IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received February 17, 2021, accepted April 6, 2021, date of publication April 13, 2021, date of current version May 6, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3072933

Cosine-Pruned Medial Axis: A New Method for Isometric Equivariant and Noise-Free Medial Axis Extraction

DIEGO PATIÑO [1,2] AND JOHN W. BRANCH [2]
[1]School of Engineering and Applied Science, University of Pennsylvania, Philadelphia, PA 19104, USA
[2]Facultad de Minas, Universidad Nacional de Colombia–Sede Medellín, Medellín 050034, Colombia

Corresponding author: Diego Patiño (diegopc@seas.upenn.edu)

This work was supported in part by the Colombian Ministry of Science, Technology and Innovation under Grant 727 of 2015.

ABSTRACT We present the CPMA, a new method for medial axis pruning with noise robustness and equivariance to isometric transformations. The CPMA leverages the discrete cosine transform to create smooth versions of a shape Ω. We use the smooth shapes to compute a score function $\mathcal{F}_\Omega$ that filters out spurious branches from the medial axis of the original shape Ω. Our method generalizes to n-dimensional shapes given the properties of the Discrete Cosine Transform. We extensively compare with state-of-the-art pruning methods to highlight the CPMA's noise robustness and isometric equivariance. We conducted experiments using two 2D datasets — Kimia216 and Animal2000 — and one 3D dataset — the Groningen benchmark. We found that our pruning approach achieves competitive results and yields stable medial axes even in scenarios with significant contour perturbations.

INDEX TERMS Discrete cosine transform, equivariance, isometric transformation, medial axis pruning, morphological skeleton.

I. INTRODUCTION

Shape analysis arises naturally in computer vision applications where geometric information plays an essential role. The shape of an object is a useful tool in fields such as: non-destructive reconstruction of archaeology and cultural heritage [1], [2]; object classification and retrieval from large collections of images [3], [4]; human action and pose recognition for gaming and entertainment [5], [6]; environment sensing in robot navigation and planning [7], [8]; and industry for automatic visual quality inspection of product defects [9], [10].

We visually perceive shape as the collections of all the features that constitute an object. However, to perform computer-based shape analysis, one must rely on an accurate discrete mathematical representation of an object's shape. This representation should exhibit the same geometric and topological properties inherent to the shape itself. Accordingly, we can think of shape representation as a way to store the shape's information in a different format, which benefits speed, compactness, and efficiency.

The associate editor coordinating the review of this manuscript and approving it for publication was Li Zhang.

Many authors have proposed a variety of shape representations such as voxel/pixel grids, point clouds, triangular meshes, medial axis, or signed distance functions [11]–[15]. These representations differ greatly in their formulation, and aim to provide a method for extracting descriptive features from objects, while also preserving their appearance and geometric properties [16]–[20]. However, these methods also have disadvantages that limit their applications. For example, medial axis representations are highly sensitive to contour noise; voxel/pixel grids are inaccurate after isometric transformations; and signed distance functions and triangular meshes are memory-consuming representations when high-frequency details of the shape want to be stored.

We focus this study on the medial axis, also called the topological skeleton. The medial axis represents shapes as a collection of one-dimensional curves that define the ''backbone'' of an object. It provides dimensionality reduction of the amount of data needed to represent an entire shape while preserving its topological structure. Moreover, the medial axis is a rotation equivariant shape representation because the medial axis of a rotated object should ideally be the rotated medial axis of the same object. The medial axis is also robust to small deformation, such as articulation, because

of its graph-like structure. For instance, a human-like shape moving only its arm will not affect all of the points in the medial axis, only the connections between the arm's nodes.

Despite its advantages, medial axis representations are extremely sensitive to noise on the object's contour [21], [22]. Even small amounts of noise can cause erroneous sections of the skeleton called *spurious branches*. Consequently, many medial axis extraction algorithms are equipped with a mechanism to avoid or remove these spurious branches. There are two main mechanisms reported in the state-of-the-art to deal with this problem: 1) prior smoothing of the curve representing the object's boundary, and 2) pruning the spurious branches after the medial axis' computation. In the former, the smoothed boundary is obtained by removing small structures along the curve or surface. It is interesting to note that smoothing curves does not always result in a simplified skeleton [23], [24]. Effective pruning techniques focus instead on criteria to evaluate the significance of individual medial axis branches. However, pruning often requires user-defined parameters that depend on the size and complexity of the object [25]–[27], which makes the pruning method domain-dependent. Moreover, some pruning strategies result in a violation of the equivariant property [21], [22]. As a result, medial axis pruning is still an open problem in computer vision, and this problem is in need of noise-robust methods that concurrently preserve the isometric equivariance of the medial axis.

This paper presents a new method for medial axis pruning that employs mechanisms from the two aforementioned branch-removal strategies. Our method works by computing a controlled set of smoothed versions of the original shape via the discrete cosine transform (DCT). We combine these smoothed shapes' medial axis to create a score function that rates points and branches by their degree of importance. We use our score function to prune spurious branches while preserving the medial axis' ability to reconstruct the original object. Our method is robust to boundary noise and exhibits isometric equivariance.

We benchmark our approach on several datasets of 2D and 3D segmented objects. We use the Kimia216 [28] and Animal datasets [29] to evaluate 2D medial axis extraction. These two datasets provide a method to assess 2D medial axis extraction in the presence of intra-class variation. We also use the University of Groningen Benchmark [30]–[32] to evaluate our approach on 3D objects. Our experiments show that our approach achieves competitive results on isometric equivariance and noise robustness compared to the state-of-the-art.

The main **contributions** of this paper are summarized as follows:

- We define a novel method to compute medial axes that are robust to several degrees of boundary noise without losing the capacity to reconstruct the original object.
- Our computation pipeline guarantees that the isometric equivariance of the medial axis is preserved.
- The definition of our score function allows for a medial axis pruning that is efficiently computed in parallel.

II. RELATED WORK

Many algorithms and strategies exist to extract the medial axis and simplify it when affected by contour noise. This section briefly reviews the most representative algorithms for medial axis computation and discusses their key advantages and disadvantages.

A. THE MEDIAL AXIS

Blum [33] first introduced the medial axis as an analogy of a fire propagating with uniform velocity on a grass field. The field is assumed to have the form of a given shape. If one ''sets fire'' at all boundary points, the medial axis is the set of quench points.

There are other equivalent definitions of the medial axis. In this work we use a geometric definition as follows:

Definition 1 (Medial Axis): *Let Ω be a connected bounded domain in $\mathbb{R}^n$, and x, x' two points such that $x, x' \in \Omega$. The medial axis of Ω is defined as all the points x where x is the center of a maximal ball B_r of radius r that is inscribed inside Ω. Formally,*

$$\mathbf{MA}(\Omega) = \left\{ x \mid B_r(x) \nsubseteq B_{r'}(x'), \forall r' > r \right\}.$$

The medial axis, together with the associated radius of the maximally inscribed ball, is called the medial axis transform ($\mathbf{MAT}(\Omega)$). The **MAT** is a complete shape descriptor, meaning that it can be used to reconstruct the shape of the original domain. In some work, **MA** and **MAT** are also referred to as shape skeletonization. Fig. 1 shows an example of a 2D shape and its medial axis as the center of maximal discs. In $\mathbb{R}^3$ definition 1 may result in a 2-dimensional medial axis sometimes called the medial surface. We will restrict our examples and analysis to only one-dimensional medial axes.

B. MEDIAL AXIS COMPUTATION

There are three primary mechanisms to compute the **MA**: 1) layer by layer morphological erosion or ''thinning methods,'' 2) extraction of the medial axis from the edges of the Voronoi diagram (VD) generated by the boundary points, and 3) detection of ridges in the distance map generated by the boundary points. In digital spaces, only an approximation to the ''true medial axis'' can be extracted.

When using thinning methods [34]–[37], points which belong to Ω are deleted from the outer boundary first. Later, the deletion proceeds iteratively inside until it results in a single-pixel wide medial axis. Medial axis extraction by thinning can be approximated in terms of erosion and opening morphological operations [38]. Thinning methods are easy to implement in a discrete setting, but they are not robust to isometric transformations.

The most well-known algorithm for thinning skeletonization is perhaps the Zhang and Suen [38] algorithm. However, other approaches have been developed using similar principles [36], [37], [39], mainly focused on parallel computation.

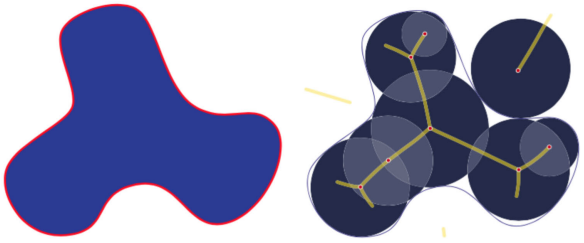


FIGURE 1. Medial Axis Transform Computation. (Left) a shape and its boundary. (Right) Medial Axis elements consisting of the centers and radius of balls inscribed in the shape [7].

Another way to estimate the medial axis works by computing the VD of a polygonal approximation of the object's contour. The contour is expressed as line segments in 2D or a polygonal mesh in 3D. The seed points for the VD are the vertices of the polygonal representation. The medial axis is then computed as the union of all of the edges e_{ij} of the VD, such that the points i, and j are not neighbors in the polygonal approximation [40].

Voronoi skeletonization methods preserve the topology of Ω. However, a suitable polygonal approximation of an object is crucial to guarantee the medial axis' accuracy. Noise in the boundary forms convex areas in the contour, which induce spurious branches on the medial axis. In general, the better the polygonal approximation, the closer the Voronoi skeleton will be to the real **MA**. Nevertheless, this is an expensive process, particularly for large and complex objects [22].

The most common methods to extract the medial axis are those based on the Euclidean distance transform (EDT). Within these methods, the medial axis is computed as the ridges of the EDT inside the object [21], [41]–[46]. This interpretation of the medial axis follows definition 1, because the centers of the maximal balls are located on points x along the ridgeline of the EDT, and the radius of the balls correspond to the distance value at x.

When computed in a discrete framework, distance-based approaches suffer from the same isometric robustness limitations as thinning and Voronoi methods [22]. Moreover, noise in the contour produced by a low discretization resolution directly affects the medial axis' computation by introducing artificial ridges and, consequently, spurious branches.

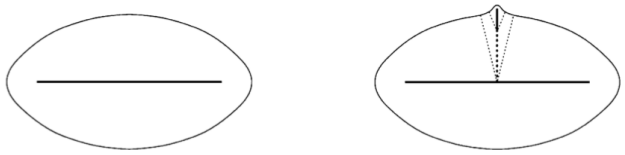


FIGURE 2. (Left) Spurious branch in medial axis. (Right) A new branch appears in the presence of a small perturbation in the contour [25].

C. MEDIAL AXIS SIMPLIFICATION

The medial axis' sensitivity to boundary noise limits its applications to real-life problems [47]. Even negligible boundary noise can cause spurious branches, as shown in Fig. 2.

One strategy for removal of spurious branches consists of computing $\mathbf{MA}(\Omega')$ instead of $\mathbf{MA}(\Omega)$. Here, Ω' is a smoothed version of Ω [48], [49]. The main disadvantage of this approach is that, in most cases, the resulting medial axis is not a good approximation. Additionally, Ω' can potentially change the topology of the original object. Miklos *et al.* [50] introduced a slightly different approach they call Scaled Axis Transform (SAT). The SAT involves scaling the EDT and computing the medial axis of the original un-scaled shape as the medial axis of the scaled one. However, in [45], the authors show that the SAT is not necessarily a subset of the medial axis of the original shape. In another work, [45] propose the Scale Filtered Euclidean Medial Axis (SFEMA), a solution to the stated problem with the SAT that guarantees a better approximation by including additional constraints on the scaled EDT.

