Math. Model. Nat. Phenom.
Vol. 9, No. 5, 2014, pp. 177–193
DOI: 10.1051/mmnp/20149512

Efficient Processing of Fluorescence Images Using Directional Multiscale Representations

D. Labate[1] *, F. Laezza[2], P. Negi[1], B. Ozcan[1], M. Papadakis[1]
[1] Dept. of Mathematics, University of Houston, Houston, Texas 77204, USA
[2] Dept. of Pharmacology and Toxicology, UT Medical Branch, Galveston, TX 77555, USA

Abstract. Recent advances in high-resolution fluorescence microscopy have enabled the systematic study of morphological changes in large populations of cells induced by chemical and genetic perturbations, facilitating the discovery of signaling pathways underlying diseases and the development of new pharmacological treatments. In these studies, though, due to the complexity of the data, quantification and analysis of morphological features are for the vast majority handled manually, slowing significantly data processing and limiting often the information gained to a descriptive level. Thus, there is an urgent need for developing highly efficient automated analysis and processing tools for fluorescent images. In this paper, we present the application of a method based on the shearlet representation for confocal image analysis of neurons. The shearlet representation is a newly emerged method designed to combine multiscale data analysis with superior directional sensitivity, making this approach particularly effective for the representation of objects defined over a wide range of scales and with highly anisotropic features. Here, we apply the shearlet representation to problems of soma detection of neurons in culture and extraction of geometrical features of neuronal processes in brain tissue, and propose it as a new framework for large-scale fluorescent image analysis of biomedical data.

Keywords and phrases: curvelets, fluorescent microscopy, image processing, segmentation, shearlets, sparse representations, wavelets

Mathematics Subject Classification: 42C15, 42C40, 92C15, 92C55

1. Introduction

High-resolution fluorescence microscopy has emerged during the last decade as a fundamental tool for the study of molecular mechanisms in biological samples. Unlike traditional optical microscopy, fluorescence microscopy is based on the principle of absorption and subsequent re-radiation of light emitted from fluorophores with separable spectral properties. Through chemical conjugation, these fluorophores can be linked with a high degree of specificity to single molecules and used as probes to track sub-cellular localization and expression pattern of any protein of interest. Hence, fluorescence microscopy can be invaluable to unravel structure-function properties at single cell level, in tissues or in whole organisms. Furthermore, combined with automated image acquisition and robotic handling, it can be applied for the systematic study of morphological changes within a large populations of cells under a variety of perturba-

*Corresponding author. E-mail: dlabate@math.uh.edu

© EDP Sciences, 2014

Article published by EDP Sciences and available at http://www.mmnp-journal.org or http://dx.doi.org/10.1051/mmnp/20149512
https://doi.org/10.1051/mmnp/20149512 Published online by Cambridge University Press

tions (e.g., drugs, compounds, gene silencing), so that automated fluorescence microscopy has become an essential technique for discovering new molecular pathways in diseases or new potential pharmacological treatments. However, the manual or semi-manual analysis of the large data sets acquired from these studies is very labor-intensive and there is a urgent need for automated methods which can rapidly and objectively process and analyze the image data. [24, 27]

The efficient analysis and processing of fluorescence images presents several additional challenges with respect to standard image processing problems. Fluorescence images are formed by detecting small amount of lights emitted by the fluorophores, a process which is affected by the randomness of photon emission, the spatial uncertainty inherent with the photon detection and photochemical reactions such as photobleaching, where a fluorophore loses its ability to fluoresce over time [28]. As a result, images acquired from fluorescent microscopy usually appear blurry and noisy. An additional difficulty stems from the complexity of the image data, which typically contain large variations in image contrast and object sizes, with features of interest varying from several tens of microns down to the resolution limit (~ 0.1 micron). Due to all these factors, standard image processing tools frequently perform poorly on fluorescent images and several specialized algorithms have been proposed in the literature, with different approaches often required to deal with different types of data [5, 22, 27].

Wavelets and multiscale methods, in particular, have been very successful in signal and image processing during the last decade, and several wavelet-based methods have been applied to fluorescent images for tasks such as image denoising, deblurring and segmentation (e.g., [3, 5, 38, 39]). Even though wavelets generally outperform Fourier-based and other traditional methods, they have serious limitations in multidimensional applications. This is due to the fact that the analyzing functions associated with the wavelet transform have poor directional sensitivity so that they are rather inefficient to represent edges and other anisotropic features which are frequently the most relevant objects in multidimensional data. To overcome this limitation, a new class of improved multiscale methods has emerged during the last decade, including most notably the shearlets and curvelets [2], which combine the advantages of multiscale analysis with the ability to efficiently encode directional information. Thanks to these properties, shearlets provide optimally sparse representations for 2D and 3D data containing edge discontinuities [2, 10, 21] and have been shown to perform very competitively for image denoising and enhancement, regularization and feature extraction [6, 20, 30, 44]. In particular, shearlet-based algorithms for denoising have been recently applied to fluorescent images in [7, 17], where they were shown to significantly outperform competing wavelet-based methods.

Motivated by the properties of directional multiscale representations, in this paper we present two algorithms which take advantage of the shearlet transform to efficiently and robustly extract the local geometry content from fluorescent imaging data. In particular, we will describe an algorithm for the automated identification of somas in neuronal cultures and an algorithm for the automated quantification of geometrical features of neurites in brain tissue.

The paper is organized as follows. In Section 2, we recall the basic properties of the shearlet transform. We next present the application of our shearlet-based approach to the identification of somas in neuronal cultures (Section 3) and to the quantification of geometrical features of neurites in neuronal networks (Section 4).

2. The Shearlet Transform

Before presenting the shearlet transform, let us briefly recall a few facts from the theory of wavelets.

2.1. The wavelet transform

Wavelets were introduced in the Eighties to provide efficient representations for functions with point discontinuities [25].

The basic idea at the core of this approach is to generate a family of waveforms ranging over several scales and locations through the actions of dilation and translation operators on a fixed 'mother' function.

https://doi.org/10.1051/mmnp/20149512 Published online by Cambridge University Press

Hence, in the two-dimensional case, let G be a subgroup of the group $GL_2(\mathbb{R})$ of 2×2 invertible matrices. The *affine systems* generated by $\psi \in L^2(\mathbb{R}^2)$ are the collections of functions of the form

$$\psi_{M,t}(x) = |\det M|^{-\frac{1}{2}} \psi(M^{-1}(x-t)), \quad t \in \mathbb{R}^2, M \in G. \tag{2.1}$$

If there is an *admissible* function ψ so that any $f \in L^2(\mathbb{R}^2)$ can be recovered via the reproducing formula

$$f = \int_{\mathbb{R}^n} \int_G \langle f, \psi_{M,t} \rangle \, \psi_{M,t} \, d\lambda(M) \, dt,$$

where λ is a measure on G, then ψ is a *continuous wavelet* and the *continuous wavelet transform* is the mapping

$$f \to W_\psi f(M,t) = \langle f, \psi_{M,t} \rangle, \quad (M,t) \in G \times \mathbb{R}^2. \tag{2.2}$$

There are many examples of continuous wavelet transforms [41]. The simplest case is when the matrices M have the form $a\,I$, where $a > 0$ and I is the identity matrix, in which case the admissible ψ need to satisfy the simple condition

$$\int_{\mathbb{R}^2} |\hat{\psi}(a\xi)|^2 \, \frac{d\xi}{a^2} = 1, \quad \text{for a.e. } \xi \in \mathbb{R}^2.$$

In this situation, one obtains the 'traditional' continuous wavelet transform

$$W_\psi f(a,t) = a^{-1} \int_{\mathbb{R}} f(x) \, \overline{\psi(a^{-1}(x-t))} \, dx, \tag{2.3}$$

where the dilation factor is the same for all coordinate directions. Note that the classical discrete wavelet transform (in 2D) is obtained by discretizing (2.3) for $a = 2^{-j}$, $t = 2^{-j}k$, $j \in \mathbb{Z}$, $k \in \mathbb{Z}^2$.

One of the most useful properties of the continuous wavelet transform is its ability to identify the singularities of functions and distributions. Namely, if f is a function or distribution which is smooth apart from a discontinuity at a point x_0, then the continuous wavelet transform $W_\psi f(a,t)$ decays rapidly as $a \to 0$, unless $t = x_0$ [15]. This property is critical to capture the local regularity and explains the ability of discrete wavelets to provide sparse approximations for functions with point singularities, which is one major reason for the success of wavelets in signal processing applications.

2.2. The shearlet transform

Despite its useful properties, the traditional continuous wavelet transform does not provide much information about the *geometry* of the singularities of a function or distribution. In many applications, edge-type discontinuities are frequently the most significant features and it is desirable not only to identify their locations but also to capture their orientations. As shown by one of the authors and their collaborators in [11,19], it is possible to combine multiscale analysis and improved geometric sensitivity by constructing a non-isotropic version of the continuous wavelet transform (2.2), called the *continuous shearlet transform*.

