TITLE:

A Shape Analysis Algorithm Quantifies Spatial Morphology and Context of 2D to 3D Cell Culture for Correlating Novel Phenotypes with Treatment Resistance

AUTHORS AND AFFILIATIONS:
David H. Nguyen, Ph.D., Tissue Spatial Geometrics Laboratory, Concord, CA, USA

Michael Bruck, Department of Medicine, University of California San Francisco, San Francisco, CA, USA

Jennifer M. Rosenbluth, M.D., Ph.D., Department of Medicine, University of California San Francisco, San Francisco, CA, USA; Chan Zuckerberg Biohub, San Francisco, CA, USA

SUMMARY:
This article explains how novel features of morphology in cells and organoids can be measured by the Linearized Compressed Polar Coordinates (LCPC) Transform, a spatial algorithm that measures what traditional metrics, such as area, volume, surface area, etc., cannot. Best practices for shape orientation and alignment are discussed.

ABSTRACT:
Numerous studies have shown that the morphological phenotype of a cell or organoid correlates with its susceptibility to anti-cancer agents. However, traditional methods of measuring phenotype rely on spatial metrics such as area, volume, perimeter, and signal intensity, which work but are limited. These approaches cannot measure many crucial features of spatial context, such as chirality, which is the property of having left- and right-handedness. Volume cannot register chirality because the left shoe and right shoe hold the same amount of volume. Though spatial context in the form of chirality, direction of gravity, and the axis of polarity are intuitive notions to humans, traditional metrics relied on by cell biologists, pathologists, radiologists, and machine learning practitioners up to this point cannot register these fundamental notions. The Linearized Compressed Polar Coordinates (LCPC) Transform is a novel algorithm that can capture spatial context unlike any other metric. The LCPC Transform translates a two-dimensional (2D) contour into a discrete sinusoid wave via overlaying a grid system that tracks points of intersection between the contour and the grid lines. It turns the contour into a series of sequential pairs of discrete coordinates, with the independent coordinate (x-coordinate) being consecutive positions in 2D space. Each dependent coordinate (y-coordinate) consists of the distance, between an intersection of the contour and gridline, to the origin or baseline of the grid system. In the form of a discrete sinusoid wave, the Fast Fourier Transform is then applied to the data. In this way, the shape of cells in 2D and 3D cell culture are represented systematically and multidimensionally, allowing for robust quantitative stratification that will reveal insights into treatment resistance.

INTRODUCTION:

Three-dimensional (3D) organoid culture has proven superior to two-dimensional (2D) cell culture in mimicking *in vivo* biology[1,2]. Organoids have become indispensable in cancer research for the purposes of screening for effective treatments and extracting insights into disease progression. While it is plain knowledge in the field of cancer biology that distinctions in morphology correlate with the distinct biological behavior of cells and tissues, the field relies on traditional shape metrics that are limited in their scope. This study describes a novel spatial algorithm that objectively quantifies subtle features of morphology in unprecedented precision compared to the traditional approaches of area, volume, surface area, etc.

Numerous studies have shown a correlation between tumor organoid morphology and treatment resistance. Patient-derived colorectal cancer organoids exhibit distinct 3D morphologies, such as solid structure, cystic structure, and budding morphology, which correlate with sensitivity or resistance to drug therapies[3]. Furthermore, colon cancer organoids derived from patients that were either sensitive or resistant to oxaliplatin chemotherapy exhibited distinct morphologies. Resistant organoids that underwent genetic knockout of the long non-coding RNA lnc-RP11-536 K7.3 become sensitive to oxaliplatin, but the surviving organoids exhibited a morphology that was distinct from that at the start of treatment[4]. Hadj Bachir et al.[5] developed tumor organoids from Pancreatic Ductal Adenocarcinoma that were resistant to FOLFIRINOX treatment. Compared to organoids that were sensitive to treatment, when the resistant organoids reached developmental maturity and exhibited folding patterns, their folding was more complex than those observed in the sensitive organoids, highlighting clear morphological differences associated with chemoresistance. Masako[1] et al. studied thirteen triple negative breast cancer cell lines and found that those what grew in 3D as spheroids were more resistant to chemotherapy than those described as grape-like. Zhao[6] et al. showed that certain mesenchymal mammary tumor organoids exhibited a spikey morphology, which changes to a jagged-edge morphology upon induction of miR-200c, a micro-RNA that induces mesenchymal-to-epithelial (MET) transition. Based on this distinct phenotypic change, the authors developed a morphology-based screening method to identify drugs that can induce MET and thereby enhance chemosensitivity.

While there is ample evidence across multiple types of cancer, beyond the studies cited here, that organoid morphology is indicative of treatment resistance, studies up to this point rely on metrics such as diameter, area, volume, and surface area. The existence of automated high-throughput cell culturing techniques that include imaging components means that morphology data is accessible as part of the research and development pipeline. However, without methods that robustly quantify subtle differences in morphology, it is difficult to discover unknown morphological subtypes and/or leverage the power of machine learning to correlate these subtypes with disease progression and treatment resistance.

The Importance of Spatial Context for Biological Morphology

Biology and nature are full of spatial contexts that are crucial to proper function. Epithelial cells have inherent morphological spatial context, such as apical-basal polarity, to execute their

function of forming barriers and directionally secreting products[7,8,9,10]. When cells migrate 2D cell culture, the leading edge of the cell has a distinct morphology compared to the trailing edge. This asymmetry in cell morphology during migration is essential for sensing the environment and for anchoring the cell for generating force in specific directions.

Another crucial spatial context in nature is the direction of gravity. Bone health, muscle health, and vascular health in astronauts become dysfunctional due to prolonged exposure to microgravity conditions[11, 12]. Not surprisingly, cancer cells behave very differently in microgravity compared to being on Earth, including alterations in their morphology as single cells and in 3D colonies[13,14,15]. A third spatial context in cancer is that in situ pre-neoplasia and neoplasia can sprout from a base, or trunk, which is its central support structure. Pedunculated colon polyps are an example of this morphology. As the lesion regresses, it changes shape towards this central support structure, which is also its bridge to the stable vascular network that carries oxygen[16,17].

The Limits of Traditional Shape Metrics

Traditional shape metrics used in cancer research and cell biology include measuring area, surface area, volume, perimeter, etc. While these metrics are informative, they cannot register spatial contexts that are obvious to humans. The apical-basal polarity of a cell in a lumen, the direction of gravity, and the direction of the central support structure is non-existent to traditional approaches that cannot register spatial context.

Traditional metrics also cannot distinguish gradations of shape change between common shape concepts, such as a square that is standing on one of its flat sides and a diamond that is the same square but rotated so that it stands vertically on one corner. Any intermediate forms resulting from rotating the square about its center to become the diamond are subjective interpretations. Are the intermediate steps “rotated squares” or “rotated diamonds”? Measuring area and perimeter to answer this question is pointless because the square and diamond are equivalent by these two metrics. Another shortcoming of traditional shape metrics is that they cannot distinguish between size and shape. A series of circles that gradually increase in size will exhibit increasing area and perimeter measures, even though the shape concept never changes, meaning the circles remain circles.

A Novel Algorithm and Its Creative Augmentations Can Register Spatial Context

This study presents the Linearized Compressed Polar Coordinates (LCPC) Transform that was invented with the intention of objectively and quantitatively describing complex morphologies observed in tissue histopathology for the purposes of enhancing the grading of colon polyps. The desire to then expand this approach to capture macroscopic MRI-based spatial features in brain pathologies, such as Bipolar Disorder and Alzheimer’s Disease led to augmentations that enhanced it even further, which are described herein.

Open-source Python scripts have been developed and made publicly available, including scripts that do pre-processing steps such as contour extractions and rotations. This study describes the step-by-step protocol and best practices for applying two versions of LCPC Transform: the

parallel grid system and the radial grid system. Advice is provided on when to use which system for obtaining optimal results. The GitHub repository containing the scripts includes multiple sets of practice images and their results, so that users can replicate the results to ensure that the scripts are working as intended.

The LCPC Transform provides an unprecedented level of precision in measuring spatial information because it represents shapes in multiple dimensions. Unlike traditional methods, like area and volume, which yield only one scalar value per shape (i.e. 25 cm^2), the results of the LCPC Transform can yield multiple indices, each correlating with a different morphological aspect of a shape (i.e. roundness vs. sharpness of corners, smoothness vs. jaggedness of edges). The LCPC Transform is poised to revolutionize shape analysis for many fields of science, allowing users to measure what might be sub-dimensions of spatial information in between the x-, y-, and z-dimensions. It is now possible to measure what is hidden to the naked eye and to the traditional approaches derived from classical geometry. Like the Ancient Greek mathematician Archimedes who realized that the volume of water displaced by an object is also a measure of the object's density, this study describes a "Eureka" moment that will allow humans to measure nature in whole new ways.

How Does the LCPC Transform Work?

The LCPC Transform describes a 2D shape by overlaying a grid system of straight lines that intersect with the contour of the shape. Each grid system has an origin (Figure 1A) or a baseline (Figure 2A) from which to measure linear distance. Every point of intersection between the gridlines and the contour of the shape is detected. The distance of each intersection is then calculated relative to a baseline or origin. In this way, the LCPC Transform converts 2D shapes into a series of consecutive x and y pairs of coordinates. The x-coordinate represents the position of the gridline from zero to infinity, while the y-coordinate represents the distance of the intersection to the baseline or origin. In this form, which is a discrete sinusoid wave, the Fast Fourier Transform (FFT) is then applied to convert the data from the "position domain" to the frequency domain. If the x-coordinate represented time, then the "position domain" would be equivalent to the "time domain" in the standard applications of the FFT.

While the radial grid system measures the distance of each intersection from the origin of the radial grid (Figure 1A), the parallel grid system always measures the distance of intersections with reference to an imaginary line at the left of the shape (Figure 2A). The open-source script determines the position of this imaginary line by finding the left-most pixel of the contour and then moving 10 pixels to the left of this position. Here, the x-coordinate of this position becomes the reference point from which all intersections are calculated. This 10-pixel rule is arbitrary but is why all contours analyzed by the open-source script for the parallel grid system must have at least 15 pixels of white space on all four sides of a contour.

For 2D shapes that have folds or multiple layers, gridlines may intersect with the contour more than once. In this case, the distances of all intersections on a gridline are summed into one value. Thus, each gridline has only one x-coordinate and one y-coordinate. This summation is represented by the term "compressed" in the name LCPC Transform. The rendering of a non-

linear 2D shape into a discrete sinusoid wave is represented by the term “linearized” in the algorithm’s name. Lastly, the term “polar coordinates” is in the algorithm’s name because the first grid system conceptualized was a 180-degree radial grid of polar coordinates[18]. Even after the realization that polar coordinates and Cartesian coordinates are interchangeable, the name of the algorithm was kept as is. The acknowledgements section describes the personal reasons that motivated the invention and augmentation of the LCPC Transform.

The frequency domain of the LCPC Transform reveals objective and quantitative features of a 2D shape that the naked eye cannot readily perceive. Very subtle differences in curvature, jaggedness, and tortuosity can be robustly quantified. However, once morphological subtypes are identified via this method, visually distinct themes can be observed that correlate with the quantitative differences. This feature makes the LCPC Transform interpretable as a method of extracting spatial information for machine learning (ML) models and makes the resulting ML models themselves more interpretable. While neural networks are a powerful ML method, their results are often not interpretable to humans, which is a significant problem for the field of medical artificial intelligence (AI). Physicians are reluctant to employ an ML model for clinical decisions if they don’t understand why the model makes bad decisions. Furthermore, if no human understands why or how a model made a biased decision, then fixing this becomes intractable.

This manuscript describes only two grid systems for the LCPC Transform, though many grid systems can be envisioned. It is important to note that while all grid systems can extract spatial information, some will provide cleaner data than others. Thus, users are encouraged to apply both the parallel and the radial grid systems to their data to determine optimal fit. For simplicity and reproducibility, the open-source Python scripts only include a 360-degree radial grid system[19] and a parallel grid system[20], which will suffice for most applications. Again, all grid systems can extract spatial information, but some are better than others depending on the spatial contexts described in this manuscript. The aforementioned GitHub repositories contain links to YouTube videos that are tutorials on how to run the scripts on the associated practice images and how to do quality control checks on the outputs. It is highly recommended that users learn how and where to do quality control checks.

The Radial Grid System vs. The Parallel Grid System

In general, the radial grid system (Figure 1B) is good for shapes that are radially symmetric, round, and closed. However, it also works well for shapes that are open (meaning the contour has at least one disconnected region) but curved in on itself (like the letter “c”). This grid system also works for contours that consist of multiple layers or even contours that have multiple non-intersecting components. The parallel grid system, on the other hand, is good for shapes that are elongated and open (Figure 2B). Though it was previously stated that the radial grid system works well for closed shapes, closed shapes that are elongated sometimes yield cleaner results when assessed by the parallel grid system. Just as with the radial grid system, the parallel grid system can measure the spatial information of objects that consist of multiple non-intersecting components.

Importance of Standardizing Shape Orientation Prior to the LCPC Transform

As discussed in the introduction, nature is inherently polar, which includes the shapes of cells and organoids. Cells can have apical-basal polarity and organoids can have asymmetries that indicate poles. Thus, orienting shapes in a systematic and objective manner before extracting spatial information via the LCPC Transform will yield cleaner, more reproducible results. Figure 3A shows two shapes that are identical but mirror images of each other. Though they are the same shapes, because the parallel grid system always measures distances of intersections from the left side of the image, the parallel grid LCPC will yield different numbers for each image. The LCPC Transform was designed to capture spatial context, so it is very sensitive to seemingly minor differences in spatial context. A better way to compare Shape 2 to Shape 1 in Figure 3 is to flip Shape 2 horizontally, which changes the spatial orientation but doesn't change the shape, and then apply the parallel grid LCPC (as shown in Figure 3B).