Another method to overcome the noise sensitivity limitation of the medial axis is *spurious branch pruning*. Pruning methods are a family of regularization processes incorporated into some medial axis extraction algorithms [25]. Effective pruning techniques focus on different criteria to evaluate the significance of medial axis branches. Thus, the algorithm decides whether to remove the branch (and its points) or not. We can say that pruning methods are adequate if the resulting **MA** is noticeably simplified while preserving the topology and its ability to reconstruct the original object. Most pruning methods rely on *ad hoc* heuristic rules, which are invented and often reinvented in a variety of equivalent application-driven formulations [25]. Some authors apply these rules while computing all medial axis points. Others do so by removing branches that are considered useless after the computation [22], [43], [51].

One of the most popular pruning methods is the Bisector Euclidean Medial Axis (BEMA), which was introduced by Couprie *et al.* [44]. They consider the angle θ formed by a point $x \in \Omega$, and its two closest boundary points denoted by the set $\Pi_\Omega(x)$. This solution removes points from the **MA** for which θ is lower than a constant threshold. This criterion allows different scales within a shape but generally leads to an unconnected medial axis.

Another pruning method found in the state-of-the-art is the work of Hesselink and Roerdink [43]. They introduce the Gamma Integer Medial Axis (GIMA), where a point belongs to the medial axis if the distance between its two closest boundary points is at least equal to γ.

Many pruning methods rely on the EDT, which we will denote as $\mathcal{D}_\Omega(x)$ in the remaining of this document. For these methods, $\mathcal{D}_\Omega(x)$ acts as a generator function for the medial axis, such that points $x \in \mathbf{MA}$ if and only if they satisfy some constraint involving their distance to the boundary. However, other authors have proposed alternative generator functions in their pruning strategies [52], [53].

In [52] and [53], the authors introduce what they call Poisson skeletons by approximating $\mathcal{D}_\Omega(x)$ as the solution of the Poisson equation with constant source function. Poisson skeletons rely on a solid mathematical formulation. Among

other concepts, they use the local minimums and maximums of the curvature along $\delta\Omega$. However, when such methodology is applied in a discrete environment, many spurious branches appear due to the need to define the size of a kernel to estimate these local extreme points.

III. METHOD

We propose a new pruning approach to remove spurious branches from the medial axis of a n-dimensional closed shape Ω. We call our method the Cosine-Pruned Medial Axis (CPMA). The CPMA works by filtering out points from the Euclidean Medial Axis $\mathbf{MAT}(\Omega)$ with a score function $\mathcal{F}_\Omega(x) : \mathbb{N}^n \mapsto [0, 1]$. We define the function $\mathcal{F}_\Omega$ by aggregating the medial axis of controlled smoothed versions of Ω. Our formulation of $\mathcal{F}_\Omega$ must yield high values at points x that belong to the real medial axis (RMA) and low values at points that belong to spurious branches. Additionally, we require $\mathcal{F}_\Omega$ to be equivariant to isometric transformations.

A. THE COSINE-PRUNED MEDIAL AXIS

Let us represent Ω as a square binary image $\mathcal{I} : \mathbb{N}^2 \mapsto \{0, 1\}$ with a resolution of $M \times M$ pixels. We start the computation of the CPMA by estimating a set of smoothed versions of $\mathcal{I}$ via the Discrete Cosine Transform (DCT) and its inverse (IDCT):

$$\mathfrak{F}(u, v) = \frac{C_u C_v}{4} \sum_{x=0}^{M-1} \sum_{y=0}^{M-1} \mathcal{I} \cdot cos\left(u\pi \frac{2x+1}{2M}\right) \cdot cos\left(v\pi \frac{2y+1}{2M}\right) \tag{1}$$

$$\mathcal{I}(x, y) = \sum_{u=0}^{M-1} \sum_{v=0}^{M-1} \frac{C_u C_v}{4} \mathfrak{F} \cdot cos\left(u\pi \frac{2x+1}{2M}\right) \cdot cos\left(v\pi \frac{2y+1}{2M}\right). \tag{2}$$

where (u, v) are coordinates in the frequency domain, and (x, y) are the spatial coordinates of the Euclidean space where Ω is defined. The values of C_u and C_v are determined by:

$$C_u = \begin{cases} \frac{1}{\sqrt{2}} & \text{if } u = 0 \\ 1 & \text{otherwise} \end{cases}$$
$$C_v = \text{(Similar to above)}$$

The DCT is closely related to the discrete Fourier transform of real valued-functions. However, it has better energy compaction properties with only a few of the transform coefficients representing the majority of the energy. Multidimensional variants of the various DCT types follow directly from the one-dimensional definition. They are simply a separable product along each dimension.

Let us now denote by $\hat{\mathcal{I}}^{(i)}$ with $i = 1, 2, \ldots, M$ the reconstructions of $\mathcal{I}$ using only the first i frequencies as per equation 2. We seek to obtain a score function $\mathcal{F}_\Omega$ acting as a sort of probability indicating how likely it is for a point x to be in the medial axis of Ω. Points on relevant branches will appear regularly in the smoothed shapes' medial axis, resulting in high score function values. In contrast, spurious branches will only appear occasionally, resulting in low values.

Definition 2 (Score Function): *Let $\mathcal{I} : \mathbb{N}^n \mapsto \{0, 1\}$ be a square binary image such that $\mathcal{I}(x) = 1\ \forall x \in \Omega$. Let $\hat{\mathcal{I}}^{(i)}$ also be the i-frequency reconstruction of $\mathcal{I}$ via the IDCT. We define $\mathcal{F}_\Omega(x) : \mathbb{N}^n \mapsto [0, 1]$ as the per pixel average over a set of estimations of the* **MAT** *on the smoothed shapes $\mathcal{I}^{(i)}$.*

$$\mathcal{F}_\Omega(x) = \frac{1}{M} \sum_{i=0}^{M-1} \left[\mathbf{MAT}\left(\hat{\mathcal{I}}^{(i)}\right)\right](x). \tag{3}$$

The score function is defined for all x in the domain of $\mathcal{I}$. Notice that we represent the **MAT** as another binary image where $\mathbf{MAT}(x) = 1$ only when x belongs to the medial axis, and 0 otherwise. Using $\mathcal{F}_\Omega$, we finally have all the elements to present our definition of the CPMA.

Definition 3 (Cosine-Pruned Medial Axis): *Given a binary image $\mathcal{I} : \mathbb{N}^n \mapsto \{0, 1\}$ representing a shape Ω, the* **CPMA**(Ω) *consist of all the pairs $(x, r) \in \mathbf{MAT}(\Omega)$ such that $\mathcal{F}_\Omega(x)$ is greater than a threshold τ:*

$$[CPMA(\Omega)](x) = \begin{cases} 1 & \mathcal{F}_\Omega(x) > \tau \\ 0 & otherwise \end{cases}$$

We empirically set the value of the threshold to $\tau = 0.47$. However, we conducted an additional experiment to show that this value is consistent across different shapes.

Although the CPMA results in a noise-free **MA**, there is no restriction in its formulation to force the CPMA to create a connected medial axis. We solve this by first finding all of the disconnected pieces of the CPMA. Later, we connect all of the pieces using a minimum distance criterion $g(x_i, x_j)$, where x_i and x_j are endpoints of two distinct pieces. However, neither the Euclidean distance nor the geodesic distance are suitable criteria because they lead to connections between nodes that do not follow the medial axis (See Fig. 4). We instead connect x_i and x_j with a minimum energy path using an energy function E_Ω. We must guarantee that $E_\Omega(x)$ has high values when x is close to $\delta\Omega$ and low values when x is close to the medial axis. This way, we enforce the paths to be close to the **MAT**. We call the result of this strategy the Connected CPMA (C-CPMA). In section III-C, we provide details for E_Ω computation.

B. ISOMETRIC EQUIVARIANCE OF THE CPMA

The distance transform-based medial axis depends only on the shape Ω, not on the position or size in the embedding Euclidean space. Therefore the medial axis should be equivariant under isometric transformations.

Proposition 1: Let $\mathbf{MAT}(\Omega)$ *be the medial axis transform of a connected bounded domain Ω embedded in $\mathbb{R}^n$, and let $R(x) = Mx + b$ be an isometric transformation in $\mathbb{R}^n$. The square matrix M is a composition of any number or rotations and reflection matrices, and b is an n-dimensional vector. We say that* $\mathbf{MAT}(\Omega)$ *is* ***equivariant*** *to any R such that* $\mathbf{MAT}(R(\Omega)) = R(\mathbf{MAT}(\Omega))$.

Proof: Recall that R is an isometric transformation and thus it is invertible and preserves the Euclidean distance. The previous statement implies that R is an isomorphic map

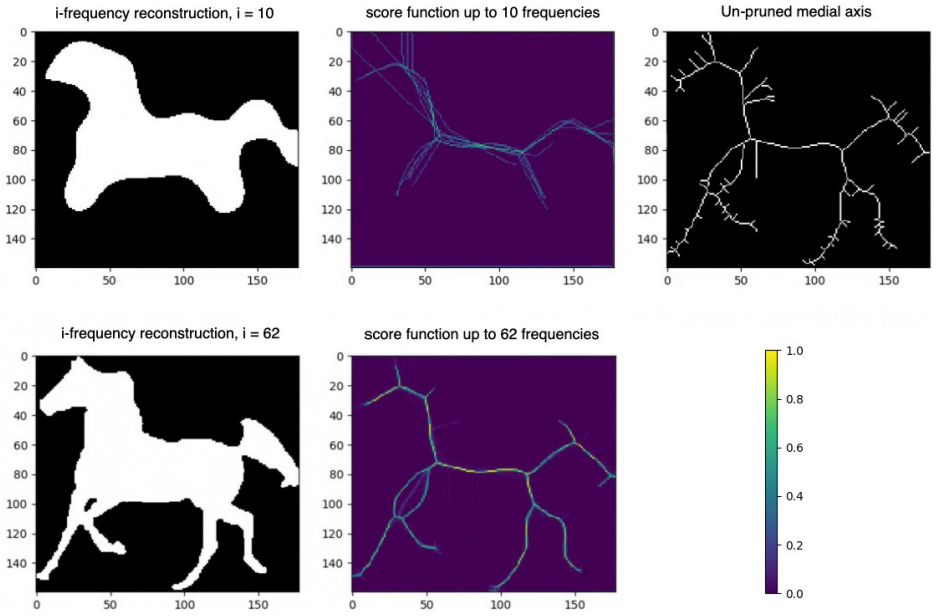


FIGURE 3. Score Function illustrative example. The rows show $\mathcal{F}_\Omega$ of an image of size 180 × 180. We computed the reconstructions up to only $M_1 = 10$ and $M_2 = 62$ of the first frequencies.

between an open ball $B_r(x)$ and $B_r(R(x))$. Consequently, if $B_r(R(x))$ is a maximal ball in Ω then from definition 1 we have that $B_r(R(x)) \nsubseteq B_{r'}(R(x')), \forall r' > r$.