For appropriate admissible functions $\psi^{(h)}, \psi^{(v)} \in L^2(\mathbb{R}^2)$ and matrices

$$M_{as} = \begin{pmatrix} a & -\sqrt{a}s \\ 0 & \sqrt{a} \end{pmatrix}, \quad N_{as} = \begin{pmatrix} \sqrt{a} & 0 \\ -\sqrt{a}s & a \end{pmatrix},$$

we define the *horizontal* and *vertical (continuous) shearlets* by

$$\psi^{(h)}_{a,s,t}(x) = |\det M_{as}|^{-\frac{1}{2}} \psi^{(h)}(M_{as}^{-1}(x-t)), \quad a > 0, s \in \mathbb{R}, t \in \mathbb{R}^2,$$

and

$$\psi^{(v)}_{a,s,t}(x) = |\det N_{as}|^{-\frac{1}{2}} \psi^{(v)}(N_{as}^{-1}(x-t)), \quad a > 0, s \in \mathbb{R}, t \in \mathbb{R}^2,$$

respectively. The reason for choosing two systems of analyzing functions is to ensure a more uniform covering of the range of directions through the shearing variable s. Indeed, rather than using a single shearlet system where s ranges over $\mathbb{R}$, we will use the two systems of shearlets defined above and let s range over a finite interval only.

https://doi.org/10.1051/mmnp/20149512 Published online by Cambridge University Press

To define our admissible functions $\psi^{(h)}, \psi^{(v)}$, for $\xi = (\xi_1, \xi_2) \in \mathbb{R}^2$, let

$$\hat{\psi}^{(h)}(\xi_1, \xi_2) = \hat{\psi}_1(\xi_1)\, \hat{\psi}_2(\tfrac{\xi_2}{\xi_1}), \quad \hat{\psi}^{(v)}(\xi_1, \xi_2) = \hat{\psi}_1(\xi_2)\, \hat{\psi}_2(\tfrac{\xi_1}{\xi_2}), \tag{2.4}$$

where $\hat{\psi}_1$, $\hat{\psi}_2 \in C^\infty(\mathbb{R})$ and satisfy the following conditions:

$$\int_0^\infty |\hat{\psi}_1(a\omega)|^2 \frac{da}{a} = 1, \text{ for a.e. } \omega \in \mathbb{R}, \text{ and } \operatorname{supp} \hat{\psi}_1 \subset [-2, -\tfrac{1}{2}] \cup [\tfrac{1}{2}, 2]; \tag{2.5}$$

$$\|\psi_2\|_2 = 1 \text{ and } \operatorname{supp} \hat{\psi}_2 \subset [-\tfrac{\sqrt{2}}{4}, \tfrac{\sqrt{2}}{4}]. \tag{2.6}$$

Observe that, in the frequency domain, a shearlet $\psi^{(h)}_{a,s,t}$ has the form:

$$\hat{\psi}^{(h)}_{a,s,t}(\xi_1, \xi_2) = a^{\frac{3}{4}} \hat{\psi}_1(a\,\xi_1)\, \hat{\psi}_2(a^{-1/2}(\tfrac{\xi_2}{\xi_1} - s))\, e^{-2\pi i \xi \cdot t}.$$

This shows each function $\hat{\psi}^{(h)}_{a,s,t}$ has the following support:

$$\operatorname{supp} \hat{\psi}^{(h)}_{a,s,t} \subset \{(\xi_1, \xi_2) : \xi_1 \in [-\tfrac{2}{a}, -\tfrac{1}{2a}] \cup [\tfrac{1}{2a}, \tfrac{2}{a}],\ |\tfrac{\xi_2}{\xi_1} - s| \le \sqrt{a}\}.$$

That is, the frequency support of $\hat{\psi}^{(h)}_{a,s,t}$ is a pair of trapezoids, symmetric with respect to the origin, oriented along a line of slope s.

The horizontal and vertical shearlets form a collection of functions ranging not only over various locations and scales, like the elements of a traditional wavelet system, but also over various orientations controlled by the variable s, and with frequency supports becoming highly anisotropic at fine scales ($a \to 0$). In space domain, these functions are not compactly supported but their supports are essentially concentrated on orientable rectangles with side-lengths a, $\sqrt{a}$ and orientation controlled by s. The support becomes increasingly thin as $a \to 0$. The frequency supports of some representative horizontal shearlets are illustrated in Figure 1.

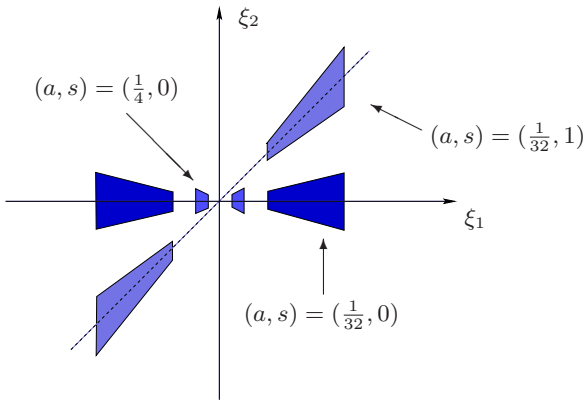


FIGURE 1. *Frequency supports of the horizontal shearlets $\hat{\psi}^{(h)}_{a,s,t}$ for different values of a and s.*

For $0 < a < \frac{1}{4}$ and $|s| \le \frac{3}{2}$, each system of continuous shearlets spans a subspace of $L^2(\mathbb{R}^2)$ consisting of functions having frequency supports in one of the horizontal or vertical cones defined in the frequency domain by

$$\mathcal{P}^{(h)} = \{(\xi_1, \xi_2) \in \mathbb{R}^2 : |\xi_1| \ge 2 \text{ and } |\tfrac{\xi_2}{\xi_1}| \le 1\}$$
$$\mathcal{P}^{(v)} = \{(\xi_1, \xi_2) \in \mathbb{R}^2 : |\xi_1| \ge 2 \text{ and } |\tfrac{\xi_2}{\xi_1}| > 1\}.$$

https://doi.org/10.1051/mmnp/20149512 Published online by Cambridge University Press

More precisely, the following proposition from [19] shows that the horizontal and vertical shearlets, with appropriate ranges for a, s, t, form continuous reproducing systems for the spaces of L^2 functions whose frequency supports are contained in $\mathcal{P}^{(h)}$ and $\mathcal{P}^{(v)}$, respectively.

Proposition 2.1. *Let $\psi^{(h)}$ and $\psi^{(v)}$ be given by (2.4) with $\hat{\psi}_1$ and $\hat{\psi}_2$ satisfying (2.5) and (2.6). Let*

$$L^2(\mathcal{P}^{(h)})^\vee = \{f \in L^2(\mathbb{R}^2) : \operatorname{supp} \hat{f} \subset \mathcal{P}^{(h)}\},$$

with a similar definition for $L^2(\mathcal{P}^{(v)})^\vee$. We have the following reproducing formulas.

(i) For all $f \in L^2(\mathcal{P}^{(h)})^\vee$,

$$f = \int_{\mathbb{R}^2} \int_{-2}^{2} \int_{0}^{\frac{1}{4}} \langle f, \psi^{(h)}_{ast} \rangle \, \psi^{(h)}_{a,s,t} \, \frac{da}{a^3} \, ds \, dt \, ,$$

(ii) For all $f \in L^2(\mathcal{P}^{(v)})^\vee$,

$$f = \int_{\mathbb{R}^2} \int_{-2}^{2} \int_{0}^{\frac{1}{4}} \langle f, \psi^{(v)}_{ast} \rangle \, \psi^{(v)}_{a,s,t} \, \frac{da}{a^3} \, ds \, dt \, ,$$

where both identities hold in the weak topology of $L^2(\mathbb{R}^2)$.

Using the horizontal and vertical shearlets, we define the *(fine-scale) continuous shearlet transform* on $L^2(\mathbb{R}^2)$ as the mapping

$$f \in L^2(\mathbb{R}^2 \setminus [-2,2]^2)^\vee \to \mathcal{SH}_\psi f(a,s,t), \quad a \in (0, \tfrac{1}{4}], s \in [-\infty, \infty], t \in \mathbb{R}^2,$$

given by

$$\mathcal{SH}_\psi f(a,s,t) = \begin{cases} \mathcal{SH}^{(h)}_\psi f(a,s,t) = \langle f, \psi^{(h)}_{a,s,t} \rangle, & \text{if } |s| \le 1 \\ \mathcal{SH}^{(v)}_\psi f(a, \frac{1}{s}, t) = \langle f, \psi^{(v)}_{a,\frac{1}{s},t} \rangle, & \text{if } |s| > 1. \end{cases} \tag{2.7}$$

In the above expression we adopt the convention that the limit value of $\mathcal{SH}_\psi f(a,s,t)$ for $s = \pm\infty$ is set equal to $\mathcal{SH}^{(v)}_\psi f(a,0,t)$.

The term *fine-scale* refers to the fact that the shearlet transform $\mathcal{SH}_\psi f$, given by (2.7), is only defined for the "fine scales" scale variable $a \in (0, 1/4]$. In fact, as it is clear from Proposition 2.1, $\mathcal{SH}_\psi f$ defines an isometry on $L^2(\mathbb{R}^2 \setminus [-2,2]^2)^\vee$, the subspace of $L^2(\mathbb{R}^2)$ of functions with frequency support away from $[-2,2]^2$ but not on $L^2(\mathbb{R}^2)$. This is not a limitation since the geometric characterization of singularities only requires asymptotic estimates as a approaches 0.

Thanks to its directional sensitivity, the continuous shearlet transform is able to capture the geometry of edges precisely, going beyond the limitations of the traditional wavelet approach. If f is a cartoon-like image – a simplified model of natural images consisting of regular regions separated by smooth (or C^m regular, $m \ge 2$) edges – then the shearlet transform of f provides a description of the location and orientations of the edges of f through its asymptotic decay at fine scales. In particular, we have the following theorem from [11].

