Quantifying spatial relationships from whole retinal images

Brian E. Ruttenberg[1,2,*], Gabriel Luna[3], Geoffrey P. Lewis[3], Steven K. Fisher[3,4] and Ambuj K. Singh[1,*]

[1]Department of Computer Science, University of California, Santa Barbara, CA 93106, [2]Charles River Analytics, Cambridge, MA 02138, [3]Neuroscience Research Institute, University of California Santa Barbara, Santa Barbara, CA 93106 and [4]Department of Molecular, Cellular, and Developmental Biology, University of California Santa Barbara, Santa Barbara, CA 93106, USA

Associate Editor: Jonathan Wren

ABSTRACT

Motivation: Microscopy advances have enabled the acquisition of large-scale biological images that capture whole tissues *in situ*. This in turn has fostered the study of spatial relationships between cells and various biological structures, which has proved enormously beneficial toward understanding organ and organism function. However, the unique nature of biological images and tissues precludes the application of many existing spatial mining and quantification methods necessary to make inferences about the data. Especially difficult is attempting to quantify the spatial correlation between heterogeneous structures and point objects, which often occurs in many biological tissues.
Results: We develop a method to quantify the spatial correlation between a continuous structure and point data in large ($17\,500 \times 17\,500$ pixel) biological images. We use this method to study the spatial relationship between the vasculature and a type of cell in the retina called astrocytes. We use a geodesic feature space based on vascular structures and embed astrocytes into the space by spatial sampling. We then propose a quantification method in this feature space that enables us to empirically demonstrate that the spatial distribution of astrocytes is often correlated with vascular structure. Additionally, these patterns are conserved in the retina after injury. These results prove the long-assumed patterns of astrocyte spatial distribution and provide a novel methodology for conducting other spatial studies of similar tissue and structures.
Availability: The Matlab code for the method described in this article can be found at http://www.cs.ucsb.edu/~dbl/software.php.
Contact: bruttenberg@cra.com or ambuj@cs.ucsb.edu
Supplementary information: Supplementary data are available at *Bioinformatics* online.

1 INTRODUCTION

The advent of high-throughput large-scale microscopy has been a boon to the biological community; technological advances have enabled the capture of whole tissue at micrometer resolution, providing opportunities of study that have previously been unavailable. These improvements have especially had a profound impact on examination of the spatial distribution and correlation of biological entities, as the locations of large amounts of cells

and other structures can be imaged *in situ*. Studying the spatial arrangement and relationships inherent in tissues can improve our understanding of the various development or pathological processes that underlie proper organ or organism function (Whitney *et al.*, 2008).

Despite the abundance of image data, mining or quantifying spatial properties in tissue can be a challenging task. Although automated microscopy has significantly improved acquisition rates, capturing the spatial layout of whole cell populations or structures at high resolution can require thousands of images that take many days to acquire. Under limited time constraints, oftentimes small samples of tissue are imaged and analyzed to infer spatial properties of a whole tissue. However, many tissues are spatially heterogeneous; application of mining or quantification methods on small images may not provide a representative sample necessary to draw conclusions about the whole tissue. As such, development of spatial mining and quantification methods that can be effectively applied to extremely large images with diverse structures are highly desirable.

Yet many spatial mining methods are not suitable to handle the complex and non-traditional spatial relationships found in these large biological images. For example, neuronal or vascular structures are pervasive in many tissues, and oftentimes are spatially correlated with other cells (Armstrong, 2003; Suematsu *et al.*, 1994). Traditional spatial quantification methods such as co-location (Shekhar and Huang, 2001), K-function or nearest neighbor methods (Cressie, 1992) cannot be applied to model the relationship between these structures and cells because they inherently operate on discrete points in Euclidean space. Methods that can model the complex relationship between biological structure and cellular spatial distribution are needed to take full advantage of these rich images. In addition, comparative analysis of spatial relationships between normal and pathological tissues is highly beneficial for understanding underlying biological processes. These existing spatial methods are difficult to apply to different specimens or conditions because the Euclidean orientation and scaling between a set of images may not be the same (Hagmann *et al.*, 2008; Kaiser, 2011).

One tissue that exhibits many of the aforementioned spatial characteristics is the mammalian retina. The nerve fiber layer (NFL) of the retina is blanketed by a cell known as the astrocyte, which performs a multitude of physiological functions (Kimelberg, 2010). Astrocyte processes physically contact the vascular structure (Yu *et al.*, 2010), and they also play a key

---

*To whom correspondence should be addressed.

functional role in the development of the retinal vasculature (Metea and Newman, 2006). An apparent spatial correlation between astrocytes and the blood vessels has been noted (Stone and Dreher, 1987), but only observational evidence of such a relationship has been provided. Strong evidence of astrocyte spatial correlation with various vascular properties would lend further support to additional hypotheses of astrocyte function, such as the suspected role of astrocytes in vasodilation and constriction (Kimelberg, 2010).

Attempts to quantify the correlation between vascular structure and the spatial placement of astrocytes will suffer from the problems previously discussed. The vasculature is a large heterogeneous structure with specific arterial and venous delineations (Dorrell and Friedlander, 2006; Gariano and Gardner, 2005) that requires us to image the *entire* retina. In addition, while astrocytes can be represented as points in Euclidean space, the vasculature is a continuous structure and cannot be represented as such. Hence, a quantification method is needed that can map these disparate and large data sources into a common space.