Let us now define $y = R(x)$. Applying R to every element of $\mathbf{MAT}(\Omega)$*, we have*:

$$\begin{aligned} R(\mathbf{MAT}(\Omega)) &= \left\{(R(x), r) \mid B_r(x) \nsubseteq B_{r'}(x'), \forall r' > r\right\} \\ &= \left\{(R(x), r) \mid B_r(R(x)) \nsubseteq B_{r'}(R(x')), \forall r' > r\right\} \\ &= \left\{(y, r) \mid B_r(y) \nsubseteq B_{r'}(y'), \forall r' > r\right\} \\ &= \mathbf{MAT}(R(\Omega)). \end{aligned}$$

■

Moreover, the CPMA depends primarily on $\mathcal{F}_\Omega$, which also holds the isometric equivariant property.

Corollary 1: Let $\mathcal{F}_\Omega$ be the score function of Ω as per definition 3, and let R be an isometric transformation. We say that $\mathcal{F}_\Omega$ is equivariant to any R such that $R(\mathcal{F}_\Omega) = \mathcal{F}_{R(\Omega)}$.

Proof: Using the results from proposition 1 and recalling that R is a linear transformation, we conclude that:

$$\begin{aligned} R(\mathcal{F}_\Omega) &= R\left(\frac{1}{M}\sum_{i=1}^{M}\mathbf{MAT}(\hat{\mathcal{I}}^{(i)})\right) \\ &= \frac{1}{M}\sum_{i=1}^{M} R\left(\mathbf{MAT}(\hat{\mathcal{I}}^{(i)})\right) \\ &= \frac{1}{M}\sum_{i=1}^{M}\mathbf{MAT}\left(R(\hat{\mathcal{I}}^{(i)})\right) = \mathcal{F}_{R(\Omega)}. \end{aligned}$$

■

However, in a discrete domain, this equivariance is only an approximation because points on both Ω and $\mathbf{MAT}(\Omega)$ are constrained to be on a fixed, regular grid. In a continuous domain, it is easy to demonstrate that the cosine transform has exact isometric equivariance.

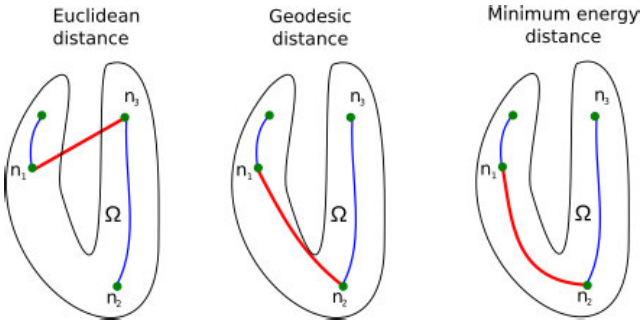


FIGURE 4. Path connectivity between CPMA segments. When using the Euclidean distance (left), two nodes can connect through a path that is not contained within Ω. The geodesic distance (center) guarantees that the path is in Ω, but does not follow the center-line. The minimum energy distance (right) E_Ω is a better alternative to enforce the path to follow the medial axis.

C. IMPLEMENTATION DETAILS

To compute the CPMA enforcing the connectivity, we create a lattice graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. A point p in the domain of $\mathcal{I}$ is a node of $\mathcal{G}$, if and only if $p \in \Omega$. The node p shares an edge with all its neighbors in the lattice only if the neighbors are also inside Ω. We used an 8-connectivity neighborhood in 2D and a 26-connectivity neighborhood in 3D.

In order to determine the minimum energy path between pairs of pixels/voxels, we compute the minimum distance path inside $\mathcal{G}$ using Dijkstra's algorithm. We assign weights to every edge with values extracted from E_Ω. Given $(x, y) \in \mathcal{E}$, we compute the energy of every edge as follows:

$$E_\Omega(x, y) = 1 - \frac{\mathcal{F}_\Omega(x) + \mathcal{F}_\Omega(y)}{2}, \forall (x, y) \in \mathcal{E}.$$

This method guarantees connectivity, but it is inefficient because of the minimum energy path's iterative computation. We sacrifice performance in favor of connectivity. We include the pseudo-code to compute the CPMA and the C-CPMA in algorithms 1 and 2 respectively.

Algorithm 1 Cosine-Pruned Medial Axis (CPMA)

Input:
$\mathcal{I}$: N-dimensional binary array representing the object
M: number of frequencies of $\mathcal{I}$ to be used in the computation
Output:
CPMA: Cosine-Pruned Medial Axis
$\tau \leftarrow 0.47$
$\mathfrak{F} \leftarrow DCT(\mathcal{I})$
$\mathcal{F}_\Omega \leftarrow 0$
$i \leftarrow 1$
while $i < M$ **do**
    `// Reconstructs` $\mathcal{I}$ `using only the first` i `frequencies`
    $\hat{\mathcal{I}}^{(i)} = IDCT(\mathfrak{F}, i)$
    $\mathcal{F}_\Omega = \mathcal{F}_\Omega + \mathbf{MAT}(\hat{\mathcal{I}}^{(i)})$
    $i \leftarrow i + 1$
end
$\mathcal{F}_\Omega \leftarrow \mathcal{F}_\Omega / M$
`// The final` $\mathcal{F}_\Omega$ `is the average of all reconstructions`

CPMA $= \mathcal{F}_\Omega > \tau$
return CPMA, $\mathcal{F}_\Omega$

The CPMA only relies on one parameter, τ. The value of τ is the threshold that determines whether a point of $\mathcal{F}_\Omega$ is a medial axis point. We empirically set the value of $\tau = 0.47$. However, in section V, we present the result of an additional experiment to show how sensitive the CPMA is to different threshold values.

Another essential consideration when computing the CPMA is the maximum frequency used to reconstruct the original shape through the IDCT. Using less than M frequencies enables a faster computation of the CPMA without losing accuracy. We found that using frequencies greater than $\frac{M}{2}$ does not yield significant improvement for the CPMA.

IV. EXPERIMENTS

In this section, we describe the experiments used to evaluate our approach compared to state-of-the-art medial axis pruning methods.

A. COMPARATIVE STUDIES

We chose seven of the most relevant methods in the scientific literature to compare with CPMA extraction results. Each method was selected based on a careful review of the state-of-the-art. These methods illustrate the variety of approaches that authors employ to prune the medial axis. We included the un-pruned **MAT** as our first baseline, in order to see how each method compares to no pruning.

Algorithm 2 Connect Skeleton Segments

Input:
CPMA: Cosine-Pruned Medial Axis representing the object
$\mathcal{F}_\Omega$: Score function
Output:
C-CPMA: Connected Cosine-Pruned Medial Axis
C-CPMA $\leftarrow$ **copy**(CPMA)
skeleton-parts $\leftarrow$ **compute-skeleton-parts**(CPMA)
max-iter $\leftarrow$ 200
it $\leftarrow$ 0
while *it* $<$ *max-iter* ***and*** $|$*skeleton-parts*$|$ > 1 **do**
    graph-i $\leftarrow$ skeleton-parts[0]
    graph-f $\leftarrow$ skeleton-parts[1]
    `// Finds the minimum geodesic path bt. two pieces of the CPMA`
    min-path $\leftarrow$ **find-min-path**(graph-i, graph-f, $\mathcal{F}_\Omega$)
    C-CPMA[min-path] $\leftarrow$ True
    skeleton-parts $\leftarrow$ **compute-skeleton-parts**(C-CPMA)
    it $\leftarrow$ it $+ 1$
end
return C-CPMA

Table 1 summarizes all of the methods in our comparative study. In many cases, the performance of a pruning method depended on its parametrization. We selected parameters for all of the methods following the best performance parametrization reported in the state-of-the-art.

B. DATASETS

We chose three extensively used datasets of 2D and 3D segmented objects to evaluate our methodology on medial axis extraction robustness. These datasets are part of the accepted benchmarks in literature, enabling us to compare our results.

1) Kimia216 DATASET [28]

This dataset consists of 18 classes of different shapes with 12 samples in each class. The dataset's images are a collection of slightly different views of a set of shapes with varying topology. Fig. 5 shows two samples from each class. Contour noise and random rotations are also present in some images in the dataset. Kimia216 has been largely used to test a wide range of medial axis extraction algorithms. Because of the large variety of shapes, we assume that this benchmark ensured a fair comparison with the state-of-the-art.

2) Animal2000 DATASET [29]

The Animal2000 dataset enables us to evaluate the properties of our approach in the presence of non-rigid transformations.

TABLE 1. Pruning methods employed for the comparative study in 2D.

Dimension	Method	Full name	Parameter description
2D	MAT [33]	Medial Axis Transform	N/A
	Thinning [38]	Zhang-Suen Algorithm	N/A
	GIMA [43]	Gamma Integer Medial Axis	γ: minimum distance between $\Pi_\Omega(x)$ and $\Pi_\Omega(y)$, $y \in N_x$, the neighborhood of x.
	BEMA [44]	Bisector Euclidean Medial Axis	θ: angle formed by the point x and the two projections $\Pi_\Omega(x)$ and $\Pi_\Omega(y)$, $y \in N_x$.
	SAT [54]	Scale Axis Transform	$\mathbf{s}$: scale factor to resize $\mathbf{M}AT(\Omega)$.
	SFEMA [45]	Scale Filtered Euclidean Medial Axis	$\mathbf{s}$: scale factor to all balls in the $\mathbf{M}AT(\Omega)$.
	Poisson skel. [53]	Poisson Skeleton	$\mathbf{w}$: window size to find the local maximum of contour curvature.
3D	Thinning [38]	Zhang-Suen Algorithm	N/A
	TEASAR [46]	Tree-structure skeleton extraction	N/A

This table shows the name and parameter description for each method. The point $x \in \mathbf{M}AT$ is an element of the MAT that can be potentially pruned. $\Pi_\Omega(x)$ refers to the set of closest boundary points of x. All the elements of $\Pi_\Omega(x)$ have the same distance from x.

FIGURE 5. Kimia 216 dataset: Two sample shapes from each class.

FIGURE 6. Sample shapes from Animal2000 dataset.

It contains 2000 silhouettes of animals in 20 categories. Each category consists of 100 images of the same type of animal in different poses (Fig. 6). Because silhouettes in Animal2000 come from real images, each class is characterized by a large intra-class variation.

3) UNIVERSITY OF GRONINGEN's BENCHMARK

This set of 3D meshes is commonly found in the literature to evaluate medial axis extraction methods in 3D [30]–[32]. It includes scanned and synthetic shapes taken from other popular datasets. It contains shapes with and without holes, shapes of varying thickness, and shapes with smooth and noisy boundaries. See Fig. 7. All meshes are pre-processed, ensuring a consistent orientation, closeness, non-duplicated vertices, and no degenerate faces.

C. SENSITIVITY TO NOISE AND EQUIVARIANCE ANALYSIS

To compare the robustness of a medial axis extraction method, we adopt an evaluation strategy similar to [32]. Consequently, we measure the similarity between the medial

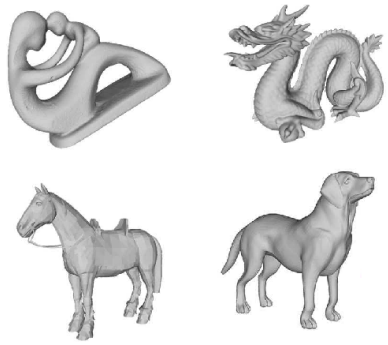


FIGURE 7. Sample shapes form the University of Groningen Benchmark.

axis of a shape Ω and Ω'. The shape Ω' derives from a "perturbation" of Ω. We are interested in evaluating how well our methodology responds to induced noise on the contour/surface. We are also interested in assessing how stable the CPMA is in the presence of rotations of Ω to test its isometric equivariance.