Theorem 2.2. *Let $f = \chi_S$, where $S \subset \mathbb{R}^2$ is a compact set whose boundary ∂S is a simple smooth curve. Then we have the following.*

(i) If $t \notin \partial S$, then

$$\lim_{a \to 0^+} a^{-N} \, \mathcal{SH}_\psi f(a,s,t) = 0, \quad \text{for all } N > 0.$$

(ii) If $t \in \partial S$ and s does not correspond to the normal direction of ∂S at t, then

$$\lim_{a \to 0^+} a^{-N} \, \mathcal{SH}_\psi f(a,s,t) = 0, \quad \text{for all } N > 0.$$

(iii) If $t \in \partial S$ and s corresponds to the normal direction of ∂S at t, then

$$\lim_{a \to 0^+} a^{-\frac{3}{4}} \, \mathcal{SH}_\psi f(a,s,t) \neq 0.$$

https://doi.org/10.1051/mmnp/20149512 Published online by Cambridge University Press

In other words, the continuous shearlet transform of $f = \chi_S$ decays rapidly, asymptotically for $a \to 0$, unless t is on the boundary of S and s corresponds to the normal direction of ∂S at t, in which case one observes the slow decay

$$\mathcal{SH}_\psi f(a, s, t) \sim O(a^{\frac{3}{4}}), \text{ as } a \to 0.$$

This result can be extended to planar regions S whose boundaries contain corner points and even to more general functions which are not necessarily characteristic functions of sets [11, 14].

The discrete shearlet transform is obtained by sampling (2.7) at $a = 2^{-2j}$, $s = 2^{-j}\ell$, $t = B_{2^{-j}\ell} A_{2^{-2j}} k$, for $j, \ell \in \mathbb{Z}$, $k \in \mathbb{Z}^2$ [13]. In this case, with an abuse of notation, we write the *discrete shearlet transform* as the mapping

$$f \to \mathcal{SH}_\psi f(j, \ell, k) = \langle f, \psi_{j,\ell,k} \rangle, \quad j, \ell \in \mathbb{Z}, k \in \mathbb{Z}^2.$$

This transform inherits some useful analytic properties from its continuous counterpart and can be implemented by fast wavelet-like algorithms. These efficient numerical implementations, combined with the property that discrete shearlet representations provide highly sparse approximations, make the discrete shearlet transform very useful for singularity detection and other image processing applications [20, 44].

3. Soma identification in fluorescent images of neuronal cultures

Recall that a neuron typically possesses two morphological structures: a round cell body called the *soma* and several branched processes called *neurites*, i.e., dendrites and axons. In many biological studies, it is often necessary to identify the individual neurons and their subcellular compartments, i.e., soma and neurites, in a culture or a brain tissue. For example, High Content Analysis applications for screening drugs normally require the extraction and quantification of several morphological features on large datasets for identifying compounds that cause phenotypic changes. Clearly, in these studies, traditional analysis methods such as manual segmentation or semi-automated segmentation with human intervention are highly impractical and automated methods for the quantification of morphological features are virtually necessary.

The automated extraction of the location and morphology of the somas in a neuronal network is essential for such applications and several ideas have been proposed in the literature. Some popular methods use intensity thresholding [37, 40], which is very effective in images acquired from phase-contrast microcopy where the somas often have higher contrast than the neurites. However, this approach is unreliable in fluorescent microscopy due to the inhomogeneity of the fluorescence intensity. Other more sophisticated methods from the literature use Laplacian-of-Gaussian and related filters [1, 32] to detect local maxima, but irregular intensity profiles inside the soma frequently lead to detect more than one soma candidate for each soma. Other more flexible methods include the use of morphological operators [37, 43].

In this paper, we propose a novel approach which takes advantage of the shearlet transform to accurately separate the somas from the neurites by classifying appropriate geometric descriptors extracted from the fluorescent image data. Our method is motivated by the properties of the shearlet transform presented in Sec. 2 in connection with the detection of the location and orientation of edges.

3.1. A shearlet-based Directional Ratio

Let us consider the cartoon-like image of a neuron $f = \chi_S$ in Figure 2, where S is a subset of $\mathbb{R}^2$ consisting of the union of a disk-shaped region and a thin and long solid rectangle. One might expect S to be compact but, to simplify the mathematical analysis, we assume that the thin and long rectangle extends to infinity. Moreover, we consider its centerline to be the y-axis and the disk-shaped area to be centered at the origin of the coordinate system and be attached to it, so that S is connected. The continuous shearlet transform provides a precise characterization of the boundaries ∂S of the region S since $\mathcal{SH}_\psi f(a, s, t)$ has rapid asymptotic decay, as $a \to 0$, for all (s, t), except for t located on ∂S and for s associated with the normal orientation to ∂S at t. Clearly, the discrete version of the shearlet transform $\mathcal{SH}_\psi f(j, \ell, k)$ can only be computed up to a minimal but finite resolution level. In this case, we expect that, at "fine scales", the largest shearlet coefficients in magnitude will be those located near the edges, with the values of the shear parameter corresponding to the normal orientation to the edges. In particular, as shown by one of the authors in [44], we can use this observation to distinguish points near the edges from points away from the edges. At the points near the edges, the local orientation of an

https://doi.org/10.1051/mmnp/20149512 Published online by Cambridge University Press

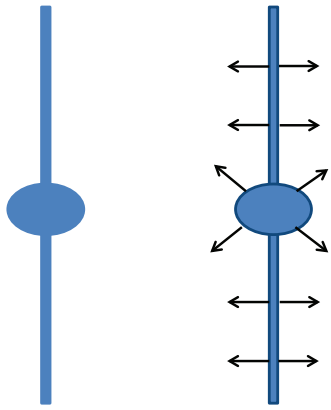


FIGURE 2. *Left: Cartoon-like model of a neurite and a soma. Right: The continuous shearlet transform of the cartoon-like image has rapid asymptotic decay everywhere except at the boundaries, for values of the shear variable corresponding to the normal orientation.*

edge remains constant and the shearlet coefficient $\mathcal{SH}_\psi f(j,\ell,k)$ corresponding to a directional subband parallel to this orientation will have a significant value. This is visible in Panel B in Fig. 3.1, showing the consistent directional sensitivity of the shearlet transform for a single scale a near the singularities. The panel illustrates that the response of the transform at higher scales is more consistent in thinner neurites, while in thicker ones the response of the shearlet transform is less consistent. The same phenomenon is observed in the interior of the neurite, where the orientations of maximum response appear to be distributed almost like white-noise, in stark contrast to the consistent strong responses shown at the boundaries of the neurite. However, at intermediate scales, depending on the relative thickness of a neurite as compared to its length (aspect ratio), the shearlet transform "sees" the entire neurite as a singularity spatially organized along its centerline. Then, the orientations of maximum response in the interior of the neurite and on its boundary become more consistent. This observation leads to the conclusion that tubular and non-tubular structures appear with different degrees of directional coherence at different scales in the shearlet transform domain.

As a measure of directional coherence, we consider a notion of *directional ratio* with respect to the continuous shearlet transform, which we use as a shape descriptor to separate isotropic from anisotropic structures. This concept was originally introduced in [29] using another set of analysis atoms. Similar directional ratios can be defined using other classes of multiscale transforms induced by appropriate 'directional' atoms.

The shearlet-based *Directional Ratio* of f at scale $a > 0$ and location $t \in \mathbb{R}^3$ is given by

$$\mathcal{D}_a f(t) = \frac{\inf_s\{|\mathcal{SH}_\psi f(a,s,t)|\}}{\sup_s\{|\mathcal{SH}_\psi f(a,s,t)|\}} \quad , f \in L^2(\mathbb{R}^2), \tag{3.1}$$

and measures the strength of directional information at scale a and location t. The next theorem predicts consistent responses of directional ratio on tubular structures in general and, thus, it formalizes the preceding discussion.

Theorem 3.1. *Let $f = \chi_S$ be a cartoon-like image of a neuron, where S contains two disjoint subsets: A, a ball centered at the origin with radius $R > 0$ and B, the cylinder $[-r, r] \times [2R, +\infty)$, $r > 0$. Let the cylinder $[-r, r] \times [0, \infty)$ be contained in S. Moreover, assume that*

1. *the Fourier transforms of both ψ_1 and ψ_2 are C^∞ and even;*
2. *the wavelet function ψ_1 has two "plateaus" in the frequency domain: $\hat{\psi_1}(\xi) = 1$ for all $3/4 \leq |\xi| \leq 3/2$;*

https://doi.org/10.1051/mmnp/20149512 Published online by Cambridge University Press

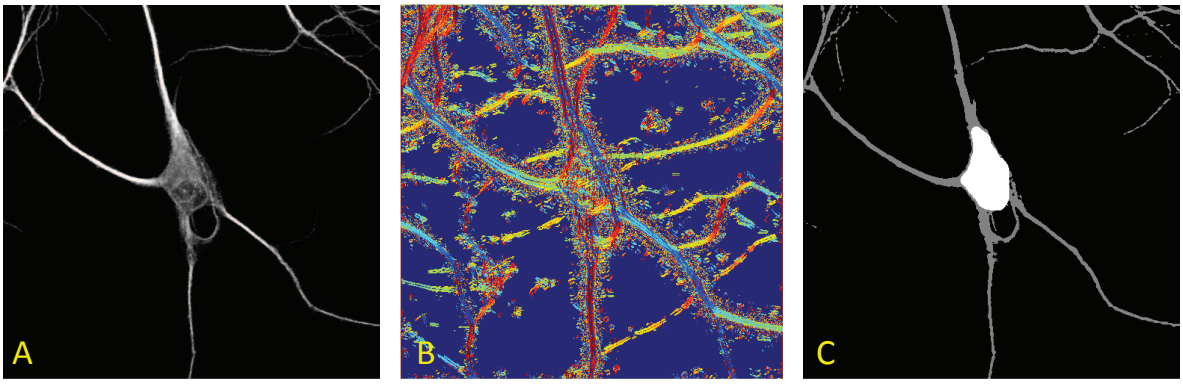


FIGURE 3. *Automatic detection of the soma. Panel A shows the maximum projection of a neuron imaged from fluorescent microscopy. Panel B illustrates the maximum response orientation components of the image (directions are color-coded) computed using the continuous shearlet decomposition at a single scale. Note that similar or slowly-changing colors are located along the neurites, unlike the soma region where there are no dominant orientations, that is, directional components vary randomly creating a white noise pattern. Using the directional ratio, we can accurately segment the soma from the rest of the structure in the image (Panel C).*

3. the "bump" function ψ_2 satisfies $\hat{\psi}_2(\xi) = c > 0$ for all $|\xi| \leq 1/4$ and that $\int_{-R}^{R} \psi_2(x)dx > 1$.

