Decomposition of a curve into arcs and line segments based on dominant point detection

Thanh Phuong Nguyen, Isabelle Debled-Rennesson

Decomposition of a curve into arcs and line segments based on dominant point detection

Thanh Phuong NGUYEN and Isabelle DEBLED-RENNESSON

ADAGIo team, LORIA, UMR 7503, Nancy University
Campus Scientifique - BP 239
54506 Vandoeuvre-ls-Nancy Cedex, France
{nguyentp,debled}@loria.fr

Abstract. A new solution is proposed to decompose a curve into arcs and straight line segments in $O(n \log n)$ time. It is a combined solution based on arc detection [1] and dominant point detection [2] to strengthen the quality of the segmentation results. Experimental results show the fastness of the proposed method.

1 Introduction

An important problem in computer vision is the extraction of meaningful features from image contour for constructing high level descriptors of images. Many existing methods use critical points or straight segments as meaningful features to construct the descriptors. Arc and straight segments are basic objects that appear often in images, specially in graphic document images. A combination of arcs and straight segments is a good solution that avoids the problem in which an arc is approximated by many straight segments or critical points.

Many methods have been proposed for decomposition of a planar curve into arcs and line segments. Rosin et al. [3] constructed firstly a polygonal description and detected fitting arcs by grouping connected lines. Chen et al. [4] proposed a method for segmenting a digital curve into lines and arcs in which the number of primitives is given. This procedure has two stages. The first stage, based on the detection of significant changes on curvature profile, is to obtain a starting set of break points and determine an initial approximation by arcs and lines based on this set of break points. The second stage is an optimization phase that adjusts the break points until the error norm is locally minimized. Horng et al. [5] introduced a curve-fitting method to approximate digital planar curves using lines and arcs based on an approach of dynamic programming. After that, Horng [6] proposed an adaptive smoothing approach for decomposition of a digital curve into arcs and lines. Firstly, a curvature profile is determined by using a Gaussian filter. Then, it is smoothed by using an adaptive smoothing technique. Finally, the input curve is segmented by arcs and lines based on the smoothed curvature representation. Similarly, Salmon et al. [7] proposed a method for decomposition of a curve into arcs and segments based on curvature profile. They used a notion of discrete curvature based on arithmetic discrete lines and blurred segments. The main idea is to construct the curvature profile of the curve and use the extracted key points for reconstruction. Tortorella [8] et al. introduced a method to approximate a curve by arcs and straight segments based

on an approach of dynamic programming. This method works in a transformed domain, called the turning function (see Arkin [9]). Bodansky [10] presented a method for the approximation of a polyline with straight segments, circular arcs and free curves. It contains two steps. The first step is the segmentation of polygonal lines into fragments (short polygonal lines) and the second step is the approximation of the fragments by geometric primitives. If some fragments can not be approximated by geometric primitives with acceptable precision, they are recognized as free curves.

In this paper, we present a novel method for decomposition of a curve into arcs and lines. It is based on a new method for circle detection [1] and dominant point detection [2]. Dominant point detection [2] is used as a preliminary step to extract the critical points of the curve. It has a complexity in $O(n \log n)$ time. So, it possesses a low processing cost. In addition, it is easy and simple to implement. The rest of this paper is organized as follows. The next section recalls a method for dominant point detection. Section 3 presents a method for circle detection. In section 4, we propose a method to split a curve into arcs and straight line segments. Section 5 presents some experimentations and applications to vectorization based on curve reconstruction.

2   Dominant point detection

Dominant points (DP) are local maximum curvature points on a curve that have a rich information content and are sufficient to characterize this curve. We recall hereafter a method of dominant point detection [2] based on an approach of discrete geometry.

