QUANTIFYING IMAGE STRUCTURES IN HIGH-THROUGHPUT MICROSCOPY WITH TOTAL VARIATION FLOW

Shekoufeh Gorgi Zadeh[⋆] *Max Hermann*[⋆] *Elisa Merklinger*[†] *Jan-Gero Schloetel*[†] *Thomas Schultz*[⋆]

[⋆]Institute of Computer Science II, University of Bonn
[†] Life and Medical Sciences Institute, University of Bonn

ABSTRACT

A recurrent problem in the analysis of microscopy images is to quantify the number and size of spots on a homogeneous background. Unfortunately, segmenting the individual spots becomes unreliable when they are close together, or when the image contains noise and artifacts. On the other hand, manual counting and line-scan measurements are prone to bias and too time-consuming to be used in high-throughput microscopy. In this work, we derive novel per-pixel measures of spot scale and density from Total Variation Flow, a partial differential equation that changes the intensities of image regions at a rate inverse to their scale. On simulated, phantom, and real-world data from Stimulated Emission Depletion (STED) microscopy, we demonstrate the robustness of our novel method relative to a standard segmentation-based approach.

***Index Terms*—** Spot size and number, scale, segmentation, partial differential equations, microscopy images

1. INTRODUCTION

Many biological microscopy images, such as those of membrane protein clusters [1] in Fig. 7, show spots on a homogeneous background. Analysis of such images requires counting and measuring the sizes of those spots. Even though methods based on wavelet multiscale product (WMP) operator [2], ellipse fitting [3], modified circular hough transform (CHT) [4], or spectral graph partitioning [5] have been introduced to detect and segment individual spots, practitioners still frequently resort to manual analysis because automated segmentation is not reliable enough in many real-world scenarios. In this work, we introduce a novel approach to this problem that derives *per-pixel* estimates of spot size and density that do not require reliable delineation of individual spots.

2. A NOVEL SCALE MEASURE FROM TOTAL VARIATION

Since our method is based on total variation (TV) regularization and flow, we will briefly introduce them in this section.

2.1. TV Regularization and TV Flow

In total variation (TV) based image regularization, the original grayscale image is modeled as a function $f : D \to \mathbb{R}$ with $D \subset \mathbb{R}^2$, and a regularized version u is obtained as the differentiable function $u : D \to \mathbb{R}$ that minimizes the energy functional

$$E(u; \alpha, f) := \iint_D \left({}^1\!/_2 \, (u - f)^2 + \alpha \|\nabla u\| \right) dx \, dy, \quad (1)$$

where α controls the degree of smoothness of u.

In Total Variation flow (TV flow) [6], the original image f is embedded into a family of increasingly smoothed images $u(\mathbf{x}, t)$, where $\mathbf{x} \in D$ and $t \in [0, \infty)$ is an artificial "time" parameter that specifies the degree of smoothness. TV flow assumes the initial value $u(\mathbf{x}, 0) = f$ and finds $u(\mathbf{x}, t)$ for $t > 0$ by solving the partial differential equation ${}^{\partial u}\!/_{\partial t} = \operatorname{div}\left({}^{\nabla_{\mathbf{x}} u}\!/_{\|\nabla_{\mathbf{x}} u\|} \right)$. We use the additive operator splitting scheme, which is a modification of the semi-implicit scheme [7] to discretize and numerically solve the TV flow equation [8]. In general, the result of TV flow $u(\mathbf{x}, t)$ at time $t = \alpha$ approximates TV regularization with parameter α [9].

2.2. Total Variation and Scale

In piecewise constant images, the absolute difference δ between the image intensity of a pixel in the original image, compared to the TV regularized image with parameter α, can be written as

$$\delta = {}^{\alpha}\!/_{s} \quad \text{with} \quad s = {}^{|\Omega|}\!/_{|\partial\Omega|}, \quad (2)$$

where Strong [10] defines scale s of a pixel as the ratio of the area $|\Omega|$ over the boundary length $|\partial\Omega|$ of the image patch the pixel belongs to.