Then, for $4r \leq a \leq 1/4$, there exists a threshold τ such that the shearlet-based Directional Ratio of f, given by (3.1)*, satisfies:*

$$\mathcal{D}_a f(y) \leq \tau \qquad y \in B. \tag{3.2}$$

On the other hand, if $\epsilon > 0$ is an arbitrary positive number, then, for every $y \in \kappa_\epsilon A$, orientation s and sufficiently fine scale a, where $0 < \kappa_\epsilon < 1$, we have that $|\mathcal{SH}_\psi f(a, s, y)| < \epsilon$.

3.2. Proof of Theorem 3.1

Before starting the proof of the theorem we remark that the constant c above must exceed 1 because the support of $\hat{\psi}_2$ is the interval $[-\frac{\sqrt{2}}{4}, \frac{\sqrt{2}}{4}]$ and $||\psi_2||_2 = 1$. Also, with no loss of generality, we can assume that $\hat{\psi}_2(0) = 1$.

Proof. We begin by making clear that the sinc-function is given by

$$\mathrm{sinc}(x) := \frac{\sin(2\pi x)}{\pi x} = \left(\chi_{[-1,1]}\right)^{\vee}(x) \quad \text{with } x \in \mathbb{R}.$$

Let $y = (y_1, y_2)$ be an arbitrary element in B which is contained in the rectangle $C = [-r, r] \times [y_2 - R, y_2 + R] \subset B$. Since the length of the cylinder B is large enough, with no loss of generality we can shift

https://doi.org/10.1051/mmnp/20149512 Published online by Cambridge University Press

the origin in order to have $y_2 = 0$. We have:

$$\begin{aligned}\mathcal{D}_a f(y) &= \frac{\inf_s\{|\mathcal{SH}_\psi f(a,s,y)|\}}{\sup_s\{|\mathcal{SH}_\psi f(a,s,y)|\}} \\ &\leq \frac{|\mathcal{SH}_\psi f(a,\infty,y)|}{|\mathcal{SH}_\psi f(a,0,y)|} \\ &= \frac{|\mathcal{SH}_\psi f^{(v)}(a,0,y)|}{|\mathcal{SH}_\psi f^{(h)}(a,0,y)|} \\ &= \frac{|\langle f, \psi^{(v)}_{a,0,y}\rangle|}{|\langle f, \psi^{(h)}_{a,0,y}\rangle|}. \end{aligned} \tag{3.3}$$

The absolute integrability of the shearlet functions implies that, for any $\epsilon > 0$, there exists a square with side of length $2R_0$, such that

$$\int_{([-R_0,R_0]\times[-R_0,R_0])^c} |\psi^{(v)}_{1,0,0}(x,y)|dxdy < \epsilon.$$

A similar relation holds for $\psi^{(h)}_{1,0,0}$, with the roles of R_0 and r_0 reversed. For technical reasons, we will assume $R_0 > 1/2$. Later in the proof we will determine ϵ. The above inequality implies that, for both the vertical and horizontal shearlets, we have

$$\int_{([-a^{1/2}R_0,a^{1/2}R_0]\times[-aR_0,aR_0])^c} |\psi^{(v)}_{a,0,0}(x,y)|dxdy < a^{3/4}\epsilon, \tag{3.4}$$

and

$$\int_{([-aR_0,aR_0]\times[-a^{1/2}R_0,a^{1/2}R_0])^c} |\psi^{(h)}_{a,0,0}(x,y)|dxdy < a^{3/4}\epsilon.$$

Our strategy is to show that the numerator in the fraction (3.3) is significantly smaller than the denominator. We begin by estimating the latter.

We stress that we have not specified the quantity ϵ yet, thus R_0 still needs to be determined. However, we postpone this task until the end of the analysis of the behavior of directional ratio on the tubular part of the cartoon-like neuron.

On the other hand, $f = \chi_S$ contains the rectangle $[-r, r] \times [-R, R]$ which, in turn, contains y. Once again, with no loss of generality, since S extends as long we wish toward the positive direction of the y-axis, we can assume $R > R_0$. Our assumptions on R_0 and on the scale a imply $aR_0 \geq 2r$. Consequently, both rectangles $-y + [-r, r] \times [-R, R]$ and $[-aR_0, aR_0] \times [-a^{1/2}r_0, a^{1/2}r_0]$ are contained in the rectangle $[-aR_0, aR_0] \times [-R, R]$.

We observe that

$$\langle f, \psi^{(h)}_{a,0,y}\rangle = \langle T_{-y}f, \chi_{[-aR_0,aR_0]\times[-R,R]}\psi^{(h)}_{a,0,0}\rangle + \langle T_{-y}f, \chi_{([-aR_0,aR_0]\times[-R,R])^c}\psi^{(h)}_{a,0,0}\rangle,$$

which, if combined with $aR_0 \geq 2r$, allows us to take $y_1 = 0$ and to conclude that

$$|\langle f, \psi^{(h)}_{a,0,y}\rangle| > |\langle f\chi_{[-aR_0,aR_0]\times[-R,R]}, \psi^{(h)}_{a,0,0}\rangle| - a^{3/4}\epsilon.$$

Thus,

$$|\langle f, \psi^{(h)}_{a,0,y}\rangle| > |\langle \chi_{[-r,r]\times[-R,R]}, \psi^{(h)}_{a,0,0}\rangle| - a^{3/4}\epsilon.$$

We also observe that

$$\begin{aligned}&|\langle \chi_{[-r,r]\times[-R,R]}, \psi^{(h)}_{a,0,0}\rangle| \\ &= \left|\int_{\frac{1}{2a}<|\xi_1|<\frac{2}{a}} \int_{|\frac{\xi_2}{\xi_1}|<\frac{\sqrt{2a}}{4}} a^{3/4} rR\,\mathrm{sinc}(r\xi_1)\mathrm{sinc}(R\xi_2)\hat{\psi}_1(a\xi_1)\hat{\psi}_2\left(\frac{\xi_2}{\sqrt{a}\xi_1}\right) d\xi_2\, d\xi_1\right| \\ &= 2a^{3/4}\left|\int_{\frac{1}{2a}<\xi_1<\frac{2}{a}} r\,\mathrm{sinc}(r\xi_1)\hat{\psi}_1(a\xi_1)\left(\int_{\frac{|\xi_2|}{\xi_1}<\frac{\sqrt{2a}}{4}} R\,\mathrm{sinc}(R\xi_2)\hat{\psi}_2\left(\frac{\xi_2}{\sqrt{a}\xi_1}\right) d\xi_2\right) d\xi_1\right|,\end{aligned}$$

https://doi.org/10.1051/mmnp/20149512 Published online by Cambridge University Press

where we drop the absolute value of ξ_1 due to the symmetry of the integrand with respect to the ξ_2-axis. First, we estimate the inner integral. To do so, we set $\eta = \frac{\xi_2}{\sqrt{a}\xi_1}$. Then we observe:

$$\int_{\frac{|\xi_2|}{\xi_1} < \frac{\sqrt{2a}}{4}} R\,\mathrm{sinc}(R\xi_2)\hat{\psi}_2\left(\frac{\xi_2}{\sqrt{a}\xi_1}\right) d\xi_2 = \int_{|\eta| < \frac{\sqrt{2}}{4}} \sqrt{a}\xi_1 R\,\mathrm{sinc}(R\sqrt{a}\xi_1\eta)\hat{\psi}_2(\eta) d\eta.$$

Note that

$$\chi_{[-R\sqrt{a}\xi_1, R\sqrt{a}\xi_1]} * \psi_2(0) = \int_{-R\sqrt{a}\xi_1}^{R\sqrt{a}\xi_1} \psi_2(x) dx.$$

Since $c = \hat{\psi}_2(0) = \int_{-\infty}^{\infty} \psi_2(x)dx$, the scale a does not exceed $\frac{1}{4}$ and $\frac{1}{2a} < \xi_1$, we conclude that both $\int_{-R\sqrt{a}\xi_1}^{R\sqrt{a}\xi_1} \psi_2(x)dx$ and $\int_{-R}^{R} \psi_2(x)dx$ are greater than 1. In conclusion,

$$\int_{\frac{|\xi_2|}{\xi_1} < \frac{\sqrt{2a}}{4}} R\,\mathrm{sinc}(R\xi_2)\hat{\psi}_2\left(\frac{\xi_2}{\sqrt{a}\xi_1}\right) d\xi_2 > 1. \tag{3.5}$$