2.1   Blurred segment

The notion of blurred segment [11] was introduced from the notion of an arithmetical discrete line. An ***arithmetical discrete line***, noted $D(a, b, \mu, \omega)$, $(a, b, \mu, \omega) \in \mathbb{Z}^4, gcd(a,b) = 1$, is a set of points $(x, y) \in \mathbb{Z}^2$ that satisfies: $\mu \leq ax - by < \mu + \omega$. A ***blurred segment*** (BS) [11] with a main vector (b, a), lower bound μ and thickness ω is a set of integer points (x, y) that is optimally bounded (see [11] for more detail) by a discrete line $D(a, b, \mu, \omega)$. The value $\nu = \frac{\omega - 1}{max(|a|, |b|)}$ is called the width of this BS. Figure 1.a shows a blurred segment (the sequence of gray points) whose the optimal bounding line is $\mathcal{D}(5, 8, -8, 11)$, the vertical distance is 1.25. Nguyen et al. proposed in [12] the notion of maximal blurred segment. A ***maximal blurred segment of width*** ν (MBS) (see figure 1.b) is a width ν blurred segment that can not be extended to the left and the right sides of a given curve. A linear recognition algorithm of width ν blurred segments is described in [11].

2.2   A method for dominant point detection

Nguyen et al. introduced some propositions utilized in [2] to locate and eliminate weak candidates as dominant points (DP). Considering a given width ν, we have:

Proposition 1. *A DP must be in a common zone of successive maximal blurred segments (see figure 2).*

Proposition 2. *The smallest common zone of successive maximal blurred segments whose slopes are monotone contains a candidate of DP (see figure 3.a).*

Proposition 3. *A maximal blurred segment contains a maximum of 2 DP candidates (see figure 3.b).*

Heuristic strategy: *In each smallest zone of successive maximal blurred segments whose slopes are increasing or decreasing, the candidate as dominant point is detected as the middle point of this zone.*

Based on the above study, Nguyen et al. proposed a method for the dominant point detection (see algorithm 1).

Algorithm 1: Dominant point detection [2].

Data: C discrete curve of n points, ν width of the segmentation
Result: D set of extracted dominant points
begin
    Build $MBS_\nu = \{MBS(B_i, E_i, \nu)\}_{i=1}^m$, $\{slope_i\}_{i=1}^m$;
    $m = |MBS_\nu|$; $p = 1$; $q = 1$; $D = \emptyset$;
    while $p \leq m$ **do**
        while $E_q > B_p$ **do** $p + +$;
        Add $(q, p - 1)$ to stack;
        q=p-1;
    while $stack \neq \emptyset$ **do**
        Take (q, p) from stack;
        Decompose $\{slope_q, slope_{q+1}, ..., slope_p\}$ into monotone sequences;
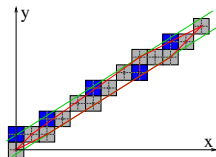   Determine the last monotone sequence $\{slope_r, ..., slope_p\}$;
        Determine the middle point DP of the last monotone sequence
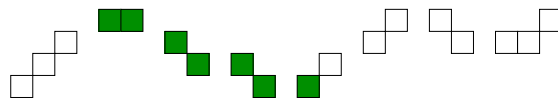 $\{slope_r, ..., slope_p\}$;
        $D = \{D \cup DP\}$;
end

(a) A blurred segment

(b) A maximal blurred segment of width 1 (in dark gray points).
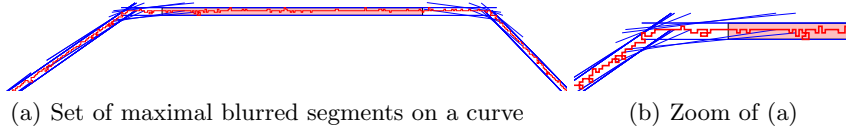
Fig. 1. Blurred segments of width ν

(a) Set of maximal blurred segments on a curve   (b) Zoom of (a)

Fig. 2. Gray zone is not a common zone of successive maximal blurred segments.

(a) Common zone in black

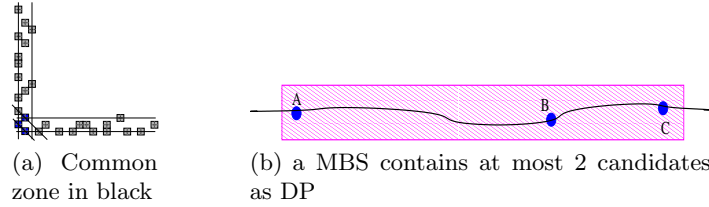(b) a MBS contains at most 2 candidates as DP

Fig. 3. MBS and dominant point

3   Arc detection

In this section, we recall a linear method [1] for the detection of digital arcs. Nguyen and Debled proposed in [1] some properties of arcs in tangent space representation that are inspired from Arkin [9] and Latecki [13].

