A Medial-Axis-Based Measure of District Compactness

J. D'Amico, E. Gasparovic, G. Malen, and M. Zhong

Abstract - An essential question for democracy is how to rigorously determine the likelihood that a congressional map has been gerrymandered. A number of state constitutions require districting plans to be "compact", yet no set definition of compactness exists, other than an oft-cited sentiment that you "know it when you see it". We introduce a novel compactness measure based on a geometric structure known as the medial axis, which has strong ties to the science of how humans perceive and process shapes. As such, we argue that our measure performs well as a mathematical quantification of "knowing it when you see it". We compare our medial-axis-based measure to a recent machine-learning-based compactness measure introduced by Kaufman, King, and Komisarchik. Specifically, we examine the performance of our measure and theirs in several case studies, including two states whose districting plans have been extensively covered in the media as well as the entire 2016 congressional district map.

Keywords : redistricting; measures of compactness; medial axis; gerrymandering

Mathematics Subject Classification (2020) **:** 68U05; 91F10

1 Introduction

Gerrymandering is the act of manipulating the boundaries of an electoral district in order to favor or disenfranchise one party or group. Gerrymandering may take different forms: partisan gerrymandering to benefit a particular political party; racial gerrymandering to dilute the voting power of a minority group; incumbent gerrymandering to create "safe" districts for re-election; and prison-based gerrymandering to increase district populations with non-voters when apportioning representatives.

General principles or (in many cases) constraints on district map drawing have arisen from a number of Supreme Court rulings. In general, it is accepted that districts should have essentially equal populations and preserve "communities of interest"; respect geographic boundaries and/or political subdivisions; and be contiguous or connected (with the exception of islands) with no holes. An additional principle without an agreed-upon definition is the notion that districts should be geometrically "compact". Former Supreme Court Justice John Paul Stevens famously stated, "Substantial divergences from a mathematical standard of compactness may be symptoms of illegitimate gerrymandering" (*Karcher v. Daggett*, 462 U.S. 725 (1983)).)

But what exactly is meant by *compactness*? Academics across diverse disciplines in the social and mathematical sciences have shown that there is not a single dimension of compactness, and as a result, a plethora of compactness measures that are often in conflict with one another have been introduced. Many measures are subject to implementation flexibility when it comes to dealing with such issues as non-contiguous districts, choice of map projection, data resolution, and other choices involved when drawing the boundaries of districts [4],[11]. These vagaries in the definition often provide a loophole for groups hoping to evade the enacting of changes, and the existence of so many competing measures makes it extremely difficult to declare that a given district fails to meet the standards of compactness. Further complicating the issue is that districts which have been gerrymandered can often still appear "compact", and fairly drawn districts may have irregular boundaries or shapes due to attempts to preserve principles such as equal population. In the end, all sides are likely to invoke the familiar trope that "you know it when you see it".

Instead of running from this blatantly subjective adage, a recent paper by Kaufman, King, and Komisarchik (hereafter referred to as Kaufman, et al.) [17] sought to build a predictive model that might serve as a proxy for the so-called eye test. Based on the premise that people have a common understanding of compactness, even without an agreed-upon definition, their model synthesized a large amount of collected data in which participants ranked a set of districts from most to least compact, and created a model that assigns compactness scores on a scale from 1 (least compact) to 100 (most compact). (Note: In their original paper [17], the authors actually used 1 to indicate most compact and 100 to indicate least compact, but to be consistent with our measure and other standard compactness score conventions, we have flipped the scale for Kaufman, et al.'s measure in all of our results.) Their model incorporated geometric features of the district shapes as well as values of some common compactness measures, such as the convex hull, Reock, and Polsby-Popper measures that will be discussed in the next section. They determined in [17] that participants found districts to be compact if they are "squarish, with minimal arms, pockets, islands, or jagged edges". The authors also conclude that no extant measures they came across "offer a simple geometric representation for what humans know when they see", and they exhort the community at large to search for such a definition.

The novel and purely mathematical measure that we introduce herein is not meant to supplant existing compactness measures in redistricting discussions, and we do not claim that it provides a perfect solution to the question of how to determine whether or not a districting map is the product of gerrymandering. Instead, our goal is to introduce a new measure of compactness grounded in the principles of human shape recognition and to illustrate its utility within this context. The measure is based on a geometric structure – the "medial axis" — that is known to be related to how the human brain perceives and processes shapes. Indeed, there is a sizable body of research in cognitive science and related fields that suggests human visual perception relies heavily on skeletal, or medial axis, representations to interpret and categorize shapes. These internal shape skeletons provide a stable and efficient framework for object recognition and analysis, even when the objects' contours are irregular, making them particularly relevant in assessing the overall form of complex geometries like political districts.

Our medial-axis–based compactness measure aligns closely with the criteria identified by Kaufman, et al. While their paper concluded that existing measures failed to offer a simple geometric representation of this intuitive understanding (what "humans know when they see it"), our measure directly responds to this challenge. By capturing the internal structure and spatial coherence of a district via its medial axis, our approach offers a promising step toward formalizing this human-centric notion of compactness. In what follows, we explore and compare this new measure with the results of Kaufman, et al. (which incorporates human assessments of compactness as well as other established measures) in various settings, such as in states where district maps have been redrawn after being declared gerrymandered in court. One of the advantages of our novel compactness measure compared to that of Kaufman, et al. is that ours does not rely on machine learning but offers many of the same strengths and generates similar results. Perhaps the main feature of our measure is that it is less sensitive to noise in the outermost segments of a district, where "outermost" is not a subjective concept but is here rigorously defined by the measure itself.

In Section 2, we introduce the structures upon which we build our compactness measure, and discuss the connections to the psychology of shape recognition. In Section 3, we construct the measure in detail, and in Section 4 we examine the performance of the measure and compare it with Kaufman, et al.'s predictive model. We use Kaufman, et al.'s compactness scores as our performance standard since their model embeds human assessments of compactness as well as several of the most popular measures in the literature as covariates in their statistical model. Finally, in Section 5 we discuss possible adjustments to our method that might be made and other ideas for future work.

2 Background

2.1 How to Quantify District Compactness: Historical Examples

There are many available options for formally computing the compactness of a given district. We note that one drawback to all of the following measures is that they only take into account district geometry and not population distribution (as opposed to measures discussed in, e.g., [14] and [25]), but our focus in this paper is to prioritize quantifying the pure shape of a district.

For instance, the *convex hull measure* computes the ratio of a district's area to the area of its convex hull (i.e., the smallest convex set containing the district, meaning any line drawn between two points within the set stays within the set's boundaries). The convex hull measure captures how "concave" or indented a shape is and it is not particularly sensitive to noise on the perimeter of a shape. One criticism is that the convex hull ratio (a number between 0 and 1) is equal to 1 on any district which is already convex, even though the district itself may not be compact (e.g., if it is a very long and skinny rectangle). The *Reock score* is the ratio of a district's area to the area of its minimum bounding circle [30]. A drawback to the Reock score is that circles receive the highest value under it, and districts are rarely circular in shape. Although it is simple and intuitive and is useful for identifying "stretched out" or sprawling shapes, it is sensitive in that a single protruding area can dramatically reduce the score, and it doesn't consider internal boundaries. A third example, the *Polsby-Popper score*, computes the (approximate) "isoperimetric ratio", which is the ratio of a district's area to the area of a circle with the same perimeter [29] [31]. Like Reock, the Polsby-Popper measure also idealizes circles; moreover, it penalizes long, jagged borders. However, the perimeter may be jagged for good reason (e.g., due to low population density, coastline geography or other natural boundaries, etc.).

The authors of [34] proposed a related method for approximating this isoperimetric ratio that depends on a traversal of a geometric object known as the *medial axis*, and they achieved more efficient and robust computational bounds on the ratio than the pre-existing state-of-the-art. The novel compactness measure we introduce in this paper is calculated directly from the medial axis, a structure that we will discuss in more detail in the next subsection.

