Fast and Accurate Approximation of the Euclidean Opening Function in Arbitrary Dimension

David Coeurjolly

Université de Lyon, CNRS, LIRIS, UMR5205, F-69622, France
david.coeurjolly@liris.cnrs.fr

Abstract

In this paper, we present a fast and accurate approximation of the Euclidean opening function which is a wide-used tool in morphological mathematics to analyze binary shapes since it allows us to define a local thickness distribution. The proposed algorithm can be defined in arbitrary dimension thanks to the existing techniques to compute the discrete power diagram.

1. Introduction

In many applications the granulometric analysis of binary images is an important task in which morphological mathematics tools pay an important role since it is based on opening computations [4, 6]. When dealing with the Euclidean metric, the Euclidean opening function or thickness distribution [3] can be simply defined as follows: we associate to each grid point x of a shape X the radius of the maximal ball B contained in X with $x \in X$. We are thus facing a geometrical problem in which we will have to deal with the discrete medial axis of a binary shape (see [2] for an overview on Euclidean Medial Axis). Existing algorithms to compute opening functions are based on opening with a structuring element and thus approximate the Euclidean metric with kinds of fixed neighborhood metrics in which unit balls are squares or diamonds for instance [7]. When we consider large binary images in 3-D as in [3] and the Euclidean case, the naive implementation of the Euclidean opening function leads to expensive computational costs. In this paper, we investigate fast approximation algorithms defined in arbitrary dimensions based on a discrete version of the well-known Power Diagram from Computational Geometry [1]. Since such a discrete power diagram mapping can be obtained in optimal time in arbitrary dimension, the proposed opening function algorithm can also be generalized to handle binary objects as subset of $\mathbb{Z}^d$.

2. Preliminaries

The notion of granulometry and granulometric functions was first introduced by Matheron in 1967 [4] in order to study images of porous materials. First of all, a generic definition of granulometry can be sketched as follows [6]: Let $\Phi = (\phi_\lambda)_{\lambda \geq 0}$ be a family of image transformations. Φ is a granulometry if and only if: for all $\lambda, \mu \geq 0$, ϕ_λ is increasing (i), ϕ_λ is anti-extensive (ii), and $\phi_\lambda \phi_\mu = \phi_{\max(\lambda,\mu)}$ (iii). From the granulometry, the opening function can be defined as follows [7]: The *granulometry function or opening function* G_Φ^X of a binary image X for granulometry Φ maps each pixel $x \in X$ to the size of the smallest λ such that $x \notin \phi_\lambda(X)$.

Granulometry can be interpreted as a sieving process with increasing mesh size. From Mathematical Morphology, the set Φ exactly corresponds to a decreasing family of algebraic openings. Furthermore, if the family of openings with the element λB is considered with B convex ($\lambda \in \mathbb{R}^+$), then the resulting set is a granulometry [5]. Dealing with the Mathematical Morphology approach, several authors have investigated fast algorithms to compute such a sequence of openings (see [7, 6] for overviews). In this paper, we consider a granulometry based on the Euclidean metric using perfect increasing disks as structuring elements.

In Digital Geometry, Distance Transformation and Medial Axis (MA for short) extraction are classical problems from decades. Briefly, the Medial Axis can be defined as the set of maximal balls of X: A ball $B \subset X$ is maximal in X if there is no ball $B' \subset X$ such that $B \subset B'$. In [2], we have presented an algorithm to extract the set of Euclidean maximal balls in linear time in arbitrary fixed dimension. More precisely, if $X \subset [0..n]^d$, the Medial Axis is obtained in $O(n^d)$ (see Fig. 1-*a* in 2-D).

We investigate in this paper the granulometry analysis and opening function based on the Euclidean metric [3]. More precisely, we are looking for an algorithm to

construct the opening function G_E^X such that:

$$G_E^X(x) = \max(\{r \mid \forall B(c,r) \subset X, x \in B(c,r)\} \quad (1)$$

where $B(c,r)$ denotes an Euclidean ball with center $c \in X$ and radius r. In other words, we want to associate to each point x in X, the maximal radius of the set of balls inscribed in X that contains x. The set of transformations Φ induced by G_E^X is clearly a granulometry in which openings are Euclidean balls with increasing radii. As discussed in [3], Eq. (1) can be optimized since it is sufficient to only consider MA balls $\{B_i(x_i, r_i)\}$. Indeed, Eq. (1) is equivalent to

$$G_E^X(x) = \max(\{r_i \mid \forall i, x \in B_i(c_i, r_i)\} \quad (2)$$

If we consider the height function $h_i(x)$ as follows

$$h_i(x) = \begin{cases} r_i^2 & \text{if } r_i^2 - \|s_i - x\|^2 > 0 \\ 0 & \text{otherwise} \end{cases}.$$