3.1   Tangent space representation

Let $C = \{C_i\}_{i=0}^n$ be a polygon, l_i - length of segment C_iC_{i+1} and $\alpha_i = \angle(\overrightarrow{C_{i-1}C_i}, \overrightarrow{C_iC_{i+1}})$. If C_{i+1} is on the right of $\overrightarrow{C_{i-1}C_i}$ then $\alpha_i > 0$, otherwise $\alpha_i < 0$.

Let us consider the transformation that associates a polygon C of $\mathbb{Z}^2$ to a polygon of $\mathbb{R}^2$ constituted by segments $T_{i2}T_{(i+1)1}, T_{(i+1)1}T_{(i+1)2}, 0 \leq i < n$ (see figure 4) with:

$T_{02} = (0,0)$,
$T_{i1} = (T_{(i-1)2}.x + l_{i-1}, T_{(i-1)2}.y)$, i from 1 to n,
$T_{i2} = (T_{i1}.x, T_{i1}.y + \alpha_i)$, i from 1 to $n - 1$.

3.2   Properties of arc in the tangent space

Nguyen et al. also proposed in [1] some properties of a set of sequential chords of a circle in the tangent space. They are resumed by proposition 4 (see also figure 5).

Proposition 4. *[1] Let $C = \{C_i\}_{i=0}^n$ be a polygon, $\alpha_i = \angle(\overrightarrow{C_{i-1}C_i}, \overrightarrow{C_iC_{i+1}})$ such that $\alpha_i \leq \alpha_{max} \leq \frac{\pi}{4}$. The length of C_iC_{i+1} is l_i, for $i \in \{1, \ldots, n\}$. We consider the polygon $T(C)$, that corresponds to its representation in the modified tangent space, constituted by the segments $T_{i2}T_{(i+1)1}, T_{(i+1)1}T_{(i+1)2}$ for i from 0 to $n - 1$. $MpC = \{M_i\}_{i=0}^{n-1}$ is the midpoint set of $\{T_{i2}T_{(i+1)1}\}_{i=0}^{n-1}$. So, C is a polygon whose vertices are on a real arc only if $MpC = \{M_i\}_{i=0}^{n-1}$ is a set of quasi collinear points.*

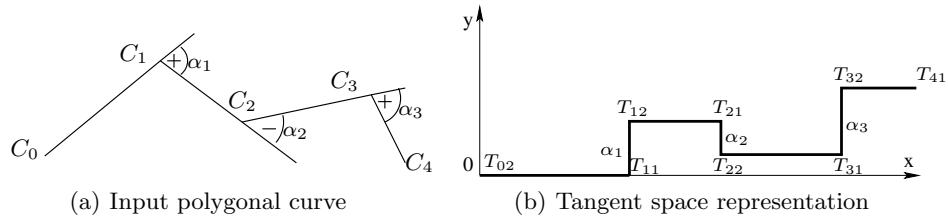From now on, MpC is called the midpoint curve.

(a) Input polygonal curve

(b) Tangent space representation

Fig. 4. Tangent space representation

(a) A set of sequential chords of an arc.

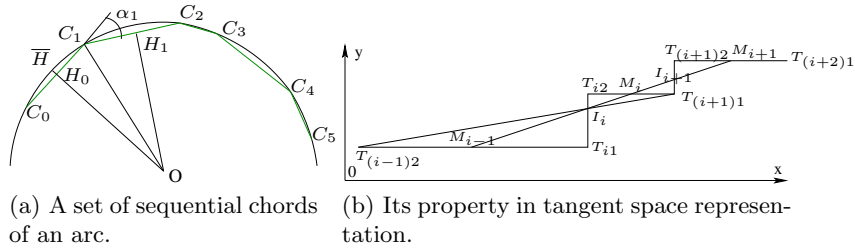(b) Its property in tangent space representation.

Fig. 5. The chords in tangent space.

3.3 Algorithm for arc (circle) detection

Thanks to proposition 4, Nguyen et al. proposed a linear algorithm in [1] (see algo. 2) for the arc/circle detection.

4 Curve decomposition into arcs and lines

Algorithm 2 allows to recognize a digital circle by detecting straight line segment in the tangent space. The parameter α_{max} is used to assure that the hypothesis of theorem 4 is valid. We have the definition below.