As mentioned in the introduction, Kaufman, et al.'s compactness measure is based on a supervised statistical model trained to replicate human judgments of district compactness [17]. The authors first collected human responses through a series of pairwise comparisons, in which participants (ranging from students to experts such as judges and lawyers) were shown images of districts and asked which of the two appeared more compact. These responses were then incorporated as features in their ensemble-based predictive model, in addition to a variety of other established geometric features, including well-known measures such as Polsby–Popper, Reock, and convex hull scores, along with additional descriptors of boundary complexity. Thus, Kaufman, et al.'s model relies heavily on boundary-sensitive features such as perimeter, convexity, and human perceptions of boundary irregularity.

Before turning to the definition of the medial axis, we have to address why the Supreme Court and other judicial bodies have generally shied away from using geometric compactness measures alone as decisive legal standards. For one, these measures reduce complex political geography and community representation to a single number, which often fails to capture the full picture. There's no agreement among experts or legal authorities on which compactness measure should be used, and each tells a slightly different story. Perhaps most importantly, a shape might be geometrically "compact" but still split communities of interest, dilute minority votes, clinch a partisan advantage, and so on. On the flip side, some oddly shaped districts may be drawn that way to, e.g., preserve minority representation or respect community boundaries, and purely geometric measures cannot account for this.

Each compactness measure captures different aspects of a district's geometry. When used together, they provide a richer picture than any one measure alone. Researchers have moved from simply ranking districts to asking "compact compared to what"? This has led to simulation-based approaches and ensemble models to assess whether a district is unusual relative to realistic alternatives. Some recent work (like that by Kaufman et al.)

tries to model human judgments of compactness, suggesting that human intuition might integrate multiple geometric cues better than any single formula. Courts' reluctance to adopt compactness measures mirrors academia's ongoing challenge: defining not just how compact a district is, but what kind of compactness matters in a representative democracy.

2.2 The Medial Axis and the Extended Distance Function

In 1967 engineer Harry Blum introduced a skeleton-like structure associated to a given shape which he called the *medial axis* [6].

Definition 2.1 *The* ***medial axis*** *of a closed and bounded region in the plane consists of the set of centers of circles that are contained entirely inside the region and that touch the boundary in two or more places.*

In Figure 1, we have an example of a shape with its medial axis and examples of circles that touch the boundary in two or three places. The medial axis effectively captures intuitive shape structure and provides a simpler representation of an object, analogous to how the human skeleton is a simpler representation of the shape of the human body.

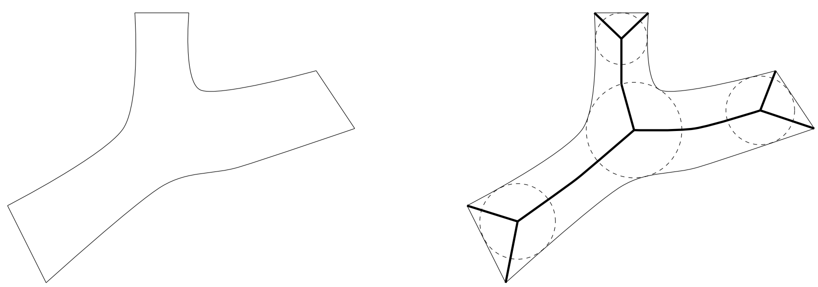

Figure 1: A shape (left) with its medial axis in bold and examples of maximally enclosed circles that touch the boundary at two or more points (right).

One computational issue with the medial axis is its instability and sensitivity to boundary noise; small, insignificant bumps on the boundary of the shape create many offshoot branches of the medial axis. To get around this issue, it is common to undertake various pruning methods to remove extraneous branches. The *extended distance function* [23] (also called the *burn time*) was first introduced as a means of pruning the medial axis, but has since been used effectively in such applications as shape description and analysis.

Definition 2.2 *Let x be a point contained in a continuous path f in a medial axis m. Let $r_f(x)$ be the sum of the shortest distance from x to an endpoint of f (via f) plus the subsequent distance to the closest boundary point. The* ***extended distance function (EDF)*** *for x is the largest such $r_f(x)$ among all paths f in m that contain x.*

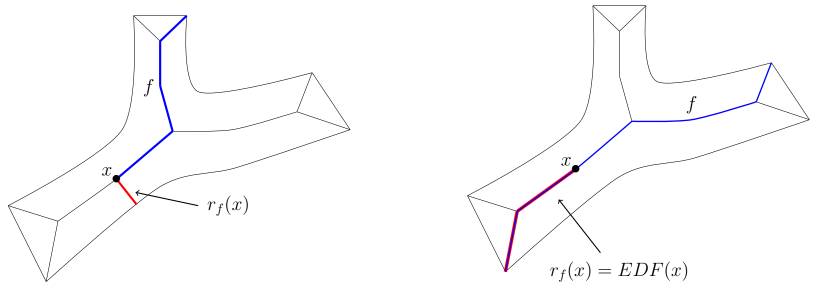

Figure 2: Left: A path f in blue containing the point x, with $r_f(x)$ indicating the length of the red line segment perpendicular to the boundary. Here, since x is one of the endpoints of f, the shortest path to the boundary via f starting at x is indeed the red segment. Right: A path f in blue containing the point x, with $r_f(x)$ in this case being the length of the red path. The length of this red path is the largest such $r_f(x)$ in this example, and therefore illustrates EDF(x).

Intuitively, the EDF provides a measure of depth within a shape, capturing the amount of "sideways" shape expansion (see Figure 2 and Figure 3). The larger the EDF value, the "deeper" a point is within the shape and the closer it is to the central "core" of the shape. One of its most significant qualities is that (in contrast to the medial axis itself) the EDF is insensitive to noise and remains stable under slight perturbations to the boundary of the region [23].

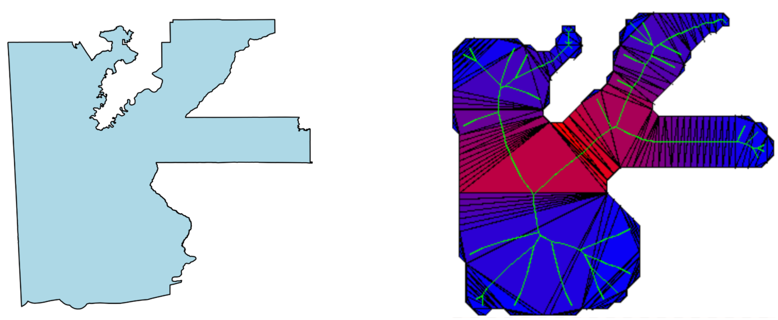

Figure 3: Left: Alabama's 1st congressional district in 2016. Right: Medial axis (green) and heat map illustrating EDF values (red indicates high values, blue low values). Thus, the "core" part of the district is characterized by higher EDF values, while the "arms" feature lower values.

A version of the EDF was introduced in [21] for the purpose of automated parts decomposition. The authors of [21] enacted a user study in which internet participants were asked to label points inside each shape in their database as belonging to one of three regions: the central "core" shape, parts, or details. Users were also given the option to add additional finer levels to this three-tiered hierarchy (core, parts, and details). However, very few chose to do so, indicating that using three levels in a shape decomposition seems to be an accurate reflection of human perception. This user study and its results influenced the definition of our novel compactness measure (see Section 3).

We note that the authors of [5] previously suggested a compactness measure based on the medial axis. They sought to quantify the degree of "meandering" for a given district by finding the ratio of the length of the medial axis of a district to the length of the medial axis of the district's convex hull. A drawback of their "medial-hull ratio" is the instability of the medial axis and its sensitivity to noise on the boundary; even one very slight jut on the boundary of a district could lead to a significant change in shape of the medial axis and thus have a large effect on its length. The authors attempted to alleviate this issue by pruning branches, but depending on the choice of pruning coefficient, the assigned level of gerrymandering may be underestimated or magnified. Moreover, the medial-hull ratio of a long and skinny rectangular district would be close to 1, but this does not mean the district was not likely gerrymandered. Nevertheless, we believe that utilizing the medial axis in the construction of a measure of compactness can be extremely effective due to the wealth of evidence that a skeletal model is used by our visual system to represent shapes. In the next subsection, we will provide examples of this research.

