1$, from equations 1,2, and 3, we obtain that the penalty in the former situation is

$$\exp -\frac{d_1(c) + d_2(c) + 2d_{overlap}(c)}{5}$$

while that in the latter is $1/\sqrt{n}$ according to equation 4. If we put $\varepsilon = d_1(c)+d_2(c)+2d_{overlap}(c)$, then a glance at table 1 rapidly tells us what happens: situation 2 is more interesting than situation 1 for a majority of cases, especially if we are concerned with thin vectors. Consequently, the m parameter of our

$W(g)$ / ε	1	2	3	4	5
1	0.368	0.607	0.717	0.779	0.819
2	0.135	0.368	0.513	0.607	0.670
3	0.050	0.223	0.368	0.472	0.549
4	0.018	0.135	0.264	0.368	0.449
5	0.007	0.082	0.189	0.287	0.368

n	1	2	3	4	5	6
Q_{fr}	1	0.707	0.577	0.500	0.447	0.408

Table 1. Left: values of $Q_v = \exp -\varepsilon/W(g)$. Right: first values of $Q_{fr} = 1/\sqrt{n}$ assuming a fragmentation ratio of $1 : n$.

method was set to 1, the smallest possible value we can supply to properly extract lines and circles without shifting.

Regarding the circular bounds $\rho_{min}, \theta_{min}, \rho_{max}$, the native implementation of our method offers to set both ρ_{min} and θ_{min} independently. For the purpose of the contest, we used a different version: the condition $\rho \geq \rho_{min} \wedge \theta \geq \theta_{min}$ was replaced by a simple test on length: to be accepted, a circular pattern must have a length of 15 pixels or more – an arbitrary, but common-sense value. We also set ρ_{max} to $\max(w/2, h/2)$, where h and w are the image's dimensions, which means that any circular pattern should always have a supporting circle fully included inside the smallest square image that contains the source image itself.

Finally, f was set automatically, following the estimation procedure detailed in [2], with no prior thin/thick layer separation.

4 A short analysis

Although our system achieved the best overall performance, it is interesting to note that concurrent systems did better in two cases: with image `8.tif` for Elliman system's, and with image `8_rn.tif` for Keysers' system. These images, as well as a rendering of the concurrent solutions, are presented at figure 1.

In both cases, the lack of accuracy of our system is only due to the fact that the default setting $\max(h/2, w/2)$ for the upper bound ρ_{max} was too small. As a result, in image `8.tif`, the largest arc is detected as 5 arcs and one fake segment. In image `8_rn.tif`, addition of noise worsens the situation (as m was set to 1), and this time it is detected mostly as segments. The same result can be observed on image `8_sp.tif`.

It is also a questionable point why other participants did not output *any* line in their solutions. As stated in section 3, even if a solution is fragmented or approximate but close to the ground truth, then better is to output it than keeping silent. For example, our system did not properly recognize the smallest arc in each of the `8*.tif` images, but reported a small segment instead. In image `8_rn.tif`, for example, if we remove this segment in the solution, then the VRI score drops from 0.693 to 0.687. If, furthermore, we remove all the remaining lines, then it drops to 0.675.