We employ the Hausdorff distance (d_H), and Dubuisson-Jain dissimilarity (d_D) as metrics between shapes. The Dubuisson-Jain similarity is a normalization of the Hausdorff distance [55], which aims to overcome d_H sensitivity to outliers. The Dubuisson-Jain similarity between point sets X and Y is defined as:

$$d_D(X, Y) = max\left\{D(X|Y), D(Y|X)\right\}, \tag{4}$$

with

$$D(X|Y) = \frac{1}{|X|}\sum_{x\in X}\min_{y\in Y}\left\{d(x, y)\right\}. \tag{5}$$

We must first choose a strategy to induce noise to the boundary in order to evaluate the noise sensitivity. We use a stochastic approach denoted by $\mathcal{E}$, where a random number of points $p \in \delta\Omega$ are deformed by a vector v in the direction orthogonal to the boundary, with a deformation magnitude that is normally distributed, $|v| \sim \mathcal{N}(0, 1)$. This noise model is recursively applied n times to every shape in our datasets. We denote as $\mathbf{MAT}(\mathcal{E}(\Omega, k))$ the medial axis of a shape Ω

after applying the noise model k times. In our experiments, we used $k = 1, \ldots, 20$ noise levels.

For every object in each dataset, we compared the medial axis $\mathbf{MAT}(\Omega)$ with the noisy versions, $\mathbf{MAT}(\mathcal{E}(\Omega, k))$, to determine how sensitive a method is to boundary noise.

The medial axis is ideally an isometric equivariant shape representation so that $\mathbf{MAT}(R(\Omega)) = R(\mathbf{MAT}(\Omega))$. Due to sampling factors, this relationship is only an approximation. However, we can measure the equivariance by comparing $\mathbf{MAT}(R(\Omega))$ with $R(\mathbf{MAT}(\Omega))$. The more similar they are, the more equivariant the method.

Because the translation equivariance is trivial, we evaluate isometric equivariance only with rotations of Ω. We do not use reflections because the properties of the rotation matrices we use in this study can be extended to reflection matrices. In our experiments, 2D rotations are counterclockwise in the range $[0, 90]$ degrees around the origin. We use 30 rotations at regular intervals. In 3D, we use a combination of azimuthal ($\theta \in [0, 90]$) and elevation ($\phi \in [0, 90]$) rotations around the origin. The angles θ and ϕ take values at intervals of 18 degrees.

D. RECONSTRUCTION VS SIMPLIFICATION TRADE-OFF

The ideal pruned medial axis should: 1) have a more simplified medial axis with less or no spurious branches, and 2) provide the best reconstruction of the given shape. In practice, there is always a trade-off between these two properties. We aim to estimate this trade-off using a probabilistic approach similar to the one introduced in [56].

Let $\hat{S}_i^k$ be the medial axis of a shape Ω_i computed with a pruning method k from Table 1. We are interested in measuring how accurately $\hat{S}_i^k$ can reconstruct the original shape, while simplifying its un-pruned medial axis, S_i.

To estimate the quality of the reconstruction, we use the area difference between Ω_i and its reconstruction $\Re(\hat{S}_i^k)$. We call this metric the reconstructed area error, and compute it as:

$$RAE(\hat{S}_i^k, \Omega_i) = \frac{\Lambda(\Omega_i) - \Lambda(\Re(\hat{S}_i^k))}{\Lambda(\Omega_i)}, \tag{6}$$

where $\Lambda(.)$ denotes the area in pixels.

Similarly, we measure how much $\hat{S}_i^k$ simplifies the un-pruned medial axis with a metric that depends on the normalized length of $\hat{S}_i^k$. We call this metric the log-simplification ratio, and compute it through:

$$LSR(\hat{S}_i^k, S_i) = log_2\left(\frac{\Gamma(\hat{S}_i^k)}{\Gamma(S_i)} + 1\right), \tag{7}$$

where $\Gamma(.)$ is the total length in pixels of the medial axis' segments.

We evaluate the trade-off between area reconstruction and medial axis simplification by computing *RAE* and *LSR* on all samples per dataset. Later, we use their empirical probability distributions (Eq. 8 and 9) to plot a ROC-like curve at increasing thresholds $t \in [0, 1]$. As it is well-established in ROC curves [57], the closer the curve gets to the upper-left corner, the better the trade-off between the underlying metrics. Notice that we use the left and right tails of $P(RAE)$ and $P(LSR)$, respectively. This is merely a convention to place the perfect trade-off on the upper-left corner, and it does not change the properties of the curves.

$$P(RAE < t) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{RAE_i < t} \tag{8}$$

$$P(LSR > t) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{LSR_i > t} \tag{9}$$

V. RESULTS

In this section, we present and discuss the results of our approach on medial axis pruning. We are evaluating results on two properties: 1) robustness to noise of the contour, and 2) isometric equivariance.

A. STABILITY UNDER BOUNDARY NOISE

We compared the stability of the CPMA under boundary noise against other approaches in Table 1. We conducted our experiments on Kimia216 and the Animal2000 Dataset for 2D images. Additionally, we used a set of 3D triangular meshes from the Groningen Benchmark for experimentation.

For our noise sensitivity experiments, we applied 20 times the noise model $\mathcal{E}(\Omega, k)$ to every object of each dataset. We then computed their **MAT** using every method in Table 1 with different parameters. Finally, each $\mathbf{MAT}(E(\Omega, k))$ was compared with $\mathbf{MAT}(\Omega)$ using both the Hausdorff distance and Dubuisson-Jain dissimilarity. We report the per method average of each metric over all the elements of each dataset.

First, we tested our medial axis pruning method on the Kimia216 dataset, the results of which are presented in Table 2. The table shows that the CPMA and the C-CPMA are competitive against state-of-the-art methods for medial axis extraction. Our results show similar performance to two state-of-the-art methods: the GIMA and SFEMA. The CPMA and C-CPMA also performed better than Poisson Skeletons, SAT, topological thinning, and the un-pruned **MAT**. For visual comparison, Fig. 8 (top row) shows both Hausdorff distance and Dubuisson-Jain dissimilarity against noise level. The figure only displays the best parametrization of every method to improve visualization. As seen here, it is clear that the CPMA and the C-CPMA are among the three best results when we use the Dubuisson-Jain dissimilarity metric. However, we observe a decrease in performance compared to the Hausdorff distance metric. We believe this occurs because of Hausdorff's distance sensitivity to outliers.

The Animal2000 dataset contains nearly ten times more shapes than Kimia216. This leads to more variation between shapes, and therefore a more challenging setting. Table 3 shows similar results compared to Kimia216, confirming that the noise invariant properties of the CPMA still hold in a more robust dataset. The GIMA and the SFEMA are still the

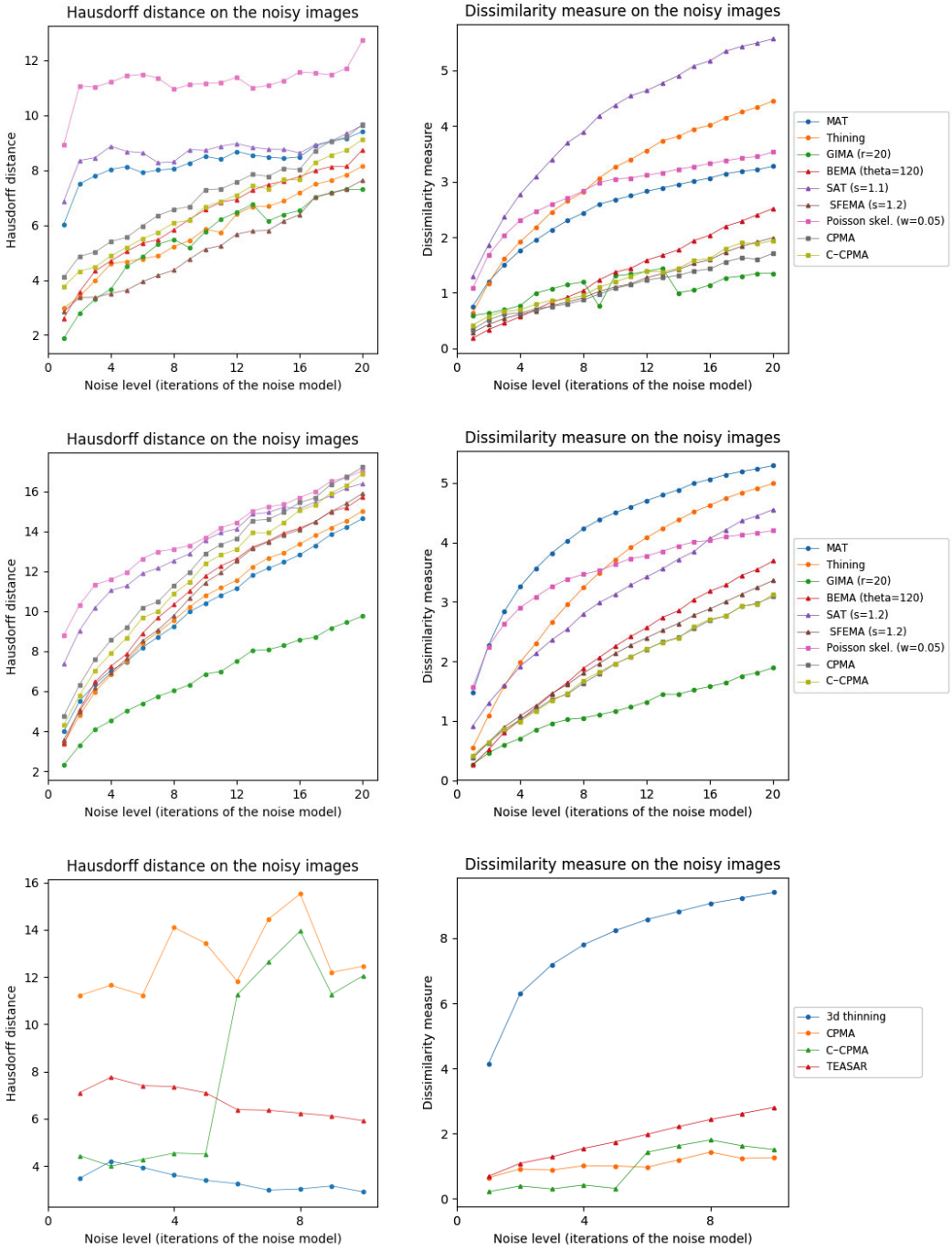


FIGURE 8. Noise sensitivity results on Kimia216 dataset (top), Animal2000 dataset (middle), and Groningen Benchmark (bottom). The figure shows the Hausdorff distance (left) and the Dubuisson-Jain dissimilarity (right) for all of the methods in Tables 2, 3, and 4. Only the best parametrization of each method is depicted for better interpretation.

best methods measured with the Dubuisson-Jain dissimilarity, closely followed by both the CPMA and the C-CPMA. Results of using the Dubuisson-Jain dissimilarity as a metric show that the CPMA is close to methods such as BEMA and SFEMA. However, the results are not as good when we compare using the Hausdorff distance metric. Fig. 8 (middle row) depicts the best performance for every method in comparison to ours. Our experiment's results suggest that the CPMA noise invariant properties generalize across different datasets.