Definition 1. *In the curve of midpoints in the tangent space,* **an isolated point** *is a midpoint satisfying that the differences of ordinate values between it and one of its 2 neighboring midpoints on this curve is higher than the threshold α_{max}. If this condition is satisfied with all 2 neighboring midpoints, it is called* **a full isolated point**

4.1 Main idea of the proposed method

We present in this section a new method for curve decomposition into arcs and straight line segments. Our principal idea is to apply a dominant point detector [2] as preprocessing step to enhance the segmentation quality. We assume that the extremities between an arc and a straight segment, or among 2 arcs, or among 2 straight line segments are also dominant points. It is true for almost all cases. Therefore, we detect firstly the dominant points on the input curve C. These points are considered as candidates for extremities between the arcs

---

Algorithm 2: Detection of a digital arc/circle [1].

Data: $C = \{C_i\}_{i=0}^{n}$ digital curve, α_{max} - maximal admissible angle, ν- width of blurred segment

Result: ARC if C is an arc, CIRCLE if C is a circle, FALSE otherwise.

begin

    Use [11] to decompose C with blurred segments of width 1: $P = \{P\}_{i=0}^{m}$;

    Represent P in the modified tangent space by $T(P)$ (see section 3.1);

    if *there exists i such that $T_{i2}.y - T_{i1}.y > \alpha_{max}$* **then return** FALSE;

    Determine the midpoint set $MpC = \{M_i\}_{i=0}^{m-1}$ of $\{T_{i2}T_{(i+1)1}\}_{i=0}^{m-1}$;

    Use the algorithm in [11] to verify if MpC is a blurred segment of width ν;

    if *MpC is a straight line segment* **then**

        if $|M_{m-1}.y - M_0.y| \simeq 2 * \pi$ **then**

            return CIRCLE;

        else return ARC;

    else return FALSE;

end

---

and the straight line segments in a decomposition of C. In the next step, we will group the points if they constitute an arc. The detection of an arc [1] is based on theorem 4 by using algo. 2. Thanks to this algorithm, an arc corresponds to a straight segment on the curve of midpoints (MpC) such that the difference of ordinate values among 2 successive midpoints is less than a threshold α_{max}. The isolated points correspond to straight line segments. This process is done in the tangent space representation of the polygon that is constructed from detected dominant points.

4.2   Analysis of configurations

Let us consider figure 6. In this example, there are all basic configurations among the primitive arc and line: arc-arc, arc-line and line-line. Figure 7 presents these configurations in detail in the tangent space. Concerning the midpoint curve (MpC) in the tangent space, we have several remarks below.

– An isolated point in MpC corresponds to an extremity among two adjacent primitives in C.
– A full isolated point in MpC corresponds to an line segment in C.
– An isolated point in MpC can be co-linear with a set of co-linear points that corresponds to an arc.

Due to the second remark, it is not appropriate to apply directly a polygonalization on the midpoint curve to extract arcs from the input curve.

4.3   Proposed algorithm

Thanks to the above remarks, we present hereafter an algorithm (see algo. 3) to decompose a curve C into arcs and straight line segments. First, the sequence of dominant points (DpC) of C is computed. DpC is then transformed in the tangent space and the MpC curve is constructed. An incremental process is
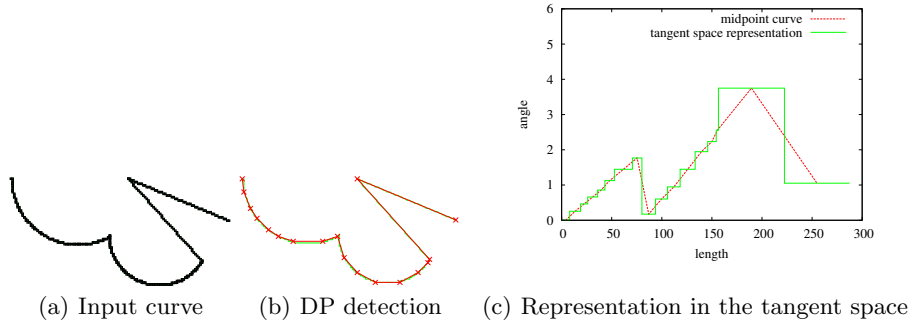
(a) Input curve   (b) DP detection   (c) Representation in the tangent space