Compared to traditional shape metrics, such as area, volume, and perimeter, that are not affected by spatial orientation, the LCPC Transform requires the user to think about how shapes should be oriented before measurement so that the optimal results are produced. The volume of liquid in an unopened can of soda does not change depending on whether the can is right-side up, upside down, or any other orientation.

When measuring shapes with the LCPC Transform, it is helpful to determine if the shapes should be oriented in a systematic way beforehand. Figure 3C shows four items that are shaped like the number "3" and arranged haphazardly. Unless there is a reason to keep each shape in its original orientation, the user should orient each shape like in Figure 3D. The rules for systematically aligning all items in Figure 3C into Figure 3D are: (1) The object is rotated such that the line that connects the opening of the "3" is vertical, (2) the wider part is on top, (3) and all objects have their concavity facing to the left so that the bulge of "3" pushes away from the left of the object where the parallel grid LCPC will place the reference baseline for measuring distance. Note that these rules require that Item 1 and Item 4 be both rotated and flipped horizontally, while Item 2 and Item 3 only need rotation. Flipping horizontally changes the spatial orientation of a shape, but not necessarily the shape. The user needs to determine what orientation rules to apply, and why, to every object, which will dramatically improve the cleanliness, and thus reproducibility, of the data.

It is important to note that the LCPC Transform can be applied to sub-compartments of a shape instead of the whole shape. Morphologies in nature are highly complex, so scientists often assume insignificance for minor features of morphology to simplify the system being studied. However, the LCPC Transform reveals that seemingly minor differences in morphology can be quantitatively measured and correlated with distinct biological function. Thus, Figure 3E shows how the wider bulge of the items in Figure 3C can be extracted and objectively compared. However, Figure 3F shows how the user can compare these wider bulges in the context of the alignment rules in Figure 3D. Figure 3F and 3D are meant to highlight the fact that the user must determine what spatial context is relevant and why.

Deriving Informative Indices from Results of the Fast Fourier Transform

Though this manuscript refers to the LCPC Transform as including the step of applying the FFT, the actual LCPC Transform happens before FFT step (Figure 1A Step 4, Figure 2A Step 3). However, the effect of applying the FFT is to reveal objective, systematic insights from the frequency domain of the output of the LCPC Transform. Thus, the LCPC Transform would not be as effective without including the FFT as the final step.

In the LCPC Transform, the FFT results in a series of bins and their magnitudes, all of which are a static multidimensional representation of the contour inputted in Step 1. An analogy of a static multidimensional representation of information is the nine-digit phone number representing a person's telephone communication channel in the United States. While cumbersome in comparison to scalar value representations of shapes (such as 85 cm in length, 56 degrees wide, etc.), multidimensional data harbors much more information for extracting insights.
The typical result of the LCPC Transform, including the last FFT step, is a histogram of magnitudes as a function of position (Figure 1A Step 5). The first bin has the largest magnitude, and subsequent bins can exhibit various repeating patterns. Figure 4 displays several hypothetical FFT results that have been observed. Each shape analyzed by the LCPC Transform results in many frequency bins and their magnitudes that represent it, so using an unsupervised dimension reduction technique, such as Principal Component Analysis (PCA), is useful for exploring which bins are responsible for stratifying the experiment groups most. However, a complementary approach is to derive a single-valued scalar index from the FFT results to represent each shape. These indices allow for the ease of employing descriptive statistics, such as the t-test or Wilcoxon Rank-Sum Test, to show differences in averages between experimental groups. They also allow for multiple distinct indices to be derived from the same FFT output, each of which *may* represent different morphological features in the shapes being analyzed. This ability to provide multiple indices that each represent a different feature of morphology is unprecedented compared to traditional metrics, such as measuring the area within an object, which provides no more than one value for area. Figure 4A-4C show common patterns in bin magnitudes that can be used to create indices in the form of ratios. Figure 4A shows that there are three indices that can be derived based on the patterns highlighted by the dotted lines. Indices can be ratios between two bins, a sum of 2 or more bins, or a coefficient of a curve fitted onto multiple consecutive bins. Figure 4D shows hypothetical shape features that can correlate with different indices derived from the same FFT profile (data from actual experiments not shown).

What is Pure Shape vs. At-Scale Shape?

One of the limitations of traditional shape metrics, such as area and perimeter, is that they measure both the shape of an object and the size of an object at the same time. Figure 5A shows a series of check mark-shaped objects that increase in size but are the same shape concept. While the area and perimeter of each check mark change, the shape concept stays the same, but this is undetectable by area or perimeter. Figure 5B shows the concept of a square that becomes a diamond if it is rotated by 45 degrees. However, for any rotation between 0 and 45 degrees, the shape concept can be subjectively interpreted as either a "tilted square" or a "tilted diamond". Through this series of rotations, the area and perimeter of each object don't

change while the shape concept does. The examples in Figure 5A and 5B highlight the inability of traditional shape metrics, such as area, volume, perimeter, and surface area in capturing notions that are obvious to humans.

When measuring shapes, the LCPC Transform can separate the effect of size from the effect of shape. In combination with an objective and systematic method of resizing objects, the LCPC Transform can measure what will be called “pure shape”, which is shape without the effect of size. Figure 6 shows two objects that are different in both shape and size. If they are not resized before being measured by the LCPC Transform, then they are being compared “at-scale”, meaning at their original scales relative to each other. However, by applying an objective set of rules for rotation and resizing, the two objects can be transformed to a similar scale while keeping their original shape. Figure 6 Step 2 shows how the longest internal line of each object is an objective morphological landmark by which the object can be rotated such that this line is horizontal. Step 3 of Figure 6 then shows how the width of both objects can be resized to be 400-pixels wide while also constraining the aspect ratio so that the object is not skewed (not constraining the aspect ratio while resizing a square turns it into a rectangle). In this way, both objects now have the same width. 400-pixels wide is an arbitrary value written into the open-source script. Lastly, Step 4 of Figure 6 shows an optional step of rotating the longest internal line within each object to be vertical, such that they get optimal coverage from the parallel grid LCPC Transform. If the radial grid LCPC Transform were applied, then this last rotation is unnecessary. However, it is suggested that the longest axes of objects be made vertical in case the user wants to also apply the parallel grid LCPC Transform on the same objects for comparison with results from the radial grid LCPC Transform. In the vertical position, the parallel grid LCPC Transform applies the greatest coverage across the object because the parallel lines are stacked vertically.

Adding Markers to Capture Spatial Orientation in Nature

Because traditional metrics, such as area, volume, and surface area, are agnostic to spatial context such as the direction of gravity, the direction of the strongest sunlight, and the axis of polarity, the LCPC Transform provides a unique opportunity to insert spatial markers that capture these contexts. As described previously, how one rotates and orients a shape can also capture spatial context. Figure 7A shows another example of this in the form of cell migration in 2D culture. Figure 7B, however, is an example of how adding rationally placed information to the mask image allows the LCPC Transform to capture spatial context in exquisite ways. Figure 7B shows a hypothetical tumor or pre-cancerous lesion that is anchored outside of its encasement. The lesion grows or regresses within the encasement, while the encasement does not change. Thus, as the lesion grows away from, or regresses back to, its anchor due to therapeutic treatment, it does so relative to an encasement that remains static. This means that measures of area, perimeter, and volume (if the imaging context is 3D) would completely miss the fact that the spatial change in the lesion is happening within the context of an encasement. Why is this significant? In disease progression, the shape of cells, tissues, and organs do not change uniformly on all sides. They often grow away from or shrivel towards their central support structure. Furthermore, the “empty” space that they create relative to their encasement can have biological significance. Traditional shape metrics cannot capture the

change of this “empty” space. Figure 7C shows two options for applying the parallel grid LCPC Transform to the contour of the lesions. The orange line in Figure 7B represents the location of the encasement directly across from the leading edge of the lesion. Including this orange line in the mask of the lesion, as shown in Figure 7C, captures the fact that there is significantly more empty space between Specimen 1 and its encasement than for Specimen 2 and its encasement. The dotted lines in Figure 7B represent an objective and rational manner to include the spatial context of the encasement without adding too little or too much extra information.

Notes on Statistical Methods

When doing PCA on data from the LCPC Transform, the data should first be median-centered instead of mean-centered, because the LCPC Transform always results in a wide range of magnitudes and thus have what seem to be outliers. Bins 1, 2, and 3 are usually much greater in magnitude than the rest of the bins. However, the frequency domain of the FFT typically exhibits data like this, so labeling the extreme values as outliers is incorrect. Proper usage of PCA on data produced by the LCPC Transform requires an awareness of the wide data scatter resulting from the FFT.

Cliff’s Delta is used in place of a t-test or Wilcoxon Rank-Sum Test due to limited sample sizes and the non-parametric nature of the data. For n = 3 per group, there are 3 x 3 = 9 possible pairwise comparisons for Cliff’s Delta: -1, -0.78, -0.56, -0.33, -0.11, 0.11, 0.33, 0.56, 0.78, 1. Delta > 0 indicates that the second set of values has shifted downward, while Delta < 0 indicates that second set has shifted upwards. Effect size labels: N = negligible ($|\delta|<0.11$), S = small ($0.11<|\delta|<0.33$), M = medium ($0.33<|\delta|<0.56$), and L = large ($|\delta|>0.56$)).

Addition to the above statistical methods, the following methods were calculated using standard libraries in Python (v3.12.4): K-means Clustering (the value for K was determined using the elbow method), Silhouette Score, Chamfer Distance, Wasserstein Distance, and Average K-nearest Neighbor Distance (the value for K was determined as the square root of the sample size).

Notes on Cell Culture and Image Acquisition of Primary Breast Organoids

Organoids were generated from human breast epithelial tissue following a previously published protocol[21,22,23], under IRB-approved tissue collection protocols at University of California, San Francisco, and at Brigham and Women’s Hospital. Surgical samples were processed on the day of surgery. In brief, tissues were digested for up to 2 hours with 1 µg/ml collagenase (Fisher, Cat. No. 17101015), then cell pellets were directly embedded in a 50 µl drop of basement membrane extract type 2 (Cultrex, Cat. No. 3532-005-02) in a dome in the center of a pre-warmed 24-well tissue culture plate (Greiner, Cat. No. M9312). This dome was then allowed to harden at 37 °C for 20 min and subsequently submerged in 500 µl of Type 2 Organoid Medium[21]. Plates were kept in humidified 37°C / 5% CO2 incubators at ambient O_2 and organoid Medium was changed every 2-3 days. Organoids were passaged when confluent, depending on their growth rate, typically every 2-4 weeks, as previously described[21]. Images were obtained on an Echo Revolve Microscope, then important into ImageJ where individual

structures were manually categorized and outlined depending on the number and density of cystic appearance of each organoid.

PROTOCOL:

1. Extracting Pure Shape of 2D Organoid Contours for the LCPC Transform

NOTE (1a): This protocol was the sequence for obtaining the data shown in Figure 8B. The step-by-step method for extracting the pure shape of objects described here has individual Python scripts that are available in the GitHub repository called "Pre-Processing-Tools-for-LCPC-Transform"[24]. Here is an example of the pre-processing steps to derive pure shape in preparation for the LCPC Transform. It is important to note that the scripts for doing either the radial grid or parallel grid LCPC Transform require that the masks of objects be blue lines on a white background. The color blue is arbitrary but assumed by the open-source scripts. It is important to note that there are many image-editing software available outside of the provided Python scripts. Users can write their own scripts or use a preferred image-editing software to attain the same pre-processing results.

1.1. Segment the outer edge of each organoid.

NOTE (1b): Segmentation can be done manually in image processing software, such as Mac's Preview software or Microsoft's Paint, or thresholding-based methods. If done manually, one of the following four colors should be chosen: blue, green, pink/magenta, red. See NOTE (1c) for the rationale for these colors.

1.2. Extract the contours of the organoids as open masks that are blue lines on a white background.

NOTE (1c): The GitHub repository called "Pre-Processing-Tools-for-LCPC-Transform" (Nguyen, 2025c) contains a folder called "color extraction scripts". This folder contains four Jupyter Notebook files that have Python code that extracts the following colors and turns them into blue masks on a white background: blue, green, pink/magenta, red (drawn on a Windows computer), and red (drawn on a Mac computer).

NOTE (1d): This process can also be done via thresholding for the outline color in an image processing software such as Fiji/ImageJ.

NOTE (1e): For images that contain contours that have touching borders or overlapping edges (like a Venn Diagram), they will need to be separated via an image processing software such that they become independent objects in the composite mask image that contains multiple organoid contours. This does not apply to images that only have one organoid or that have multiple organoids that don't touch each other.

1.3. Isolate each blue contour onto its own white canvas.

NOTE (1f): This step is only necessary if an image contains more than one contour. The goal is to measure each organoid separately, so mask images that have multiple blue contours on them need each contour to be transferred to its own white canvas. There is a Jupyter Notebook written in Python called “Isolate Objects Empty Blue.ipynb” that searches for each independent blue object in an image and places each on its own white canvas.

1.4. Crop each contour to have a one-pixel margin on all four sides of the blue object.

NOTE (1g): There is a Jupyter Notebook written in Python called “Trim margin to 1-pixel border.ipynb” that trims the borders around blue objects on a white background.