2.3 Evidence for a Skeletal/Medial Axis Model in Human Perception

The question of exactly how shape is represented in human visual perception remains unresolved, yet there is extensive evidence that a medial axis-like model is utilized in the processes underlying human vision. For instance, [18] provided evidence that the human vision system extracts and utilizes a medial axis/skeleton model for object shape representation and memory coding. The authors of [8] and [9] proposed and experimentally justified a model for visual processing whereby shapes are identified and represented via "cores" connecting and relating object boundaries. The papers [16] and [20] both found concrete evidence from neurons and the visual cortex that a medial axis structure is used by the brain in perceptual processes. A recent paper [7] reaffirmed that shape representations are potent organizing principles in high-level visual cortex processing.

A related concept to the medial axis is the shock graph; the authors of [32] demonstrated through a series of experiments that shock graph-based medial descriptions predict performance in shape perception. The authors of [13] showed participants a shape on a tablet device, asked them to tap the shape anywhere, and aggregated their touches. The resulting structure essentially formed the skeleton of the shape. One study [22] investigated the utility of "medialness" for capturing and describing works of art, and how it might even be put to use in computations for directing artificial "arms" to create or modify artwork. The results in [1] and [2] demonstrated that a "skeletal similarity" model accounted for the largest amount of variance involved in their study participants' shape dissimilarity judgment and identification/recognition tasks, as opposed to other visual similarity models under consideration. Another paper [3] involving functional magnetic resonance imaging once again supported the role of a skeletal model in human perception and object recognition.

We emphasize that the work discussed in this subsection is merely a sampling of the diverse research investigating the role of the medial axis and related structures in human visual processing.

3 A Novel Medial-Axis-Based Compactness Measure

Since the medial axis plays a pivotal role in how humans perceive shape, it is both natural and evidence-supported to utilize a medial-axis-based measure for quantifying compactness and for effectively accessing that intangible essence that you “only know when you see it”. Our novel compactness measure computes the ratio of the area of an automatically determined “essential” region within a given district to the area of the entire district. The formal definition and details of its computation constitute this section.

We obtained the district shapefiles from the TIGER/Line databases supplied by `census.gov` [10] and saved them as images in order to compute the medial axes and EDF measures from those images. Given an input image of a congressional district, we segment the image into foreground and background using a simple k-means clustering with $k = 2$ clusters. If the image is not of sufficiently high resolution and there are high levels of pixelation noise, we perform a very gentle smoothing of the boundary of the district. Using the same methods as in [21], we compute a robust triangulation of the district and use that to construct a discrete approximation of the district's medial axis. Each triangle contributes one point to the approximate medial axis, and these triangles are then associated to their corresponding medial points.

Once we obtain the medial axis, we use 3-means clustering to group the points on the medial axis by their EDF value. This clustering by EDF value leads to a separation of the medial points into a central “core” group C, a “parts” group P, and an “extra details” group X. As mentioned in Section 2, our rationale for using three clusters stems from the parts decomposition results in [21]. For a given district D, we then refer to the region consisting of the triangles associated to the medial points in both the core C and parts P as the “essential” region E.

Definition 3.1 *The* ***extended distance function (EDF) measure*** *for assessing the compactness of a district D is defined to be* $\dfrac{area(E)}{area(D)}$.

In Figure 4, we give an example of one of Colorado's congressional districts in 2016 accompanied by the medial axis and decomposition used for computing the EDF measure.

A final note is that one may think it more natural to use only two clusters, rather than three, to define the EDF measure. In this case, we would have medial points belonging to a central core shape cluster C and the rest classified as extra details X. The ratio would then be given as $\dfrac{\text{area}(C)}{\text{area}(D)}$. Our reasoning for choosing to use three clusters, rather than two, is twofold. First, as previously mentioned, the user study in [21] established strong evidence for the use of three clusters to reflect human perception. Second, the paper by Kaufman, et al. confirmed the unsurprising notion that people deem “squarish” districts to be the most compact. As seen in Figure 5, using two clusters to compute the EDF compactness measure for a nearly-square region results in a lower compactness value than when three clusters are utilized.

One final note: One significant departure of our measure from that of Kaufman, et

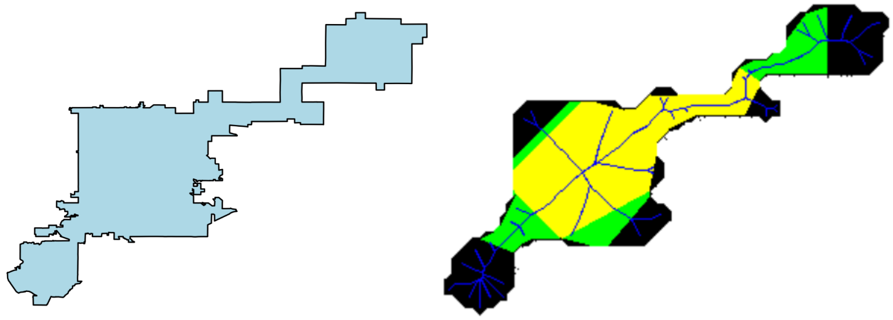

Figure 4: Left: Colorado's first congressional district in 2016 (the 114th congress). Right: The yellow region is formed by the triangles associated to the medial axis points in the core C, the green regions correspond to the parts group P, and the black regions correspond to the extra details X.

al. is that their measure gives greater importance to noise on the boundaries of regions. More generally, according to the authors of [11], one drawback of contour-based compactness measures is that they tend to be quite sensitive to the "physical geography" of the boundary. Given the EDF's stability and robustness to minor boundary perturbations, the measure tends to underweight jagged features along district edges and is therefore not as susceptible to small fluctuations in the physical geography of the district boundaries. Since such boundary noise may reflect either gerrymandering tactics or legitimate geographic and community-based delineations, the extent to which this is either a feature or a bug depends on the context in which the measure is applied.

4 Illustrating the EDF Measure via Examples

In this section, we take an in-depth look at two states that have seen their districting maps go to the Supreme Court: Pennsylvania and North Carolina. We compare their districts' compactness scores across two different districting cycles, focusing on similarities and differences between the EDF measure and Kaufman, et al.'s. We also compare the EDF measure and Kaufman, et al.'s measure for all congressional districts in the United States in 2016 under the 114th congress. In order to analyze the similarities and differences between the EDF measure and Kaufman, et al.'s measure, we focused on correlation analysis and rankings comparison. We reiterate that we are not attempting to assert whether or not gerrymandering has occurred; rather, our focus will be on comparing the compactness scores of the districts relative to one another, as well as analyzing how the results of our measure align with or diverge from Kaufman, et al.'s findings.

4.1 Pennsylvania District Maps in 2011 and 2018

In late 2017, a legal challenge was made by the League of Women Voters of Pennsylvania alleging that the congressional districts drawn in 2011 in response to the 2010 U.S. Census were the result of an illegal partisan gerrymander. After a lower court initially ruled that

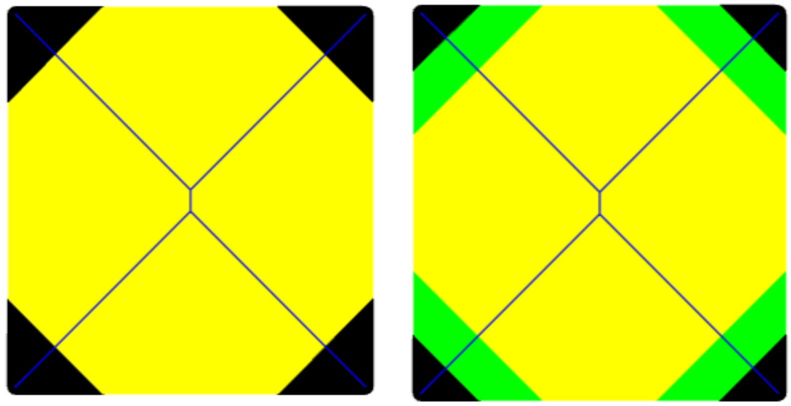

Figure 5: Left: For a nearly-square shape (with its medial axis plotted in dark blue), the core cluster based on EDF value is in yellow and the extra details cluster is in black. The EDF compactness score when using two clusters is 0.87. Right: The core cluster is in yellow, the parts cluster is in green, and the extra details cluster is in black. The EDF score for three clusters is 0.94.

the districts were not illegally gerrymandered, the case was appealed to the Pennsylvania Supreme Court, which then ruled that the drawing "achiev[ed] unfair partisan advantage". The majority opinion offered a strong condemnation of the map drawing and the impact that it has on fair elections, and ordered Pennsylvania lawmakers to redraw the map. Following this ruling, the Pennsylvania General Assembly failed to agree upon a redrawn map prior to the deadline for submission, after which a Court-appointed third party redrew the congressional districts and the Court adopted the new drawing in 2018 for that year's election. The map was met with praise from Pennsylvania Democrats, but both the process by which a new map was decided upon and the map itself was criticized by Pennsylvania Republicans. Despite unsuccessful attempts by Pennsylvania Republicans to appeal the case to the United States Supreme Court, the Court-drawn map was in place from the 2018 election through the 2022 midterm elections [19].