Fig. 6. An example of curve

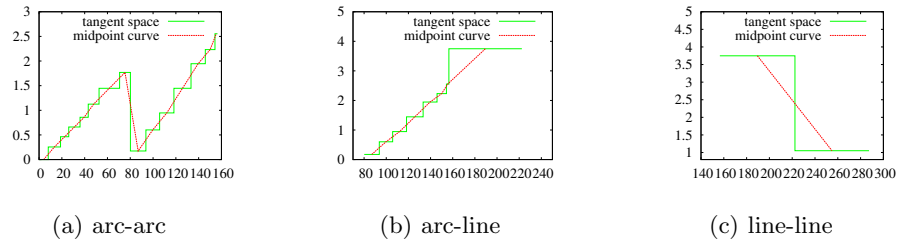(a) arc-arc   (b) arc-line   (c) line-line

Fig. 7. Configurations on tangent space.

then used and each point of MpC is tested: if it is not an isolated point (in this case, it corresponds to a segment in C), the blurred segment recognition algorithm [11] permits to test if it can be added to the current blurred segment (which corresponds to an arc in C). If it is not possible, a new blurred segment starts with this point.

Complexity: As shown in [2], the detection of dominant points can be done in $O(n \log n)$ time. The transform to the tangent space is done in linear time. The recognition process in MpC is also done in linear time [11]. So, the proposed method is done in $O(n \log n)$ time.

5 Experimentations

5.1 Experimental results and comparisons

This method is rapid and simple to implement. Figures 8, 9 and 10 show some experimental results of the proposed methods.

Moreover, figures 11, 12 and table 1 show some comparisons with other methods [5–7]. Salmon et al. [7] proposed a method for the same purpose based on curvature profile that is constructed by using the determination of left and right

---

[1] By default, $\alpha_{max} = \frac{\pi}{4}$, $\nu = 0.2$ (see algo. 3).

Algorithm 3: Curve decomposition into arcs and lines

Data: $C = \{C_1, \ldots, C_n\}$-a digital curve, α_{max}- maximal angle, ν-width of
   blurred segments [1]

Result: $ARCs$- set of arcs, $LINEs$- set of lines

begin
  Use [2] to detect the set of dominant points: $DpC = \{D_0, \ldots, D_m\}$; $BS = \emptyset$;
  Transform DpC in the tangent space and construct the midpoint curve
  $MpC = \{M_i\}_{i=0}^{m-1}$;
  for $i=0$ **to** $m\text{-}1$ **do**
   $C_{b_i} C_{e_i} = \{C_i\}_{b_i}^{e_i}$- part of C wich corresponds to M_i;
   if $\left((|M_i.y - M_{i-1}.y| > \alpha_{max}) \&\& (|M_i.y - M_{i+1}.y| > \alpha_{max}) \right)$ **then**
    Push $C_{b_i} C_{e_i}$ to $LINEs$;
   else
    if $BS \cup M_i$ *is a blurred segment of width* ν *[11]* **then**
    $BS = BS \cup M_i$;
    else
     C'- part of C corresponding to BS;Push C' to $ARCs$;
     if $\left((|M_i.y - M_{i-1}.y| > \alpha_{max}) || (|M_i.y - M_{i+1}.y| > \alpha_{max}) \right)$ **then**
      Push $C_{b_i} C_{e_i}$ to $LINEs$;$BS = \emptyset$;
     else
      $BS = \{M_i\}$;