1.5. Add 100 pixels of white space on all four sides to create new margins.

NOTE (1h): This centers the object in the canvas and ensures that there are at least 15 pixels of white space on all four sides, which is a requirement for the LCPC Transform scripts. This step also allows for subsequent rotation steps to not cause the object to be cut-off due to rotations, which is a common problem when rotating an elongated object about its center. There is a Jupyter Notebook written in Python called “Add 100 pixels to margins.ipynb”.

1.6. Rotate the objects such that the longest internal length is horizontal.

NOTE (1i): This step turns the longest internal line with each object into the object’s width. There are two scripts available for this purpose. For closed shapes, use the Jupyter Notebook named “Rotate CLOSED Object Horizontal by Longest Internal Length_v2.ipynb”. For open shapes or shapes with multiple components, use the Jupyter Notebook named “Rotate Green Line_v3.ipynb”. The Rotate Green Line tool has a video tutorial that is linked in the ReadMe file contained within the same folder as the .ipynb file.

1.7. Crop each image to have a one-pixel margin on all four sides of the blue object.

NOTE (1j): This is the same as Step 1.4. The purpose of trimming this time, however, is because the resizing step that comes next resizes the entire canvas, not just the blue object in it. Thus, by making the blue object nearly the same width and height as the canvas itself, resizing the canvas to 400-pixels wide also resizes the object to be nearly 400-pixels wide. There is a Jupyter Notebook named “Trim margin to 1-pixel border.ipynb”.

1.8. Resize the canvas width of each image to be 400-pixels wide while also constraining the aspect ratio.

NOTE (1k): This step makes each object have the same width. Constraining the aspect ratio prevents skewing the object during the resizing. There is a Jupyter Notebook named “Resize width to 400 pixels but constrain aspect ratio.ipynb”.

1.9. Add 100 pixels of white space on all four sides to create new margins.

NOTE (1l): This is the same as Step 1.5 and ensures that there are at least 15 pixels of white space on all four sides. There is a Jupyter Notebook named “Add 100 pixels to margins.ipynb”.

1.10. Rotate each image either 90-degrees clockwise or counterclockwise to make the longest internal length vertical, being consistent with the choice for all images.

NOTE (1m): This step prepares the shapes to be optimally covered when using the parallel grid LCPC Transform, but has no effect when using the radial grid LCPC Transform. There are Jupyter Notebooks named “Rotate images 90CW.ipynb” (CW = clockwise) and “Rotate images 90CCW.ipynb” (CCW = counter clockwise).

NOTE (1n): In the above sequence, pure shape is obtained by sequentially implementing Steps 1.6, 1.7, 1.8, and 1.9. Adding margin space in Step 1.9 is required to avoid errors in the LCPC Transform scripts that will be done next but is not involved in extracting pure shape. Users who desire to measure shapes “at-scale” need not perform Steps 1.6, 1.7, and 1.8 in the above sequence, though they should still pre-process their mask images to attain optimal orientation prior to applying the LCPC Transform.

NOTE (1o): The GitHub repository that contains the pre-processing scripts also contains scripts that do the following: (1) flipping images into their mirror image [“Flip images VERTically.ipynb”, “Flip images HORIZontally.ipynb”], (2) thickening the blue lines on a white background to attain optimal thickness, which is 5-10 pixels thick [“Thicken Blue Lines.ipynb”].

1.11. Perform the LCPC Transform on each contour.

NOTE (1p): For the parallel grid LCPC Transform use the script called “Master Script_Parallel Grid LCPC_DN_071025.ipynb” in the GitHub repository called Parallel-Grid-LCPC-Transform[20]. For the radial grid LCPC Transform use the script called “Master Script_Radial Grid LCPC_bBoxCentered.ipynb” in the GitHub repository called Radial-Grid-LCPC-Transform[19].

2. Extracting At-Scale Shape of 2D Organoid Contours for the LCPC Transform

NOTE (2a): This protocol was the sequence for obtaining the data shown in Figure 8C & 8D. The step-by-step method for extracting the at-scale shape of objects is like that for extracting pure shape in Protocol 1 above. Thus, the steps for this protocol, Protocol 2, will have no notes describing the steps. The steps are named identically in both protocols for ease of matching. Please refer to the notes in Protocol 1 for each step. The main difference between pure shape and at-scale shape is that that at-scale shape does not require Steps 1.6, 1.7, and 1.8, which is the sequence for resizing objects.

2.1. Segment the outer edge of each organoid.

2.2. Extract the contours of the organoids as open masks that are blue lines on a white canvas.

2.3. Isolate each blue contour onto its own white canvas.

2.4. Crop each contour to have a one-pixel margin on all four sides of the blue object.

2.5. Add 100 pixels of white space on all four sides to create new margins.

2.6 Rotate each contour such that its longest internal line is vertical.

2.7. Perform the LCPC Transform on each contour.

REPRESENTATIVE RESULTS:

The Radial Grid LCPC Transform Applied to Breast Cancer Organoids

3D organoids and tumors can adopt a variety of shapes, some of which are obviously different to the human eye even though traditional metrics yield statistically insignificant differences. On the other hand, organoids can also exhibit shapes that seem heterogeneous, and thus insignificant, to the human eye, masking recurring subtle morphologies that represent distinct subtypes.

Human primary breast organoids were extracted from patients and cultured in 3D following a previously published protocol[21,22,23]. Figure 8A shows three morphological categories accessible based on visual inspection: a cystic phenotype, a polycystic phenotype, and a dense polycystic phenotype. Figure 8B shows that the outer contours of these organoids were manually segmented and their pure shape was extracted, which removes the effect of size on their shapes. Because the contours were circular in nature, the radial grid LCPC Transform was applied to the pure shapes followed by PCA. Figure 8B shows a dot plot of the first two principle components (PCs) and that K-means clustering identified three clusters: 0, 1, and 2. Silhouette Scores (SS) confirmed the likely existence of two separate clusters between Clusters 0 vs. 1 (SS = 0.68) and Clusters 1 vs. 2 (SS = 0.69), while indicating a weak separation between Clusters 0 vs 2 (SS = 0.27). Despite the weak Silhouette Score between Clusters 0 and 2, labeling the data points with the pure shape of each data point justifies that Cluster 2 represents the most distinct phenotype: highly asymmetrical, non-circular contours (the thickness of the contours does not matter in this case).

While pure shape removes the effect of size on shape, which is impossible when measuring area and volume, measuring the organoid contours at-scale also reveals interesting insights. Figure 8C shows a 3D dot plot of the first three PCs resulting from the radial grid LCPC Transform performed on contours as they were segmented from the images without any

resizing. This approach means that the LCPC Transform is measuring both the shape and size of the organoids simultaneously. Figure 8C shows that the polycystic and dense polycystic organoids are less densely packed compared to the cystic organoids. To quantify this expansion, Figure 8D shows the scatter of each phenotype in their shared PCA space. The average K-nearest Neighbor Distance (KNND) is calculated for each phenotype, which supports the visual assessment of decreasing density going from cystic organoids (KNND = 1.061) to polycystic organoids (KNND = 1.575) to dense polycystic organoids (KNND = 2.247); this decreased density can also be interpreted as increased morphological heterogeneity. Lastly, for a comparison of the LCPC Transform to the traditional metric of area, Figure 8E shows that while there is a statistical significance between the area of cystic organoids vs. dense polycystic organoids ($p = 0.016$, Wilcoxon Rank-Sum Test), there is no indication based on area that there are three distinct morphological groups across the phenotypes as identified by K-means Clustering after pure shape LCPC analysis (Figure 8B) or that there is increasing morphological heterogeneity as show by at-scale LCPC analysis (Figure 8C). Figure 8E shows a decreasing scatter of the data points going from cystic to polycystic to dense polycystic organoids, which is the opposite of what is observed by at-scale LCPC analysis (Figure 8C), highlighting that the spatial information captured by the LCPC Transform is different from what area captures.

The same approach applied to individual 2D shapes in Figures 7 and 8 also apply to the contour of cells in 2D cell culture: single-cells (such as red blood cells) or clusters of cells (such as in cervical Pap smears). Pilot studies for each of these cases are described on YouTube:

- Shape analysis of brain organoids that exhibit varying degrees of folding in their outer contour[25].
- Shape analysis of lung alveoli *in situ*, the same principles can be applied to analyzing 2D cell culture[26].
- Shape analysis of cervical tissue from Pap smears in which both the contour of the nucleus and of the cell membrane are simultaneously quantified via the radial grid LCPC Transform[27].
- Shape analysis of red blood cells infected by malaria and cells adjacent to infected cells but are not themselves infected[28].
- Shape analysis of red blood cells affected by Sickle Cell Disease[29].

The Parallel Grid System for Measuring Density or Texture as an Index of Spatial Information

Spatial information involves more than the usual notions of shape, such as squareness, roundness, and curvature, because it encompasses concepts such as texture and density. Complex tubular networks, such as blood vessels, or branching networks, such as cytoskeletal filaments, do not fit the notion of shape in the same sense of our geometry-based intuition of shapes, such as circles and polygons. However, density and texture are very useful notions of spatial information for understanding the structure and function of cells, organoids, and tumors. In dealing with how to quantify complex networks in fluorescence imaging, cell biologists measure features such as the average signal intensity within an image or the percent

612 of the total area of an image that contains a signal. These approaches recognize that there is
613 more to spatial information than what is described in terms of geometry motifs.
614
615 The pattern of vascular networks within tissues, tumors, and organs on a chip are not easily
616 defined by traditional geometry motifs, but benefit from the interpretation as texture and
617 density. Vascular remodeling in tumor development and treatment resistance is a well-
618 established feature of cancer (reviewed in[30]). Tadeo et al.[31] analyzed histopathological sections
619 of 458 primary neuroblastic tumors, characterizing the density, size and shape of total blood
620 vessels and vascular segments. They found that blood vessels were larger, more abundant, and
621 more irregularly-shaped in tumors of patients with poor prognostic factors compared to the
622 favorable cohort. Interestingly, Fang et al.[32] coined a term called "vessels that encapsulated
623 tumor clusters" (VETC), which describes a pattern of blood vessels that surround clusters of
624 small tumors within a larger hepatocellular carcinoma and are associated with higher rates of
625 metastasis and recurrence. The same group later founded that VETC-positive hepatocellular
626 carcinoma correlated with a positive response to treatment with the kinase inhibitor
627 Sorafenib[33]. Thus, the ability to quantify vascular patterns is crucial for prognosis and
628 understanding mechanisms of resistance.
629
630 For quantifying dense networks of tubes, islands, and branches, the parallel grid system of the
631 LCPC Transform is useful for measuring what can be called texture-density, instead of shape. As
632 an example of how the parallel grid LCPC Transform can do this, Figure 9A shows cardiac
633 microvasculature patterns from Gkontra et al.[34]: to the naked eye, objectively and
634 quantitatively distinguishing the patterns between the infarcted areas and the areas at a region
635 remote from the site in injury is very difficult. By extracting the mask of the blood vessels
636 (Figure 9A Step 1) and applying the parallel grid LCPC Transform, quantitative differences can
637 be observed. First, to enhance the cleanliness of the data, each image was cropped with a circle
638 of uniform diameter and that is centered at the middle of the square image (Figure 9A Step 2),
639 which ensures that each image has the same height and width regardless of rotations. Next, the
640 cropped images are rotated such that the main axis of vessels (red double arrows in Figure 9A
641 Step 1) is vertical (Figure 9A Step 3). Then, a blue line is applied to the circumference of the
642 image to ensure that even images with sparse vascular patterns have the same diameter in
643 terms of blue pixels (Figure 9A Step 4). Steps 3 and 4 were performed using the Fiji/ImageJ
644 software (v2.14.0). Lastly, the parallel grid LCPC Transform is applied (Figure 9A Step 5).
645 Analyzing the results via PCA and plotting the first three PCs reveals that the infarcted regions
646 cluster separately from the remote regions (Figure 9B). Calculations of Chamfer Distance (CD)
647 quantify the visual patterns exhibited by the graphs. At 1-day after injury, the infarcted site
648 exhibits a vascular pattern that is more distant from the basal condition (Basal vs. 1Day_infarct:
649 CD = 83.89) than is the non-injured site (Basal vs 1Day_remote: CD = 53.95). At 1-day after
650 injury, the non-injured remote site is closer to the injured site (1Day_remote vs 1Day_infarcxt:
651 CD = 34.48) than it is to the basal condition. At 3-days after injury, the vascular patterns of the
652 injured site and associated non-injured remote site exhibit drastic deviation from the basal
653 condition and from each other (Basal vs. 3Days_Infarct: CD = 240.63; Basal vs. 3Days_Remote:
654 CD = 67.28; 3Days_Remote vs. 3Days_Infarct: CD = 168.35). Very interestingly, at 7-days after
655 injury, the vascular patterns of the injured site and the non-injured remote site occupy a 3D

space that is approximately orthogonal to the basal condition, while aligned with each other at a distance that is closer than they were at 1-day post injury (CD = 24.76 for day 7 compared to 34.48 for day 1). The data suggest two things: (1) as healing progresses, the vascular pattern of the injured site and non-injured remote site adopt patterns that are distinct from the basal condition; (2) the reconstruction program activated at the injured site also triggers similar morphological changes at their associated non-injured remote sites.