Eq. (2) is thus equivalent to

$$G_E^X(x) = \sqrt{\max_{B_i(c_i, r_i) \in MA(x)} (h_i(x))} \quad (3)$$

In a computational point of view, the overall computational cost of the algorithm derived from Eq. (2) is $O(n^d + \sum_{B_i \in MA(X)} |B_i|)$. Indeed, $O(n^d)$ is required to obtain the MA, then we have to scan all MA ball grid points and store the maximal radius at each point (see Fig. 2-b). As illustrated in Sect. 5, such a brute-force algorithm is very expensive since many balls of the MA may overlap and thus their associated grid points would be scanned several times.

3. Power Diagram as an Opening Function

In [2], we have also demonstrated links between the MA extraction and the computation of a *Power diagram* (also known as the *Laguerre diagram*) [1]. First, we consider a set of sites $\mathcal{S} = \{c_i\}$ such that each point c_i is associated with a radius r_i. The power $\sigma_i(p)$ of a point p in the plane according to the site c_i is given by $\sigma_i(p) = \|p - c_i\|^2 - r_i^2$. If $\sigma_i(p) < 0$, p belongs to the disk of center v_i and radius r_i. If $\sigma_i(p) > 0$, p is outside the disk. The *power diagram* is a kind of Voronoi diagram based on the metric induced by σ. Hence, the power diagram $\mathcal{V}_\mathcal{S}$ is a decomposition of the plane into cells $F = \{f_i\}$ associated with each site c_i such that $f_i = \{p \in \mathbb{R}^2 : \text{argmin}_i\{\sigma_i(p)\}\}$ [1] (see Figure 2-c).

In [2], we have demonstrate that during the Medial Axis extraction, a Power diagram mapping $\Pi_X(x)$ of MA balls can be obtained without changing the overall computational cost. The mapping associates to each
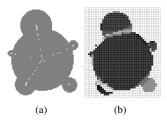

(a)   (b)

Figure 1. $(a) - (b)$ **discrete medial axis and discrete power diagram of a 2-D shape.**

point $x \in X$ the ball $B_i = (c_i, r_i)$ such that x belongs to the open cell f_i. If x belongs to a cell boundary, we arbitrarily choose to associate x to one of the balls of its adjacent cells. If we consider now a power function σ' such that $\sigma_i'(p) = -\sigma_i(p)$, we do not change the power diagram cells geometry and we can interpret the diagram construction as the upper envelope of a set of elliptic paraboloids in 2-D (see [2]).Let us consider the power diagram of balls in MA(X). We define the function G_Π^X such that $G_\Pi^X(x) = r_i$, with $i = \Pi_X(x)$. In other words, $G_\Pi^X(x) = \text{argmax}_{r_i}\{\sigma_i'(x)\}$ with $B_i(c_i, r_i) \in MA(X)$.

Lemma 1 *We have:* (i) $G_\Pi^X(x)$ *is a granulometry function;* (ii) $G_\Pi^X(x) \leq G_E^X(x)$.

Proof: From $G_\Pi^X(x)$, the transformations ϕ_λ are such that $\phi_\lambda(X) = \{x \in X \mid G_\Pi^X(x) > \lambda\}$ with $\lambda \in [r_{min}, r_{max}]$. Hence it clear that functions ϕ_λ satisfy the statements (i), (ii), (iii) in Sect. 2. To prove the second statement, let us consider a ball $B_i(c_i, r_i)$, thus we have for all $x \in X$, $\sigma_i'(x) \leq h_i(x)$. Since $G_\Pi^X(x)$ and $G_E^X(x)$ are obtained at a point x as the ball radius r_i maximizing the sets $\{\sigma_i'(x)\}_i$ and $\{h_i(x)\}$, we obtain that $G_\Pi^X(x) \leq G_E^X(x)$.□

Beside the properties presented in Lemma 1, we can demonstrate that $G_\Pi^X(x)$ is an *intuitive local thickness* measurement on classical shapes as defined in [3]. As detailed above and from [2], an implementation of the $G_\Pi^X(x)$ map computation can be designed in $O(n^d)$ for objects in dimension d. As illustrated in Fig. 2-c, the differences between $G_\Pi^X(x)$ and $G_E^X(x)$ are located on balls intersections. In the following section we provide a quasi-linear but fast algorithm to correct some of these differences.