In this article, we describe a method to quantify the spatial correlation between features of heterogeneous biological structures and Euclidean points in tissue, and apply it to the retina. The proposed method is capable of capturing differences between the features of heterogeneous structures and spatial distribution of points. The method implements a geodesic coordinate system that is based on the vascular structure. The vasculature and astrocytes are converted to this geodesic feature space such that the position of astrocytes relative to the vasculature is preserved. Astrocyte distribution in the geodesic feature space is then empirically compared with the vascular distribution. We show that the distribution of astrocytes often spatially correlates with changes in the vascular structure. Finally, we apply our methodology to a set of injured retinas and demonstrate that long-term detached retina do not deviate from the spatial patterns observed on normal retinas.

Our methodology uses a geodesic feature space that is independent of Euclidean space: the scale of the images and rotation of the retina about the optic nerve have no impact on the results. The scale and rotation invariance facilitates comparison between retinas and different cell types. For example, retinal neurons, pericytes or Müller cells can be converted into this same feature space for comparative analysis. In addition, this methodology can be used on spatial studies between other structures and cells, such as the relationship between neuronal networks and supporting cells. Our results demonstrate that we can successfully quantify the spatial correlation between these disparate biological entities, and provide a foundation for future research aimed at studying the spatial distribution of various biological components in large tissue images.

2 MATERIALS AND METHODS

The goal of our method is to determine if astrocyte distribution in the retina correlates with changes in vascular structure. For instance, as some feature of the vascular structure becomes more frequent, we want to determine if astrocyte distribution increases in close proximity to the feature. To accomplish this, we extract astrocyte and blood vessel data from large image mosaics of the retinal NFL. We then convert this retinal data into a geodesic feature space that is based on the structure of the

vasculature, preserving the relative spatial relationship between astrocytes and the vasculature. Histograms are then created for vascular and astrocyte data in the geodesic space. Finally, a distance is computed between the vascular and astrocyte histograms, and bootstrapping is performed to determine the correlation between the vascular structure and astrocyte spatial distribution. We now detail each part of our spatial quantification method.

2.1 Images and data

2.1.1 Image acquisition   Our method is designed to quantify the spatial relationship between large biological structures and point data such as cellular locations. Hence, we acquired images of the entire retinal NFL, which allows us to capture the complete retinal vasculature and all astrocytes present in the retina.

Images of mouse retinal NFL were viewed and collected on a laser scanning confocal microscope using an automated stage to capture optical sections at $0.5\,\mu m$ intervals in the z-axis and pixel resolution of 1024×1024 in the x–y direction, with 20% overlap in the x–y plane. Approximately 350–400 3D images were acquired per retina, which were then used to create maximum-intensity projections. The resulting projections were then stitched together to create a single mosaic on the order of ~17 500 pixels by ~17 500 pixels ($\sim 5400^2\,\mu m^2$).

A total of nine mosaics were created for this study, four normal and five detached. The normal retinas are denoted as $N1$, $N2$, $N3$ and $N4$. Three 1 month detached retinas, a 2 month detached retina and a 4 month detached retina are denoted as $D1$, $D1A$, $D1B$, $D2$ and $D4$, respectively. All retinas were stained with anti-GFAP and anti-collagen IV. Astrocytes express glial fibrillary acidic protein (GFAP), outlining the cytoskeleton of each astrocyte in the retina. The retinal vasculature was captured by examining the anti-collagen IV labeling.

An example mosaic is shown in Figure 1A. The retinal vasculature is clearly visible as the tree-like structures in the images. Individual astrocytes are visible as small star-shaped cells and can be seen in detail in Figure 1B. Color images of astrocytes in Figure 1B can be found in the Supplementary Material, in addition to further details on the tissue preparation and the image-acquisition process.

2.1.2 Data extraction   Each retina contains five or six arteries and veins arranged in an alternating pattern around the optic nerve head. The veins are easily identified on the images by their 'conveying' type branching (Ganesan *et al.*, 2010). As explained in the Supplementary Material, the images are slightly modified so that there exists only one path from any point on the vasculature to the optic nerve head, thus enforcing a tree structure on each artery and vein. The vascular network in each retina is traced using Neuronstudio (Rodriguez *et al.*, 2008), an automated tracing tool for network structures. The trace is discretized into $2\,\mu m$ segments of vasculature, and the tree structure of each of the veins and arteries is recorded. That is, the parent and ancestor of each $2\,\mu m$ segment are recorded.

Anti-GFAP staining of the retinal tissue labels the cytoskeleton of all astrocytes in the retina. As the nucleus of the cell often contained the highest concentration of GFAP, the presumed center of each cell is manually selected from the image, and the x–y coordinates recorded. There are ~3500–4000 astrocytes in each image.

2.2 Geodesic feature space

The core of our method is the transformation of all retinal data to a 2D geodesic feature space that is based on the structure of the vasculature. The first dimension measures the distance from a location in the retina to the optic nerve head *through* the vasculature (i.e. geodesic distance on a blood vessel). The second dimension represents physical blood vessel width.