As a reference for the versatility of the LCPC Transform, the traditional metrics of percent area (Figure 10A) and total perimeter (Figure 10B) of blood vessels was calculated. While area and perimeter reveal changes between points in the time course, they do not provide coherent quantitation of distinct spatial arrangements of vessels. Area and perimeter are abstractions that can represent many different vascular arrangements and thus are not informative about how the spatial arrangement of vessels are changing across the treatment conditions and time course. With an n=3 for each group, the t-test and Wilcoxon test are unreliable. Cliff's Delta is nonparametric and works well to show a shift in the effect size. The major difference revealed by measuring percent area of blood vessels is between the basal condition and the 7-day time points; for total perimeter of blood vessels, the differences are between the basal condition vs. 1-day and the basal condition vs. 7-days. However, area and perimeter provide no indication that there are differences between the infarcted sites and their non-injured remote sites, which is very surprising given that the infarcted site was purposefully injured. Here is where the versatility of the LCPC Transform emerges compared to traditional metrics. For each sample, the results of the FFT as part of the LCPC Transform produces a set of magnitudes associated with a set of frequencies. The magnitudes of each frequency bin can be averaged and then compared across the treatment groups (Figure S1). This allows for a simple assessment of which frequency bins exhibit the biggest differences between treatment groups. Figure 10C shows six bins that exhibit clear effect size differences between various treatment groups, especially between the infarcted sites and their associated non-injured remote sites, which area and perimeter indicated had no difference. The independent values in these six bins provide multiple indices that show quantitative differences between treatment conditions. Interestingly, the pattern between the groups for Bin 1 of the LCPC Transform results is similar to the result obtained from measuring total perimeter. The other indices, however, reveal quantitative differences that are hidden to both area and perimeter. Indices unique to the basal condition: Bin 49, Bin 56, Bin 21. Indices unique between the Day 1 groups: Bin 1, Bin 6, Bin 49. Indices unique to the Day 3 groups: Bin 6, Bin 112, Bin 49. Indices unique to the Day 7 groups: Bin 6, Bin 49, Bin 112. These data highlight the versatility and precision of the LCPC Transform over the traditional approaches of area and perimeter.

FIGURE AND TABLE LEGENDS:

Figure 1. Description of the Radial Grid System of the LCPC Transform.
(A) Step 1: The shape of interest is segmented and extracted as a blue mask. Step 2: A bounding box (a.k.a. the smallest rectangle encasing and touching the object on all four sides) is determined for the mask. A system of radial grid lines originating from the center of the

bounding box is overlaid onto the mask. The decision to pick the center of the bounding box instead of the centroid of the object as the origin was arbitrary. Step 3: The distance of each intersection between a gridline and the shape is calculated relative to the origin. Grid lines that have more than one intersection sum all those intersection distances into one value for that grid line. Step 4: The FFT is applied to turn the data in Step 3 into a frequency plot, which is a static multidimensional representation (Step 5) of the shape in Step 1. Step 6: A scalar value can be created from Step 5 via an index. See the Figure 4 for a discussion on types of indices to calculate. (B) Examples shape themes for which the radial grid system is optimal.

Figure 2. Description of the Parallel Grid System of the LCPC Transform.
(A) Step 1: The shape of interest is manually segmented and extracted as a binary mask. Step 2: A parallel grid system is overlaid onto the shape, which is rotated such that the base is horizontal. Step 3: The distance of each intersection between a gridline and the shape is calculated relative to the baseline. Grid lines that have more than one intersection sum all those intersection distances into one value for that grid line. Step 4: The FFT is applied to turn the data in Step 3 into a frequency plot, which is a static multidimensional representation (Step 5) of the shape in Step 1. Step 6: A scalar value can be created from Step 5 via an index. See the Figure 4 for a discussion on types of indices to calculate. (B) Examples shape themes for which the parallel grid system is optimal.

Figure 3. The Importance of Orienting Shapes Systematically for Extracting Clean Data.
(A) An example of identical shapes that are mirror images of each other and how the parallel grid system determines their distinct orientation and thus will produce different results. (B) An example of how to flip Shape 2 horizontally, such that it's orientation now matches Shape 1 and becomes in a fairer comparison to Shape 1. In this configuration, the parallel grid LCPC Transform will yield the same results for both Shape 1 and Shape 2. This alteration of orientation should only be done if there is no spatial context that suggests that Shape 2 should not be modified. (C) An example of four variations of objects that look like the number "3", with some that face the wrong direction. (D) An example of how to orient each object to maximize the cleanliness of the data that will result from the parallel grid LCPC Transform. (E) An example of measuring the wider bulge of the items in Panel C based on the same orientation rules as for Panel D. (F) An example of measuring the large bulbs of the items in Panel C but keeping their alignment context as in Panel D. The motivation of keeping the alignment rules in Panel D is to dissect the contribution of the spatial information in the wider bulge relative to the contribution of the narrower bulge within these rules.

Figure 4. Hypothetical FFT Results and Common Indices That Can Be Derived From Them.
(A) The frequency plot resulting from the FFT of two hypothetical shapes is displayed. Since the magnitudes in Bin 3 and Bin 4 exhibit the opposite trend, a ratio between these two bins will be a good index representing a difference between these two shapes. The red and green dotted lines represent other frequency bins that could be used to derive good indices. (B) A histogram of two hypothetical shapes whose FFT profiles exhibit a stepwise decreasing trend as a representation of their morphological differences. (C) A histogram of two hypothetical shapes whose FFT profiles differ most in the central region along the span of frequency bins. (D)

Examples of hypothetical morphological features that may yield indices that correlate with their presence.

Figure 5. The Inability of Traditional Shape Metrics to Register Obvious Context.
(A) The series of checkmarks are the same shape but increase in size from left to right. In this series, the area of each checkmark is different while the shape concept remains the same: that is, they remain the same style of checkmarks. (B) A square is rotated about its center for 45-degrees until its shape concept becomes what humans call a diamond. While the shape concept of square is very distinct from that of a diamond, the in-between rotations are subjectively interpreted as either “tilted squares” or “tilted diamonds”. The LCPC Transform can objectively and precisely measure the difference in each step in this series.

Figure 6. Explanation of Pure Shape vs. At-Scale Shape.
Measuring objects “At-Scale” means measuring the shape of objects according to their original sizes relatively to each other. “Pure Shape”, on the other hand, means measuring the shape of different objects at a similar scale, which removes the effect of size on the shape. It is important to find an objective manner in which to resize the objects being compared. The longest internal line is often an optimal choice (Step 1). Both objects are rotated such that the longest internal line is horizontal (Step 2). In this case, the two objects are also rotated such that the indented side faces down. Next, the width of each object, which is now the same magnitude as the longest internal line of each object, is resized to be 400-pixels wide while also constraining the aspect ratio to prevent skewing the objects (Step 3). Lastly, it is optional to rotate the resized objects 90-degrees (Step 4) (in this case, counterclockwise) before applying the LCPC Transform.

Figure 7. Example of Spatial Markers for Cancer Research.
(A) A hypothetical example of cells migrating in 2D culture by expanding their lamellipodia in the direction of migration. In the context of migratory morphology of epithelial cells, orienting the contour of each cell such that the longest internal line that aligns with the axis of migration is vertical makes the most sense for capturing the spatial context. (B) Hypothetical examples of tumors or pre-cancerous lesions that grow within an encasement that does not change shape. The lesion grows from a stalk, which is the central support structure. As the lesion grows or shrivels, its contour moves towards or away from the internal wall of the encasement. The dotted lines represent a rational and objective way to define how much of the encasement should be traced (orange line) such that not too little or too much extra information is added. (C) Option 1 shows the parallel grid LCPC Transform applied to the contour of the lesions without context of where the encasement is located. Option 2 shows the same but including spatial context of where the encasement is located.

Figure 8. The LCPC Transform Objectively Identifies Morphological Subtypes of Breast Cancer Organoids.
(A) Primary human breast organoids were derived and cultured as previously reported[21,22,23]. Three phenotypes could be observed and classified based on visual inspection: cystic, polycystic, and dense polycystic. (B) The outer contours of the organoids were manually

segmented and the pure shape of these contours were obtained. The radial grid LCPC Transform was then applied, and the results were analyzed via PCA. K-means Clustering identified three clusters (0, 1, and 2) and Silhouette Scores (SS) were calculated to confirm the existence of these clusters. (C) The outer contours of the organoids were also measured via the radial grid LCPC Transform on their at-scale shapes, meaning the shapes as they are in the images relative to each other without any resizing. The inset is enlarged to show the looser packing of polycystic and dense polycystic organoids compared to the cystic organoids. For B and C: Green = cystic, Purple = polcystic, Yellow = dense polycystic. (D) The first three PC values are plotted for each phenotype separately and the average KNND is calculated to show that the cystic phenotype is most densely packed, while the polycystic and dense polycystic phenotypes exhibit less density (meaning more heterogeneity). Though plotted separately, all three phenotypes share the same PCA space. For the cystic phenotype, average KNND was calculated on the inset because the outliers that exist only in this phenotype skew the calculation.

Figure 9. Measuring Texture-Density with the Parallel Grid LCPC Transform.
(A) Photomicrographs of microvasculature in cardiac tissue previously obtained as described in Gkontra et al.[34,35]. Images from the remote sites that were non-injured in each time point are not shown. Step 1: A 2D mask is generated and the general axis of alignment of vessels is determined (red double-arrow). Step 2: A circular region centered on the midpoint of the image is obtained in order to force each image to have the same height after the images are rotated to varying degrees in the next step (rotating square images to varying degrees will alter the height of the image depending on the degree of rotation). Step 3: Each image is rotated such that the general axis of alignment determined in Step 1 is vertical. Step 4: A uniform blue border is applied around each image to ensure that each image has the same diameter, because images with sparser densities can be shorter (example: see “7 Days Post Myocardial Infarction” in Step 3). Step 5: The parallel grid LCPC Transform is applied. (B) Dot plots of the first three PCs of the output from the parallel grid LCPC Transform (Purple: Basal condition, Green: infarcted site, Yellow: non-injured site remote to the infarcted site).

Figure 10. The LCPC Transform Measures Features Hidden to Area and Perimeter.
For a comparison to traditional shape metrics, the percent area (A) and the total perimeter (B) of the vascular networks were measured. (C) To show the precision that can be obtained from the result of the LCPC Transform (Figure S1), the value of individual frequency bins resulting from the FFT were compared. Six frequency bins were chosen as indices that represent the objective quantitation of differences between treatment groups. For all panels, due to the limited sample size of each group (n=3), the nonparametric Cliff’s Delta for effect size was calculated instead of a p-value from the Wilcoxon Rank-Sum Test.

Figure S1. Box-Whisker Plots Comparing the Magnitudes of Various Frequency Bins Across the LCPC Transform Results of Vascular Networks.
The LCPC Transform of the vascular networks described in Figures 9 and 10, results in a set of frequency bins, and their magnitudes, that represent each image. Box-whisker plots of the magnitudes in each bin are used to compare all treatment groups, which identifies which bin

best serves as an index for stratifying texture-density that was measured via the LCPC Transform. Each box-whisker plot consists of n=3 independent images.

DISCUSSION:

As with any computational tool, the quality of the output depends on whether the inputs adhere to the expected rules around which the tool was designed. The following rules should be followed and used as quality control checks when utilizing the Python scripts located on the GitHub repositories cited in this study.

1) Make sure that that open masks are blue lines on a white background. Blue was an arbitrary decision and has no significance, but the scripts that perform the LCPC Transform were written to search for blue pixels on a white background.
2) Do not run the LCPC Transform scripts on images that have less than 10-pixels of white space on all four sides. See this video for an explanation of why: "The Margin Thickness Parameter - How to Prepare Shape Outlines"[36].
3) This script assumes that the outlines of your object are blue pixels on a white background. The thickness of the line should be between 5-10 pixels thick. See this video for an explanation of why: "The Line Thickness Parameter - How to Prepare Shape Outlines"[37].
4) Do not run the LCPC Transform scripts on blank images. This will cause an error in the script.
5) Do not run the LCPC Transform scripts on images that consist of only a speck of debris. This will cause an error because the grid system may not detect the speck and thus attempt to analyze a blank image.
6) Do not run the LCPC Transform scripts on images that contain specks of unwanted debris. The debris will alter the height and width of the open mask object and thus cause inaccurate results.
7) See this video to understand why it is important to understand the scale of your objects before analyzing them with the LCPC Transform: "The Scale Parameter - How to Prepare Shape Outlines"[38].

While the scripts for the LCPC Transform are open-source and thus available for recreation and integration into existing computational pipelines, users who create their own version or modify the preset parameters should check to see if their results match those that are provided as part of this study. The GitHub repositories for the parallel grid and radial grid LCPC Transform scripts contain several sets of practice images and their associated output files. The purpose of including both input and output files is so that users can check to make sure that their version of the script produces identical data as those provided in this study. Most users will adopt the preset parameters in the existing publicly available scripts, so those who alter or integrate the scripts for their own purposes should be aware that their data may no longer match everyone else's.

It is important to note that the results of the LCPC Transform cannot be reconstructed in reverse to derive the original shape of the input object. While the results of the FFT on typical

discrete sinusoid data does allow for reconstruction, doing so on FFT data what was obtained on LCPC Transform results does not make sense. Reminder: the LCPC Transform technically finishes before the FFT is applied as a last step, though the phrase “LCPC Transform” in this manuscript is meant to include the FFT as part of the LCPC Transform’s routine operation unless specifically stated otherwise.