For our 3D experiments, we selected 14 objects from the Groningen Benchmark, reflecting the most common shapes used in the literature. Each object was voxelized to a binary voxel grid with resolution $150 \times 150 \times 150$. This resolution offered sufficient details as well as a sufficiently low computational cost. In contrast to the 2D

TABLE 2. Noise sensitivity results on Kimia216.

Method	Hausdorff				Dissimilarity			
	5	10	15	20	5	10	15	20
MAT	8.13	8.50	8.43	9.41	1.95	2.67	3.01	3.27
Thinning	4.68	5.85	6.88	8.15	2.18	3.26	3.94	4.45
GIMA (r=5)	5.46	6.50	7.37	8.84	0.87	1.31	1.60	1.88
GIMA (r=10)	5.40	7.12	8.35	9.18	**0.68**	1.08	1.35	1.58
GIMA (r=20)	4.49	5.76	6.39	**7.30**	1.00	1.30	**1.05**	**1.35**
BEMA (theta=90)	5.22	6.55	7.11	8.30	0.99	1.60	2.07	2.53
BEMA (theta=120)	5.05	6.56	7.60	8.74	0.70	1.37	1.94	2.52
BEMA (theta=150)	6.68	7.69	7.89	9.40	0.99	1.80	2.50	3.37
SAT (s=1.1)	8.68	8.73	8.76	9.64	3.09	4.37	5.08	5.57
SAT (s=1.2)	9.61	10.05	9.79	10.20	2.50	3.22	3.89	4.47
SFEMA (s=1.1)	4.15	5.35	6.18	7.53	0.84	1.37	1.92	2.50
SFEMA (s=1.2)	**3.64**	**5.13**	**6.15**	7.64	0.68	1.11	1.53	1.99
Poisson skel. (w=0.05)	11.43	11.16	11.26	12.73	2.46	3.05	3.27	3.53
Poisson skel. (w=0.10)	15.60	15.48	16.07	17.35	3.62	4.07	4.19	4.56
Poisson skel. (w=0.20)	17.71	18.02	19.54	21.35	5.00	5.38	5.63	6.20
CPMA	5.55	7.28	8.07	9.66	0.71	**1.07**	1.39	1.71
C-CPMA	5.19	6.68	7.66	9.12	0.80	1.20	1.58	1.94

This table shows the average Hausdorff distance and Dubuisson-Jain dissimilarity for different noise levels (5-20) over each element of the dataset.

TABLE 3. Noise sensitivity results on Animal2000.

Method	Hausdorff				Dissimilarity			
	5	10	15	20	5	10	15	20
MAT	7.47	10.39	12.47	14.64	3.56	4.50	5.00	5.29
Thinning	7.51	10.79	12.94	15.03	2.30	3.71	4.52	4.99
GIMA (r=5)	7.96	11.00	13.23	15.27	1.24	1.88	2.35	2.68
GIMA (r=10)	6.78	8.56	10.21	11.54	0.89	1.29	1.62	1.89
GIMA (r=20)	**5.02**	**6.86**	**8.29**	**9.77**	**0.85**	**1.16**	**1.52**	**1.89**
BEMA (theta=90)	7.84	10.61	12.76	14.78	1.45	2.37	3.06	3.57
BEMA (theta=120)	7.86	11.76	13.93	15.74	1.22	2.25	3.04	3.69
BEMA (theta=150)	8.88	12.38	14.39	16.51	1.68	3.00	3.95	4.72
SAT (s=1.1)	9.44	11.80	13.47	15.21	2.80	4.13	5.02	5.49
SAT (s=1.2)	11.28	13.57	15.21	16.40	2.13	3.13	3.85	4.55
SFEMA (s=1.1)	7.44	11.00	13.30	15.35	1.33	2.27	3.03	3.69
SFEMA (s=1.2)	7.64	11.43	13.83	15.90	1.26	2.13	2.78	3.36
Poisson skel. (w=0.05)	11.94	13.68	15.35	17.03	3.08	3.63	4.01	4.20
Poisson skel. (w=0.10)	14.55	17.11	18.64	20.10	3.55	4.08	4.68	4.94
Poisson skel. (w=0.20)	17.37	20.35	21.67	23.69	3.94	4.69	5.33	5.73
CPMA	9.20	12.88	14.96	17.22	1.18	1.96	2.55	3.09
C-CPMA	8.67	12.39	14.45	16.88	1.17	1.96	2.58	3.13

Noise sensitivity results on Animal2000. This table shows the average Hausdorff distance and Dubuisson-Jain dissimilarity for different noise levels (5-20) over each element of the dataset.

case, we applied $\mathcal{E}(\Omega, k)$ only 10 times to the 3D object. We did this for two reasons: 1) to reduce computational complexity, and 2) because contour noise tends to be more extreme in 3D at the chosen resolution. The results on the Groningen dataset are shown in Table 4 and Fig. 8 (bottom row). Notice that both the CPMA and C-CPMA achieved the best results among the other methods when compared with the dissimilarity measure. These results show that our methodology has noise-invariance properties, and it is stable in the presence of small surface deformation. However, the results show unusual patterns when compared with the Hausdorff distance. In fact, for some methods, the metric decreases when the noise level increases. We attribute this behavior to the outlier sensibility of the Hausdorff distance.

We complete the noise stability analysis showing examples of the **MAT** computed with our methodology in Fig. 9, and compare them to the other methods in this study.

B. SENSITIVITY TO ROTATIONS

We measured the dissimilarity between $\mathbf{MAT}(R(\Omega))$ and $R(\mathbf{MAT}(\Omega))$ across different shapes in the datasets. The lower this dissimilarity, the more stable the method is under rotation. The rotation sensitivity analysis on the Kimia216 dataset is summarized in Table 5 and Fig. 10 (top row). The results show that the curves of the CPMA and the C-CPMA fall near the average of the rest of the methods achieving state-of-the-art performance. The results also surpassed several methods, including Poisson skeleton, SAT, and thinning. Notice that when using the dissimilarity metric the CPMA, the GIMA, the SFEMA, and the BEMA form a subgroup that performs significantly better compared to the others. Moreover, the performance of these methods oscillates around a value of dissimilarity of around 1 pixel on average. The intuition for this result is that regardless of the rotation, skeletons computed with these methods vary only at one pixel. Consequently, we can claim that they exhibit rotation equivariance.

We applied the same analysis to the Animal2000 dataset achieving similar results. In this case, the CPMA and the C-CPMA ranked third and fourth, respectively, among all methods when we used the dissimilarity metric. The results for all methods and parameters are presented in Table 6. As before, we also present a summary with the best parametrization for each method in Fig. 10 (middle row) to facilitate the interpretation. Notice that due to the larger number of objects in the Animal2000 dataset, the curves for every method appear to be smoother, highlighting stability across different rotation angles and shapes.

Finally, we conducted the rotation sensitivity analysis on the 3D dataset. The results are summarized in Fig. 10 (bottom row). The image shows the four 3D medial axis extraction methods we compared in our study for combinations of azimuthal and elevation angles. This figure shows how both the Hausdorff distance and the Dubuisson-Jain dissimilarity become higher when the rotation is more extreme, except in the case of C-CPMA. We believe this behavior is due to the connectivity enforcement mitigating the gaps in the medial axis, and reducing the metrics.

C. RECONSTRUCTION VS SIMPLIFICATION TRADE-OFF

This experiment aims to estimate the trade-off between reconstruction accuracy and medial axis simplification of the methods in Table 1. To evaluate the trade-off, we plot ROC curves, using the empirical probability distributions of the *RAE* and *LSR* metrics defined in section IV-D. For the sake of readability, we use only the best parametrization of each method as we did in previous experiments.

Additionally, we only include methods whose resulting medial axis is a single connected piece. Unconnected medial axes artificially decrease the *LSR* metric because it depends

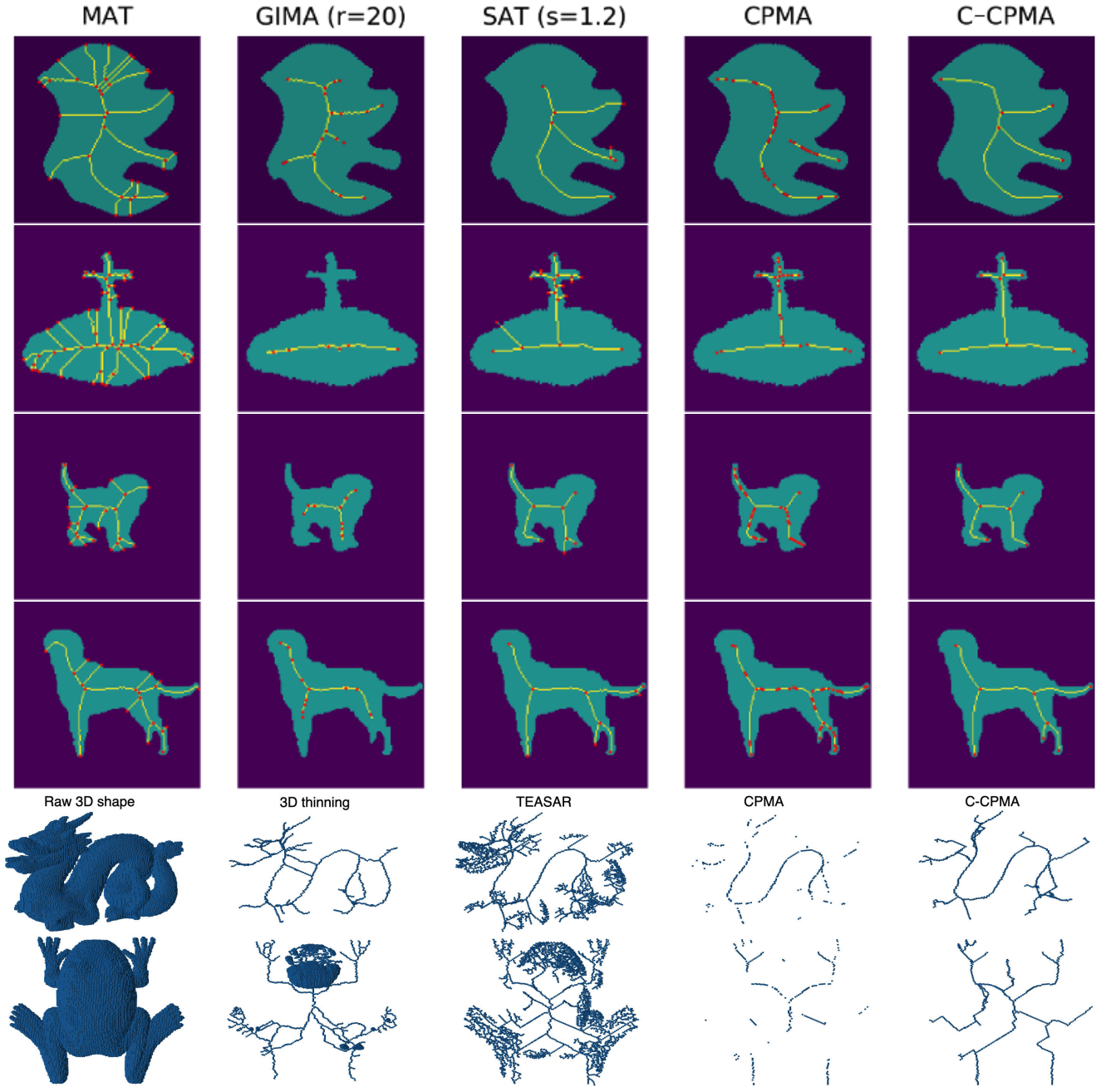


FIGURE 9. The images show the un-pruned MAT and the results of four different pruning methods. Rows one and two are objects from Kimia216 dataset. Rows three and four are objects from Animal2000. Rows five and six are objects from the Groningen Benchmark. Notice how the CPMA and the C-CPMA yield medial axes with less spurious branches while preserving the topology.