This vascular centric feature space is central to our quantification methodology because it allows us to express the physical area and
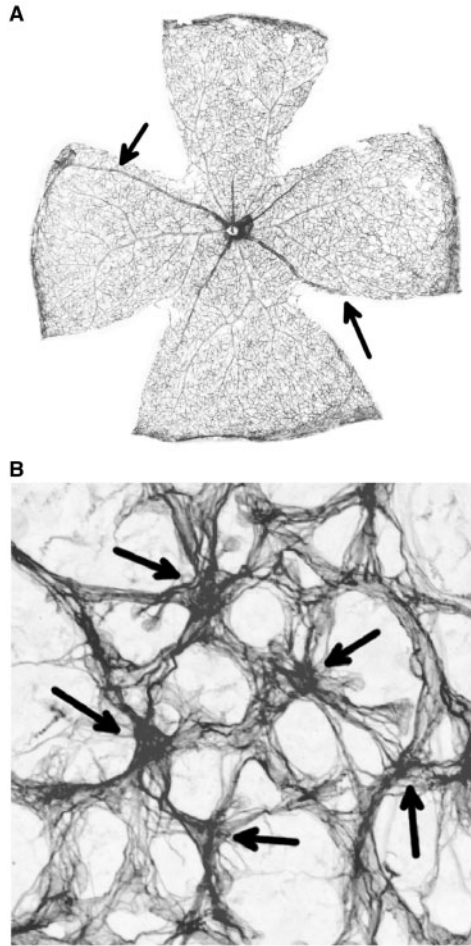
Fig. 1. (**A**) An example retinal mosaic used in the study. The blood vessels are easily seen alternating around the optic nerve head in the center of the retina. The arrows indicate two large veins that are present in each retina, discussed in Section 3. (**B**) Magnified view of a retinal mosaic, with the arrows indicating the center of the astrocytes

extent of the vasculature in structural terms. For example, vessels tend to become thinner as they branch outward from the optic nerve, and increased branching also leads to increased geodesic distances. Using the geodesic distance to the optic nerve head allows us to orient the data to a single point, while still maintaining the spatial relationships in the original retina. In addition, this space is not specific to any one retina; all retinas possess an optic nerve head, thus enabling direct comparison of the vasculature from different retinas.

Finally, each retina has been dissected to enable proper image acquisition of the naturally curved tissue. These incisions (easily seen in Fig. 1A) frequently disconnect portions of the vascular plexus, and distort Euclidean distances in the retina. When such a disconnection occurs, the vascular network is manually reconnected across the incision, and the geodesic distance across the incision is set to zero, which easily accounts for these Euclidean distortions.

2.2.1 Retinal data conversion   The vasculature and astrocytes in a retina must be converted to data in the geodesic feature space to quantify their spatial relationship. Converting the vasculature is a simple process, as any point on the vasculature can be directly converted to a point in the geodesic feature space. Converting the astrocytes is a more complicated endeavor, as astrocytes are cells that are not natively associated with vascular distance or width.
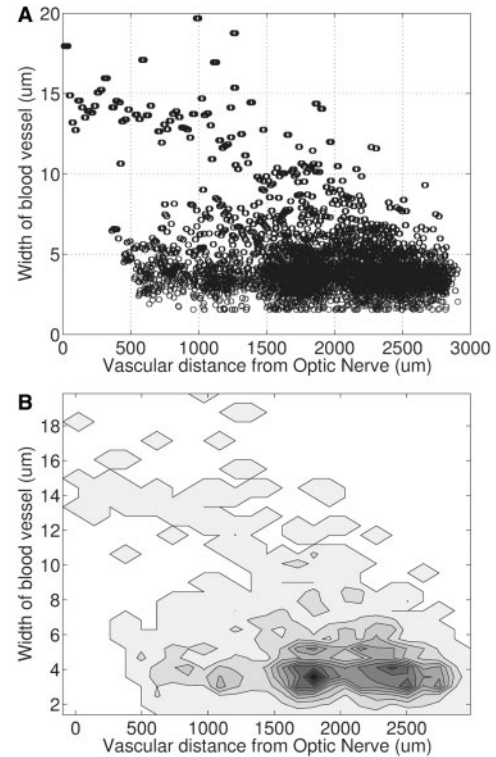

Fig. 2. (**A**) Scatter plot of an artery in geodesic space. Each circle represents a $2\,\mu$m section of the vasculature. (**B**) Contour plot of the normalized histogram of the same artery in geodesic space

The discrete vascular segments for each of the veins and arteries are converted to a feature vector $\vec{V_i}$ ($\vec{A_i}$), where each $V_i[k]$ is a tuple (d, w, x, y) that represents the geodesic distance and width for each $2\,\mu$m segment of vessel, as well as the Euclidean coordinates for the segment. The Euclidean coordinates are only used to aid in astrocyte conversion. The closest segment in each artery or vein to the optic nerve head is considered the root of the vessel tree, and the geodesic distance to the optic nerve head from the root is set as zero. The geodesic distance for all other $2\,\mu$m segments is simply computed as the distance from a segment's parent (the previously connected segment on the vasculature) to the optic nerve head, plus two microns. Figure 2A shows an example visual representation of $V_i[k].d$ and $V_i[k].w$ for a single artery. As can be seen, conversion of a blood vessel tree into the feature space produces a distinct pattern that describes the structure of the tree. For testing, the distance and width values are normalized by the mean and standard deviation of the vector to put the data on the same scale.