We computed compactness scores using the EDF measure for the 18 congressional districts in Pennsylvania in 2011, then again for the redrawn districts in 2018. (See Figures 8 and 9 in the Appendix for the two maps.) More than half of the districts in 2011 had EDF compactness scores below 0.75, whereas in 2018, all values were greater than 0.8 (see Table 1). Similar results were obtained using Kaufman, et al.'s measure; see Table 4 in the Appendix.

It is worth noting that the district numbers in 2011 do not correspond to the district numbers in 2018, so direct comparisons should not be made between, e.g., district 10 in 2011 and district 10 in 2018. Instead, notice that for both the EDF measure and Kaufman, et al.'s measure, the average scores increased from 2011 to 2018, indicating a higher degree of overall compactness in the redrawn map. Furthermore, using population overlap information obtained from [33], Table 2 shows the approximate correspondence between Pennsylvania's 2011 and 2018 congressional district numberings and the changes in their

Pennsylvania EDF Scores		
District #	**2011**	**2018**
1	0.68	0.89
2	0.83	0.89
3	0.69	0.84
4	0.82	0.91
5	0.82	0.86
6	0.74	0.90
7	0.59	0.89
8	0.79	0.86
9	0.61	0.80
10	0.71	0.82
11	0.68	0.88
12	0.48	0.81
13	0.68	0.81
14	0.85	0.87
15	0.71	0.88
16	0.79	0.85
17	0.67	0.87
18	0.76	0.80
Average	**0.72**	**0.86**

Table 1: EDF compactness scores for Pennsylvania district map in 2011 and 2018.

EDF scores. (We caution that this is certainly not a perfect correspondence because redistricting drastically changed boundaries.) The 2018 redistricting led to a substantial improvement in EDF scores across nearly all districts. In particular, notice that all but two of the corresponding districts showed a positive increase in their compactness scores from 2011 to 2018, with one district remaining constant and one decreasing by 0.05. Similar outcomes were seen for Kaufman, et al.'s measure; see Table 5 in the Appendix.

In 2011, districts 2, 4, 5, and 8 were rated highly compact under both measures. At the bottom, districts 7, 12, and 17 had the three lowest compactness scores under Kaufman, et al.'s measure, and three of the four lowest scores under the EDF measure. If we break up the districts into three equally-sized categories (high, medium, and low compactness) within each of the two compactness measures, 12 of the 18 districts belong to the same group (e.g., highly compact for both EDF and Kaufman) and the other 6 differ by one group (e.g., medium for EDF vs. low for Kaufman). Overall, the Spearman correlation coefficient for the two compactness measures for the Pennsylvania districts in 2011 is 0.7, indicating a strong positive relationship between the two measures (i.e., districts tend to maintain relatively similar positions across both rankings). In general, a correlation coefficient of ≥ 0.7 is considered indicative of a strong positive relationship; see, e.g., [12], [15]. The results are not as consistent for the 2018 districts, with a Spearman coefficient of 0.51 (indicating only a moderate relationship). Nevertheless, 10 out of 18 districts

Pennsylvania EDF Score Comparisons				
New #	Old #	EDF 2011	EDF 2018	Δ
1	8	0.79	0.89	+0.10
2	13	0.68	0.89	+0.21
3	2	0.83	0.84	+0.01
4	13/6	0.71	0.91	+0.20
5	7	0.59	0.86	+0.27
6	6	0.74	0.90	+0.16
7	15	0.71	0.89	+0.18
8	17	0.67	0.86	+0.19
9	9	0.61	0.80	+0.19
10	4	0.82	0.82	0.00
11	16	0.79	0.88	+0.09
12	10	0.71	0.81	+0.10
13	9	0.61	0.81	+0.20
14	18	0.76	0.87	+0.11
15	5	0.82	0.88	+0.06
16	3	0.69	0.85	+0.16
17	12	0.48	0.87	+0.39
18	14	0.85	0.80	-0.05

Table 2: Approximate correspondence between Pennsylvania's 2018 and 2011 (new and old) congressional district numberings based on significant population overlap [33], together with the changes in their EDF scores.

landed in the same group, 7 districts differed by one group, and only 1 district (district 6; see Figure 7) differed by two groups.

We now dive deeper into significant qualitative differences between the EDF measure and Kaufman, et al's "know it when you see it" measure. In 2011, under the EDF measure district 14 (Figure 6, left) had the highest compactness score. However, it landed near the median score in Kaufman, et al.'s ranking, and similarly for convex hull and Polsby-Popper. District 9 stood out as a major discrepancy (see Figure 6, right), accounting for the EDF measure's third lowest score, while falling in the top half of Kaufman, et al.'s scores, and similarly for Polsby-Popper and Reock. The two "arms" in district 9 (culminating in Indiana and Fayette counties) constitute a significant portion of the area of the shape, resulting in a low EDF score.

Moving to 2018, the biggest discrepancies between the EDF measure and Kaufman, et al.'s came down to district 17 (near the median for the EDF measure, second-to-lowest score for Kaufman, et al.'s) and district 6 (the EDF measure's second highest score and Kaufman, et al.'s third lowest); see Figure 9 in the Appendix. Also, note the relative similarity in shape between districts 1 and 4 in 2018. The EDF measure placed both districts in the top three with similar scores, yet Kaufman, et al.'s measure ranked district 1 high and district 4 near the median (see Figure 7 for images of districts 6, 1, and

4). These differences in relative compactness can most likely be attributed to the jagged nature of the boundaries, a feature that plays a much more prominent role in Kaufman, et al's measure than the EDF measure.

In summary, the observed differences in compactness values in Pennsylvania between the Kaufman and EDF measures (particularly for the 2018 congressional districts) likely stem from distinct sensitivities of each measure to specific geometric features. Specifically, the divergences suggest that while the districts in 2018 may exhibit thickened interiors (or “cores”), contributing to higher EDF scores (as the EDF measure tends to be more sensitive to a district's overall “squarishness”), they also possess more jagged or irregular boundaries. Kaufman, et al.'s measure's sensitivity to these deviations means it penalizes these boundary irregularities more heavily than the EDF score does. This difference in measure sensitivity explains why a district might score relatively well for the EDF measure due to its “fullness” but poorly for Kaufman's measure due to a complicated perimeter. Essentially, the 2018 redistricting appears to have resulted in shapes that are internally dense but externally complex, causing the Kaufman and EDF measures to register different degrees of compactness due to their differing algorithmic approaches to measuring a district's geometry.