Next, we examine the outer integral

$$\int_{\frac{1}{2a}}^{\frac{2}{a}} r\,\mathrm{sinc}(r\xi_1)\hat{\psi}_1(a\xi_1) d\xi_1 = \int_{\frac{1}{2}}^{2} \frac{r}{a}\mathrm{sinc}\left(\frac{r}{a}\xi\right)\hat{\psi}_1(\xi) d\xi\ .$$

The selection of the scale a implies $0 \le \frac{r}{a}\xi_1 \le \frac{1}{2}$ for every value of ξ_1 in the interval of integration. Since the sinc function takes positive values in the open interval $(-\frac{1}{2}, \frac{1}{2})$ and $\hat{\psi}_1(\xi) = 1$ for all $\xi \in [\frac{3}{4}, \frac{3}{2}]$, using (3.5) we finally conclude that

$$|\langle f, \psi^{(h)}_{a,0,y}\rangle| > 2a^{3/4}\frac{3}{4}\frac{\sin(\frac{3r\pi}{a})}{\frac{3r\pi}{2a}} - a^{3/4}\epsilon \ge 2a^{3/4}\frac{3}{4}\frac{\sin(\frac{3\pi}{4})}{\frac{3\pi}{8}} - a^{3/4}\epsilon = a^{3/4}\frac{2\sqrt{2}}{\pi} - a^{3/4}\epsilon. \tag{3.6}$$

Now we turn our attention to the numerator of (3.3); we need to show that it is relatively much smaller than the denominator. We have that

$$\langle f, \psi^{(v)}_{a,0,y}\rangle = \langle T_{-y}(f\chi_{[-r,r]\times[-R,R]}), \psi^{(v)}_{a,0,0}\rangle + \langle T_{-y}(f\chi_{([-r,r]\times[-R,R])^c}), \psi^{(v)}_{a,0,0}\rangle.$$

Regardless of how y_1 varies between $-r$ and r, the rectangle $[-\sqrt{a}R_0, \sqrt{a}R_0] \times [-aR_0, aR_0]$ always stays inside $\mathbb{R} \times [-R, R]$, so $|\langle T_{-y}(f\chi_{([-r,r]\times[-R,R])^c}), \psi^{(v)}_{a,0,0}\rangle| \le a^{3/4}\epsilon$. Therefore, we have that

$$|\langle f, \psi^{(v)}_{a,0,y}\rangle| \le |\langle T_{-y}\chi_{[-r,r]\times[-R,R]}, \psi^{(v)}_{a,0,0}\rangle| + a^{3/4}\epsilon. \tag{3.7}$$

Note that

$$\begin{aligned}&\langle T_{-y}\chi_{[-r,r]\times[-R,R]}, \psi^{(v)}_{a,0,0}\rangle \\ &= a^{3/4}\int_{\frac{1}{2a}<|\xi_2|<\frac{2}{a}} \int_{|\frac{\xi_1}{\xi_2}|<\frac{\sqrt{2a}}{4}} Rr\,\mathrm{sinc}(r\xi_1)\mathrm{sinc}(R\xi_2)e^{2\pi i\xi_1 y_1}\hat{\psi}_1(a\xi_2)\hat{\psi}_2\left(\frac{\xi_1}{\sqrt{a}\xi_2}\right) d\xi_1\, d\xi_2.\end{aligned}$$

Since $\hat{\psi}_2$ is even, we obtain that the latter double integral is equal to

$$a^{3/4}\int_{\frac{1}{2a}<|\xi_2|<\frac{2}{a}} \int_{\frac{-\sqrt{2a}|\xi_2|}{4}}^{\frac{\sqrt{2a}|\xi_2|}{4}} \hat{\psi}_1(a\xi_2)R\,\mathrm{sinc}(R\xi_2)e^{2\pi i\xi_1 y_1} r\,\mathrm{sinc}(r\xi_1)\hat{\psi}_2\left(\frac{\xi_1}{\sqrt{a}|\xi_2|}\right) d\xi_1\, d\xi_2. \tag{3.8}$$

https://doi.org/10.1051/mmnp/20149512 Published online by Cambridge University Press

We now estimate the inner integral. First, using the change the variable $\eta = \frac{\xi_1}{\sqrt{a}|\xi_2|}$, we obtain that the inner integral in (3.8) is equal to

$$\int_{\frac{-\sqrt{2}}{4}}^{\frac{\sqrt{2}}{4}} e^{2\pi i\eta\sqrt{a}|\xi_2|y_1} r\sqrt{a}|\xi_2|\mathrm{sinc}(r\sqrt{a}|\xi_2|\eta)\hat{\psi}_2(\eta)\,d\eta = \chi_{[-r\sqrt{a}|\xi_2|,r\sqrt{a}|\xi_2|]} * \psi_2(\sqrt{a}|\xi_2|y_1)$$
$$= \int_{\sqrt{a}|\xi_2|y_1-r\sqrt{a}|\xi_2|}^{\sqrt{a}|\xi_2|y_1+r\sqrt{a}|\xi_2|} \psi_2(x)dx.$$

By applying the Cauchy-Schwartz inequality to ψ_2, which has norm equal to 1, we deduce $||\psi_2||_\infty \leq ||\hat{\psi}_2||_1 < 1$. Therefore, the inner integral in (3.8) is bounded above $2r\sqrt{a}|\xi_2|$. Thus, the absolute value of the integral in (3.8) is bounded above by

$$a^{3/4}\int_{\frac{1}{2a}<|\xi_2|<\frac{2}{a}} R|\mathrm{sinc}(R\xi_2)\hat{\psi}_1(a\xi_2)|2r\sqrt{a}|\xi_2|d\xi_2 = 2a^{3/4}\int_{\frac{1}{2a}}^{\frac{2}{a}} R\frac{|\sin(2\pi R\xi_2)|}{\pi R\xi_2}2r\sqrt{a}\xi_2 d\xi_2$$
$$\leq \frac{12ra^{3/4}}{2\pi\sqrt{a}} \leq \frac{3\sqrt{a}a^{3/4}}{2\pi},$$

where the last inequality follows from our assumption on the scale a. Using (3.7) and the last inequality, we obtain that

$$|\langle f, \psi^{(v)}_{a,0,y}\rangle| \leq \frac{3\sqrt{a}a^{3/4}}{2\pi} + a^{3/4}\epsilon.$$

Next, using (3.6), we derive the following upper bound for $\mathcal{D}_a f(y)$:

$$\mathcal{D}_a f(y) \leq \frac{\frac{3\sqrt{a}}{2\pi}+\epsilon}{\frac{2\sqrt{2}}{\pi}-\epsilon}.$$

Taking $\epsilon = \frac{1}{10\pi}$, we get that $\mathcal{D}_a f(y) < \tau := \frac{15\sqrt{a}+1}{20\sqrt{2}-1} < 1/3$, for all $y \in B$.

To complete the proof we turn our attention to the isotropic part of the cartoon neuron, the ball A. Since A is centered at the origin and it is a ball, it is enough to prove the statement for horizontal shearlets only. As we did in the first part of the proof, for each arbitrary $\epsilon > 0$, we can find a square with side of length $2R_{0,\epsilon}$ such that

$$\int_{([-R_{0,\epsilon},R_{0,\epsilon}]\times[-R_{0,\epsilon},R_{0,\epsilon}])^c} |\psi^{(h)}_{1,0,0}(x,y)|dxdy < \epsilon.$$

The shearing parameter s varies between -1 and 1 and, as it does so, it "shears" the square $Q := [-R_{0,\epsilon}, R_{0,\epsilon}] \times [-R_{0,\epsilon}, R_{0,\epsilon}]$. Therefore, if the scale is sufficiently fine, it is not hard to see that there exists $0 < \kappa_\epsilon < 1$ such that

$$y + M_{as}(Q) \subset A,$$

for every $y \in \kappa_\epsilon A$. Now, let $-1 \leq s \leq 1$. Then $|\mathcal{SH}_\psi f(a,s,y)| = |\langle f, \psi^{(h)}_{a,s,y}\rangle|$. However, $\int_{\mathbb{R}}\int_{\mathbb{R}} \psi^{(h)}_{a,s,y} = 0$. Thus,

$$\langle f, \psi^{(h)}_{a,s,y}\rangle = \int_{\mathbb{R}^2} (f-1)\overline{\psi^{(h)}_{a,s,y}} = \int_{(y+M_{as}(Q))^c} (\chi_S - 1)\overline{\psi^{(h)}_{a,s,y}},$$

because $y + M_{as}(Q) \subseteq A$. Therefore, we conclude that

$$|\mathcal{SH}_\psi f(a,s,y)| < \int_{(y+M_{as}(C))^c} |f-1||\psi^{(h)}_{a,s,y}| < 2a^{3/4}\epsilon < \epsilon.$$

□

https://doi.org/10.1051/mmnp/20149512 Published online by Cambridge University Press

The last part of the proof of Theorem 3.1 illustrates the behavior of the shearlet transform in isotropic regions. We don't have to take the ϵ very small because, as the scale becomes finer and finer, the quantity $2a^{3/4}\epsilon$ decreases as well. In any case, the main implication of the second part of Theorem 3.1 is that the shearlet transform will attain very small values everywhere in the isotropic part of the cartoon-like neuron, as long as the location variable of the transform is away from the boundaries. As a result, the values of the Directional Ratio $\mathcal{D}_a f$ will vary wildly in isotropic regions (see Fig. 3.2) and, consequently, the Directional Ratio is not useful to detect such regions. As we discuss in the next paragraph, there is a method to detect and segment isotropic regions such as somas.