TABLE 4. Noise sensitivity results on Groningen Benchmark.

Method	Hausdorff					Dissimilarity				
	2	4	6	8	10	2	4	6	8	10
3D thinning	4.20	**3.61**	**3.25**	**3.03**	**2.90**	6.30	7.80	8.58	9.07	9.41
TEASAR	7.76	7.36	6.39	6.24	5.92	1.09	1.55	1.98	2.44	2.80
CPMA	11.66	14.10	11.83	15.52	12.46	0.91	1.02	**0.97**	**1.44**	**1.27**
C-CPMA	**3.99**	4.54	11.25	13.95	12.06	**0.40**	**0.43**	1.43	1.81	1.51

This table shows the average Hausdorff distance and Dubuisson-Jain dissimilarity for different noise levels (5-20) over each element of the dataset.

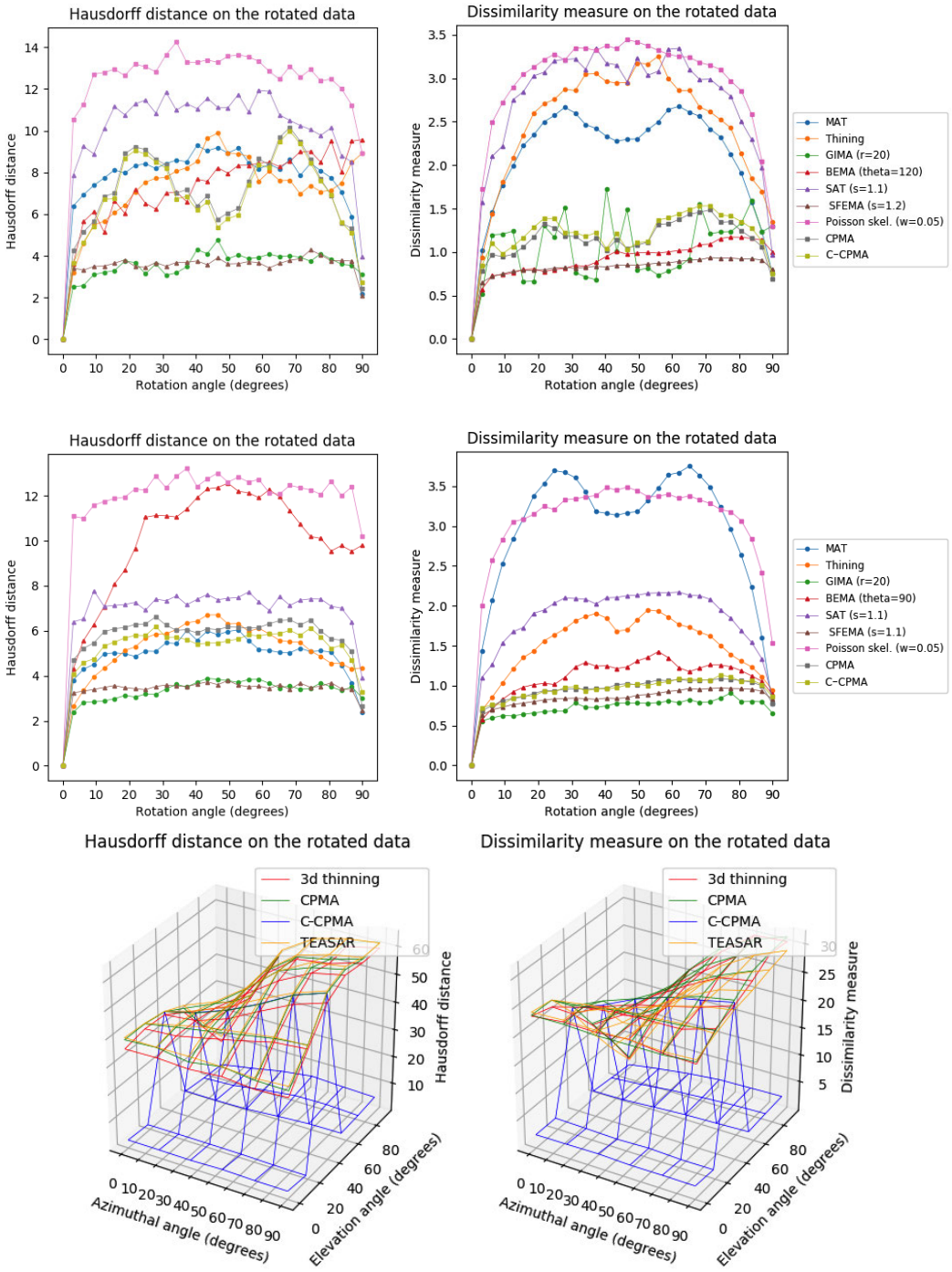
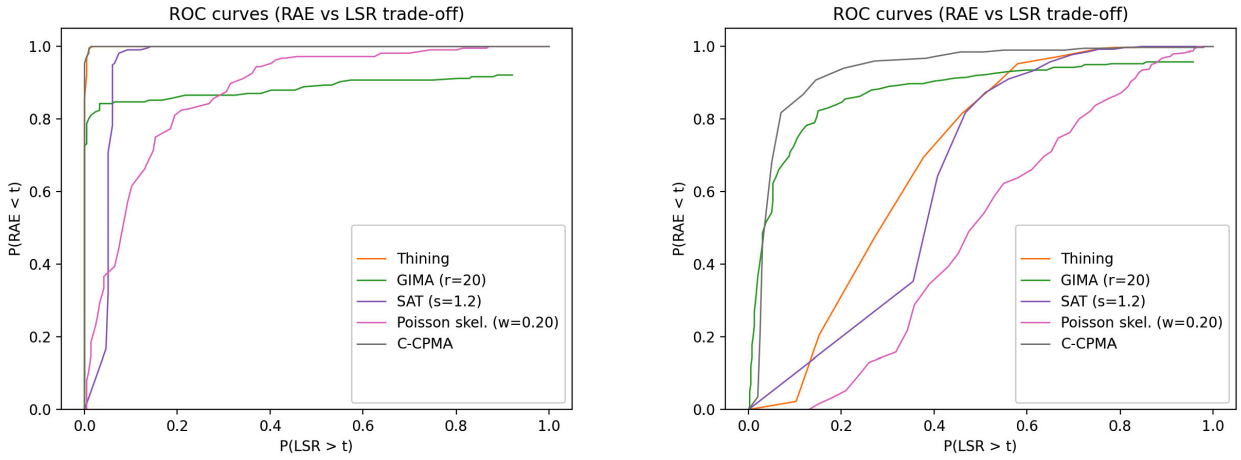


FIGURE 10. Rotation equivariance results on Kimia216 dataset (top), Animal2000 (middle), and Groningen Benchmark (bottom). The top row shows the Hausdorff distance and Dubuisson-Jain dissimilarity for all the methods in Tables 5 and 6.

on the total medial axis' length. Unconnected medial axes, therefore, affect the trade-off and result in biased ROC curves.

Consequently, we exclude SFEMA, BEMA, and CPMA from our trade-off experiments. In contrast, we can include the C-CPMA because of its connectivity enforcing strategy. The energy function we use to enforce connectivity in the C-CPMA is a result of the score function's formulation from the CPMA definition. Note that such energy function cannot be obtained from the SFEMA and the BEMA definitions. Fig. 12 provides a qualitative comparison of unconnected medial axis pruning methods with the C-CPMA. We also exclude the Groningen dataset because most 3D methods only provide the **MA** and not the **MAT**, necessary for the object's reconstruction.

We summarize our trade-off experiments in Fig. 11. The vertical axis shows the ratio of shapes whose reconstruction



(a) Kimia216 dataset **(b)** Animal2000 dataset

FIGURE 11. ROC curves for *RAE* and *LSR* trade-off. The trade-off performance improves as it is closer to the upper-left corner, where the number of objects with good reconstruction is high, while the number of objects with poor medial axis simplification is low.

TABLE 5. Rotation equivariance results on Kimia216.

Method	Hausdorff			Dissimilarity		
	30°	60°	90°	30°	60°	90°
MAT	8.18	8.17	2.17	2.67	2.64	0.75
Thinning	7.72	7.58	8.92	2.87	2.99	1.35
GIMA (r=5)	6.16	6.03	5.54	1.02	1.11	0.85
GIMA (r=10)	5.54	6.25	5.04	0.83	1.02	0.72
GIMA (r=20)	**3.62**	3.93	3.12	1.51	**0.78**	1.30
BEMA (theta=90)	12.35	12.76	11.72	1.31	1.60	1.06
BEMA (theta=120)	6.24	8.44	9.57	0.81	1.00	1.00
BEMA (theta=150)	9.14	10.09	10.27	1.11	1.35	1.36
SAT (s=1.1)	10.84	11.93	3.96	3.22	3.34	0.97
SAT (s=1.2)	11.44	12.40	4.45	2.64	2.87	0.97
SFEMA (s=1.1)	3.86	3.98	2.52	0.92	1.01	0.83
SFEMA (s=1.2)	3.68	**3.66**	**2.08**	**0.82**	0.88	0.80
Poisson skel. (w=0.05)	12.83	13.32	8.93	3.21	3.28	1.30
Poisson skel. (w=0.10)	16.36	17.03	10.17	4.24	4.30	1.78
Poisson skel. (w=0.20)	18.94	19.90	9.65	5.61	5.66	2.69
CPMA	8.63	8.65	2.42	1.18	1.33	**0.70**
C-CPMA	8.51	8.36	2.72	1.22	1.39	0.75

Rotation equivariance results on Kimia216. This table shows the average Hausdorff distance and Dubuisson-Jain dissimilarity for different rotations of each element in the dataset.

TABLE 6. Rotation equivariance results on Animal2000.