The FFT in the scripts that do the LCPC Transform, a sampling rate of 40 was chosen because the original way the sampling rate was picked was by dividing a 200-pixel-tall image by 5 pixels resulting in 40 horizontal lines. While changing the sampling rate changes the resulting frequencies, it does not change the resulting magnitudes. Since the magnitude of each frequency bin is the important feature describing a shape, the sampling rate does not matter. For those who desire to rewrite the scripts to better fit their computational workflow, the scripts calculate a single-sided magnitude for the FFT. The GitHub repositories that contain the two LCPC scripts also contain practice input files and their associated output files. This is so that users and those who re-write the code have images that can be used to benchmarks to determine if their script produces the same results as the original scripts.

It is worth adding to the discussion of measuring texture-density (Figures 9 and 10) that the human eye is easily confused by optical illusions or disorderly patterns. Haphazard arrangement of complex patterns is often interpreted as being random, not measurable, or not worth measuring. The LCPC Transform can measure reproducible, objective, quantitative differences in seemingly random patterns. However, unless the user is trained to not be misled by chaotic, disorderly patterns in nature that masquerade as randomness, no attempt will be made to measure the disorder to uncover predictable patterns. Predictable patterns is a clear indication that the system is not completely random.

The LCPC Transform was invented for the purpose of measuring shapes that are subtly different or inconceivably different to the naked eye. Because many morphologies in nature exhibit minor differences that seem more like meaningless chaotic heterogeneity than predictable subtypes of disorder – that each correlate with distinct biological functions – the following “Four Step” conceptual framework was developed to be used in conjunction with the LCPC Transform. The reason being that though the LCPC Transform was designed to measure features that are hidden to the naked eye and traditional geometric metrics, if the user assumes that a complex system is random then there will be no attempt to apply the LCPC Transform.

Four Steps to Overcoming the Trap of Thinking That a Chaotic System is Random
(1) Resist the notion that a complicated-looking system is random.
(2) Call the chaotic heterogeneity “disorder” and then find a type of order that can serve as a rational reference condition. Hypothesize that the disorder is deviating from the reference.
(3) Find a way of objectively measuring the pattern that is the reference, such as the LCPC Transform.
(4) Apply this same objective method to measuring what is considered the disorder in system.

ACKNOWLEDGMENTS:

DHN would like to acknowledge Duane Nichols, a high school science teacher who inspired him to pursue a career in science. Nichols passed away from an aggressive form of colon cancer that kept returning after three surgeries. His passing inspired the DHN to invent the LCPC Transform to better characterize the morphology of colon polyps. DHN would also like to acknowledge Thuan Trinh, his college roommate and friend, who suffered from Bipolar Disorder 2 and took his own life during his late twenties. In desiring to apply the LCPC Transform to the complex morphologies of the human brain in honor of Trinh, DHN was inspired to develop ways of enhancing the LCPC Transform, such as adding markers to capture spatial context surrounding the shape of interest. Trinh was an electrical and computer engineer whose discussions with DHN at the dining table about the intersection of engineering and biology is part of the inspiration that would one day lead to the DHN's computational biology research program. The LCPC Transform is informally referred to as the Nguyen-Nichols-Trinh (NNT) Transform to honor the memory of Nichols and Trinh. DHN would also like to acknowledge Paul Leal, the lead angel investor who supported DHN's startup that sought to commercialize the LCPC Transform. Though unsuccessful, the five years during which the startup operated allowed DHN to prove the utility of the LCPC Transform for many fields of science: microbiology, pathology, materials science, agricultural science, and radiology. Thus, this manuscript is the first of many insightful studies to come. Lastly, DHN would like to thank Polyxeni Gkontra, Ph.D., for sharing photomicrographs of cardiac vasculature for this study. Additionally, authors MB and JMR would like to thank the DF/HCC Breast SPORE: Specialized Program of Research Excellence (SPORE) (NCI 1P50CA168504), the UCSF Breast Care Center surgical and pathology teams and the Breast Care Center interns for assistance with tissue specimen, and funding support from the National Institutes of Health (R01CA281361).

DISCLOSURES:

The authors have no conflicts of interest to disclose. The LCPC Transform was commercialized as a software called "Shape Genie" via a company called BrainScanology, Inc., but the patent applications were withdrawn and the company was dissolved.

REFERENCES:

1. Masako Muguruma, Saeko Teraoka, Kana Miyahara, Ai Ueda, Mariko Asaoka, Miki Okazaki, Takahiko Kawate, Masahiko Kuroda, Yohei Miyagi, & Takashi Ishikawa. Differences in drug sensitivity between two-dimensional and three-dimensional culture systems in triple-negative breast cancer cell lines. *Biochemical and Biophysical Research Communications* 533(3), 268–274 (2020).
2. Abbas, Z.N., Al-Saffar, A.Z., Jasim, S.M. et al. Comparative analysis between 2D and 3D colorectal cancer culture models for insights into cellular morphological and

962 transcriptomic variations. *Sci Rep* 13, 18380 (2023). https://doi.org/10.1038/s41598-
963 023-45144-w
964 3. Betge, J., Rindtorff, N., Sauer, J. et al. The drug-induced phenotypic landscape of Betge,
965 Johannes & Rindtorff, Niklas & Sauer, Jan & Rauscher, Benedikt & Dingert, Clara &
966 Gaitantzi, Haristi & Herweck, Frank & Srour-Mhanna, Kauthar & Miersch, Thilo &
967 Valentini, Erica & Boonekamp, Kim & Hauber, Veronika & Gutting, Tobias & Frank,
968 Larissa & Belle, Sebastian & Gaiser, Timo & Buchholz, Inga & Jesenofsky, Ralf & Härtel,
969 Nicolai & Boutros, Michael. (2022). The drug-induced phenotypic landscape of colorectal
970 cancer organoids. Nature Communications. 13. 3135. 10.1038/s41467-022-30722-9.
971 4. Li, Q., Sun, H., Luo, D. et al. Lnc-RP11-536K7.3/SOX2/HIF-1α signaling axis regulates
972 oxaliplatin resistance in patient-derived colorectal cancer organoids. *J Exp Clin Cancer*
973 *Res* 40, 348 (2021). https://doi.org/10.1186/s13046-021-02143-x
974 5. Hadj Bachir, E., Poiraud, C., Paget, S., Stoup, N., Moghrabi, S. E., Duchêne, B., Jouy, N.,
975 Bongiovanni, A., Tardivel, M., Weiswald, L.-B., Vandepeutte, M., Beugniez, C., Escande,
976 F., Leteurtre, E., OrgaRES consortium, Poulain, L., Lagadec, C., Pigny, P., Jonckheere, N.,
977 Renaud, F., Truant, S., Van Seuningen, I., & Vincent, A. A new pancreatic
978 adenocarcinoma-derived organoid model of acquired chemoresistance to FOLFIRINOX.
979 *Biology of the Cell* 114, 32–55 (2022). https://doi.org/10.1111/boc.202100003
980 6. Zhao, N., Powell, R. T., Yuan, X. et al. Morphological screening of mesenchymal
981 mammary tumor organoids to identify drugs that reverse epithelial–mesenchymal
982 transition. *Nat Commun* 12, 4262 (2021). https://doi.org/10.1038/s41467-021-24545-3
983 7. Roignot, J., Peng, X., & Mostov, K. Polarity in mammalian epithelial morphogenesis. *Cold*
984 *Spring Harb Perspect Biol* 5(2), a013789 (2013).
985 https://doi.org/10.1101/cshperspect.a013789
986 8. Tharp, K. M., & Weaver, V. M. Modeling Tissue Polarity in Context. *J Mol Biol* 430(19),
987 3613–3628 (2018). https://doi.org/10.1016/j.jmb.2018.07.015
988 9. Campanale, J. P., Sun, T. Y., & Montell, D. J. Development and dynamics of cell polarity
989 at a glance. *J Cell Sci* 130(7), 1201–1207 (2017). https://doi.org/10.1242/jcs.188599
990 10. Martin, E., Girardello, R., Dittmar, G., & Ludwig, A. New insights into the organization
991 and regulation of the apical polarity network in mammalian epithelial cells. *FEBS J* 288,
992 7073–7095 (2021). https://doi.org/10.1111/febs.15710
993 11. Khune, A. J., et al. Physiological Responses to Space Travel: A Systematic Review.
994 *European Journal of Cardiovascular Medicine* 14(5), 450–454 (2024).
995 12. Tomsia, M., et al. Long-term space missions' effects on the human organism: what we
996 do know and what requires further research. *Front. Physiol.* 15 (2024).
997 https://doi.org/10.3389/fphys.2024.1284644
998 13. Kim, S. C., Kim, M. J., Park, J. W. et al. Effects of simulated microgravity on colorectal
999 cancer organoids growth and drug response. *Sci Rep* 14, 25526 (2024).
1000 https://doi.org/10.1038/s41598-024-76737-8
1001 14. Grimm, D., Corydon, T. J., Sahana, J., González-Torres, L. F., Kraus, A., Marchal, S., Wise,
1002 P. M., Simonsen, U., & Krüger, M. Recent studies of the effects of microgravity on cancer
1003 cells and the development of 3D multicellular cancer spheroids. *Stem Cells Transl Med*
1004 14(3) (2025). https://doi.org/10.1093/stcltm/szaf008

1005 15. Yuan, J., Zhang, S., Xu, Y. et al. Microgravity and cancer cells: from molecular
1006 mechanisms to therapeutic strategies. *Cell Death Discov.* 12, 9 (2026).
1007 https://doi.org/10.1038/s41420-025-02836-8
1008 16. Ciocalteu, A., Gheonea, D. I., Saftoiu, A., Streba, L., Dragoescu, N. A., & Tenea-Cojan, T.
1009 S. Current strategies for malignant pedunculated colorectal polyps. *World J Gastrointest*
1010 *Oncol* 10(12), 465–475 (2018). https://doi.org/10.4251/wjgo.v10.i12.465
1011 17. Teo, N. Z., et al. Management of malignant colonic polyps. *Journal of Gastrointestinal*
1012 *Oncology* 11(3) (2020).
1013 18. Nguyen, D. H. Translating the architectural complexity of the colon or polyp into a
1014 sinusoid wave for classification via the fast Fourier transform [Preprint].
1015 arXiv:1801.06752 [Preprint]. 2018 Jan 21. Available from:
1016 https://arxiv.org/abs/1801.06752
1017 19. Nguyen, D. H. Radial-Grid-LCPC-Transform [Internet]. GitHub; 2025 May [cited 2026 Jan
1018 25]. Available from: https://github.com/dh2nguyen/Radial-Grid-LCPC-Transform
1019 20. Nguyen, D. H. Parallel-Grid-LCPC-Transform [Internet]. GitHub; 2025 May [cited 2026
1020 Jan 25]. Available from: https://github.com/dh2nguyen/Parallel-Grid-LCPC-Transform
1021 21. Dekkers JF, van Vliet EJ, Sachs N, Rosenbluth JM, Kopper O, Rebel HG, Wehrens EJ, Piani
1022 C, Visvader JE, Verissimo CS, Boj SF, Brugge JS, Clevers H, Rios AC. Long-term culture,
1023 genetic manipulation and xenotransplantation of human normal and breast cancer
1024 organoids. Nat Protoc. 2021 Apr;16(4):1936-1965. doi: 10.1038/s41596-020-00474-1.
1025 22. Rosenbluth JM, Schackmann RCJ, Gray GK, Selfors LM, Li CM, Boedicker M, Kuiken HJ,
1026 Richardson A, Brock J, Garber J, Dillon D, Sachs N, Clevers H, Brugge JS. Organoid
1027 cultures from normal and cancer-prone human breast tissues preserve complex
1028 epithelial lineages. Nat Commun. 2020 Apr 6;11(1):1711. doi: 10.1038/s41467-020-
1029 15548-7.
1030 23. Ye J, Goldhammer N, Vohra P, et al. FNA biopsy of breast specimens effectively harvests
1031 cells for patient-derived organoids modeling ductal carcinoma in situ. Cancer
1032 Cytopathol. 2025; e70052. doi:10.1002/cncy.70052
1033 24. Nguyen, D. H. Pre-Processing-Tools-for-LCPC-Transform [Internet]. GitHub; 2025 May
1034 [cited 2026 Jan 25]. Available from: https://github.com/dh2nguyen/Pre-Processing-
1035 Tools-for-LCPC-Transform
1036 25. Nguyen, D. H. Shape Genie Quantifies Brain Organoid Morphology in Ways That Area
1037 Cannot. YouTube. 2025. Available from:
1038 https://youtu.be/_FBtY8a5zmc?si=siT0Lom04dCvBxnM
1039 26. Nguyen, D. H. Shape Genie Identifies Sub-Groups of Lung Alveoli. YouTube. 2025.
1040 Available from: https://youtu.be/9c030muLNws?si=T7TJFL23ZQvbd8eU
1041 27. Nguyen, D. H. Shape Analysis of PAP Smears by Shape Genie. YouTube. 2025. Available
1042 from: https://youtu.be/iNcHc2Zb9eM?si=cx67NY3iveNymOID
1043 28. Nguyen, D. H. Shape Analysis Detects Malaria Without Giemsa Staining. YouTube. 2025.
1044 Available from: https://youtu.be/8hCEVxjE_Rg?si=D-XfgsEsAer-R2bK
1045 29. Nguyen, D. H. Shape Genie Precisely Quantifies Sickle Cell Shapes. YouTube. 2025.
1046 Available from: https://youtu.be/E-BfgFHVFl0
1047 30. Ribatti D, Annese T and Tamma R (2023) Vascular co-option in resistance to anti-
1048 angiogenic therapy. Front. Oncol. 13:1323350. doi: 10.3389/fonc.2023.1323350

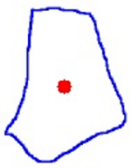
https://www.frontiersin.org/journals/oncology/articles/10.3389/fonc.2023.1323350/full