Similarly, Brox and Weickert [9] previously derived a scale measure for texture segmentation from TV flow. They compute an average scale $\bar{m}$ over a time window $t \in [0, T]$:

$$\bar{m} = 4 \frac{T - \int_0^T 1_{\partial_t u = 0} \, dt}{\int_0^T |\partial_t u| \, dt} \quad (3)$$

where $1_{\partial_t u = 0}$ equals to 1 when $\partial_t u = 0$ and it equals to zero otherwise. Compared to Strong's measure s, $\bar{m}$ replaces δ,

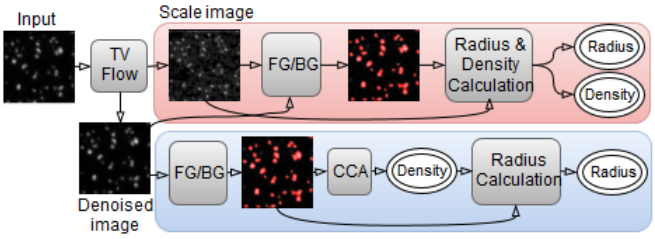

Fig. 1. The top row shows how our method finds the size of structures in images and estimates their density (red background). The row below shows how standard connected component analysis (CCA) could be used to extract scale and density information from the input image (blue background).

the overall intensity change, by $\int_0^T |\partial_t u|\, dt$, the integrated amount of absolute change during $t \in [0, T]$. Given the approximate equivalence of TV flow after time T and TV regularization with parameter $\alpha = T$, $\int_0^T |\partial_t u|\, dt \approx \delta$ if the sign of $\partial_t u$ is the same throughout the interval $[0, T]$. Second, $\bar{m}$ replaces the regularization parameter α with $T - \int_0^T 1_{\partial_t u = 0}\, dt$, i.e., it subtracts periods during which pixel intensities did not change. Finally, $\bar{m}$ contains an additional factor of 4.

2.3. Proposed Scale Measure

In this work, we propose the following new scale measure:

$$\sigma = \gamma \frac{\int_{T_{\text{start}}}^{T_{\text{stop}}} 1_{|\partial_t u| > \Theta}\, dt}{\int_{T_{\text{start}}}^{T_{\text{stop}}} |\partial_t u| 1_{|\partial_t u| > \Theta}\, dt} \tag{4}$$

Our σ is inspired by $\bar{m}$, but differs in three ways: First, it allows for a "burn-in" interval $[0, T_{\text{start}}]$ before we start the measurement. Originally, each individual pixel acts as a small region [9]. $T_{\text{start}} > 0$ allows for formation of more meaningful regions, and elimination of noise. Second, we introduce a threshold $|\partial_t u| > \Theta$ to reduce the impact of higher-order structures, which result when spots in the original image merge. Θ can be set based on Eq. (2): Since $\partial_t u \approx s^{-1}$, an upper bound r_{max} of relevant spot sizes leads to $\Theta = 2l/r_{max}$. Note that we cannot simply pick a shorter T_{stop} to reduce the impact of higher-order structures, since low-contrast spots vanish more quickly than high-contrast ones.

For selecting T_{start} and T_{stop}, we provide a graphical interface in which the user can alter the value of t using a slider and is immediately shown the resulting $u(x, t)$. Based on such visual inspection, we found it easy to select points where the noise is gone, but the interesting spots are still visible (T_{start}) and where the spots almost vanished (T_{stop}). We found that, given the use of Θ, the exact choice of T_{stop} is not very critical. Setting it to a larger value increases the computational effort, but does not reduce accuracy.

Finally, we replace the fixed factor 4 by a variable shape factor γ, which can be used, depending on the physical size of a pixel and the shape of the expected spots, to calibrate σ to the physical units of length, specifying the radius of the spot. According to Eq. (2), for circular spots with constant intensity and pixel edge length l, $\gamma = 2l$.

2.4. Proposed Density Measure

A second parameter of interest is the number of spots per area. The area covered by a circular spot of radius σ is $A = \pi\sigma^2$. Thus, a pixel of area l^2 and scale σ accounts for the fractional part $\nu = l^2/\pi\sigma^2$ of a spot. The sum $\Sigma\nu$ of fractional spot counts of all foreground pixels in an image region provides an estimate of the contained number of circular spots. Normalizing by the overall region size A_R provides a spot density measure $\rho = \Sigma\nu/A_R$.

2.5. Foreground and Background Segmentation

Our method estimates scale and density per pixel, without having to segment individual spots. We still need to segment foreground vs. background (FG/BG) so that we can aggregate our measurements only over the FG. However, in contrast to Connected Component Analysis (CCA) or other segmentation-based approaches, small changes in the segmentation will change our results only slightly and gradually, making it much more robust to the merging and splitting of clusters that can occur due to segmentation errors. Fig. 1 compares the overall strategy of quantifying image structures using our method and CCA.

In our method, the denoised image and the TV scale image are both used for FG/BG segmentation. First, we compute an adaptive Otsu threshold from the denoised image, which accounts for the fact that intensity contrast may not be uniform all over the image. We slide a window over the image and we compute the Otsu threshold in each window separately. In order to tell if a pixel is part of the FG or not, the average threshold of the overlapping windows is then used (Fig. 2). The size of the sliding window is the only parameter of this thresholding, and we set it empirically to 35×35 pixels.

