RANVEC and the Arc Segmentation Contest: Second Evaluation

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 concludes with some clues for possible, future improvements regarding the system.

1   Introduction

1.1   A Short Recall

The system presented in this paper was initially designed to vectorize architectural drawings (with no particular emphasis given on arc detection). It is fully described in [5, 4], and this section only aims to recall a very brief description of it.

An overview of our system is available at Fig. 1. The following steps are involved (in which τ, m, q, ρ_{min}, ρ_{max}, and θ_0 are parameters supplied by the user, see the discussion in section 3 regarding the values used for the contest):

- *Binarization and filtering*, which are optional, preprocessing steps. The filtering step removes all the black connected components whose diameter is lower than f (described below), fills the holes having the same property, and performs a mathematical closure.
- *Text elimination*, which aims at removing text (if any) in the image. This step implements the method of Tombre et al. [8], and is used with default parameters.
- *Thin/thick separation*, which finds the q modes in the thickness histogram that best explain the image in terms of thickness. Each of these layers, with upper estimated thickness f, is then processed independently by the system.
- *Skeletonization*, which computes the (3,4)-distance transform map of the source image, and deduces a skeleton from it following Sanniti di Baja's algorithm [2].
- *Segmentation*, which partitions the skeleton into a set of meaningful primitives (lines and arcs) by resorting to random sampling. The algorithm ensures that all the primitives are correctly found with a probability greater
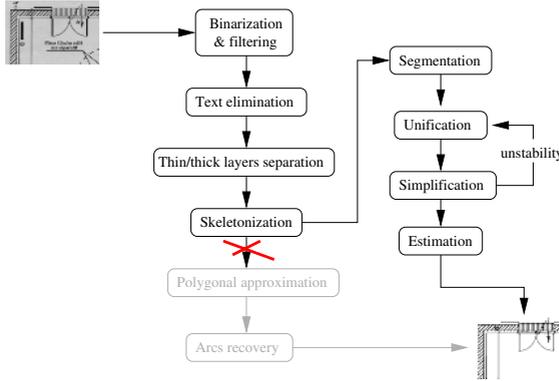
Fig. 1. An overview of the vectorization system presented at the contest

than τ, and that the pixels which define a primitive do not lie away from it further than m units. Circular arcs must also have a radius in the interval $[\rho_{min}, \rho_{max}]$, and an opening angle greater than θ_0.

- *Unification and simplification*, which are applied in turn in a loop, unify or delete primitives in order to simplify the solution. This loop removes spurious arcs or lines (which appeared either because of noise, or because of the skeletonization itself). These steps not only ensure a strict geometric equivalence at a representation point of view, but also permit to compute the preimages of all the detected primitives.
- *Estimation*, which uses the previous preimages to compute optimal estimates of the parameters of each primitive.

1.2   Some Historical Notes

The method described above was evaluated for the first time during the GREC'2001 contest on arc segmentation [3, 9]. At that time, the results were not really enthusiastic: the average VRI value did not exceed 0.63, many arcs were misdetected, and false alarms were also numerous. Indeed, most of these faults are simply explained by an early and poor implementation of the method – the program even crashed on a image, and obtained a score of zero.

To the exception of what is presented in the next section, the system used during the contest this year still follows the description of [5] and [4]. It has been reimplemented only recently[1] as a 64-bit PowerPC application, and runs on any Apple computer equipped with a G5 processor.

2   The Changes

Strictly speaking, there have been only two changes made on the vectorization method since its first presentation in [5]: the first is related to the thickness

---

[1] It was not possible to produce a new implementation in due time for the GREC'2003 contest.

evaluation, while the second is a revision of the reconstruction procedure. In this section, we only discuss the first change: the modification brought to the reconstruction procedure has no influence at all on the result as long as the question of the concurrency of more than two primitives does not arise, which was the case for the contest's images.

To explain the revision brought to the thickness estimation method, 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 complicates the problem. In [4], it has been proven, using Kanungo's document degradation model [6], 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 neighbors with value v, then add v, else add $v + 3/2$ to E;
2. Robustly estimate the thickness from E: $\hat{w} = \lfloor 2LMS(E)/3 \rfloor$-1, 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 $\hat{w}$ of the thickness, and then checks whether it is more interesting to reconstruct the shape using a ring with thickness $\hat{w}$ or with thickness $\hat{w} + 1$.