31. Tadeo I, Bueno G, Berbegall AP, Fernández-Carrobles MM, Castel V, García-Rojo M, Navarro S, Noguera R. Vascular patterns provide therapeutic targets in aggressive neuroblastic tumors. Oncotarget. 2016 Apr 12;7(15):19935-47. doi: 10.18632/oncotarget.7661.
32. Fang JH, Zhou HC, Zhang C, Shang LR, Zhang L, Xu J, Zheng L, Yuan Y, Guo RP, Jia WH, Yun JP, Chen MS, Zhang Y, Zhuang SM. A novel vascular pattern promotes metastasis of hepatocellular carcinoma in an epithelial-mesenchymal transition-independent manner. Hepatology. 2015 Aug;62(2):452-65. doi: 10.1002/hep.27760.
33. Fang, J.-H., Xu, L., Shang, L.-R., Pan, C.-Z., Ding, J., Tang, Y.-Q., Liu, H., Liu, C.-X., Zheng, J.-L., Zhang, Y.-J., Zhou, Z.-G., Xu, J., Zheng, L., Chen, M.-S. and Zhuang, S.-M. (2019), Vessels That Encapsulate Tumor Clusters (VETC) Pattern Is a Predictor of Sorafenib Benefit in Patients with Hepatocellular Carcinoma. Hepatology, 70: 824-839. https://doi.org/10.1002/hep.30366
34. Gkontra, P., Norton, K. A., García-Canadilla, P., et al. Deciphering microvascular changes after myocardial infarction through 3D fully automated image analysis. *Scientific Reports* 8, 15841 (2018). https://pubmed.ncbi.nlm.nih.gov/29382844/
35. Gkontra, P., Żak, M.M., Norton, K.A., Santos, A., Popel, A.S., Arroyo, A.G. (2015). A 3D Fractal-Based Approach towards Understanding Changes in the Infarcted Heart Microvasculature. In: Navab, N., Hornegger, J., Wells, W., Frangi, A. (eds) Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. Lecture Notes in Computer Science, vol. 9351. Springer, Cham. https://doi.org/10.1007/978-3-319-24574-4_21
36. Nguyen, D. H. The Margin Thickness Parameter - How to Prepare Shape Outlines. YouTube. 2025. Available from: https://youtu.be/byo6Hg6jkiw
37. Nguyen, D. H. The Line Thickness Parameter - How to Prepare Shape Outlines. YouTube. 2025. Available from: https://youtu.be/pozjMVwsLz4
38. Nguyen, D. H. The Scale Parameter - How to Prepare Shape Outlines. YouTube. 2025. Available from: https://youtu.be/AIZLg32JgHg

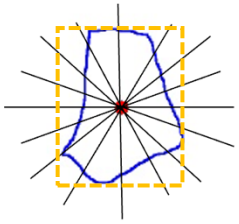
Figure 1

A

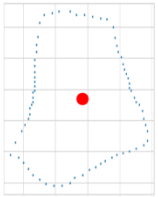
Step 1 - Create Mask of Outline & Find Center (red)



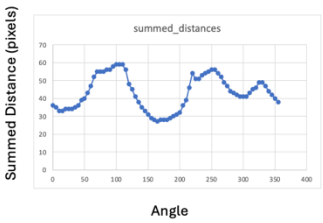
Step 2 - Overlay 360 Degree Radial Grid (Spaced Every 5 Degrees)



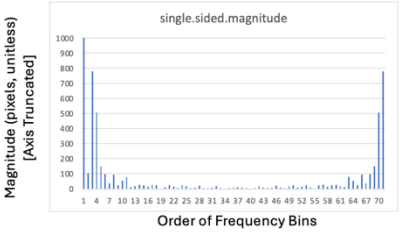
Step 3 - Calculate Distance of Intersections from Center



Step 4 - Linearize the Radial System



Step 5 - Do Fast Fourier Transform to Obtain a Multi-Dimensional Representation of the Shape



Step 6 – Dissect the Data

Derive an index as a scalar value representing the shape. Or compare different shapes in multiple dimensions with statistical tools.

B

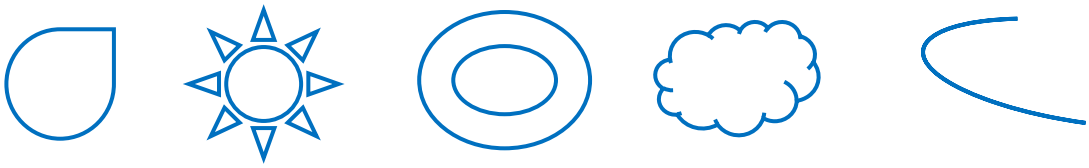
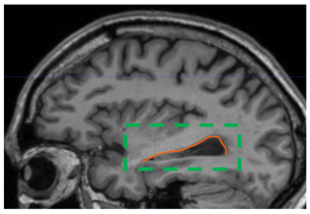


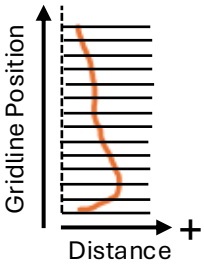
Figure 2

A

Step 1. Segment the Shape of Interest



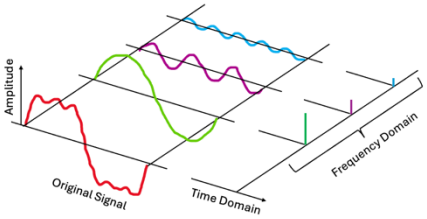
Step 2. Overlay Parallel Gridlines



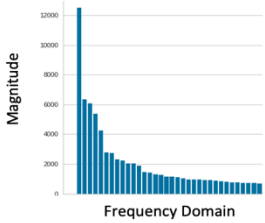
Step 3. Calculate Distance of Intersections (the Dots) at Every Position

Position	Distance
5	23
10	46
15	84
Etc.	Etc.

Step 4. Do Fast Fourier Transform on Position vs. Distance



Step 5. Obtain a Frequency Plot That Represents the Shape!



Step 6. Dissect the Data

Derive an index as a scalar value representing the shape. Or compare different shapes in multiple dimensions with statistical tools.

B

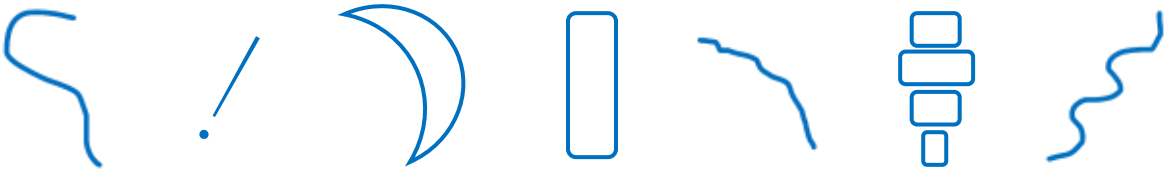


Figure 3

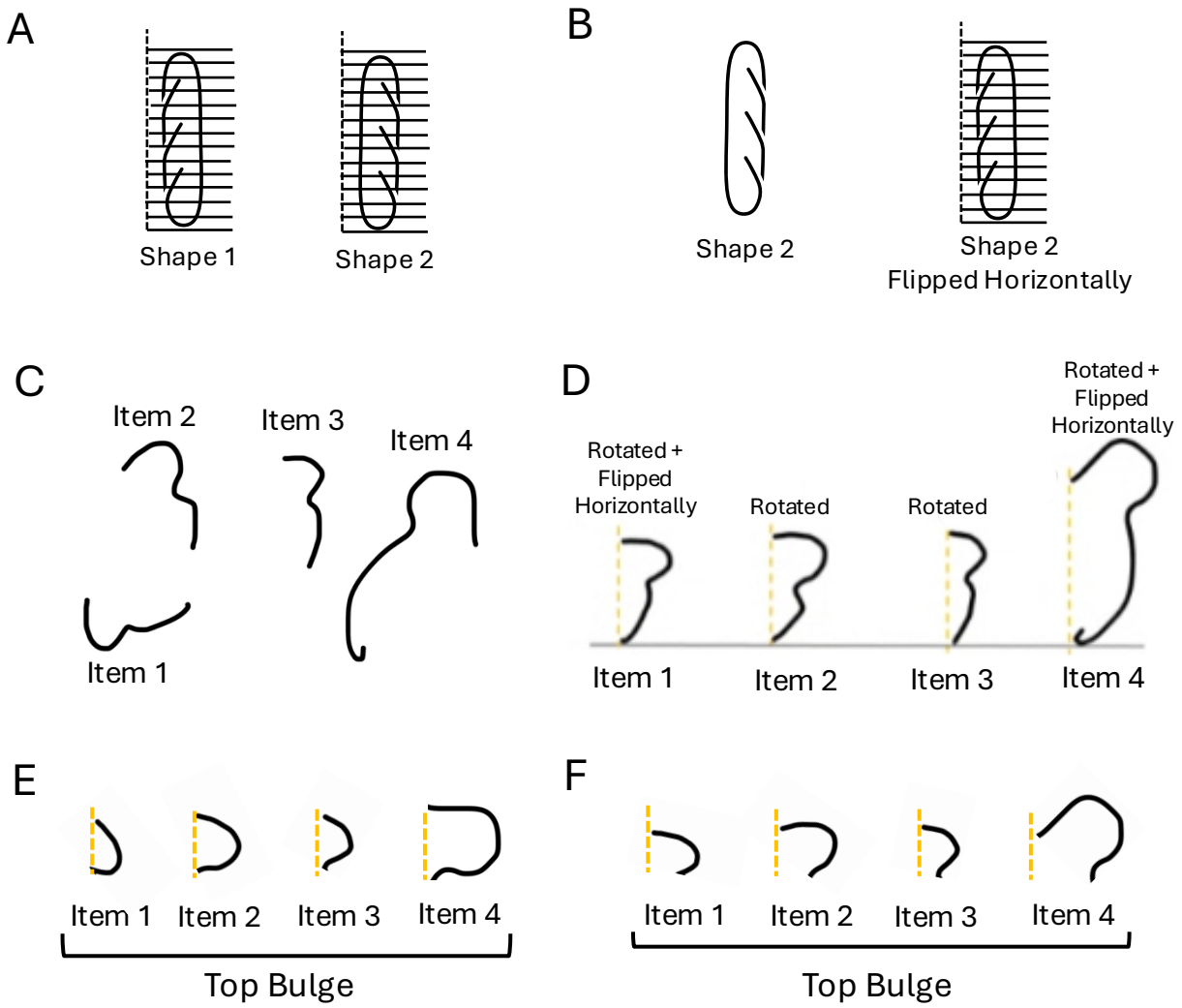

A
Shape 1
Shape 2
B
Shape 2
Shape 2
Flipped Horizontally
C
Item 2
Item 3
Item 4
Item 1
D
Rotated + Flipped Horizontally
Rotated
Rotated
Rotated + Flipped Horizontally
Item 1
Item 2
Item 3
Item 4
E
Item 1
Item 2
Item 3
Item 4
Top Bulge
F
Item 1
Item 2
Item 3
Item 4
Top Bulge