Method	Hausdorff			Dissimilarity		
	30°	60°	90°	30°	60°	90°
MAT	5.08	5.17	2.38	3.67	3.64	0.77
Thinning	5.85	6.11	4.35	1.71	1.86	0.94
GIMA (r=5)	5.58	5.54	5.62	0.96	1.08	0.83
GIMA (r=10)	5.42	5.50	4.29	0.78	0.88	0.74
GIMA (r=20)	**3.17**	3.84	3.02	**0.68**	**0.81**	**0.66**
BEMA (theta=90)	11.12	11.92	9.80	1.10	1.35	0.89
BEMA (theta=120)	5.45	6.10	6.80	0.77	0.90	0.86
BEMA (theta=150)	7.60	8.88	9.91	1.13	1.36	1.40
SAT (s=1.1)	7.41	7.27	3.90	2.10	2.16	0.86
SAT (s=1.2)	9.82	9.62	4.85	1.77	1.90	0.95
SFEMA (s=1.1)	3.52	**3.54**	2.45	0.84	0.93	0.83
SFEMA (s=1.2)	3.54	3.63	**2.26**	0.77	0.87	0.82
Poisson skel. (w=0.05)	12.88	12.72	10.18	3.33	3.40	1.54
Poisson skel. (w=0.10)	15.02	15.41	10.58	3.67	3.89	1.89
Poisson skel. (w=0.20)	16.83	17.20	8.67	4.07	4.33	1.85
CPMA	6.62	6.14	2.64	0.96	1.06	0.77
C-CPMA	6.19	5.77	3.30	0.98	1.05	0.86

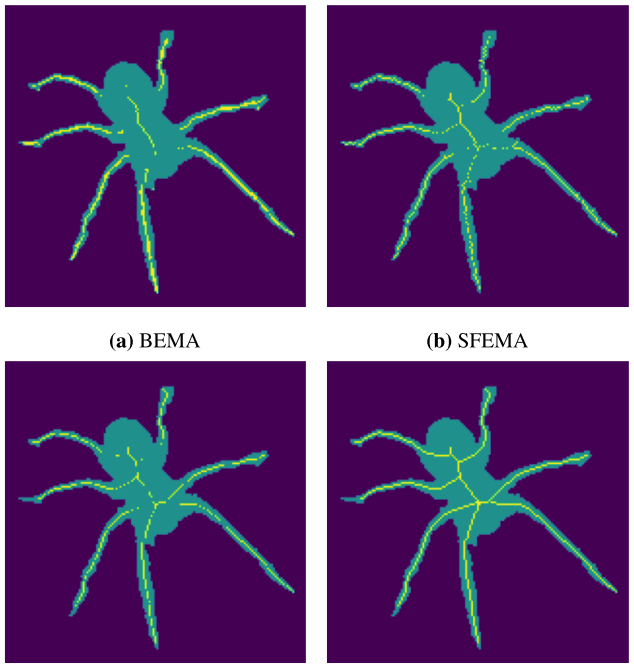
Rotation equivariance results on Animal2000. This table shows the average Hausdorff distance and Dubuisson-Jain dissimilarity for different rotations of each element in the dataset.

area error is less than a threshold t. The horizontal axis shows the ratio of shapes whose normalized length is greater than the threshold. Notice that we can compare both metrics with the same threshold because their range is bounded by the interval $[0, 1]$. As it is well-established in ROC curves [57] the trade-off is better as it approaches the upper-left corner of the graph. The curves in Fig. 11 show how the C-CPMA performs better, surpassing all methods in the comparative study. This result follows from Fig. 9 where it is clear that methods like GIMA do not achieve a satisfactory reconstruction, as it oversimplifies the medial axis. Moreover, Thinning, Poisson Skeleton, and SAT have lower reconstruction rates and tend to leave more spurious branches; this explains their ROC curves' shift to the right.

D. HYPER-PARAMETER SELECTION

Many medial axis pruning methods depend on hyper-parameters to accurately estimate the medial axis [43]–[45], [54]. These parameters usually have a physical meaning in the context of the object whose medial axis they seek to estimate. Often, the parameters are distances or angles formed between points inside the object. In other work, some authors create score functions like ours, intending to use its values as a filter parameter to remove points on spurious branches of the **MAT**. In most cases, however, such parameters are subject to factors like resolution or scale. Thus, we conducted another experiment to test the sensitivity of the CPMA to the pruning parameter τ at different scale factors of the input object.

Fig. 13 shows the Jaccard index's average, also known as Intersection over Union, as the reconstruction metric vs.



(a) BEMA (b) SFEMA

(c) CPMA (d) C-CPMA

FIGURE 12. Unconnected medial axes from the methods not included in trade-off experiments. The second row shows the CPMA and C-CPMA for visual comparison.

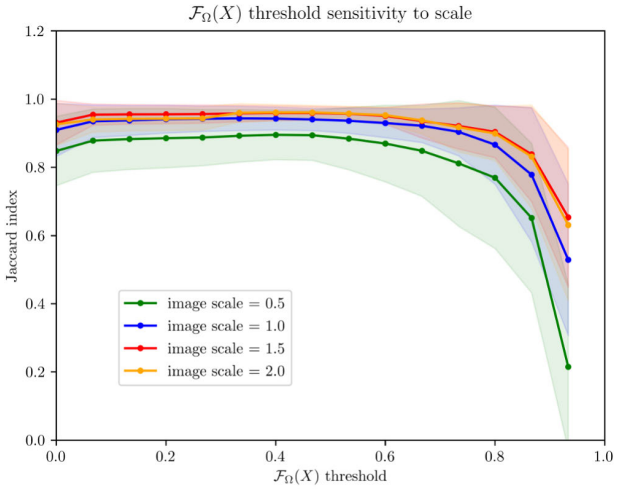


FIGURE 13. Sensitivity analysis of threshold τ to different scales of the input images. The graph shows the average Jaccard index of the reconstructed shape with respect to the original object for the CPMA computed with different values of τ. Higher values of the threshold lead to less spurious branches. We also show the standard deviation error bands.

the values of τ. We compared an object Ω against its reconstruction $\hat{\Omega}$ over all images in the Kimia216 dataset. The figure shows how high values of τ deteriorate the reconstruction, whereas lower values do not prune enough spurious branches. From the figure, we can infer that values around $\tau = 0.47$ offer a good trade-off between reconstruction and branch pruning. Moreover, around this value, the standard deviation reaches its minimum value suggesting optimal performance regardless of the object. Because the value of τ is stable for different scale factors, we conclude that scale does not affect the selection of the threshold.

VI. CONCLUSION AND FUTURE WORK

We presented the CPMA, a new method for medial axis pruning with noise robustness and equivariance to isometric transformations. Our method leverages the discrete cosine transform to compute a score function that rates the importance of individual points and branches within the medial axis of a shape.

Our pruning approach delivers competitive results compared to the state-of-the-art. Our experiments show that our method is robust to boundary noise and offers a satisfactory trade-off between reconstruction and medial axis simplification. Additionally, it is equivariant to isometric transformations, and it is capable of producing a stable and connected medial axis even in scenarios with significant perturbations of the contour.

The CPMA can be efficiently computed in parallel because it depends on an aggregation of reconstructions of the original shape. Each reconstruction is independent of the others, which allows the parallelism.

We believe our work leaves room for improvement, and thus we have identified the following potential for future work.

All of the 3D objects we used come as 3D triangular meshes. To compute the CPMA, we discretize the meshes to fix a resolution of 150^3 voxels. The discretization introduces two issues: 1) the objects lose small details in their structure, which affects the overall performance, and 2) the isometric equivariance decreases because rotated voxels will not perfectly align with non-rotated voxels.

Our algorithm for connectivity enforcement relies on iterative computations of Dijkstra's algorithm for finding the minimum energy path between two pieces of the unconnected medial axis. Better strategies to compute the paths could increase the CPMA's efficiency.

ACKNOWLEDGMENT

The authors want to thank the GRASP Laboratory at the University of Pennsylvania for providing the computational resources to conduct this investigation.

REFERENCES

[1] A. Tal, "3D shape analysis for archaeology," in *3D Research Challenges in Cultural Heritage*. Berlin, Germany: Springer, 2014, pp. 50–63, doi: 10.1007/978-3-662-44630-0_4.

[2] L. J. P. van der Maaten, P. J. Boon, G. Lange, H. Paijmans, and E. O. Postma, "Computer vision and machine learning for archaeology," in *Proceedings of the Computer Applications in Archaeology*, D. H. Kamermans, Ed. Leiden, The Netherlands: Leiden Univ., 2006. [Online]. Available: https://dbiref.uvt.nl/iPort?request=full_record&db=wo&language=eng&query=194155

[3] H. Li, L. Sun, X. Wu, and Q. Cai, "Scale-invariant wave kernel signature for non-rigid 3D shape retrieval," in *Proc. IEEE Int. Conf. Big Data Smart Comput. (BigComp)*, Jan. 2018, pp. 448–454, doi: 10.1109/bigcomp.2018.00072.

[4] M. H. Safar and C. Shahabi, *Shape Analysis and Retrieval of Multimedia Objects*. Boston, MA, USA: Springer, 2003. 10.1007/978-1-4615-0349-1.

[5] R. Chaudhry, F. Ofli, G. Kurillo, R. Bajcsy, and R. Vidal, "Bio-inspired dynamic 3D discriminative skeletal features for human action recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops*, Jun. 2013, pp. 471–478, doi: 10.1109/cvprw.2013.153.

[6] M. Li, S. Chen, X. Chen, Y. Zhang, Y. Wang, and Q. Tian, "Actional-structural graph convolutional networks for skeleton-based action recognition," 2019, *arXiv:1904.12659*. [Online]. Available: https://arxiv.org/abs/1904.12659

[7] R. Peters and H. Ledoux, "Robust approximation of the medial axis transform of LiDAR point clouds as a tool for visualisation," *Comput. Geosci.*, vol. 90, pp. 123–133, May 2016, doi: 10.1016/j.cageo.2016.02.019.

[8] R. Li, G. Bu, and P. Wang, "An automatic tree skeleton extracting method based on point cloud of terrestrial laser scanner," *Int. J. Opt.*, vol. 2017, pp. 1–11, Jan. 2017, doi: 10.1155/2017/5408503.

[9] T. Qiu, Y. Yan, and G. Lu, "A medial axis extraction algorithm for the processing of combustion flame images," in *Proc. 6th Int. Conf. Image Graph.*, Aug. 2011, pp. 182–186.

[10] G. Aufort, R. Jennane, R. Harba, and C. L. Benhamou, "A new 3D shape-dependant skeletonization method. Application to porous media," in *Proc. 14th Eur. Signal Process. Conf.*, Sep. 2006, pp. 1–5.

[11] K. Siddiqi, A. Shokoufandeh, S. J. Dickinson, and S. W. Zucker, "Shock graphs and shape matching," *Int. J. Comput. Vis.*, vol. 35, no. 1, pp. 13–32, 1999.

[12] A. Toshev, B. Taskar, and K. Daniilidis, "Shape-based object detection via boundary structure segmentation," *Int. J. Comput. Vis.*, vol. 99, no. 2, pp. 123–146, Mar. 2012, doi: 10.1007/s11263-012-0521-z.

[13] R. Marie, O. Labbani-Igbida, and E. M. Mouaddib, "The delta medial axis: A fast and robust algorithm for filtered skeleton extraction," *Pattern Recognit.*, vol. 56, pp. 26–39, Aug. 2016.

[14] O. Freifeld and M. J. Black, "Lie Bodies: A Manifold Representation of 3D Human Shape," in *Proc. Eur. Conf. Comput. Vis.* (Lecture Notes in Computer Science), B. Leibe, J. Matas, N. Sebe, and M. Welling, Eds. Cham, Switzerland: Springer, Oct. 2012, pp. 1–14. [Online]. Available: https://link.springer.com/10.1007/978-3-319-46448-0. and https://link.springer.com/10.1007/978-3-642-33718-5_1

[15] D. Zhang and G. Lu, "Review of shape representation and description techniques," *Pattern Recognit.*, vol. 37, no. 1, pp. 1–19, Jan. 2004. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0031320303002759

[16] S. Belongie, J. Malik, and J. Puzicha, "Shape matching and object recognition using shape contexts," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 4, pp. 509–522, Apr. 2002.

[17] S. Abbasi, F. Mokhtarian, and J. Kittler, "Curvature scale space image in shape similarity retrieval," *Multimedia Syst.*, vol. 7, no. 6, pp. 467–476, Nov. 1999, doi: 10.1007/s005300050147.