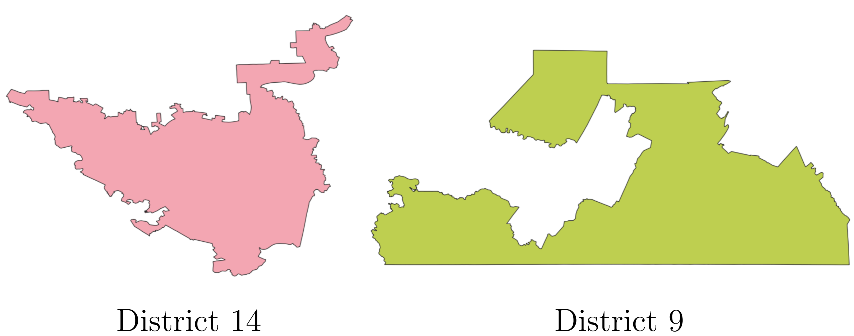

Figure 6: Discrepancies between the EDF measure and Kaufman, et al.'s measure. Left: Pennsylvania's district 14 in 2011. Although ranked highly compact under the EDF measure, its score under Kaufman, et al.'s measure landed it near the median score. This is likely due to the fact that it has a large central core but a great deal of jaggedness on the border. Right: Pennsylvania's district 9 in 2011. It received the third lowest score from the EDF measure (likely attributed to the thickness of the “arms” surrounding the hole) while Kaufman, et al.'s measure put it in the top half of scores.

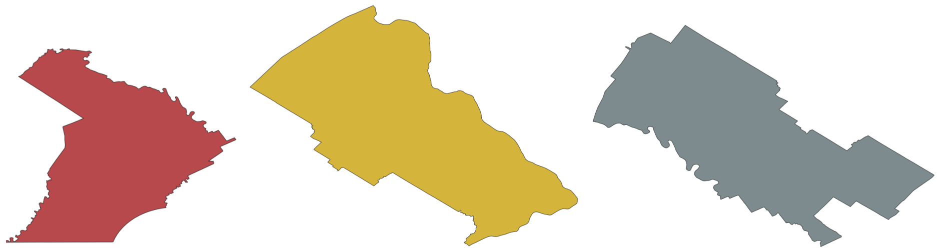

Figure 7: More discrepancies between the EDF measure and Kaufman, et al.'s measure. Left: Pennsylvania's district 6 in 2018. It received the second highest score under the EDF measure but the third lowest for Kaufman, et al.'s. Center and Right: Pennsylvania's districts 1 and district 4 in 2018; note the relative similarity in their shapes. The EDF measure placed both districts in the top three with similar scores, yet Kaufman, et al.'s measure ranked district 1 high and district 4 around the median value.

4.2 North Carolina District Maps in 2010 and 2011

For our analysis, we looked at North Carolina congressional districts used in the 2010 and 2011 elections (see Table 3, and see Figures 10 and 11 in the Appendix for the maps). In the context of the legal timeline of North Carolina's districts, the 2010 map (which had a 7-to-6 Democratic edge [26]) marks the final map before intense legal battles. The redrawn map in 2011 by a Republican-led legislature was ultimately ruled a racial gerrymander and was eventually thrown out. The two redistricting cycles involved significant changes to district boundaries, making direct comparisons between districts challenging; thus, we have not attempted to do so here.

The EDF and Kaufman, et al.'s average scores for North Carolina both decreased from 2010 to 2011 (see Tables 6 and 7 in the Appendix). The scores were qualitatively fairly similar for the North Carolina congressional district map in 2010, as 8 out of 13 districts were categorized into the same compactness group (high, medium, low) by both measures. The remaining 5 districts differed by only one group, indicating fairly consistent agreement across these broad categories. The Spearman correlation between the two measures was 0.77, indicating a strong positive relationship. Both measures identified district 12 (see Figure 10 in the Appendix) as the least compact district, which is often referred to as the "I-85 district" due to its elongated and convoluted shape along the highway and is notoriously non-compact. A significant divergence between the two measures is seen for district 9 (again, we refer the reader to Figure 10 in the Appendix), which received the second-lowest EDF compactness score but was more highly ranked according to Kaufman, et al.'s measure (namely, 7th out of 13). This is likely attributed to the fact that the shape is essentially two chunky areas attached at the narrow, central "core" portion, and the EDF measure quantifies compactness from the center outwards.

In 2011, the Spearman correlation between the EDF and Kaufman, et al. measures was more moderate at 0.59. Here, 6 of the 13 districts fell into groups one level apart

North Carolina EDF Scores		
District #	**2010**	**2011**
1	0.81	0.75
2	0.71	0.72
3	0.72	0.82
4	0.91	0.69
5	0.81	0.81
6	0.75	0.73
7	0.84	0.71
8	0.86	0.83
9	0.70	0.54
10	0.84	0.82
11	0.81	0.77
12	0.62	0.54
13	0.69	0.68
Average	**0.77**	**0.72**

Table 3: EDF compactness scores for North Carolina district map in 2010 and 2011.

according to the two compactness measures (with no districts differing by two groups). Districts with a high degree of agreement across the two measures include districts 10 (highly compact) and 12 (very low compactness); see Figure 11 in the Appendix. When it comes to the largest discrepancies between the two measures, district 8 received the highest EDF compactness score but earned near the median score under Kaufman, et al.'s measure. On the other hand, districts 2, 6, and 11 were three of the most compact districts according to Kaufman, et al.'s measure but all were near the median under the EDF measure. The EDF measure's center-outward perspective and prioritization of centralized robustness sometimes is in direct contrast with Kaufman, et al.'s more heavily weighted emphasis on minimizing boundary complexity.

4.3 United States Congressional District Map in 2016

We computed EDF compactness scores as well as the scores under Kaufman, et al.'s measure for the 435 United States congressional districts in 2016, ignoring non-contiguous districts. These non-contiguous districts are rare outliers in the data, as in fact most states require by law that all congressional districts be contiguous. There are various ways in which one might adapt the EDF measure to districts with disconnected medial axes but no canonical way; thus, we exclude non-contiguous districts in our analysis and focus only on single, unified geographic entities. There were 10 districts that landed in the top 25 most compact districts under both the EDF measure and Kaufman, et al.'s measure, and these districts are shown (in order by district number, not by compactness) in Figure 12 in the Appendix. Recall that Kaufman, et al.'s measure is said to prioritize districts that are "squarish, with minimal arms, pockets, islands, or jagged edges". All of the districts in Figure 12 do indeed contain minimal and/or small "arms" or "pockets".

The Spearman correlation between the EDF scores and Kaufman, et al.'s scores was 0.82, indicating a strong to very strong relationship and pointing to the likelihood of a very high correlation between Kaufman, et al.'s measure and the EDF measure upon the incorporation of the EDF scores as features into their model (as was done with other common compactness measures such as convex hull, Polsby-Popper, and Reock).

After breaking down the districts into three groups for each of the two measures (high, medium, and low compactness)[1], roughly 66% belonged to the same group level according to both measures, while 32% differed by one group and 9 districts (roughly 2%) differed by two groups across the two measures. These 9 districts can be seen in Figure 13 in the Appendix. The two districts in Figure 13 (a) are in the top third for the EDF measure and the bottom third for Kaufman, et al.'s measure, and vice versa for the districts in Figure 13 (b). It's particularly notable that the districts in Figure 13 (b) can all be categorized as elongated rectangular. These have lower EDF compactness scores since the central "core" portions tend to be more concentrated or compressed deep within the shape, while the "extra details" portions tend to be more substantial. Districts with a high aspect ratio such as elongated rectangles may display a desirable degree of geometric regularity, but could potentially be drawn that way to, e.g., capture a narrow voter corridor, or be less coherent in terms of preserving communities of interest. On the other hand, both districts in Figure 13 (a) appear to fall behind in Kaufman, et al.'s measure due to the offshoot arms and more jagged edges. We leave it to the reader to "know it when they see it" and determine for themselves which degree of compactness they deem more appropriate for these outlier districts.

5 Discussion

The novel medial-axis-based compactness measure introduced herein is not intended to replace all extant compactness measures in the conversation surrounding (re)districting. Rather, our intent is to provide a new mathematical measure built directly on the framework of what is understood about the process of human shape recognition, and demonstrate its utility in this arena. On the whole, the EDF measure favors many of the same general shape features as noted by Kaufman, et al., i.e., "squarish, with minimal arms, pockets, islands, or jagged edges". The most significant difference is that the EDF measure devalues perturbations occurring in the "extra details" portions of the shape and along the district boundaries.