4. Fast Quasi-linear Approximation

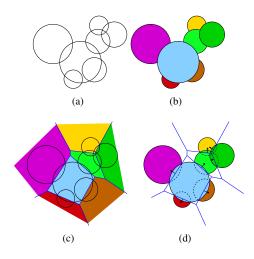As illustrated in Fig. 2-d, the main idea of this algorithm is to start from the $G_\Pi^X(x)$ result and then

Figure 2. Overview of the proposed method: (a) **an input set of maximal balls,** (b) **is Euclidean opening function,** (c) **a first approximation based on the power diagram of balls, and** (d) **the approximation based on active border propagation.**

Alg. 1: Euclidean Opening Function Approx.

Input: a digital object X in dimension d
Output: the map $G^X_{\Pi'}$
Compute $MA(X)$ and Π_X; $G^X_{\Pi'} \leftarrow G^X_\Pi$;
Initialize the $\mathcal{N}(i)$ empty stacks;
foreach *grid point* $x \in X$ **do**
  if x *is an active border* **then**
    $\mathcal{N}(\Pi_X(x)) \leftarrow \mathcal{N}(\Pi_X(x)) \cup \{x\}$

Sort the balls in $MA(X)$ with decreasing radii;
foreach *ball with index i in the sorted list* **do**
  while $\mathcal{N}(i)$ *is not empty* **do**
    $x \leftarrow$ Pop_element($\mathcal{N}(i)$);
    foreach $y \in N(x)$ **do**
      if $\Pi_X(x) = i \neq \Pi_X(y) = j$ **and**
      $y \in B_i(x_i, r_i)$ **and** $r_j < r_i$ **then**
        $G^X_{\Pi'}(y) = r_i$;
        Remove the element y from $\mathcal{N}(j)$;
        Push_element(y,$\mathcal{N}(i)$);

---

to displace the power diagram boundary between two balls taking into account the relative order between their radii. Let us define the *active border* of a discrete power diagram cell i as $\mathcal{N}(i) = \{x \in X \mid \Pi_X(x) = i, \Pi_X(N(x)) \neq i\}$ where $N(x)$ denotes the set of direct neighbors of x in X (4-neighbors in 2-D, 6-neighbors in 3-D...).

The fast approximation algorithm can be sketched as follows (see Alg. 1 for details): we first scan the grid points to detect active borders. Then, we process the active borders of cells with decreasing radii and use a *breath-first* approach to move active border of a cell i with the constraints that we stay inside the ball B_i and that we rewrite values $G^X_{\Pi'}(y)$ only if the radius of B_i is greater than the radius associated to y.

Lemma 2 *Alg. 1 computes the function $G^X_{\Pi'}$ in $O(n^d + b \cdot \log b)$ with b the number of balls in $MA(X)$. Furthermore, $G^X_{\Pi'}$ also satisfies all the statements of Lemma 1.*

Proof: The computational cost of Alg. 1 is quasi-linear since it runs in $O(n^d + b \cdot \log b)$ with b the number of balls in the $MA(X)$ ($b \ll n^d$). Indeed, as discussed above, both $MA(X)$ and G^X_Π are obtained in $O(n^d)$. Then we have to sort the balls in $O(b \cdot \log b)$. Finally, it is clear that the overall size of the active border stacks ($\{\mathcal{N}(i)\}_{i=1...b}$) is bounded by the size of X. In fact, since $\{\mathcal{N}(i)\}_{i=1...b}$ contains only boundary points of Π_X cells, its size is generally of an order on magnitude lower than $|X|$. To conclude the complexity anal-

ysis, we can see that if a point x in $\mathcal{N}(i)$ is popped from the stack i, and since we process the stacks by decreasing radii, the point will not be considered anymore (thanks to the constraint $r_j < r_i$). The proof of Lemma 1's statement on $G^X_{\Pi'}$ is very similar. We just mention that $G^X_{\Pi'}$ is still an approximation of G^X_E (and thus $G'^X_\Pi \leq G^X_E$) as illustrated in the 1-D example depicted in Fig. 3: the discrete power diagram labeling may not intersect the labeling induced by the opening function (pixels associated to the ball b). Hence, during the boundary propagation, the ball a will rewrite the pixels associated to b and thus the active border of b that would have been used to propagate b onto c. Hence when the ball b is processed, no more active border point exists for b and thus we have to process with ball c leading to an incorrect result (at $x = 8$). $\square$
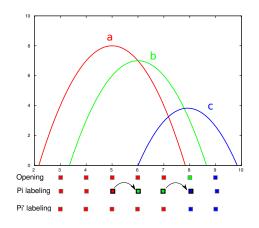

Figure 3. A illustration of 1-D configuration in which $G^X_{\Pi'}$ may contain incorrect results.