Each astrocyte is embedded into the feature space by *spatially sampling* the vasculature in close proximity. That is, an astrocyte is converted into a series of points in the geodesic feature space based on the structure of the vasculature in some local region near the astrocyte. This allows us to preserve astrocyte spatial distribution relative to the vasculature while putting both entities in the same feature space.

Let c_i denote an astrocyte in the retina, and $c_i.x$ and $c_i.y$ denote the Euclidean coordinates of the astrocyte. A feature vector for c_i is denoted as $\vec{c_i}$, where $c_i[k]$ is a tuple (d, w, n, p), representing the geodesic distance, width, primary vessel (the nth vein or artery) and a weight. For each astrocyte in the retina, we sample S_{bv} vascular segments in the retina (i.e. the length of $\vec{c_i}$ is S_{bv}). The Euclidean distance from astrocyte c_i to all $2\,\mu$m segments of vasculature in the retina is computed,

i.e. all of $\vec{V_j}$ and $\vec{A_j}$. Let D be the Euclidean distance between c_i and the S_{bv}^{th} nearest $2\,\mu$m segment to c_i. The vector $\vec{c_i}$ is then

$$\vec{c_i} = (V_j[k].d, V_j[k].w, j, \mathcal{G}_\Sigma(\vec{e})) \qquad (1)$$
$$\forall j, k : ||V_j[k] - c_i||_2 \le D$$

where $\vec{e} = (c_i.x - V_j[k].x, c_i.y - V_j[k].y)$, the difference in Euclidean coordinates of astrocyte c_i and vessel segment $V_j[k]$. Note that $\vec{c_i}$ may be composed of a combination of vein and artery segments; only the vein notation was used in Equation (1) for explanation purposes. Essentially, an astrocyte is converted into the geodesic feature space by assigning it a vector of the geodesic distance and width of the S_{bv} nearest vascular segments to the cell. Note we also record the origin of the sampled artery/vein segments.

$\mathcal{G}_\Sigma(\cdot)$ is a zero mean 2D Gaussian function with a diagonal covariance matrix Σ, applying a weight proportional to the distance between the sample and the astrocyte. Vascular segments in close proximity to astrocytes are given more importance over distant blood vessel segments. After a feature vector is constructed for an astrocyte, the weights the spatial samples in $\vec{c_i}$ are normalized to sum to one. That is, $c_i[k].p = c_i[k].p / (\sum_{j=1}^{S_{bv}} c_i[j].p)$.

2.3 Quantifying spatial relationships

2.3.1 Histograms We compose the vascular and astrocyte feature vectors into histograms so that we can compare the spatial distribution of astrocytes relative with the vasculature. One major advantage of our method is the ability to explore astrocyte distribution at different vascular scales. Three different levels of histograms are implemented: retinal, arterial/venous and individual blood vessels. We compose astrocyte histograms at the same resolution as the vasculature histograms so that we can test astrocyte spatial distribution at different granularities.

Construction of a histogram for a vein V_i and its accompanying astrocyte histogram is the same for any other individual artery/vein histogram. A vein histogram is denoted as $\mathbb{V}_i$, and the accompanying astrocyte histogram as $\mathbb{AV}_i$, where $\mathbb{V}_i = (B, N)$ and $\mathbb{AV}_i = (B, M)$. B is a common set of bins in the geodesic space, and N and M are the counts for the respective histogram bins.

The construction of $\mathbb{V}_i$ is simple. The count N_j for some bin B_j is computed as

$$N_j = \sum_{k=1}^{|V_i|} I_v(V_i[k], B_j)$$

where

$$I_v(V_i[k], B_j) = \begin{cases} 1, & \text{if } (V_i[k].d, V_i[k].w) \in B_j \\ 0, & \text{otherwise} \end{cases}$$

The process for constructing $\mathbb{AV}_i$ is similar, and the count M_j is

$$M_j = \sum_{z=1}^{|C|} \sum_{k=1}^{|\vec{c_z}|} I_c(c_z[k], B_j, i) \cdot c_z[k].p$$

where C is the set of all astrocytes, and

$$I_c(c_z[k], B_j, i) = \begin{cases} 1, & \text{if } (c_z[k].d, c_z[k].w) \in B_j \wedge c_z[k].n = i \\ 0, & \text{otherwise} \end{cases}$$

In other words, the astrocyte histogram $\mathbb{AV}_i$ is the weighted count of all astrocytes that contain a spatial sample from vein V_i.

To construct higher level histograms, histograms from the lower levels are combined. For example, the venous histogram $\mathbb{V} = \bigcup_{i=1}^{k} \mathbb{V}_i$, where k is the number of veins, and the top-level vascular histogram is simply $\mathbb{V} \cup \mathbb{A}$ (similarly for astrocyte histograms).