We are currently pursuing various avenues for expanding our analysis of this measure. First, it would be interesting to conduct an additional user study where participants are asked to partition districts into "core", "parts", and "details" regions, in line with the

[1]We considered splitting the districts into three groups using 3-means clustering instead of tertiles. The issue is that the low/medium/high group sizes end up being quite different for the EDF measure and Kaufman, et al.'s measure when broken up by 3-means clustering. For instance, for the 2016 congressional districts, one of the measures had 44 districts in the top group and 231 in the bottom group, whereas the other had 149 districts in the highest group and 97 in the lowest group. This made it more challenging to do an accurate side-by-side comparison vs. using tertiles, which is why we went with the latter.

study in [21]. Comparing the divisions made by users to those determined by the EDF measure will inform future decisions about adjusting the computations involved in our method and give further insight into how humans innately process the meaning of the word "compact" with respect to redistricting. Second, in ongoing future work, we plan to augment Kaufman, et al.'s predictive model by building in the EDF compactness scores as features and examining the impact on the model's performance. This impact could be assessed via a user study to obtain human judgments of compactness on a specific data set, and subsequently comparing the results with the model's predictions both with and without the EDF scores included as features. It would be interesting to see whether or not the inclusion of the EDF measure in Kaufman, et al.'s method leads to greater predictive accuracy via improved agreement rates between their model and human responses.

Third, the use of 3-means clustering of the medial axis points by EDF values, as opposed to 2-means, has the effect of condensing the range of the output compactness scores, with the majority of values falling between 0.75 and 0.95. The introduction of the third cluster captures a smoother gradient of shape variation and was done to match the "core, parts, details" perspective, but perhaps this could be adjusted in some way to generate values that more readily convey the differences between districts. For instance, the smaller range in EDF scores when using 3-means clustering vs. 2-means can make the scale harder to parse, and enhances the fuzziness about where the "right" numerical threshold for sufficient compactness should be. Whether a larger or smaller range of compactness scores is preferable depends on the intended application. A broader range enhances discriminatory power and increases contrast between highly regular and highly irregular districts. On the other hand, a narrower range may improve robustness, reduce noise sensitivity, and discourage over-interpretation of minor shape differences. We view this tradeoff as a design choice, and future work may explore adaptive or weighted clustering schemes that balance sensitivity and robustness more dynamically.

Appendix

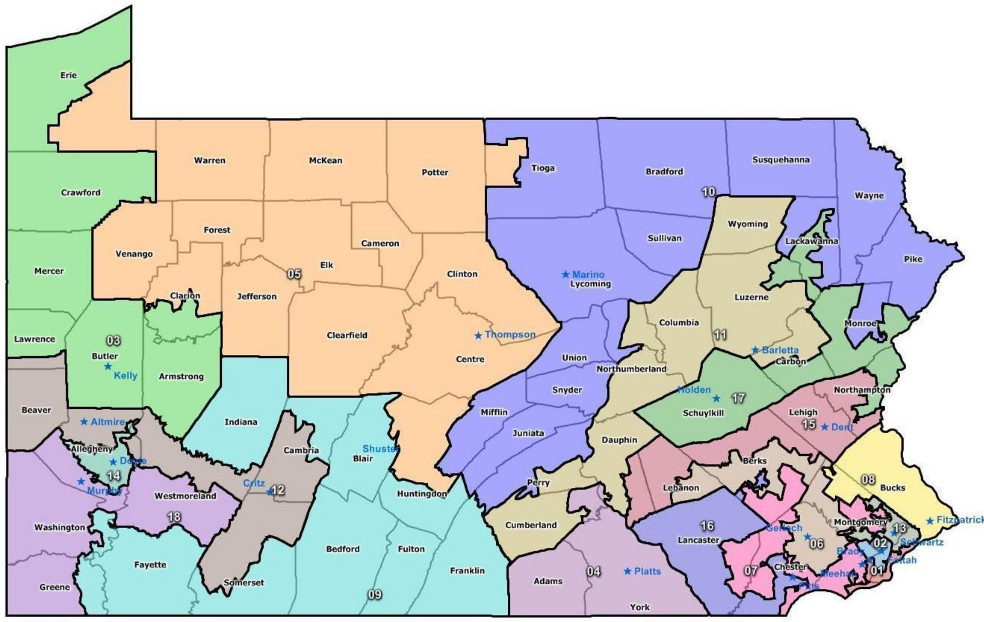

Figure 8: Pennsylvania congressional district map in 2011 [27].

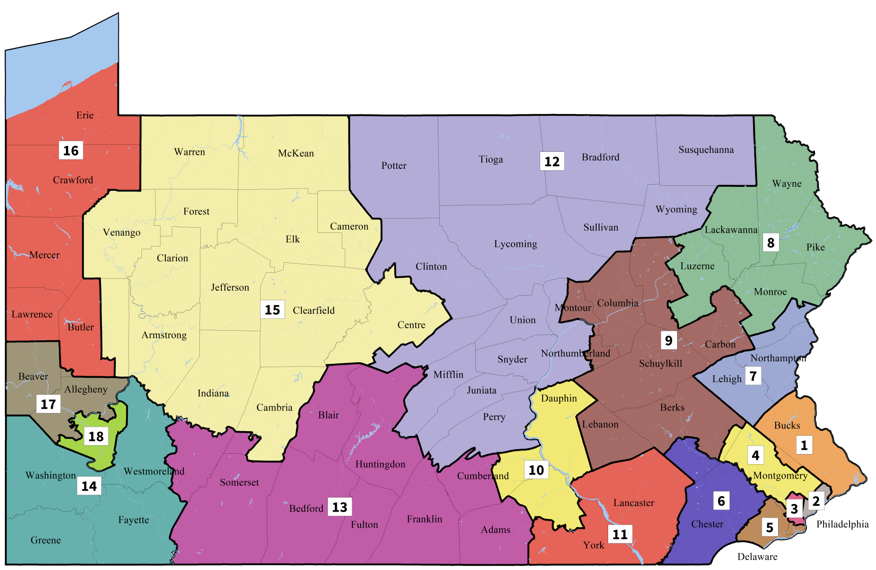

Figure 9: Pennsylvania congressional district map in 2018 [28].

Pennsylvania Kaufman, et al. Scores		
District #	**2011**	**2018**
1	12.47	67.33
2	63.92	70.26
3	38.33	55.37
4	63.33	53.08
5	45.92	62.70
6	11.95	47.76
7	9.09	64.71
8	53.39	51.24
9	26.99	48.65
10	24.19	51.80
11	17.53	65.70
12	11.66	48.90
13	15.54	49.50
14	19.68	56.33
15	19.93	68.23
16	19.50	63.22
17	10.96	47.39
18	29.47	33.82
Average	**27.43**	**55.89**

Table 4: Kaufman, et al.'s compactness values for Pennsylvania congressional districts in 2011 and 2018.

Pennsylvania Kaufman, et al. Score Comparisons				
2018 District	**2011 District**	**2011 Kaufman**	**2018 Kaufman**	**Δ**
1	8	53.39	67.33	+13.94
2	13	15.54	70.26	+54.72
3	2	63.92	55.37	−8.55
4	13/6	13.75	53.08	+39.34
5	7	9.09	62.70	+53.61
6	6	11.95	47.76	+35.81
7	15	19.93	64.71	+44.78
8	17	10.96	51.24	+40.28
9	9	26.99	48.65	+21.66
10	4	63.33	51.80	−11.53
11	16	19.50	65.70	+46.20
12	10	24.19	48.90	+24.71
13	9	26.99	49.50	+22.51
14	18	29.47	56.33	+26.86
15	5	45.92	68.23	+22.31
16	3	38.33	63.22	+24.89
17	12	11.66	47.39	+35.73
18	14	19.68	33.82	+14.14

Table 5: Change in Kaufman, et al. compactness scores from 2011 to 2018 in Pennsylvania.