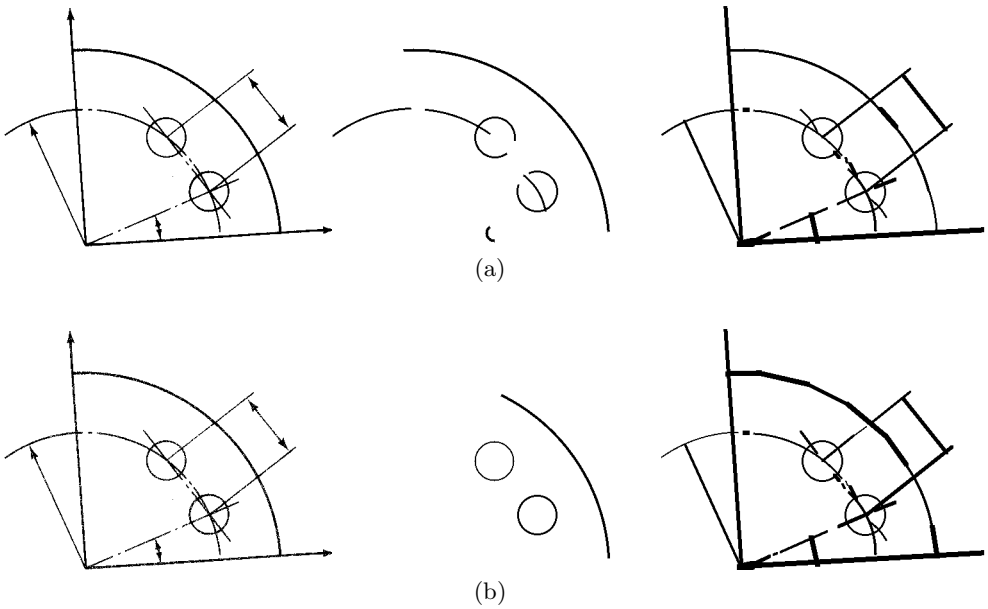

Fig. 1. Comparison of results for two particular images. (a), from left to right: source image `8.tif`, Elliman's result, our result; (b), from left to right: source image `8_rn.tif`, Keysers' result, our result.

Finally, figure 2 illustrates the best case, which occured for image `9.tif`, and leaded to a VRI score of 0.970. The noisy versions `9_rn.tif` and `9_sp.tif` also achieve the best relative performance compared to other images. In this case, the system was well parametrized, and the result typically reflects the level of accuracy the user can expect after some suitable, circular bounds have been provided.

5 Concluding remarks

The system we presented is actually able to extract arcs with an average VRI slightly greater than 0.8. To the best of our knowledge, it is the first time that such a result is reached since the first arc recognition contest, organized in 2001.

Besides, we believe there is still room for enhancement in future versions: although the system achieves optimal parameter estimation once the primitives are identified, the risk that the primitives have not been correctly extracted is still not null. Also, the system relies on skeletonization, and there are obvious situations in which it is still impossible to provide a correct solution given that fact. These are the two tracks currently followed to perfect the system.

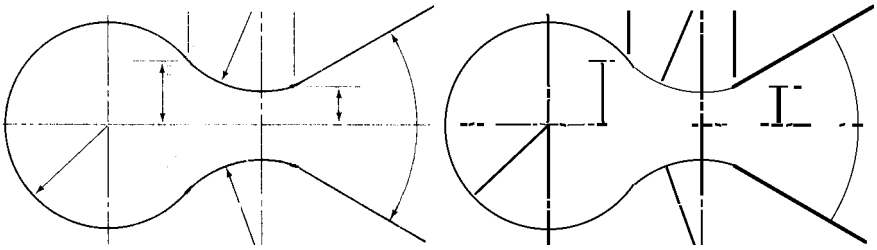

Fig. 2. The best case obtained with our system: (a) source image, (b) recognized arcs.

Acknowledgments

Most of the research presented here has been jointly supported by the french National Agency for Research and Technology (ANRT) and FS2i Corp.[1] under a CIFRE grant.

References

1. G. Borgefors. Distance Transforms in Digital Images. *Computer Vision, Graphics and Image Processing*, 34:344–371, 1986.
2. X. Hilaire. *Segmentation robuste de courbes discrètes 2D et applications à la rétroconversion de documents techniques.* Thèse de doctorat, Institut National Polytechnique de Lorraine, 2004.
3. X. Hilaire and K. Tombre. Improving the Accuracy of Skeleton-Based Vectorization. In D. Blostein and Y.-B. Kwon, editors, *Graphics Recognition – Algorithms and Applications*, volume 2390 of *Lecture Notes in Computer Science*, pages 273–288. Springer-Verlag, 2002.
4. T. Kanungo, R. M. Haralick, H. S. Baird, W. Stuezle, and D. Madigan. A Statistical, Nonparametric Methodology for Document Degradation Model Validation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(11):1209–1223, November 2000.
5. W.Y. Liu and D. Dori. A protocol for performance evaluation of line detection algorithms. *Machine Vision and Applications*, 9(5-6):240–250, 1997.
6. L. Wenyin, J. Zhai, and D. Dori. Extended Summary of the Arc Segmentation Contest. In D. Blostein and Y.-B. Kwon, editors, *Graphics Recognition – Algorithms and Applications*, volume 2390 of *Lecture Notes in Computer Science*, pages 343–349. Springer-Verlag, 2002.

[1] FS2i – www.fs2i.fr – 8 impasse de Toulouse, 78000 Versailles, France.