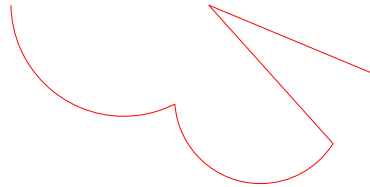**end**

(a) Decomposition into arcs and lines

(b) Vectorization

Fig. 8. Test on the curve in figure 6. Parameters: $\alpha_{max} = \frac{\pi}{4}$, $\nu = 0.2$.
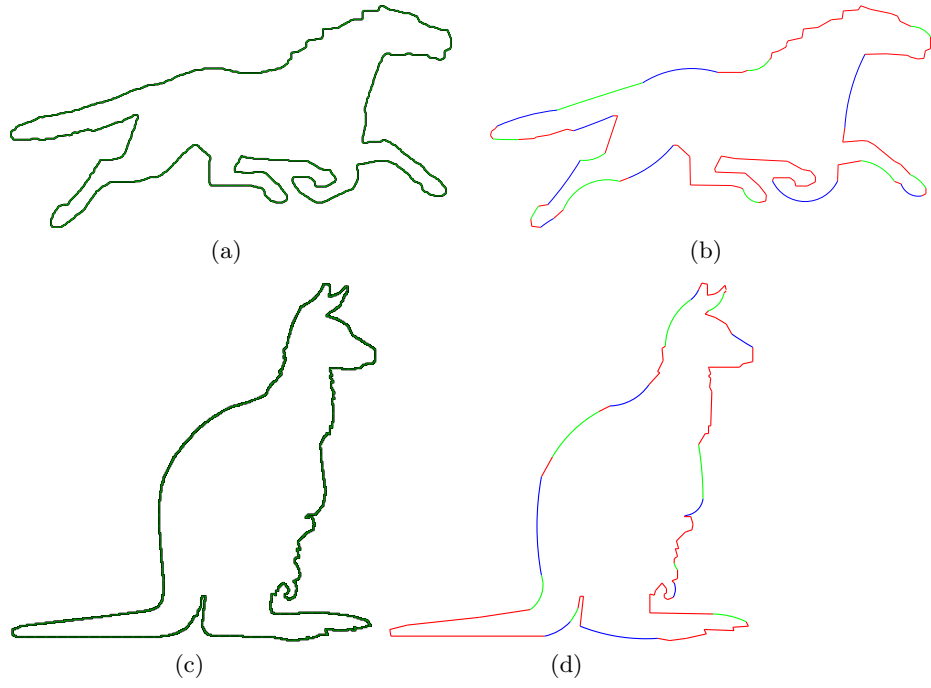
Fig. 9. Curve reconstruction by using the proposed method. (a) (resp. (c)): Input curve, (b) (resp. (d)): Reconstructed curve. Parameters: $\alpha_{max} = \frac{\pi}{4}$, $\nu = 0.2$.

(a) Input image   (b) Extracted edge   (c) Decomposition   (d) Reconstruction

(e) Input image   (f) Extracted contour   (g) Decomposition   (h) Reconstruction
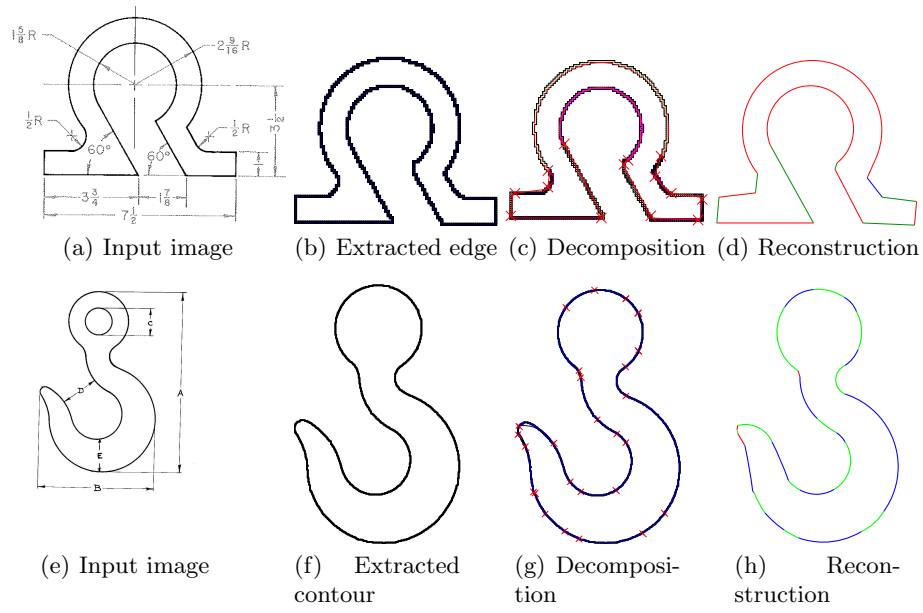
Fig. 10. Test on technical images. Parameters: $\alpha_{max} = \frac{\pi}{4}$, $\nu = 0.2$.