Figure 4

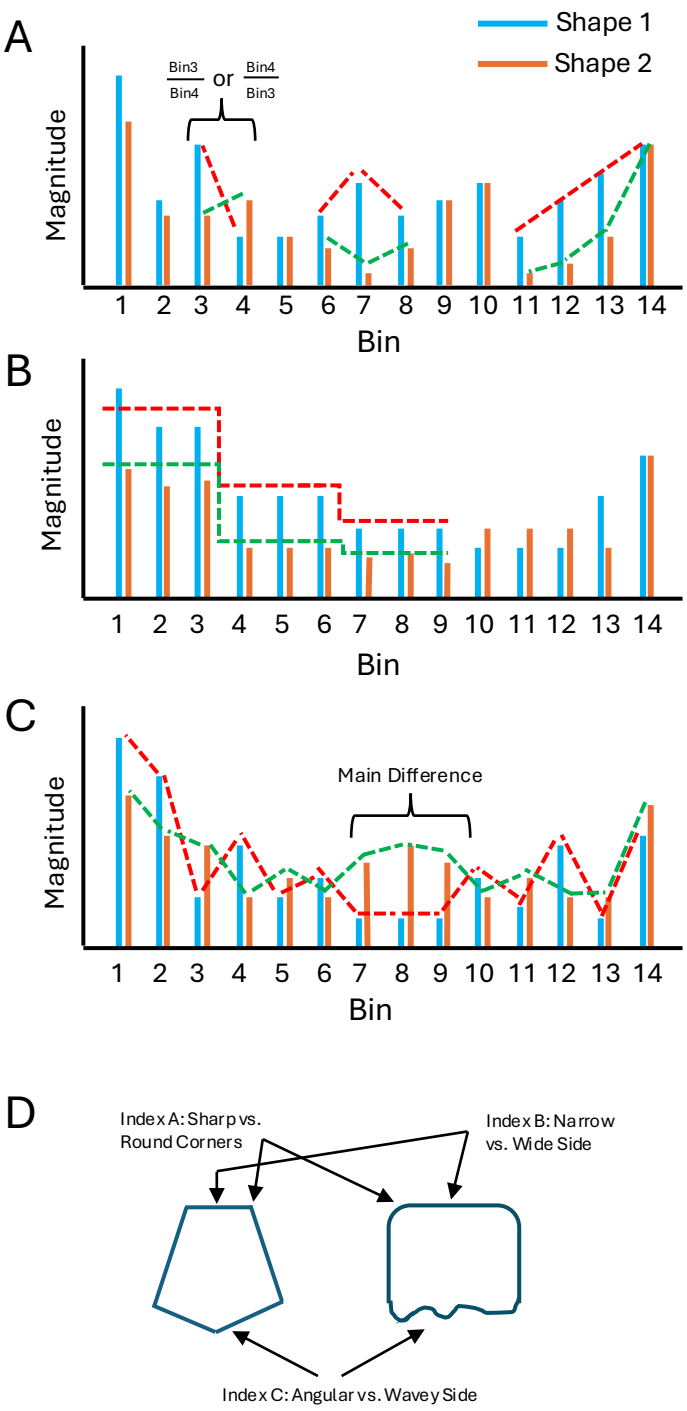


Figure 5

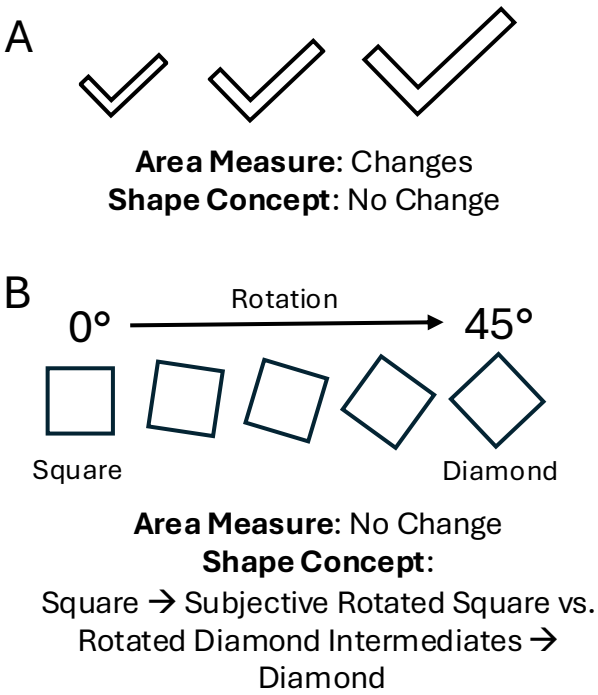


Figure 6

At-Scale: Measures Shape & Size Simultaneously
Pure Shape: Measures Shape Apart From Effect of Size

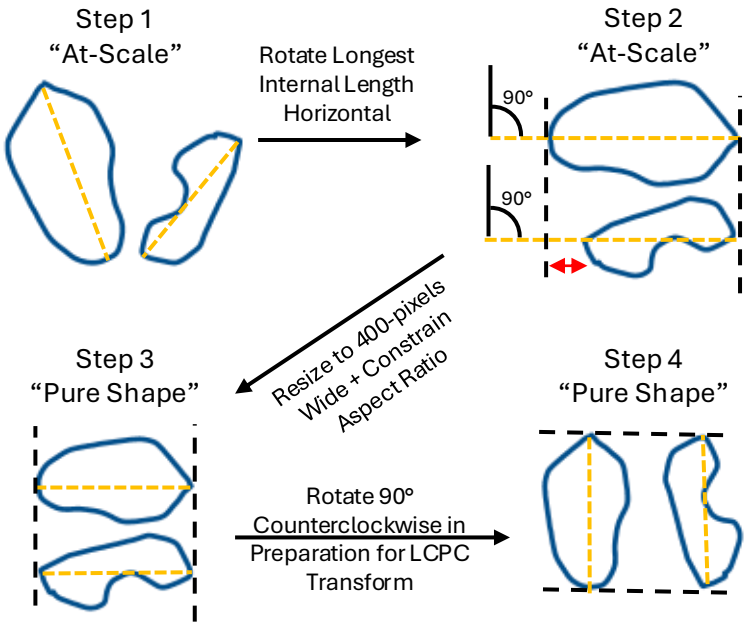


Figure 7

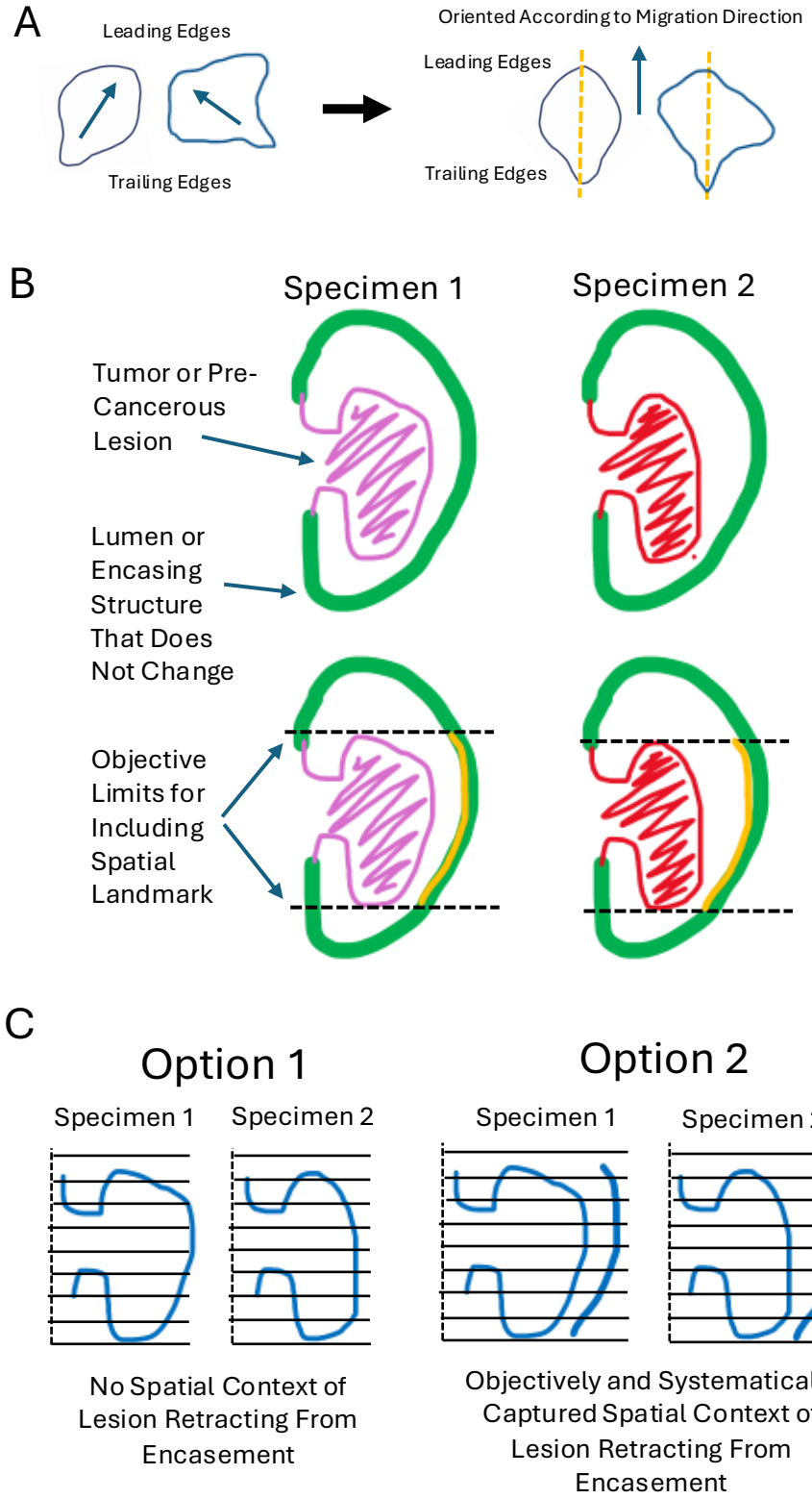

A
Leading Edges
Trailing Edges
Oriented According to Migration Direction
Leading Edges
Trailing Edges
B
Specimen 1
Specimen 2
Tumor or Pre-Cancerous Lesion
Lumen or Encasing Structure That Does Not Change
Objective Limits for Including Spatial Landmark
C
Option 1
Specimen 1
Specimen 2
No Spatial Context of Lesion Retracting From Encasement
Option 2
Specimen 1
Specimen 2
Objectively and Systematically Captured Spatial Context of Lesion Retracting From Encasement

Figure 8

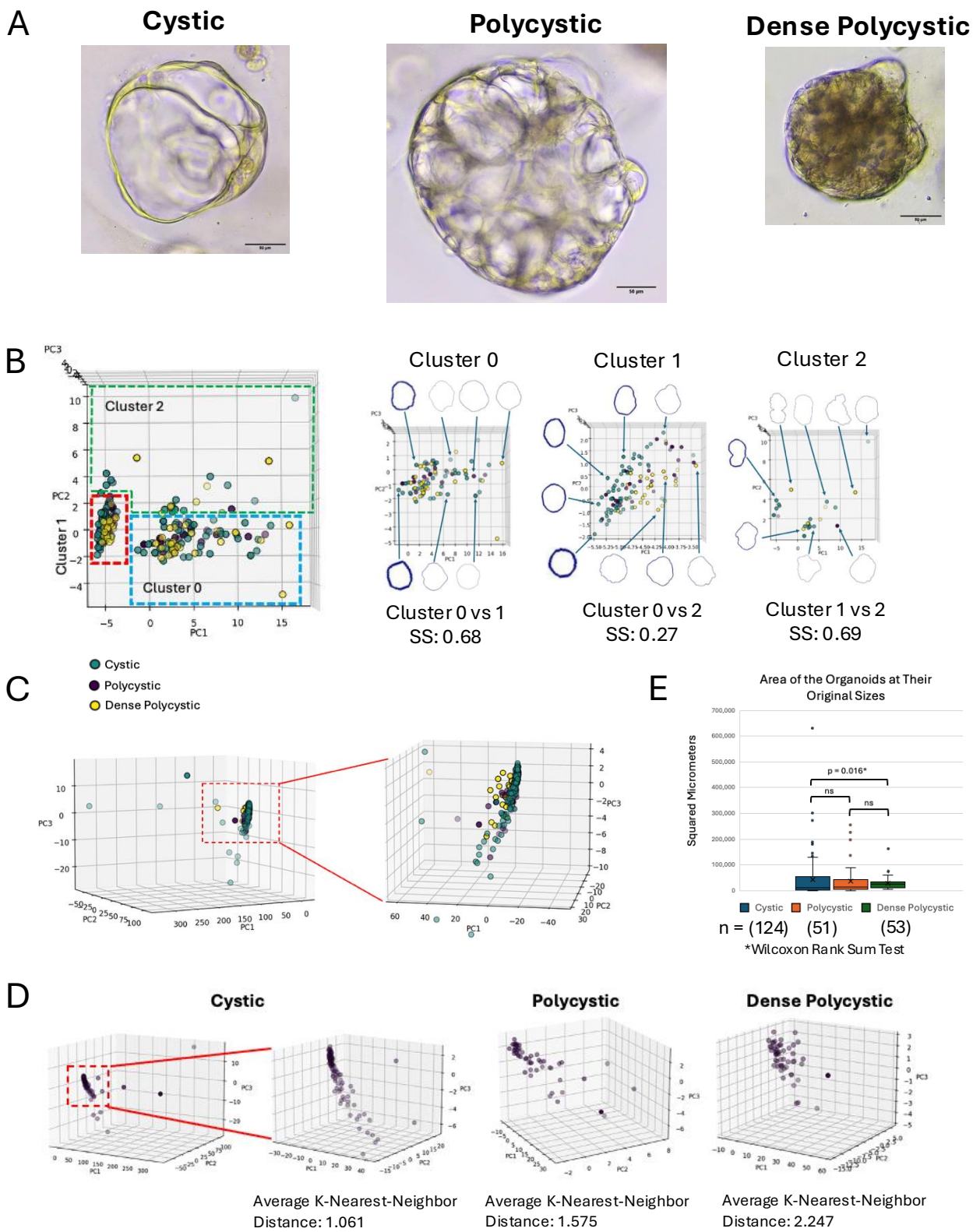
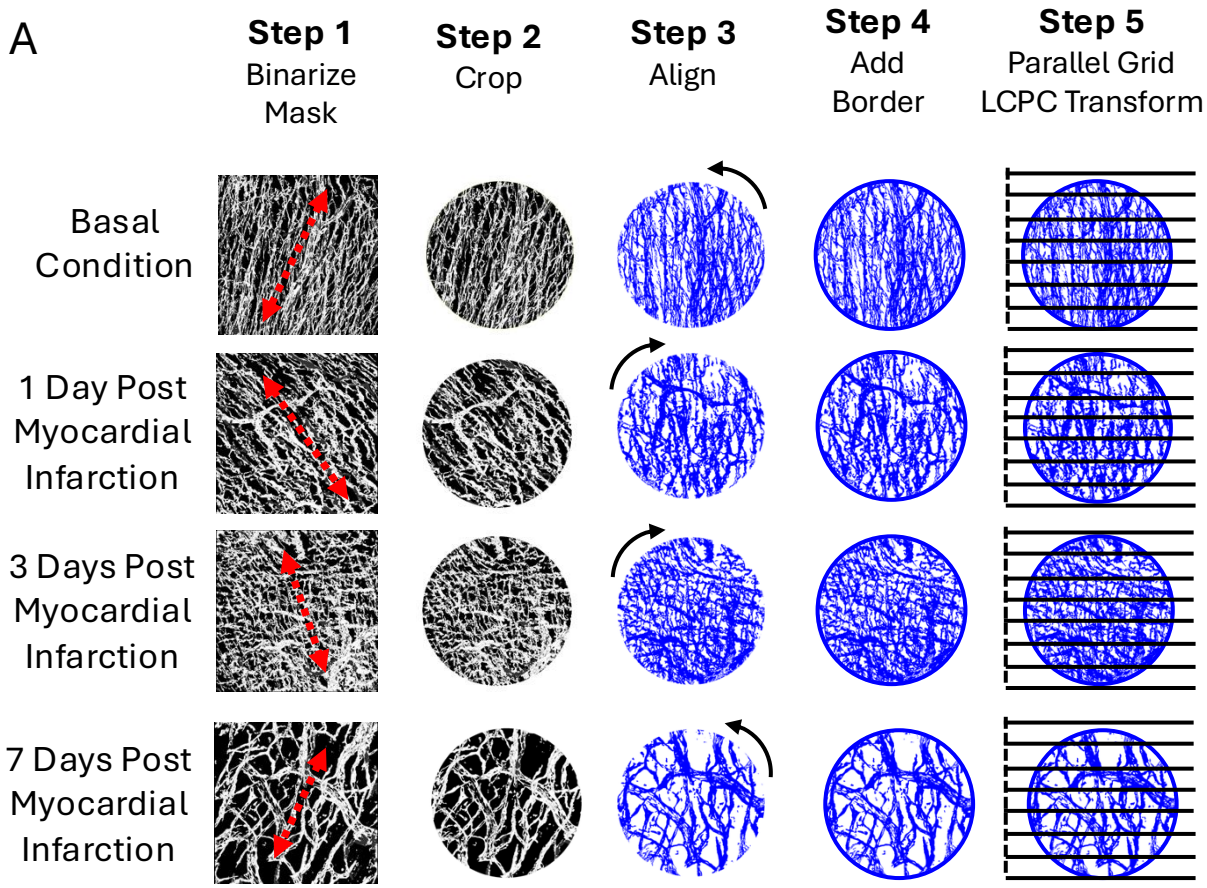
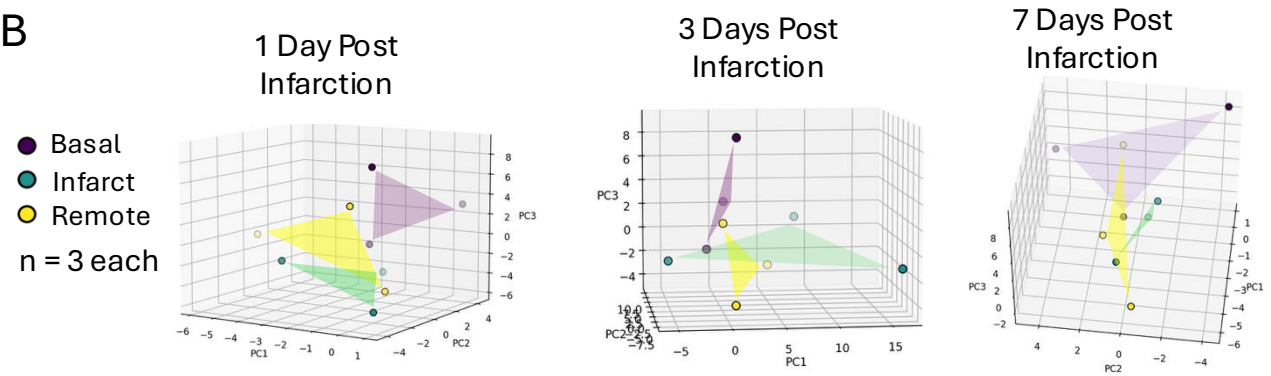