The bin set B is set so that each B_j covers a 0.1×0.1 region (recall that the distance and width are normalized by the mean and variance). The histogram counts are also normalized to sum to one to enable direct comparison of histograms between retinas.

2.3.2 Histogram distance We quantify the spatial relationship between astrocytes and the vasculature by computing the probability that an astrocyte histogram is a random sample from the corresponding vascular histogram. Because the astrocyte histograms are constructed by spatially sampling the vasculature, we hypothesize that the two histograms should be similar if astrocyte distribution is highly correlated with the structure of the vasculature.

A statistical test of probability distribution similarity is used, using the Mallows distance (Munk and Czado, 1998). For two distributions X and Y, the test statistic is

$$F(X, Y) = \sqrt{n}\{M(X_n, Y) - M(X, Y)\}$$

where X_n is an empirical sample of size n from X. Note that our histograms are normalized to sum to one, equivalent to a probability distribution.

$M(\cdot)$ is the Mallows distance between probability distributions. The Mallows distance has several desirable statistical properties (Bickel and Freedman, 1981; Mallows, 1972). It is convergent up to the lth moment, as well as convergent in distribution. These properties ensure that differences in distribution between astrocyte and vascular histograms are accurately captured. Additional details on the Mallows distance can be found in the Supplementary Material.

Bootstrapping is performed to compute a prior distribution of the test statistic F for a given vascular histogram (Shao and Tu, 1995). A vascular histogram is re-sampled multiple times, and the prior distribution of the test statistic is computed using the re-sampled data. The hypothesis that astrocyte distribution is correlated with a vascular histogram is unlikely to be correct if the test statistic between an astrocyte and vascular histogram is low compared with the re-sampled histograms.

Distance and width values are uniformly sampled from the blood vessel histogram under test (e.g. $\mathbb{V}_i$), and a sample histogram is constructed. The number of samples used is equal to the number of astrocytes that have a spatial sample from V_i. For some astrocyte $\vec{c_j}$ that has a spatial sample from V_i (i.e. $c_j[k].n = i$), sample j contributes to the sample histogram the total weight of $\vec{c_j}$ in V_i. That is, for each sample j, $\sum_k c_j[k].p, \forall c_j[k].n = i$, is added to the sample histogram (into the appropriate bin).

Note that the method of significance testing does not depend on the number of retinal mosaics that are collected. The bootstrapping procedure provides a statistically sound method to compare two histograms and is only used to determine if an astrocyte histogram is a random sample from a vascular histogram. Hence even with a single retina, the significance testing procedure still provides accurate and statistically valid hypothesis testing of the astrocyte histograms. Extrapolating the results to a general population does, of course, require more than a single retina.

3 RESULTS

The Mallows distance quantification method was performed on the three histogram scenarios for the nine retinas. In addition, the method was also tested on synthetic retinal data to demonstrate its effectiveness at capturing the spatial correlation between vascular structure and astrocyte distribution. Σ is initially set to be $80\,\mu$m, which is approximately the width of an average retinal astrocyte. S_{bv} is varied from 1 to 40. Each bootstrapping test used 500 random samples from a histogram. Exploration of the effect of Σ on the statistical testing can be found in the Supplementary Material.

3.1 Synthetic retinal data

Synthetic data was generated and tested to determine the method's ability to capture spatial correlation. An artificial retina was
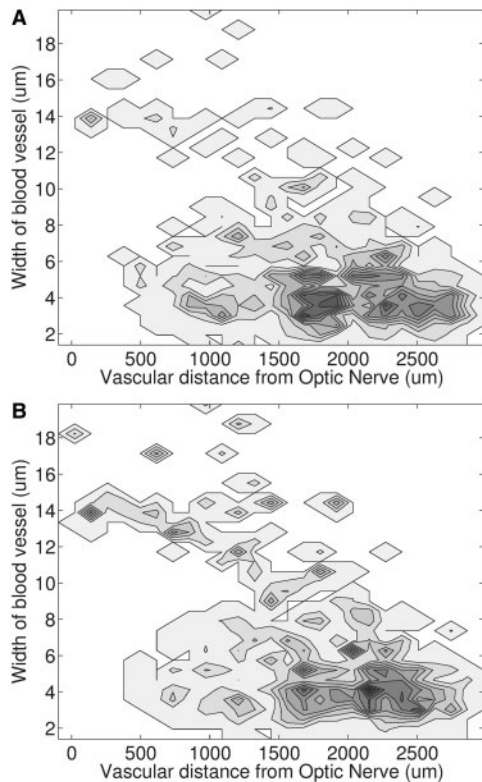
Fig. 3. (A) Contour plot of the normalized histogram of astrocytes using random placement. (B) Contour plot of the normalized histogram of astrocytes using width weighted placement

Fig. 4. Average *P*-values of astrocyte histograms for the artificial retinas. The *P*-values decrease as the width and distance weighted placement deviates farther from the distribution of the vasculature

created for this task; the vasculature of the artificial retina was the vasculature for retina $N1$. Artificial astrocytes were placed in the retina in one of three methods. In the random method, a $2\,\mu m$ vascular segment was uniformly selected at random from the vasculature, and the artificial astrocyte's image coordinates were set as the $2\,\mu m$ segment's image coordinates, added to a 2D gaussian distribution with standard deviation $10\,\mu m$. In the width weighted method, a vascular segment i was randomly selected with probability $P(w_i) = \frac{w_i^z}{\sum w_j^z}$, where the normalization is the sum of the widths for all segments in the vasculature, for some z. Finally, in the distance weighted method, a vascular segment was chosen with probability $P(d_i)$, where d_i is the geodesic distance of the segment. $P(\cdot)$ is a normal distribution centered around d_{mean}, the mean geodesic distance of the vasculature, and a standard deviation of $\frac{d_{mean}}{c}$, for some c. For both the width and distance methods, the artificial astrocyte coordinates were determined the same as the random method (with different random selection methods).