blurred segment at each point in $O(n^2)$ time. So, this method is less efficient than our method. Moreover, the use of filtering and least square fitting methods on the curvature profile cause a distortion in the results of reconstruction. On the contrary, our method is based on a dominant point detector, so the extremities among sequential primitives are well located. Table 1 compares qualitatively the proposed method with other methods. We adapt the criterion of Sarkar [14] that is used in polygonal approximation to calculate the quality of each method. CR is the compression ratio between the number of points and the number of extracted primitives, ISE is the integral square error between the curve and the reconstructed curve and FOM is the ratio between CR and ISE to balance these two aspects. The proposed method is a little less efficient than Horng et al. [5], Horng [6] but it is more rapid than these ones.

5.2 Application to vectorization

Thanks to this method, figure 8.b presents the reconstruction of the curve based on extracted arcs and lines. That is also the main idea for an application to vectorization of a curve based on the reconstruction of the curve from extracted arcs and lines. An arc is constructed simply based on 2 extremities and the middle point of the digital arc. Figure 10.d and figure 9.b, d show some examples of vectorization by using the proposed method with other curves.

Table 1. Comparison with others methods: Horng et al. [5] and Horng [6].

Curve	No of point	Method	No of primitives	ISE	CR	FOM	CPU Time (s)
Fig. 12.b	605	Proposed	22	449.828	27	0.060	0.05
		Horng et al.	15	489.7	40.333	0.0824	1274.75
		Horng	29	329.9	20.862	0.0632	3.23
Fig. 12.c	413	Proposed	22	139.746	22	0.1288	0.03
		Horng et al.	13	175.4	31.769	0.1811	511.77
		Horng	26	107.7	15.885	0.1475	0.94

6 Conclusion

We have presented a new method for decomposition of a curve into arcs and lines in $O(n \log n)$ time. A preprocessing based on dominant point detector [2]
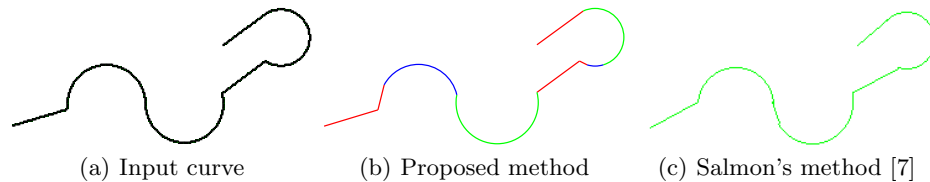

(a) Input curve   (b) Proposed method   (c) Salmon's method [7]

Fig. 11. Comparison with Salmon et al. [7].

(a) Input image

(b) Curve 1

(c) Curve 2

(d) Curve 1, proposed method

(e) Curve 1, Horng et al. [5]

(f) Curve 1, Horng [6]

(g) Curve 2, proposed method

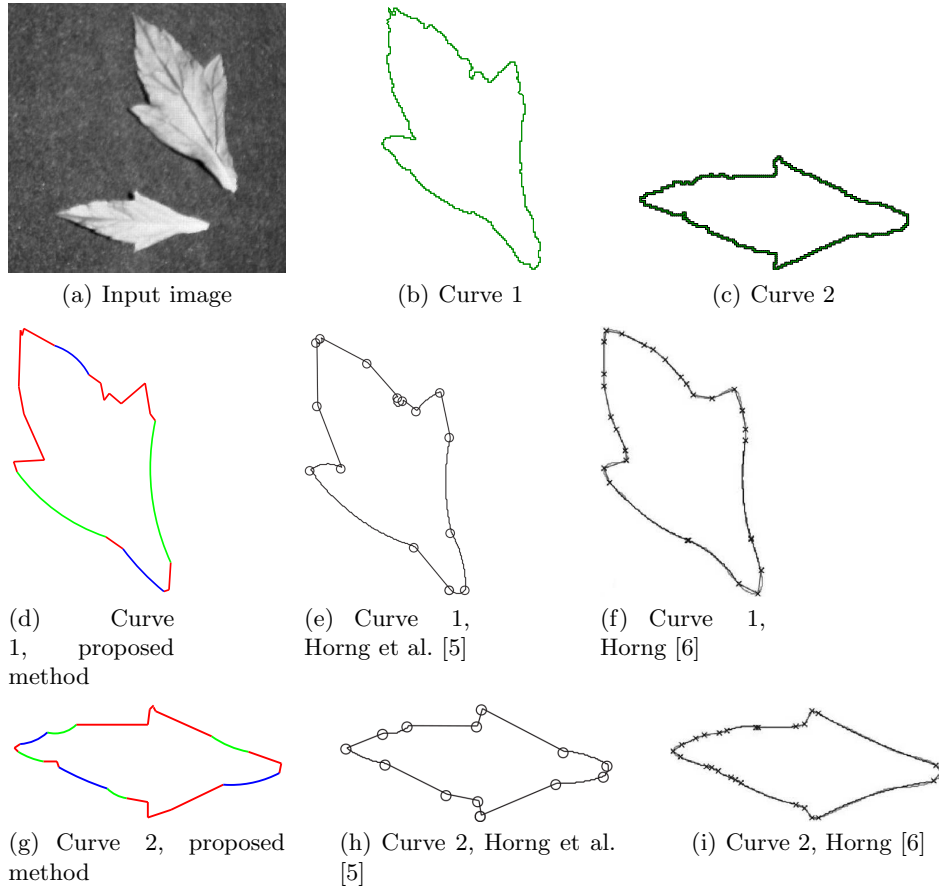(h) Curve 2, Horng et al. [5]