A
Cystic
Polycystic
Dense Polycystic
B
Cluster 2
Cluster 1
Cluster 0
PC1
PC2
PC3
Cluster 0
Cluster 1
Cluster 2
Cluster 0 vs 1
SS: 0.68
Cluster 0 vs 2
SS: 0.27
Cluster 1 vs 2
SS: 0.69
C
Cystic
Polycystic
Dense Polycystic
E
Area of the Organoids at Their Original Sizes
Squared Micrometers
p = 0.016*
ns
ns
Cystic
Polycystic
Dense Polycystic
n = (124) (51) (53)
*Wilcoxon Rank Sum Test
D
Cystic
Polycystic
Dense Polycystic
Average K-Nearest-Neighbor Distance: 1.061
Average K-Nearest-Neighbor Distance: 1.575
Average K-Nearest-Neighbor Distance: 2.247

Figure 9

A



B



Chamfer Distance

- Basal vs. 1Day_Infarct: 83.89
- Basal vs. 1Day_Remote: 53.95
- 1Day_Remote vs. 1Day_Infarct: 34.48

Chamfer Distance

- Basal vs. 3Day_Infarct: 240.63
- Basal vs. 3Days_Remote: 67.28
- 3Days_Remote vs. 3Days_Infarct: 168.35

Chamfer Distance

- Basal vs. 7Day_Infarct: 44.84
- Basal vs. 7Days_Remote: 39.54
- 7Days_Remote vs. 7Days_Infarct: 24.76

Figure 10

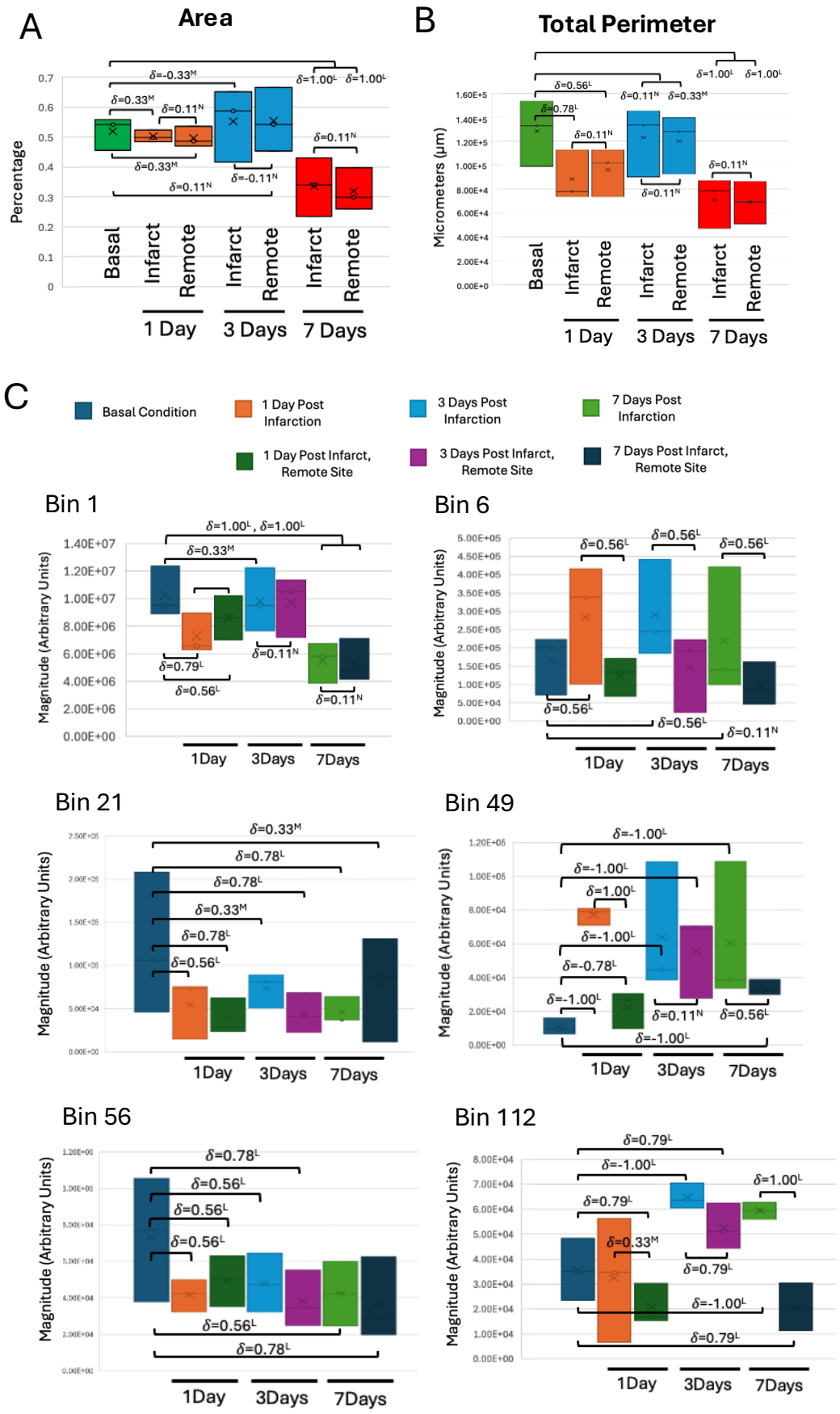
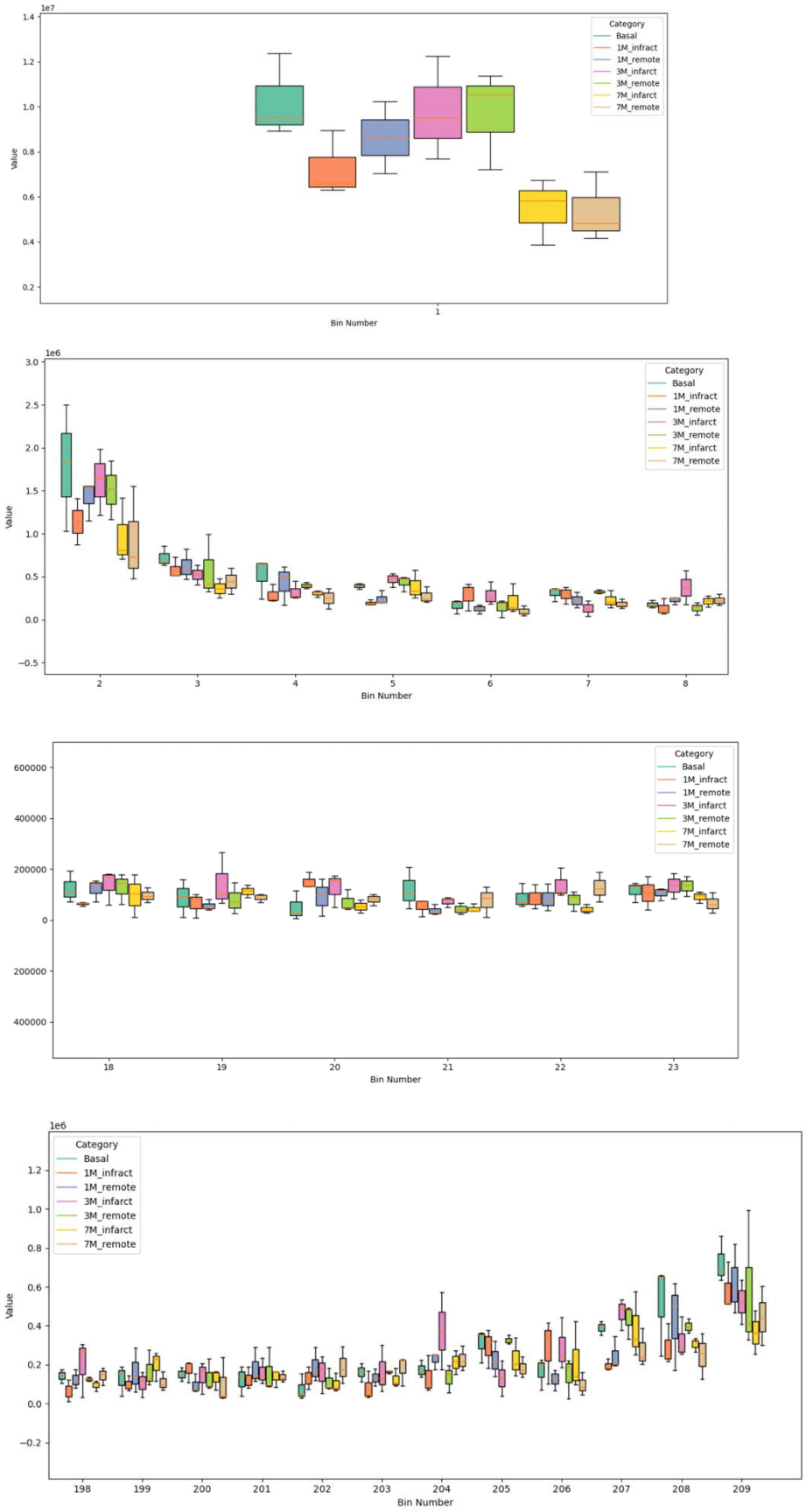

A
Area
Percentage
Basal
Infarct
Remote
1 Day
3 Days
7 Days
B
Total Perimeter
Micrometers (µm)
C
Basal Condition
1 Day Post Infarction
3 Days Post Infarction
7 Days Post Infarction
1 Day Post Infarct, Remote Site
3 Days Post Infarct, Remote Site
7 Days Post Infarct, Remote Site
Bin 1
Bin 6
Bin 21
Bin 49
Bin 56
Bin 112
Magnitude (Arbitrary Units)
1Day
3Days
7Days

Figure S1


1e7
Value
Bin Number
Category
Basal
1M_infract
1M_remote
3M_infarct
3M_remote
7M_infarct
7M_remote
1e6
Value
Bin Number
Category
Basal
1M_infract
1M_remote
3M_infarct
3M_remote
7M_infarct
7M_remote
Bin Number
Category
Basal
1M_infract
1M_remote
3M_infarct
3M_remote
7M_infarct
7M_remote
1e6
Value
Bin Number