Figure 3A and B show the normalized astrocyte histograms for the random and width weighted methods. The original normalized vascular histogram is the same one shown in Figure 2B. Visually, the random placement histogram is similar to the vascular histogram from Figure 2B. However, in contrast to the random placement method, the width weighted histogram places much more density in the regions of the space with large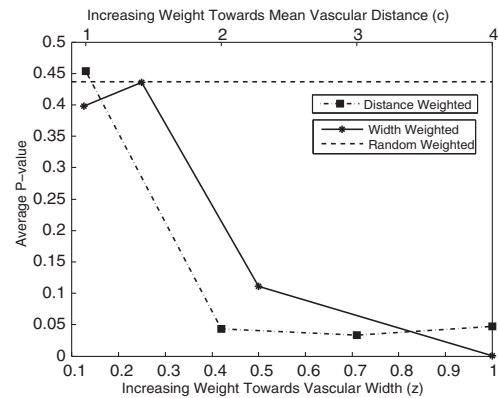r vascular width. While the differences between these two methods are subtle, they become immediately apparent using our method. An example image of a synthetic retina is available in the Supplementary Material, as well as the normalized histogram for the distance weighted method.

After creating the artificial retinas, the testing procedure as previously described was applied. Figure 4 shows the average *P*-value of the individual artery/vein histogram tests for the synthetic data. Both z and c are varied so that the width and distance weighted placements initially are close to the random method, yet are increased over various tests. The *P*-value is defined as the probability that a random test statistic is greater than the one computed between the astrocyte and blood vessel histograms. A low *P*-value indicates the hypothesis that astrocytes are spatially correlated with the corresponding portion of the vasculature is unlikely to be correct. As can be seen, as both z and c increase, the average *P*-values converge toward zero, as astrocyte distribution becomes increasingly influenced by the width/distance of the vasculature.

Clearly, there are many other models of randomly distributing synthetic astrocytes in the retina (spatially uniformly distributed, for example), and it is unrealistic to assume that our method could discern differences in spatial distribution for all models. However, the results do demonstrate that our method can detect differences in astrocyte distribution when two important structural properties of the vasculature (width and distance) are assumed to exert a significant influence on spatial placement. Testing of additional models of astrocyte distribution is an attractive target for future research.

3.2 Vascular/arterial/venous histograms

The artery and vein histograms are combined into a single retinal histogram for testing. Tables of the detailed *P*-value results are available in the Supplementary Material. The normal retinas generally have very low *P*-values, with a slight increase observed on retinas $N3$ and $N4$ as the number of spatial samples is increased. The detached retinas show a similar pattern of very low *P*-values at the whole retina level. With a single spatial sample, it is unlikely that the retinal level histograms are a random sample from the vasculature, be it a normal or detached retina. This result is expected for the retinal level histograms; the

P-values reflect all local differences in spatial distribution between the astrocytes and vasculature anywhere in the retina.

All individual artery and vein histograms are then combined into separate arterial and venous histograms (tables of *P*-values available in Supplementary Material). Most of the venous histograms have very low *P*-values. In contrast, most of the arteries have high *P*-values, or at least higher *P*-values than their corresponding venous histograms. Two basic patterns appear to be emerging: arterial astrocytes are spatially distributed as random samples from the arterial structure, and venous astrocytes spatially deviate from the venous structure.

3.3 Individual vein/artery histograms

The same pattern from the coarser scale histograms is evident on the individual blood vessel tests. With a few exceptions, most of the individual arteries have non-significant *P*-values. Even *N*1, where the arteries at coarser scales had low *P*-values, contains high *P*-values. This demonstrates that astrocytes in close proximity to the arteries are spatially distributed similarly to structure of the arteries. This pattern is evident at the arterial scale and the individual vessel scale, for both the normal and detached retinas. The *P*-values for the individual blood vessel tests can be found in the Supplementary Material.

$V(N1, 1)$ is examined more closely to determine why the *P*-values are low. Figure 5A shows the difference between the astrocyte histogram and vein histogram from $V(N1, 1)$ $(\mathbb{A}\mathbb{V}_1 - \mathbb{V}_1)$. The lighter contours indicate the regions of feature space where the astrocyte histogram is more dense than the vein histogram, and conversely for darker contours. As a comparison, the histogram difference for $A(N2, 4)$ is shown in Figure 5B, which has a very high *P*-value. There are large regions where the densities of the two histograms significantly differ. In Figure 5A, the astrocyte histogram is more dense on thicker blood vessels that are closer to the optic nerve head, indicated by the two light regions in the figure. Note how in Figure 5B, the regions of difference between the astrocyte and artery histogram are generally small and local.