[18] J. Sun, M. Ovsjanikov, and L. Guibas, "A concise and provably informative multi-scale signature based on heat diffusion," *Comput. Graph. Forum*, vol. 28, no. 5, pp. 1383–1392, Jul. 2009, doi: 10.1111/j.1467-8659.2009.01515.x.

[19] C. Esteves, C. Allen-Blanchette, A. Makadia, and K. Daniilidis, "Learning SO(3) equivariant representations with spherical CNNS," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, V. Ferrari, M. Hebert, C. Sminchisescu, and Y. Weiss, Eds. Cham, Switzerland: Springer, 2018, pp. 54–70.

[20] D. Maturana and S. Scherer, "VoxNet: A 3D convolutional neural network for real-time object recognition," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Sep. 2015, pp. 922–928.

[21] P. K. Saha, G. Borgefors, and G. Sanniti di Baja, "A survey on skeletonization algorithms and their applications," *Pattern Recognit. Lett.*, vol. 76, pp. 3–12, Jun. 2016, doi: 10.1016/j.patrec.2015.04.006.

[22] G. B. K. P. Saha and G. S. de Baja Eds., *Skeletonization: Theory, Methods and Applications*, 1st ed. Amsterdam, The Netherlands: Elsevier, 2017. [Online]. Available: https://www.elsevier.com/books/skeletonization/saha/978-0-08-101291-8

[23] L. Younes, *Shapes and Diffeomorphisms* (Applied Mathematical Sciences), 2nd ed. Berlin, Germany: Springer, 2019.

[24] J. August, A. Tannembaum, and S. W. Zucker, "On the evolution of the skeleton," in *Proc. ICCV*, Sep. 1999, pp. 315–322.

[25] D. Shaked and A. M. Bruckstein, "Pruning medial axes," *Comput. Vis. Image Understand.*, vol. 69, no. 2, pp. 156–169, 1998. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1077314297905981

[26] X. Bai, L. Latecki, and W.-Y. Liu, "Skeleton pruning by contour partitioning with discrete curve evolution," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 3, pp. 449–462, Mar. 2007.

[27] W. Shen, X. Bai, R. Hu, H. Wang, and L. Jan Latecki, "Skeleton growing and pruning with bending potential ratio," *Pattern Recognit.*, vol. 44, no. 2, pp. 196–209, Feb. 2011, doi: 10.1016/j.patcog.2010.08.021.

[28] T. B. Sebastian, P. N. Klein, and B. B. Kimia, "Recognition of shapes by editing shock graphs," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 00, no. 528, pp. 755–762, 2004.

[29] X. Bai, W. Liu, and Z. Tu, "Integrating contour and skeleton for shape classification," in *Proc. IEEE 12th Int. Conf. Comput. Vis. Workshops, (ICCV Workshops)*, Sep. 2009, pp. 360–367.

[30] A. Sobiecki, A. Jalba, and A. Telea, "Comparison of curve and surface skeletonization methods for voxel shapes," *Pattern Recognit. Lett.*, vol. 47, pp. 147–156, Oct. 2014. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0167865514000269

[31] A. Sobiecki, H. C. Yasan, A. C. Jalba, and A. C. Telea, "Qualitative Comparison of Contraction-Based Curve Skeletonization Methods," in *Mathematical Morphology and Its Applications to Signal and Image Processing* (Lecture Notes in Computer Science), vol. 7883. Berlin, Germany: Springer, 2013, pp. 425–439. [Online]. Available: https://link.springer.com/10.1007/978-3-642-38294-9_36

[32] J. Chaussard, M. Couprie, and H. Talbot, "Robust skeletonization using the discrete λ-medial axis," *Pattern Recognit. Lett.*, vol. 32, no. 9, pp. 1384–1394, Jul. 2011. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0167865510003004

[33] H. Blum, "A transformation for extracting new descriptors of shape," in *Models for the Perception of Speech and Visual Form*, W. Wathen-Dunn, Ed. Cambridge, U.K.: MIT Press, 1967, pp. 362–380.

[34] P. Dlotko and R. Specogna, "Topology preserving thinning of cell complexes," *IEEE Trans. Image Process.*, vol. 23, no. 10, pp. 4486–4495, Oct. 2014.

[35] C. Lohou and G. Bertrand, "A 3D 12-subiteration thinning algorithm based on P-simple points," *Discrete Appl. Math.*, vol. 139, nos. 1–3, pp. 171–195, Apr. 2004. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0166218X03005390

[36] C. Lohou and J. Dehos, "Automatic correction of ma and Sonka's thinning algorithm using P-simple points," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 6, pp. 1148–1152, Jun. 2010.

[37] G. Németh, P. Kardos, and K. Palágyi, "Thinning combined with iteration-by-iteration smoothing for 3D binary images," *Graph. Models*, vol. 73, no. 6, pp. 335–345, Nov. 2011.

[38] T. Y. Zhang and C. Y. Suen, "A fast parallel algorithm for thinning digital patterns," *Commun. ACM*, vol. 27, no. 3, pp. 236–239, Mar. 1984, doi: 10.1145/357994.358023.

[39] G. K. Viswanathan, A. Murugesan, and K. Nallaperumal, "A parallel thinning algorithm for contour extraction and medial axis transform," in *Proc. IEEE Int. Conf. Emerg. Trends Comput., Commun. Nanotechnol. (ICECCN)*, Mar. 2013, pp. 606–610.

[40] R. Ogniewicz and M. Ilg, "Voronoi skeletons: Theory and applications," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Jun. 1992, pp. 63–69.

[41] C. Arcelli, G. S. di Baja, and L. Serino, "Distance-driven skeletonization in voxel images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 4, pp. 709–720, Apr. 2011. [Online]. Available: https://ieeexplore.ieee.org/document/5551146/

[42] C. Arcelli and G. S. D. Baja, "Finding local maxima in a pseudo-euclidian distance transform," *Comput. Vis., Graph., Image Process.*, vol. 43, no. 3, pp. 361–367, Sep. 1988. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0734189X88900898

[43] W. H. Hesselink and J. B. Roerdink, "Euclidean skeletons of digital image and volume data in linear time by the integer medial axis transform," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 12, pp. 2204–2217, Feb. 2008.

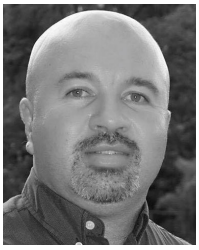
[44] M. Couprie, D. Coeurjolly, and R. Zrour, "Discrete bisector function and Euclidean skeleton in 2D and 3D," *Image Vis. Comput.*, vol. 25, no. 10, pp. 1543–1556, Oct. 2007.

[45] M. Postolski, M. Couprie, and M. Janaszewski, "Scale filtered Euclidean medial axis and its hierarchy," *Comput. Vis. Image Understand.*, vol. 129, pp. 89–102, Dec. 2014.

[46] M. Sato, I. Bitter, M. A. Bender, A. E. Kaufman, and M. Nakajima, "TEASAR: Tree-structure extraction algorithm for accurate and robust skeletons," in *Proc. 8th Pacific Conf. Comput. Graph. Appl.*, Oct. 2000, pp. 281–449.

[47] A. Beristain and M. Grana, "Pruning algorithm for Voronoi skeletons," *Electron. Lett.*, vol. 46, no. 1, pp. 39–41, Jan. 2010.
[48] M. Rumpf and T. Preusser, "A level set method for anisotropic geometric diffusion in 3D image processing," *SIAM J. Appl. Math.*, vol. 62, no. 5, pp. 1772–1793, Jan. 2002, doi: 10.1137/s0036139901384662.
[49] F. Mokhtarian and A. K. Mackworth, "A theory of multiscale, curvature-based shape representation for planar curves," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 14, no. 8, pp. 789–805, Aug. 1992. [Online]. Available: https://www2.scopus.com/inward/record.uri?eid=2-s2.0-0026903040&doi=10.1109%2f34.149591&partnerID=40&md5=a75f7cb55d21d22f322bfd1d93954df2
[50] B. Miklos, J. Giesen, and M. Pauly, "Discrete scale axis representations for 3D geometry," *ACM Trans. Graph.*, vol. 29, no. 4, pp. 101:1–101:10, Jul. 2010, doi: 10.1145/1778765.1778838.
[51] F. Gao, G. Wei, S. Xin, S. Gao, and Y. Zhou, "2D skeleton extraction based on heat equation," *Comput. Graph.*, vol. 74, pp. 99–108, Aug. 2018, doi: 10.1016/j.cag.2018.05.005.
[52] L. Gorelick, M. Galun, E. Sharon, R. Basri, and A. Brandt, "Shape representation and classification using the Poisson equation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 12, pp. 1991–2005, Dec. 2006.
[53] G. Aubert and J.-F. Aujol, "Poisson skeleton revisited: A new mathematical perspective," *J. Math. Imag. Vis.*, vol. 48, no. 1, pp. 149–159, Jan. 2014.
[54] J. Giesen, B. Miklos, M. Pauly, and C. Wormser, "The scale axis transform," in *Proc. 25th Annu. Symp. Comput. Geometry (SCG)*, New York, NY, USA, Jun. 2009, pp. 106–115. [Online]. Available: https://portal.acm.org/citation.cfm?doid=1542362.1542388
[55] M.-P. Dubuisson and A. K. Jain, "A modified Hausdorff distance for object matching," in *Proc. 12th Int. Conf. Pattern Recognit.*, vol. 1, Oct. 2002, pp. 566–568. [Online]. Available: https://ieeexplore.ieee.org/document/576361/
[56] W. Shen, X. Bai, X. Yang, and L. J. Latecki, "Skeleton pruning as trade-off between skeleton simplicity and reconstruction error," *Sci. China Inf. Sci.*, vol. 56, no. 4, pp. 1–14, Apr. 2013, doi: 10.1007/s11432-012-4715-3.
[57] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognit. Lett.*, vol. 27, no. 8, pp. 861–874, Jun. 2006. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S016786550500303X

DIEGO PATIÑO was born in Medellín, Colombia, in 1988. He received the B.S., M.S., and Ph.D. degrees in computer engineering from the Universidad Nacional de Colombia, in 2010, 2012, and 2020, respectively. In 2014, he was a Visiting Researcher with the University of Wisconsin–Madison. He is currently a Postdoctoral Researcher with the GRASP Laboratory of General Robotics, Automation, Sensing and Perception, University of Pennsylvania. His research interests include machine learning and geometric approaches to computer vision, focusing on 3D vision, shape analysis, symmetry detection, and single-image to 3D reconstruction.



JOHN W. BRANCH received the B.S. degree in mines and metallurgic engineering and the M.Sc. and Ph.D. degrees in computer engineering from the Universidad Nacional de Colombia, Medellín, in 1995, 1997, and 2007, respectively. Since 2000, he has been a full-time Professor with the Universidad Nacional de Colombia. He has directed 40 master's theses and 15 Ph.D. theses and currently directs five Ph.D. theses and ten master's. In the last 20 years as a Professor, he has more than 200 scientific publications in national and international journals and conferences, mostly related to his research theme computer vision and related areas. He has received the following prizes and distinctions, namely the Meritorious degree Work on mining and metallurgy engineering from the Universidad Nacional de Colombia, in 1995, the Fourth Place Best Grade Contest VI Version from the Universidad Nacional de Colombia, in 1996, the Thesis of Meritorious degree, Postgraduate in computer engineering from the Universidad Nacional de Colombia, in 1997, and the Meritorious Doctorate Thesis from the Universidad Nacional de Colombia, in 2007.