To identify somas with respect to neurites, we will use Theorem 3.1, which predicts that there is a threshold below which, at appropriate scales, the directional ratio in the neurites does not exceed this threshold. The existence of such a threshold is numerically verifiable as one can see in Fig. 3.2

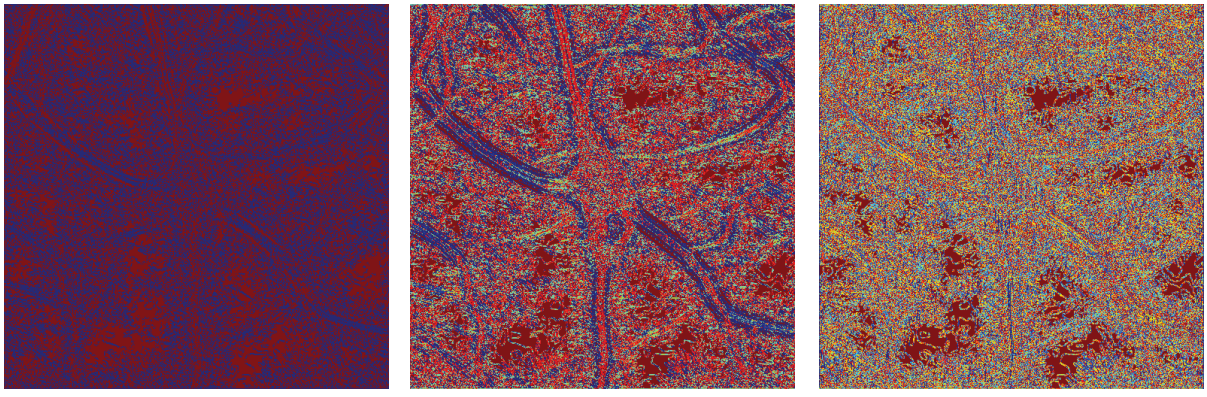


FIGURE 4. *Values of directional ratio (color-coded) computed at various discrete scales; from left to right: $a = 2^{-4}2^{\frac{1}{4}}$, $a = 2^{-3}2^{\frac{1}{4}}$, and $a = 2^{-1}2^{\frac{1}{4}}$ according to the conventions of this paper. Using the convention of the discrete shearlet transform, from left to right we show the directional ratio at two, three and five scales below the resolution level of the original. As the scale becomes coarser, the values of the directional ratio in the interior of neurites become consistently more uniform and do not exceed a certain low threshold, whose existence is predicted by Theorem 3.2. In the interior of the soma and of thicker neurites, the values of directional ratio vary wildly, once again as predicted by Theorem 3.2.*

Thus, given a segmented neuron, we can identify within this subvolume the grid points corresponding to somas. In fact, Proposition 3.2 below indicates that, in the interior of tubular structures modeling the (thin) neurites, the function $\mathcal{D}_a$ is weakly differentiable. This property can be used to circumscribe a boundary between a neurite and a soma via a level-set method. This method can be significantly more accurate than plain thresholding to locate this boundary.

Recall that level set methods propagate a surface or a curve which evolves until it coincides with a desired level set. In our case, this level set would be the boundary between a neurite and a soma. The gradient field required to evolve the interface curve to the desired level set can be defined by taking finite differences of the directional ratio, since this function is weakly differentiable. To initialize the evolution we can use the curve which is the boundary of the set of points inside the soma whose values of $\mathcal{D}_a f$, for a suitable a, are less than a threshold predicted by Theorem 3.2. In [29], we have proposed a similar method to segment somas in fluorescent images of neurons.

3.3. Weak differentiability

Let us now turn our attention to the weak differentiability of the directional ratio of f, where f is a cartoon-like neuron. To establish that $\mathcal{D}_a f$, given by the ratio (3.1), satisfies a Lipschitz condition, it

https://doi.org/10.1051/mmnp/20149512 Published online by Cambridge University Press

is enough to verify this property for the numerator and the denominator of that ratio. To control how small the denominator can become, we assume that in an open set Ω there exists $\alpha > 0$ such that $|\mathcal{SH}_\psi f(a,s,y)| > \alpha$ for all $y \in \Omega$. Select $y \in \Omega$ and a vector h such that $y + h \in \Omega$. Using the triangle inequality we have

$$|\langle f, T_{y+h}\psi^{(v)}_{a,s,0}\rangle| \leq |\langle f, T_y\psi^{(v)}_{a,s,0}\rangle| + |\langle f, (T_h - I)\psi^{(v)}_{a,s,0}\rangle|,$$

for every scale a and shear variable s. A similar inequality holds for horizontal shearlets. We will assume Ω to be convex. Since the Fourier transforms of $\psi^{(v)}$ and $\psi^{(h)}$ are C^∞ and compactly supported, the $\psi^{(v)}$ and $\psi^{(h)}$ are in the Schwartz space. Hence, all of their partial derivatives are uniformly bounded.

By applying the Mean Value Theorem of differential calculus we obtain that

$$\begin{aligned}|\langle f, (T_h - I)\psi^{(v)}_{a,s,0}\rangle| &= \left|\int_{\mathbb{R}^2} f(x)(\psi^{(v)}_{a,s,0}(x-h) - \psi^{(v)}_{a,s,0}(x))dx\right| \\ &\leq \int_{\mathbb{R}^2} |f(x)||\psi^{(v)}_{a,s,0}(x-h) - \psi^{(v)}_{a,s,0}(x)|dx \\ &\leq C||f||_1\|h\|,\end{aligned}$$

where C is a constant derived by the L^∞-bounds of the partial derivatives of the horizontal and vertical shearlets. Since the shear variable s is taken in $[-1,1]$, the constant C holds uniformly in s. Therefore, for every $s \in [-1,1]$ we have that

$$|\langle f, T_{y+h}\psi^{(v)}_{a,s,0}\rangle| \leq \sup_s\{|\langle f, T_y\psi^{(v)}_{a,s,0}\rangle|\} + C||h|| \leq \sup_{s\in\mathbb{R}}\{|\mathcal{SH}_\psi f(a,s,y)|\} + C\|h\|.$$

By combining horizontal and vertical shearlets, we deduce that

$$\sup_{s\in\mathbb{R}}\{|\mathcal{SH}_\psi f(a,s,y+h)|\} \leq \sup_{s\in\mathbb{R}}\{|\mathcal{SH}_\psi f(a,s,y)|\} + C\|h\|.$$

By swapping the points y and $y+h$ in the previous inequality, we obtain that

$$|\sup_s\{|\mathcal{SH}_\psi f(a,s,y)|\} - \sup_s\{|\mathcal{SH}_\psi f(a,s,y+h)|\}| \leq C\|h\|$$

On the other hand, a similar argument establishes that

$$|\inf_s\{|\mathcal{SH}_\psi f(a,s,y)|\} - \inf_s\{|\mathcal{SH}_\psi f(a,s,y+h)|\}| \leq C\|h\|$$

Combining these two observations we derive the following result.

$$\begin{aligned}&|\mathcal{D}_a f(y+h) - \mathcal{D}_a f(y)| \\ &= \left|\frac{\inf_s\{|\mathcal{SH}_\psi f(a,s,y+h)|\}}{\sup_s\{|\mathcal{SH}_\psi f(a,s,y+h)|\}} - \frac{\inf_s\{|\mathcal{SH}_\psi f(a,s,y)|\}}{\sup_s\{|\mathcal{SH}_\psi f(a,s,y)|\}}\right| \\ &\leq \frac{|\inf_s\{|\mathcal{SH}_\psi f(a,s,y+h)|\} - \inf_s\{|\mathcal{SH}_\psi f(a,s,y)|\}|}{\sup_s\{|\mathcal{SH}_\psi f(a,s,y+h)|\}} \\ &+ \frac{|\sup_s\{|\mathcal{SH}_\psi f(a,s,y)|\} - \sup_s\{|\mathcal{SH}_\psi f(a,s,y+h)|\}|}{\sup_s\{|\mathcal{SH}_\psi f(a,s,y+h)|\}} \frac{\inf_s\{|\mathcal{SH}_\psi f(a,s,y)|\}}{\sup_s\{|\mathcal{SH}_\psi f(a,s,y)|\}} \\ &\leq \left(\frac{C}{\alpha} + \frac{C}{\alpha}\right)\|h\|.\end{aligned}$$

The above discussion yields the next statement:

Proposition 3.2. *Assume that, for every point y in the open and convex set $\Omega \subset \mathbb{R}^2$, there exists $\alpha > 0$ such that $|\mathcal{SH}_\psi f(a,s,y)| > \alpha$. Then the shearlet-based Directional Ratio $\mathcal{D}_a f$, at scale a, restricted on Ω, satisfies a Lipschitz condition and, thus, it is weakly differentiable on Ω.*

The weak differentiability of $\mathcal{D}_a f$ follows from the fact that $\mathcal{D}_a f$ is absolutely continuous since it satisfies a Lipschitz condition.

https://doi.org/10.1051/mmnp/20149512 Published online by Cambridge University Press

4. Quantitative measure of alteration in biological neural networks

Structural and morphological changes of neurites often precede functional deficits or neuronal death associated with many brain disorders and several methods have been proposed to quantify these phenotypes. Popular methods include the Sholl analysis [23,36], the Schoenen Ramification Index [35] and the fractal analysis [26], which are used to measure the branching characteristics of neurons, and the *tortuosity* index, defined as the ratio of the arc length over the end-point distance of a curve, which is used to describe the degree of 'straightness' of a given neurite [31,42]. While these methods are very effective at single neuron level and have been often used to classify neuron populations, they are unable to provide a quantitative measure of the structural properties of a biological neural network. Emerging evidence indicates that many brain disorders are associated with the disruption of the network of neurites in defined brain areas [4,16]. Thus, the ability to quantify the loss of regularity and structural organization in these networks is potentially very helpful to unravel fundamental structure-function properties involved in the insurgence of brain disorders.