4 DISCUSSION

In this work, we developed a method to quantify the spatial relationship between a heterogeneous structure and discrete points in Euclidean space. We used a novel geodesic feature space and studied the relationship between astrocytes and the vasculature in the retinal NFL. The vasculature and astrocytes are converted to this feature space, which facilitates comparison of astrocyte spatial distribution relative to the vasculature. Empirical quantification using the Mallows distance reveals that astrocytes are spatially distributed independently from the vasculature, with the exception of increased astrocyte density on thick portions of veins.

Our method of quantifying the spatial relationship between large structures and point objects has many advantages. The used feature space is based on the geodesic distance to a common object in all the images, thus enabling direct comparison of all data. Furthermore, conversion of the vascular data to geodesic distance allows different resolutions of spatial relationships to be tested; such varying resolutions could be applied to other domains as well; for instance, from individual neurons to larger circuits. Lastly, use of the Mallows distance to compare histograms provides an accurate reflection of the differences between vascular structure and astrocyte spatial distribution. Clustering of point objects around specific regions of large continuous structures in diverse tissues can be easily identified using our method.

Our method and subsequent results already have an immediate impact on our understanding of the underlying biological processes occurring in the retina. It is clear from our results that detachment appears to have no discernible effect on the distribution of astrocytes with respect to the vasculature. This is an unexpected result, as it has been shown in various animals that astrocytes react in response to trauma, such as the up-regulation of GFAP in astrocytes (Luna *et al*., 2010; Sakai *et al*., 2003), or by proliferation (Fisher *et al*., 1991; Panagis *et al*., 2005). These previously observed injury responses by astrocytes appear to have little to no long-term impact on their spatial distribution relative to the vasculature, indicating that a specific spatial configuration of astrocytes may be required for healthy function of the retina. These results lend insight into biological processes that are occurring in an adult retina.

There are significant broader implications of our proposed method as well. The geodesic distance in our method is not restricted to radial vascular structures; it only requires a reference point. Therefore, it can be used on other vascular networks, such as the brain vascular network, or even complete organisms where the entire vasculature is observable. In addition, many other
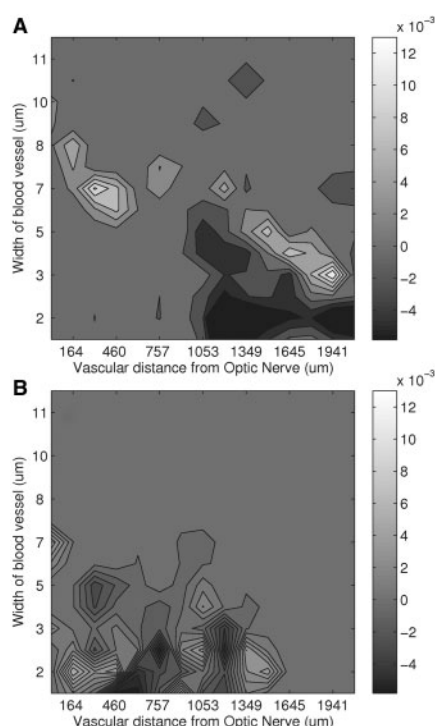

Fig. 5. (**A**) Differences between the vascular and astrocyte histograms of the low *P*-value vein $V(N1, 1)$. (**B**) Differences in the high *P*-value artery $A(N2, 4)$. Light areas indicate regions where the astrocyte histogram was denser. In (**A**), there are large regions where the two histograms significantly differ, as seen by the intense light and dark

biological tissues exhibit similar structure and relationships to those encountered in the retina. Our feature space and quantification methodology can be used to study spatial relationships in these domains as well.

For instance, connectomes of neurons and glial cells (LoTurco, 2000; Anderson *et al.*, 2011) contain large continuous structures similar to the vasculature and are often spatially correlated with other cells or entities (Leergaard *et al.*, 2012; Sporns, 2011). The geodesic feature space can be adapted for neuronal axons or circuitry; a common physiological or anatomical reference point can be used in lieu of the optic nerve (e.g. cell soma), along with axonal width or any other spatially correlated structural feature. The locations of other nearby cells can then be converted into this geodesic space (now based on neurons), and the spatial correlation between cells and neuronal structure can be quantified in a similar manner to astrocytes and the vasculature.

There have also been efforts to quantify the structure of neurons using methods similar to our proposed method. These methods attempt to characterize the morphology of neuronal structures using topological features, such as diameter, path length or branching factors (Brown *et al.*, 2008; Cuntz *et al.*, 2007; Vonhoff and Duch, 2010). While these approaches are not specifically designed to measure spatial correlation, they can be incorporated into our method as an additional histogram dimension. For instance, Sholl analysis (Sholl, 1953) could be used in conjunction with soma distance and dendrite width to measure spatial correlation between neurons and other biological features using our method.

Finally, our methodology will provide further insight into the biological processes in the retina. Spatial correlation can be studied using our method in new experimental conditions, such as development or disease. These additional spatial studies of retinal conditions, as well as spatial studies on tissues with similar structure, are promising avenues of future research efforts.