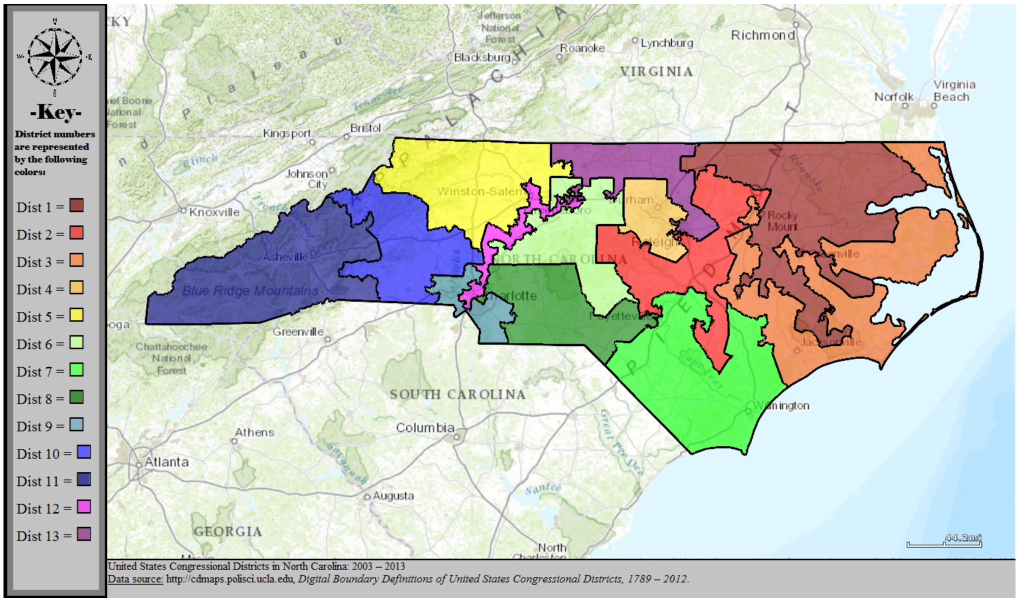

Figure 10: North Carolina congressional district map in 2010 [24].

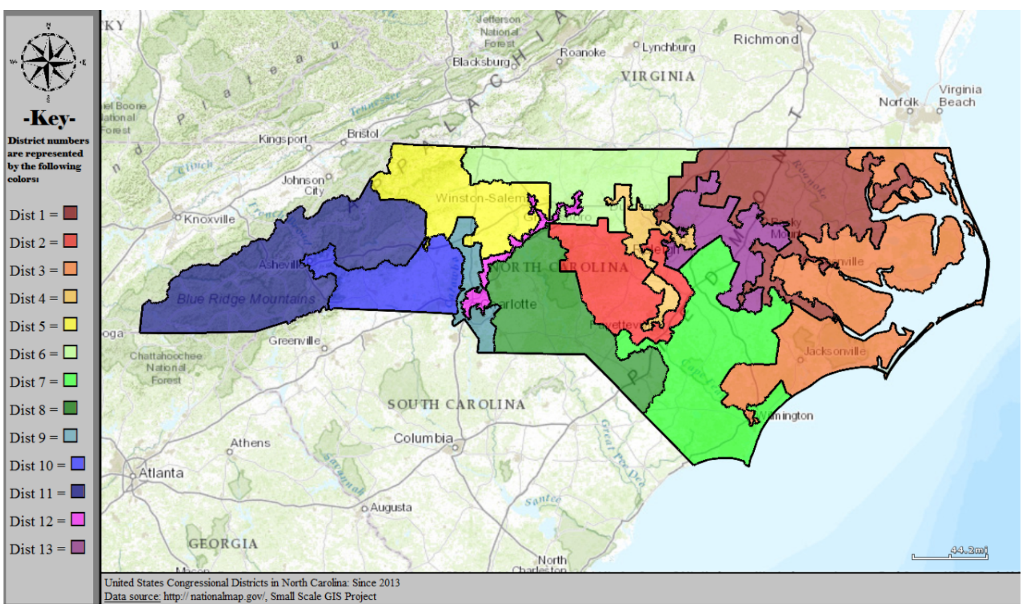

Figure 11: North Carolina congressional district map in 2011 [24].

North Carolina 2010 Compactness Values		
District #	**EDF**	**Kaufman**
1	0.81	20.21
2	0.71	13.43
3	0.72	15.38
4	0.91	47.17
5	0.81	47.82
6	0.75	16.48
7	0.84	49.13
8	0.86	39.04
9	0.70	18.75
10	0.84	33.34
11	0.81	50.31
12	0.62	8.45
13	0.69	16.28
AVG	0.77	28.91

Table 6: Compactness values for the 2010 North Carolina congressional district map.

North Carolina 2011 Compactness Values		
District #	**EDF**	**Kaufman**
1	0.75	20.01
2	0.72	28.99
3	0.82	28.47
4	0.69	16.18
5	0.81	28.18
6	0.73	33.23
7	0.71	25.40
8	0.83	24.71
9	0.54	17.76
10	0.82	29.22
11	0.77	34.08
12	0.54	15.15
13	0.68	20.52
AVG	0.72	24.76

Table 7: Compactness values for the 2011 North Carolina congressional district map.

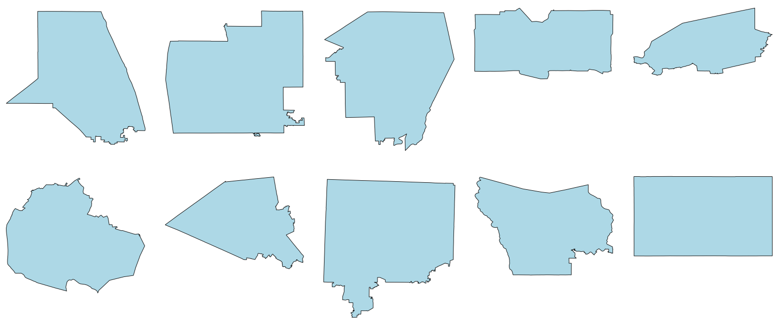

Figure 12: Districts in the top 25 most compact according to both EDF and Kaufman score.

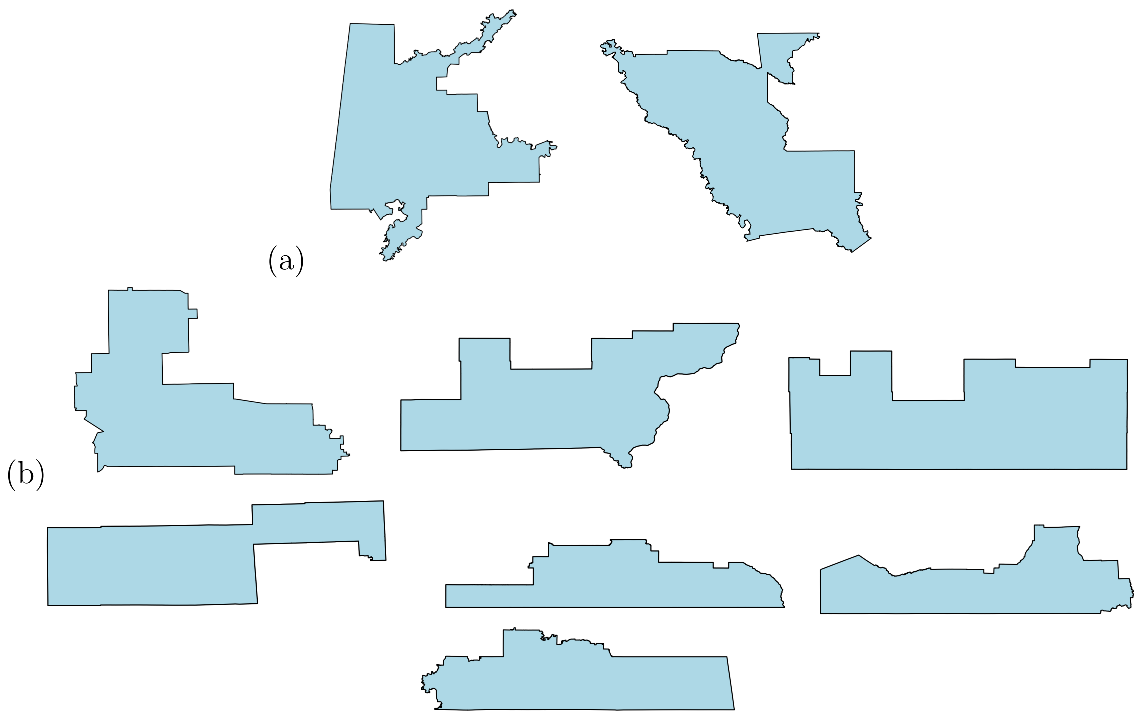

Figure 13: Districts that differed by two compactness group levels according to the EDF and Kaufman measures. (a): Those in the highest compactness group for EDF and the lowest for Kaufman, et al. (b): Those in the highest compactness group for Kaufman, et al. and the lowest compactness group for EDF.