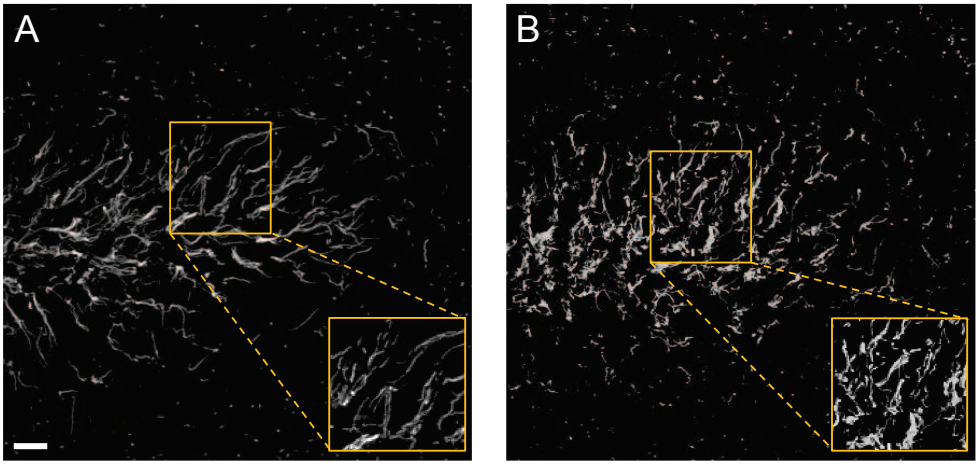


FIGURE 5. *Representative confocal images of axonal processes in the dentate gyrus, a brain area important for cognitive function, in a control mouse (A) and in a transgenic mouse model of neurodegeneration (B). Axonal processes are labeled with an anti-beta-IV-spectrin antibody and visualized with an Alexa 488-conjugated secondary antibody. Note that the pattern distribution of beta-IV-spectrin is disrupted in the dentate gyrus of the transgenic animal. The boxed region is enlarged on the bottom right of each panel to highlight the orientation pattern of the axonal processes at a higher resolution. Scale bar = 10 μm.*

An example of the type of data for which we wish to provide a quantitative analysis is illustrated in Fig. 5, where axonal processes in a brain tissue derived from a control mouse (Fig. 5(a)) and from a transgenic mouse which models neurodegeneration (Fig. 5(b)) are visualized with confocal fluorescent microscopy. The figure shows that, while in the control specimen the axonal processes are rather straight and tend to run parallel to one another, in the transgenic animal neighbouring processes are not very aligned and tortuous, leading to an overall disrupted and chaotic pattern. To deal with these types of data, we introduce a new approach which models a network of neurites as a system of tubular structures. Our observation is that one can measure the overall structural organization and regularity of this system by estimating the statistical spread of orientations of the tubular structures sampled over the region they occupy.

A direct approach to measure the spread of orientations would require to extract the local directional information of the tubular structures and use the notion of circular variance from directional statistics

https://doi.org/10.1051/mmnp/20149512 Published online by Cambridge University Press

[8]. In an ideal tubular structure, this can be obtained by detecting the edges and then estimating their orientations. However, in the case of fluorescent images, it is difficult to extract the edges reliably due to the presence of noise. Therefore, we introduce an alternative procedure which uses the shearlet transform to estimate the local directional content of an image. As illustrated in Section 2, the continuous shearlet transform maps a function f into the elements $\mathcal{SH}_\psi f(a,s,t)$ measuring the energy content of f near t at scale and orientations associated with a and s, respectively. Hence, for $a_j = 2^{-2j}$ and $t_k = B_s A_{2^{-2j}} k$ fixed, the functions $h_{j,k}(s) = |\mathcal{SH}_\psi f(a_j, s, t_k)|^2$ measure the distribution of the directional content of f at scale 2^{-2j} and location t_k. We will use these functions to build an histogram of the directional content of an image.

4.1. Shearlet-based histogram of directional content

To define an histogram of the directional content of an image I using the shearlet transform, we consider the functions $h_{j,k}(s) = |\mathcal{SH}_\psi I(a_j, s, t_k)|^2$ evaluated at the discrete directional values $s(\ell) = 2^{-j}\ell$, $\ell \in N_j$, where N_j is a finite set depending on j. We assume that the image is normalized so that $\|I\|_2 = 1$. For each fixed scale j, we compute the quantities

$$H_j(\ell) = \sum_k h_{j,k}(s(\ell)) = \sum_k |\mathcal{SH}_\psi I(j,\ell,k)|^2,$$

where $\mathcal{SH}_\psi I(j,\ell,k)$ is the discrete shearlet transform introduced in Sec.2. Hence, H_j describes the energy amount of I in each directional band indexed by ℓ.

Using the vectors H_j, we define a measure of the directional spread by computing a distance between the image I and an appropriate reference image I_0. That is, the image I_0 models the situation where all neurites are perfectly straight and parallel to one another, so that the corresponding feature vectors $H_j^{(0)}$ have all components equal to 0, except for one component (with $\|I_0\|_2 = 1$). We define the *directionality spread index* as the quantity

$$\nu_2(I) = \sum_j \min_{0\le d\le n_j} \sum_\ell |H_j([\ell+d]_{n_j}) - H_j^{(0)}(\ell)|^2, \tag{4.1}$$

where $[\ell+d]_{n(j)}$ is the sum $\ell+d$ modulo n_j and is used to ensure that I_0 is compared to all possible rotated versions of I. That is, $\nu_2(I)$ is the rotation-invariant ℓ^2 distance between the feature vectors of I and I_0. Note that ν_2 ranges between $\nu_2(I) = 0$, corresponding to the case where all neurites are perfectly straight and parallel to one another in the image I, and $\nu_2(I) \approx 2$, corresponding to the case of maximum misalignment, where the directional content is uniformly distributed among all directional bands.

However, while the use of the ℓ^2-norm in (4.1) is a natural mathematical choice, it is not very 'physical' (the inadequacy of ℓ^2-norm to measure distances between features in well-known in computer vision, e.g., [9,34]). As a more effective alternative, we can consider the Earth Mover's Distance (EMD), known in mathematics as the Wasserstein metric, which is based on a solution of the transportation problem and was originally proposed to provide a consistent measure of distance or dissimilarity between two distributions of mass [33,34]. This metric is particularly appealing in our context. In fact, the histogram associated with the vector H_j can be viewed as the amount of directional content piled over the various directionals bins. The directionality spread index associated with the EMD gives a measure of the minimum cost of turning these piles into the single pile corresponding to $H_j^{(0)}$. In other words, it is measuring the cost of turning a network of possibly misaligned neurites into an ideal configuration where all neurites run parallel to one another. With respect to other notions of distances between histograms such as the information theoretic Kullback-Leibler divergence [18] and the entropy, which account only for the correspondence between bins with the same index and do not use information across bins, EMD is a cross-bin distance measure, is not affected by binning differences and meaningfully matches the physical notion of closeness.

To illustrate the use of our new definition, we have computed the directionality spread index ν_2 and its variants associated with the EMD and other types of distances on a set of confocal fluorescent images similar to those in Fig. 5 in order to compare the axonal processes in brain tissues from control mice versus transgenic mice. The values of the directionality spread index associated with the ℓ^2-norm, EMD,

https://doi.org/10.1051/mmnp/20149512 Published online by Cambridge University Press

Sample tissue	EMD	KLD	ℓ^2	ENTROPY
Wild type	28.1 ± 3.4	32.8 ± 3.4	1.9 ± 0.07	4.7 ± 0.11
Transgenic	37.3 ± 1.5	45.8 ± 1.9	1.5 ± 0.02	4.9 ± 0.01

TABLE 1. Mean tortuosity scores using different metrics. Data are in the form: mean ± SD; $p < 0.0001$. The values are obtained from 8 sample images for each population.

Kullback-Leibler distance (KLD) and the entropy distance are reported in Table 1. For each population (control animal, i.e., wild type, vs. transgenic animal), the table shows the mean directionality spread index computed over 8 samples together with the standard deviation. Note that only the values of the directionality spread index associated with the EMD have passed the normality test. The numbers that we have obtained show that the difference between the directionality spread index of the two populations is statistically significant.

5. Conclusion

In this paper, we have shown that directional multiscale representations and, in particular, the shearlet representation offer a powerful and flexible framework for the analysis and processing of fluorescent images of complex biomedical data. The power of the shearlet representation comes for its ability to combine multiscale data analysis and high directional sensitivity, making this approach particularly effective for the representation of objects defined over a wide range of scales and with highly anisotropic features, such as the somas and neurites which are the main morphological constituents of imaged neurons. We have applied the shearlet representation to detect somas in fluorescent images of neuronal cultures and extract geometrical features from complex networks of neuronal processes in brain tissue. Our theoretical observations and numerical results show that the shearlet approach offers a very competitive analytical and computational framework for the geometric quantification of morphological features from large-scale fluorescent images of neurons.