5. Experiments and Discussion

In this section we provide a complete experimental evaluation of the proposed algorithms. Table delow details the overall results on a set of digital objects presented in Fig. 4. We have computed the elapsed time for each algorithm and the speed-up of the proposed approximation algorithm. The column MSE indicates the Mean Square Error between the G_E^X distance map and both G_Π^X and $G_{\Pi'}^X$. Also note that the number of grid points considered is equal to n^3.

Object	n	Time (in sec.)			Speed-up (w.r.t. G_E^X)		MSE	
		$G_E(X)$	G_Π^X	$G_{\Pi'}^X$	G_Π^X	$G_{\Pi'}^X$	G_Π^X	$G_{\Pi'}^X$
cat.	40	0.05	0.05	**0.06**	1	1.2	1.85	**0.26**
cat.	200	11.94	4.41	**4.96**	2.71	2.41	75.62	**1.44**
Snow	111	4.22	2.8	**3.64**	1.5	1.16	5.25	**0.34**
Al	450	248.86	56.62	**61.04**	4.4	4.08	65.1	**3.1**

To summarize the table, we can see both G_Π^X and $G_{\Pi'}^X$ outperform G_E^X. Furthermore, the active border propagation step in $G_{\Pi'}^X$ considerably increases the accuracy of the approximation while preserving the time efficiency. In Fig. 5, we have considered the object AL with different increasing resolutions ($n \in \{50, 100, 150, \dots, 450\}$). If we perform a least-square fitting analysis in log-scale space , we can observe that the behavior of the G_E^X curve is in $O(n^4)$ whereas the behavior the $G_{\Pi'}^X$ curve stays in $O(n^3)$. Figure 5 also presents a comparison of the opening function value distributions on the object Al ($n = 450$). Again, $G_{\Pi'}^X$ distribution is an accurate approximation of G_E^X.
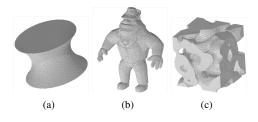

Figure 4. Objects used in the experiments: (a) catenoid, (b) Al, (c) Snow

6. Conclusion

In this paper, we have proposed a fast and accurate approximation of the Euclidean opening function based a geometrical structure from computational geometry (the Power diagram) and digital tools such as the Euclidean Medial Axis and the discrete power diagram. The proposed technique defined in arbitrary dimension
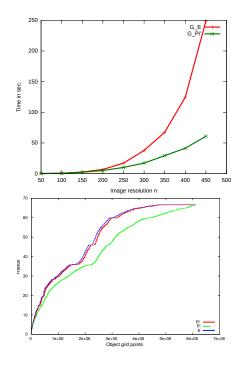

Figure 5. Opening function distribution on Al ($n = 450$) and time efficiency in log-scale on the multi-resolution object Al.

allows us to extract many geometrical parameters on large volumetric data. The intended application of this framework is to use these techniques to efficiently analyze X tomographies of snow sample with resolution up to 2048^3.

References

[1] F. Aurenhammer. Power Diagrams: Properties, Algorithms, and Applications. *SIAM Journal on Computing*, 16:78–96, 1987.

[2] D. Coeurjolly and A. Montanvert. Optimal separable algorithms to compute the reverse euclidean distance transformation and discrete medial axis in arbitrary dimension. *IEEE Transactions on PAMI*, 29(3):437–448, 2007.

[3] T. Hildebrand and P. Ruegsegger. A new method for the model-independent assessment of thickness in three-dimensional images. *Journal of Microscopy*, 185(1):67–75, janvier 1997.

[4] G. Matheron. *Eléments pour une Théorie des Milieux Poreux*. Masson, Paris, 1967.

[5] G. Matheron. *Random sets and integral geometry*. John Wiley & Sons, New York, 1975.

[6] P. Soille. *Morphological Image Analysis*. Springer-Verlag, Berlin, Heidelberg, New York, Etats-unis, 1999.

[7] L. Vincent. *Granulometries and Opening Trees*, pages 57—-90. IOS Press, 40 edition, 2000.