(i) Curve 2, Horng [6]

Fig. 12. Comparison with Horng et al. [5] and Horng [6]. Parameters: $\alpha_{max} = \frac{\pi}{4}$, $\nu = 0.2$.

allows us to locate the extremities among primitives such as lines and arcs well. By detecting the *isolated points*, the arc and line primitives can be located in the input curve. The use of 2 primitives (arc and line) allows us to obtain a good description of curves in relation with other techniques based on corner points and polygonalization.

References

1. Nguyen, T.P., Debled-Rennesson, I.: A linear method for curves segmentation into digital arcs. Technical report, LORIA, Nancy university (2010) http://www.loria.fr/~nguyentp/pubs/arcSegmentation.pdf.
2. Nguyen, T.P., Debled-Rennesson, I.: A discrete geometry approach for dominant point detection. Pattern Recognition **44** (2011) 32–44
3. Rosin, P.L., West, G.A.W.: Segmentation of edges into lines and arcs. Image Vision Comput. **7** (1989) 109–114
4. Chen J.-M., Ventura J.A., W.C.: Segmentation of planar curves into circular and line segments. Image Vision and Computing **14** (1996) 71–83
5. Horng, J.H., Li, J.T.: A dynamic programming approach for fitting digital planar curves with line segments and circular arcs. Pattern Recognition Letters **22** (2001) 183–197
6. Horng, J.H.: An adaptive smoothing approach for fitting digital planar curves with line segments and circular arcs. Pattern Recognition Letters **24** (2003) 565–577
7. Salmon, J.P., Debled-Rennesson, I., Wendling, L.: A new method to detect arcs and segments from curvature profiles. In: ICPR. Volume 3. (2006) 387–390
8. Tortorella, F., Patraccone, R., Molinara, M.: A dynamic programming approach for segmenting digital planar curves into line segments and circular arcs. In: ICPR. (2008) 1–4
9. Arkin, E.M., Chew, L.P., Huttenlocher, D.P., Kedem, K., Mitchell, J.S.B.: An efficiently computable metric for comparing polygonal shapes. PAMI **13** (1991) 209–216
10. E. Bodansky, A.G.: Approximation of a polyline with a sequence of geometric primitives. In: ICIAR. Volume 4142 of LNCS. (2006) 468–478
11. Debled-Rennesson, I., Feschet, F., Rouyer-Degli, J.: Optimal blurred segments decomposition of noisy shapes in linear time. Computers & Graphics **30** (2006) 30–36
12. Nguyen, T.P., Debled-Rennesson, I.: Curvature estimation in noisy curves. In: CAIP. Volume 4673 of LNCS. (2007) 474–481
13. Latecki, L., Lakamper, R.: Shape similarity measure based on correspondence of visual parts. PAMI **22** (2000) 1185–1190
14. Sarkar, D.: A simple algorithm for detection of significant vertices for polygonal approximation of chain-coded curves. Pattern Recognition Letters **14** (1993) 959–964