Acknowledgments

Thank you to the anonymous reviewer whose very helpful suggestions helped us improve the paper.

References

[1] V. Ayzenberg, Y. Chen, S.R. Yousif, S.F. Lourenco, Skeletal representations of shape in human vision: Evidence for a pruned medial axis model, *J. Vis.*, **19** (2019), 1–21.

[2] V. Ayzenberg, S.F. Lourenco, Skeletal descriptions of shape provide unique perceptual information for object recognition, *Sci Rep*, **9** (2019), 1–13.

[3] V. Ayzenberg, F.S. Kamps, D.D. Dilks, S.F. Lourenco, Skeletal representations of shape in the human visual cortex, *Neuropsychologia*, **164** (2022), 108092.

[4] R. Barnes, J. Solomon, Gerrymandering and compactness: Implementation flexibility and abuse, *Political Anal.*, **29** (2021), 448–466.

[5] E. Blanchard, K. Knudson, Measuring Congressional District Meandering, available online at the URL: `https://arxiv.org/abs/1805.08208`.

[6] H. Blum, A transformation for extracting new descriptions of shape, *Models for the perception of speech and visual form* (1967), 362–380.

[7] V. Bougou, M. Vanhoyland, P. Janssen, T. Theys, Neuronal tuning and population representations of shape and category in human visual cortex, *Nat. Commun.*, **15** (2024).

[8] C.A. Burbeck, S.M. Pizer, Object representation by cores: Identifying and representing primitive spatial regions, *Vis. Res.*, **35** (1995), 1917–1930.

[9] C.A. Burbeck, S.M. Pizer, B.S. Morse, D. Ariely, G.S. Zauberman, J.P. Rolland, Linking object boundaries at scale: a common mechanism for size and shape judgments, *Vis. Res.*, **36** (1996), 361–372.

[10] The United States Census Bureau TIGER/Line Shapefiles database, available online at the URL: `https://www.census.gov/cgi-bin/geo/shapefiles/index.php`

[11] M. Duchin, B.E. Tenner, Discrete Geometry for Electoral Geography, *Pol. Geogr.*, **109** (2024), 103040.

[12] J.D. Evans, *Straightforward Statistics for the Behavioral Sciences*, Brooks/Cole Publishing Company, Pacific Grove, CA (1996).

[13] C. Firestone, B.J. Scholl, "Please tap the shape, anywhere you like": Shape skeletons in human vision revealed by an exceedingly simple measure, *Psychol. Sci.*, **25** (2014), 377–386.

[14] R.G. Fryer Jr., R. Holden, Measuring the Compactness of Political Districting Plans, *J.L. & Econ.*, **54** (2011), 493–535.

[15] D.E. Hinkle, W. Wiersma, S.G. Jurs, *Applied Statistics for the Behavioral Sciences*, Houghton Mifflin, Boston, MA (2003).

[16] C.-C. Hung, E.T. Carlson, C.E. Connor, Medial axis shape coding in macaque inferotemporal cortex, *Neuron*, **74** (2012), 1099–1113.

[17] A.R. Kaufman, G. King, M. Komisarchik, How to measure legislative district compactness if you only know it when you see it, *Am. J. Political Sci.*, **65** (2021), 533–550.

[18] I. Kovacs, B. Julesz, Perceptual sensitivity maps within globally defined visual shapes, *Nature*, **370** (1994), 644–646.

[19] League of Women Voters of Pennsylvania v. the Commonwealth of Pennsylvania, Ballotpedia, available online at the URL: `https://ballotpedia.org/League_of_Women_Voters_of_Pennsylvania_v._the_Commonwealth_of_Pennsylvania`

[20] T.S. Lee, D. Mumford, R. Romero, V.A.F. Lamme, The role of the primary visual cortex in higher level vision, *Vis. Res.*, **38** (1998), 2429–2454.

[21] K. Leonard, G. Morin, S. Hahmann, A. Carlier, A 2D shape structure for decomposition and part similarity, *23rd ICPR*, (2016), 3216–3221.

[22] F.F. Leymarie, P. Aparajeya, Medialness and the perception of visual art, *ARTP*, **5** (2017), 169–232.

[23] L. Liu, E.W. Chambers, D. Letscher, T. Ju, Extended grassfire transform on medial axes of 2D shapes, *Comput. Aided Des.*, **43** (2011), 1496–1505.

[24] Wikipedia contributors, North Carolina's congressional districts, available online at the URL: `https://en.wikipedia.org/wiki/North_Carolina%27s_congressional_districts`

[25] R.G. Niemi, B. Grofman, C. Carlucci, T. Hofeller, Measuring Compactness and the Role of a Compactness Standard in a Test for Partisan and Racial Gerrymandering, *JOP*, **52** (1990), 1155–1181.

[26] M. Wines, R. Fausset, North Carolina Is Ordered to Redraw Its Gerrymandered Congressional Map. Again., The New York Times, Aug. 27, 2018, available online at the URL: `https://www.nytimes.com/2018/08/27/us/north-carolina-congressional-districts.html`

[27] PennLive, Can Pennsylvania's policy-makers really build a better Congressional map in 17 days?, available online at the URL: `https://www.pennlive.com/politics/2018/01/mapquest_can_pennsylvanias_pol.html`

[28] Pennsylvania Department of State, 2018 Remedial Congressional Districts, available online at the URL: `https://www.dos.pa.gov/VotingElections/CandidatesCommittees/RunningforOffice/Pages/2018-Remedial-Congressional-Districts.aspx`

[29] D.D. Polsby, R.D. Popper, The third criterion: Compactness as a procedural safeguard against partisan gerrymandering, *YLPR*, **9** (1991), 301–353.

[30] E.C. Reock, A note: Measuring compactness as a requirement of legislative apportionment, *Midwest J. of Political Sci.*, **5** (1961), 70–74.

[31] J.E. Schwartzberg, Reapportionment, gerrymanders, and the notion of compactness, *Minn. L. Rev.*, **50** (1965), 443–443.

[32] K. Siddiqi, B.B. Kimia, A. Tannenbaum, S.W. Zucker, On the psychophysics of the shape triangle, *Vis. Res.*, **41** (2001), 1153–1178.

[33] S. Wolf, Daily Kos Elections Live Digest 4/16, available online at the URL: `https://www.dailykos.com/stories/2018/4/16/1757278/-Daily-Kos-Elections-Live-Digest-4-16`

[34] P. Zhang, D. DeFord, J. Solomon, Medial axis isoperimetric profiles, *CGF*, **39** (2020), 1–13.

Jason D'Amico
Catherine Cutler Institute at the University of Southern Maine
Wishcamper Center, 34 Bedford Street
Portland, ME 04104, USA
E-mail: `jasondamico26@gmail.com`

Ellen Gasparovic
Union College
807 Union Street
Schenectady, NY 12309, USA
E-mail: `gasparoe@union.edu`

Greg Malen
Skidmore College
815 North Broadway
Saratoga Springs, NY 12866, USA
E-mail: `gmalen@skidmore.edu`

Mushan Zhong
Simon-Kucher
225 Liberty St, Brookfield Place 41st Floor
New York, NY 10281, USA
E-mail: zhongmushan98@gmail.com

Received: June 20, 2025 **Accepted**: September 5, 2025
Communicated by Matthias Beck