Acknowledgements. This work was partially supported by NSF-DMS 1320910 (MP), 1008900 (DL), 1005799 (DL), by NHARP-003652-0136-2009 (MP,DL) and by NIH R01 MH095995 (FL).

References

[1] Y. Al-Kofahi, W. Lassoued, W. Lee, B. Roysam. *Improved automatic detection and segmentation of cell nuclei in histopathology images.* IEEE Trans. Biomed. Eng., 57 (2010), no. 4, 841-852.

[2] E. J. Candès, D. L. Donoho. *New tight frames of curvelets and optimal representations of objects with piecewise C^2 singularities.* Comm. Pure and Appl. Math., 56 (2004), 216-266.

[3] C. W. Chang, M. A. Mycek. *Total variation versus wavelet-based methods for image denoising in fluorescence lifetime imaging microscopy.* J Biophotonics, 5 (2012), 449-457.

[4] B. Delatour, V. Blanchard, L. Pradier, C. Duyckaerts. *Alzheimer pathology disorganizes cortico-cortical circuitry: direct evidence from a transgenic animal model.* Neurobiol Dis., 16 (2004), no. 1, 41-47.

[5] A. Dima, M. Scholz, K. Obermayer. *Automatic segmentation and skeletonization of neurons from confocal microscopy images based on the 3-D wavelet transform.* IEEE Trans. Image Process., 11 (2002), no. 7, 790-801.

[6] G. Easley, D. Labate, W. Lim. *Sparse directional image representations using the discrete shearlet transform.* Appl. Comput. Harmon. Anal., 25 (2008), 25-46.

[7] G. Easley, D. Labate, P. S. Negi. *3D data denoising using combined sparse dictionaries.* Math. Model. Nat. Phen., 8 (2013), no. 1, 60-74.

[8] N.I. Fisher. Statistical analysis of circular data. Cambridge University Press, 1993.

[9] C. Grigorescu, N. Petkov. *Distance sets for shape filters and shape recognition.* IEEE Trans. on Image Processing, 12 (2003), no. 10, 1274-1286.

[10] K. Guo, D. Labate. *Optimally sparse multidimensional representation using shearlets.* SIAM J. Math. Anal., 39 (2007), 298-318.

[11] K. Guo, D. Labate. *Characterization and analysis of edges using the continuous shearlet transform.* SIAM Journal on Imaging Sciences, 2 (2009), 959-986.

[12] K. Guo, D. Labate. *Analysis and detection of surface discontinuities using the 3D continuous shearlet transform.* Appl. Comput. Harmon. Anal., 30 (2010), 231-242.

https://doi.org/10.1051/mmnp/20149512 Published online by Cambridge University Press

[13] K. Guo, D. Labate. *The construction of smooth Parseval frames of shearlets.* Math. Model. Nat. Phen., 8 (2013), no. 1, 82-105.

[14] K. Guo, D. Labate, W. Lim. *Edge Analysis and identification using the continuous shearlet transform.* Appl. Comput. Harmon. Anal., 27 (2009), 24-46.

[15] M. Holschneider. Wavelets. Analysis tool. Oxford University Press, Oxford, 1995.

[16] B. Jacobs, H. Praag, F. Gage. *Adult brain neurogenesis and psychiatry: a novel theory of depression.* Mol. Psychiatry, 5 (2000), no. 3, 262-269.

[17] T. F. James, J. Luisi, M. N. Nenov, N. Panova-Electronova, D. Labate, F. Laezza. *The Nav1.2 channel is regulated by glycogen synthase kinase 3 (GSK3).* To appear in Neuropharmacology (2014).

[18] S. Kullback. Information theory and statistics. John Wiley and Sons, NY, 1959.

[19] G. Kutyniok, D. Labate. *Resolution of the wavefront set using continuous shearlets.* Trans. Amer. Math. Soc., 361 (2009), 2719-2754.

[20] G. Kutyniok, D. Labate. Shearlets: multiscale analysis for multivariate data. Birkhäuser, Boston (2012).

[21] D. Labate, W. Lim, G. Kutyniok, G. Weiss. *Sparse multidimensional representation using shearlets.* Wavelets XI (San Diego, CA, 2005), 254-262, SPIE Proc. 5914, SPIE, Bellingham, WA, (2005).

[22] M.R. Lamprecht, D.M. Sabatini, A.E.Carpenter. *CellProfiler: free, versatile software for automated biological image analysis.* Biotechniques, 42 (2007), no. 1, 71-75.

[23] C. G. Langhammer, P. M. Previtera, E. S. Sweet, S. S. Sran, M. Chen, B. L. Firestein. *Automated Sholl analysis of digitized neuronal morphology at multiple scales: Whole cell Sholl analysis versus Sholl analysis of arbor subregions.* Cytometry A, 77 (2010), no. 12, 1160-1168.

[24] F. Li, Z. Yin, G. Jin, H. Zhao, S.T. Wong. *Bioimage informatics for systems pharmacology.* PLoS Comput Biol. 9 (2013), no. 4, Chapter 17.

[25] S. Mallat. A wavelet tour of signal processing. Academic Press, San Diego, CA, 1998.

[26] N.T. Milosevic, D. Ristanovic, J.B. Stankovic. *Fractal analysis of the laminar organization of spinal cord neurons.* Journal of Neuroscience Methods, 146 (2005), no. 2, 198-204.

[27] R. F. Murphy, E. Meijering, G. Danuser *Special issue on molecular and cellular bioimaging.* IEEE Trans. Image Process., 14 (2005), no. 9, 1233-1236.

[28] V. Ntziachristos. *Fluorescence molecular imaging.* Annu. Rev. Biomed. Eng., 8 (2006), 1-33.

[29] B. Ozcan, D. Labate, D. Jimenez, M. Papadakis. *Directional and non-directional representations for the characterization of neuronal morphology.* Wavelets XV (San Diego, CA, 2013), SPIE Proc. 8858 (2013).

[30] V. M. Patel, G. R. Easley, D. M. Healy. *Shearlet-based deconvolution.* IEEE Trans. Image Proc., 18 (2009), no. 12, 2673-2685.

[31] C. Portera-Cailliau, R.M. Weimer, V. De Paola, P. Caroni, K. Svoboda. *Diverse modes of axon elaboration in the developing neocortex.* PLoS Biol., 3 (2005), no. 8, 1473-1487.

[32] X. Qi, F. Xing, D. Foran, L. Yang. *Robust segmentation of overlapping cells in histopathology specimens using parallel seed detection and repulsive level set.* IEEE Trans Biomed Eng, 59 (2012), no. 3, 754-765.

[33] Y. Rubner, C. Tomasi, L. J. Guibas. *A metric for distributions with applications to image databases.* Proceedings ICCV, (1998), 59-66.

[34] Y. Rubner, C. Tomasi, L. J. Guibas. *The earth mover's distance as a metric for image retrieval.* International Journal of Computer Vision, 40 (2000), no. 2, 99-121.

[35] J. Schoenen. *The dendritic organization of the human spinal cord: the dorsal horn.* Neuroscience, 7 (1982), 2057-2087.

[36] D. A. Sholl. *Dendritic organization in the neurons of the visual and motor cortices of the cat.* J. Anat., 87 (1953), 387-406.

[37] P. Vallotton, R. Lagerstrom, C. Sun, M. Buckley, D. Wang. *Automated analysis of neurite branching in cultured cortical neurons using HCA-vision*, Cytom. Part A, 71 (2007), no. 10, 889-895.

[38] C. Vonesch, M. Unser. *A fast thresholded Landweber algorithm for wavelet-regularized multidimensional deconvolution.* IEEE Trans. Image Proc., 17 (2008), no. 4, 539-549.

[39] C. Vonesch, M. Unser. *A fast multilevel algorithm for wavelet-regularized image restoration.* IEEE Trans. Image Proc., 18 (2009), no. 3, 509-523.

[40] C. Wählby, I. M. Sintorn, F. Erlandsson, G. Borgefors, E. Bengtsson. *Combining intensity, edge and shape information for 2D and 3D segmentation of cell nuclei in tissue sections.* J. Microsc., 215 (2004), 67-76.

[41] G. Weiss, E. Wilson. *The mathematical theory of wavelets.* Proceeding of the NATO–ASI Meeting. Harmonic Analysis 2000. A Celebration. Kluwer Publisher, (2001).

[42] Q. Wen, A. Stepanyants, G.N. Elston, A. Y. Grosberg, D. B. Chklovskiia. *Maximization of the connectivity repertoire as a statistical principle governing the shapes of dendritic arbors.* PNAS, 106 (2009), no. 30, 12536-12541.

[43] C. Yan, A. Li, B. Zhang, W. Ding, Q. Luo, H. Gong. *Automated and accurate detection of soma location and surface morphology in large-scale 3D neuron images.* PLoS One, 8 (2013), no. 4.

[44] S. Yi, D. Labate, G. R. Easley, H. Krim. *A shearlet approach to edge analysis and detection.* IEEE Trans. Image Process., 18 (2009), no. 5, 929-941.

https://doi.org/10.1051/mmnp/20149512 Published online by Cambridge University Press