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 [5, 4], these parameters are: the thickness f, the noise tolerance m, a lower bound τ on 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 [7]. To summarize this clue, let us simply recall some equations from [7]: 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

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

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

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

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

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 less penalizing than situation 1 for a majority of cases, especially if we are concerned with thin vectors. Consequently, the m parameter of our 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 the 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

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 $1 : n$.

$W(g)$	1	2	3	4	5
ε					
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

fully included inside the smallest square image that contains the source image itself.

Finally, f was set automatically, following the estimation procedure detailed in [4], with no prior thin/thick layer separation (q set to 1).

4   A Short Analysis

Although our system achieved the best overall performance, it is interesting to note that the concurrent systems did better in two cases: with image `8.tif` for Elliman's system, and with image `8_rn.tif` for Keysers' system. These images, as well as a rendering of the concurrent solutions, are presented at figure 2.
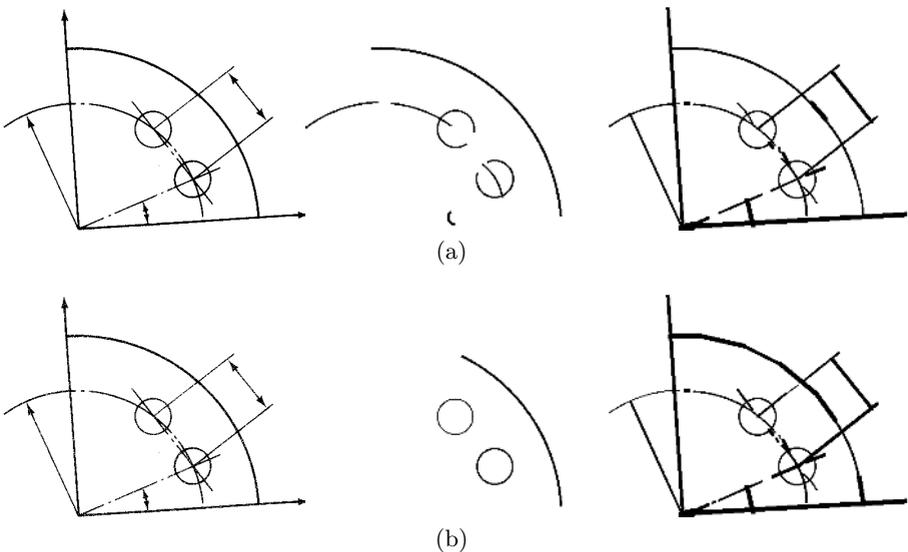

(a)

(b)

Fig. 2. 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.

In both cases, the lack of accuracy of our system is 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` too.

It is also a noticeable point that 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.
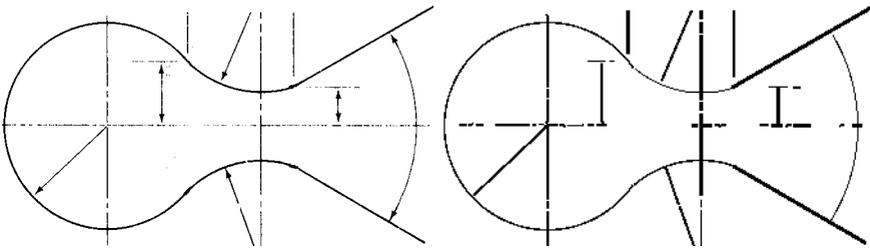

Fig. 3. The best case obtained with our system: (*a*) source image, (*b*) recognized arcs

Finally, figure 3 illustrates the best case, which occurred 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 may 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 it.

Acknowledgments

References

1. G. Borgefors. Distance Transforms in Digital Images. *Computer Vision, Graphics and Image Processing*, 34:344–371, 1986.
2. G. S. di Baja. Well-Shaped, Stable, and Reversible Skeletons from the (3,4)-Distance Transform. *Journal of Visual Communication and Image Representation*, 5(1):107–115, 1994.
3. X. Hilaire. RANVEC and 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 359–364. Springer-Verlag, 2002.
4. 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.
5. 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.
6. 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.
7. 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.
8. K. Tombre, S. Tabbone, L. Pélissier, B. Lamiroy, and Ph. Dosch. Text/Graphics Separation Revisited. In D. Lopresti, J. Hu, and R. Kashi, editors, *Proceedings of 5th IAPR International Workshop on Document Analysis Systems, Princeton (NJ, USA)*, volume 2423 of *Lecture Notes in Computer Science*, pages 200–211. Springer-Verlag, August 2002.
9. 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.

---

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