Conflict of Interest: none declared.

REFERENCES

Anderson,J. *et al.* (2011) Exploring the retinal connectome. *Mol. Vis.*, **17**, 355.

Armstrong,R. (2003) Measuring the degree of spatial correlation between histological features in thin sections of brain tissue. *Neuropathology*, **23**, 245–253.

Bickel,P. and Freedman,D. (1981) Some asymptotic theory for the bootstrap. *Ann. Stat.*, **9**, 1196–1217.

Brown,K. *et al.* (2008) Quantifying neuronal size: summing up trees and splitting the branch difference. *Semin. Cell. Dev. Biol.*, **19**, 485–493.

Cressie,N. (1992) Statistics for spatial data. *Terra Nova*, **4**, 613–617.

Cuntz,H. *et al.* (2007) Optimization principles of dendritic structure. *Theor. Biol. Med. Model.*, **4**, 21.

Dorrell,M. and Friedlander,M. (2006) Mechanisms of endothelial cell guidance and vascular patterning in the developing mouse retina. *Prog. Retinal Eye Res.*, **25**, 277–295.

Fisher,S. *et al.* (1991) Intraretinal proliferation induced by retinal detachment. *Invest Ophthalmol. Vis. Sci*, **32**, 1739–1748.

Ganesan,P. *et al.* (2010) Development of an image-based network model of retinal vasculature. *Ann. Biomed. Eng.*, **38**, 1566–1585.

Gariano,R. and Gardner,T. (2005) Retinal angiogenesis in development and disease. *Nature*, **438**, 960–966.

Hagmann,P. *et al.* (2008) Mapping the structural core of human cerebral cortex. *PLoS Biol.*, **6**, e159.

Kaiser,M. (2011) A tutorial in connectome analysis: topological and spatial features of brain networks. *Neuroimage.*, **57**, 892–907.

Kimelberg,H. (2010) Functions of mature mammalian astrocytes: a current view. *Neuroscientist*, **16**, 79–106.

Leergaard,T. *et al.* (2012) Mapping the connectome: multi-level analysis of brain connectivity. *Front. Neuroinform.*, **6**, 14.

LoTurco,J. (2000) Neural circuits in the 21st century: synaptic networks of neurons and glia. *Proc. Natl Acad. Sci.*, **97**, 8196.

Luna,G. *et al.* (2010) Expression profiles of nestin and synemin in reactive astrocytes and muller cells following retinal injury: a comparison with glial fibrillar acidic protein and vimentin. *Mol. Vis.*, **16**, 2511–2523.

Mallows,C.L. (1972) A note on asymptotic joint normality. *Ann. Math. Stat.*, **43**, 508–515.

Metea,M. and Newman,E. (2006) Glial cells dilate and constrict blood vessels: a mechanism of neurovascular coupling. *J. Neurosci.*, **26**, 2862–2870.

Munk,A. and Czado,C. (1998) Nonparametric validation of similar distributions and assessment of goodness of fit. *J. R. Stat. Soc. B*, **60**, 223–241.

Panagis,L. *et al.* (2005) Unilateral optic nerve crush induces bilateral retinal glial cell proliferation. *Eur. J. Neurosci.*, **21**, 2305–2309.

Rodriguez,A. *et al.* (2008) Automated three-dimensional detection and shape classification of dendritic spines from fluorescence microscopy images. *PLoS One*, **3**, e1997.

Sakai,T. *et al.* (2003) Cone photoreceptor recovery after experimental detachment and reattachment: an immunocytochemical, morphological, and electrophysiological study. *Invest Ophthalmol. Vis. Sci.*, **44**, 416–425.

Shao,J. and Tu,D. (1995) *The Jackknife and Bootstrap*. Springer, London.

Shekhar,S. and Huang,Y. (2001) Discovering spatial co-location patterns: a summary of results. *Adv. Spat. Temporal Databases*, **2121**, 236–256.

Sholl,D. (1953) Dendritic organization in the neurons of the visual and motor cortices of the cat. *J. Anat.*, **87** (Pt 4), 387.

Sporns,O. (2011) The human connectome: a complex network. *Ann. NY Acad. Sci.*, **1224**, 109–125.

Stone,J. and Dreher,Z. (1987) Relationship between astrocytes, ganglion cells and vasculature of the retina. *J. Comp. Neurol.*, **255**, 35–49.

Suematsu,M. *et al.* (1994) Spatial and temporal correlation between leukocyte behavior and cell injury in postischemic rat skeletal muscle microcirculation. *Lab. Invest.*, **70**, 684.

Vonhoff,F. and Duch,C. (2010) Tiling among stereotyped dendritic branches in an identified drosophila motoneuron. *J. Comp. Neurol.*, **518**, 2169–2185.

Whitney,I. *et al.* (2008) Spatial patterning of cholinergic amacrine cells in the mouse retina. *J. Comp. Neurol.*, **508**, 1–12.

Yu,P. *et al.* (2010) The structural relationship between the microvasculature, neurons, and glia in the human retina. *Invest. Ophthalmol. Vis. Sci.*, **51**, 447–458.