Segmentation based on intensity alone is challenging when given sparse images with non-uniform intensity contrast. However, in such images, the BG usually has a very large scale, which can help to reliably identify it. Hence, we are using scale images in a second segmentation step. First, we exploit the results from the previous segmentation to set background pixels to zero in the scale image. Then, we apply Hierarchical Cluster Analysis (HCA) image thresholding [11] to the masked scale image. HCA is a histogram based segmentation. It considers each gray level as one cluster, then by computing the distance measure between adjacent clusters, it combines clusters with smallest distance. It repeats this process until only two clusters remain. We chose HCA over simple Otsu thresholding since TV flow tended to overestimate the scale of pixels at the boundary of spots in our

Fig. 2. In contrast to standard global Otsu thresholding, our proposed adaptive thresholding successfully deals with non-uniformly lit images.

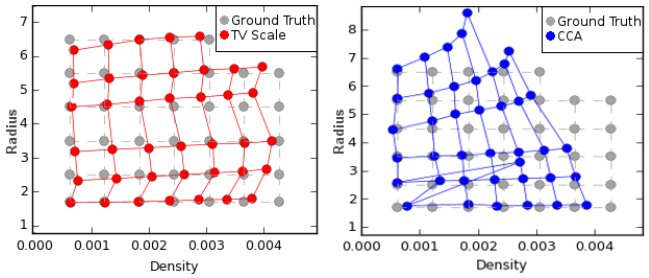

Fig. 3. Results of testing our method and CCA on simulated images, examples of which are shown in Fig. 4. In contrast to CCA, our approach is more robust to denser images, in which the spots tend to get very close to each other and overlap.

experiments, and HCA was more robust to that effect. Future work will investigate the exact source of these edge artifacts, and attempt to remove them.

For dense images, the Otsu thresholding finds a threshold larger than it should be. As a result, the boundaries of spots that usually have lower intensities will be considered as BG. We correct for this in a final step. By replacing the values of each spot that is known to be part of the FG with the average of intensity values of pixels in the boundary of that spot and then applying the first two steps again, we find a new set of pixels as FG. Pixels that are marked as FG in the first two steps are considered as definitely FG. If newly detected FG pixels from the second round are connected to the regions that are definitely FG, then they will be considered as FG as well.

2.6. Experiments

2.6.1. Calibration

In order to calibrate the scales and to evaluate our approach, we simulated different images of size 128×128 pixel with radii equal to $\{1.7, 2.5, 3.5, 4.5, 5.5, 6.5\}$ pixel size and with different number of spots equal to $\{10, 20, 30, 40, 50, 60, 70\}$. We added a highlight to the background of these images in order to simulate non-uniform intensity contrast and we added Gaussian noise. Spots are defined according to the Cauchy distribution, to approximate the expected spot shape in STED

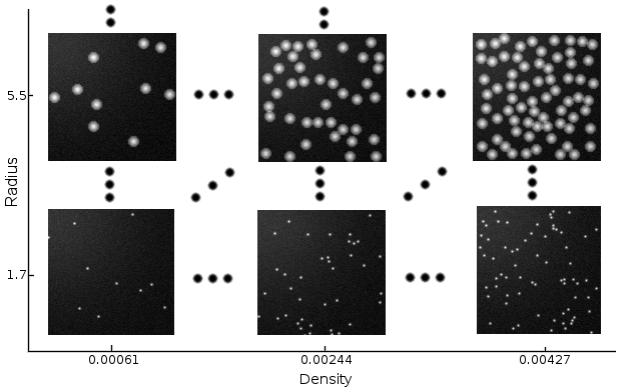

Fig. 4. These example images illustrate the range of scales and densities in the data we simulated for validation.

microscopy. Our scale estimates are calibrated to physical units of length by linear fitting the known radius of the spots in each image to the computed scale from our approach.

2.6.2. Experiments on Simulated Images

In order to evaluate our approach, we created 100 images with random layout of spots for each (radius, density) combination mentioned in Sec. 2.6.1 with $SNR \approx 6.5dB$. Then, for each combination, we computed the average of estimated scales and densities of that group, comparing the results from our method to those from CCA.

Given the number n of spots computed by CCA, we estimate the radius as $r = \sqrt{\frac{p}{n\pi}}l$, where p is the number of foreground pixels computed from the segmentation, and l is the pixel edge length. Since CCA alone gave very poor results, we pre-smoothed its input using TV flow, and used the same FG/BG segmentation as in Sec. 2.5, but without using the scale image, which is specific to our new approach. When using CCA, the lack of scale information required us to adapt the window size of the adaptive Otsu thresholding for sparse images. Despite this, segmentation failure could not be avoided in one very sparse example, as can be seen in Fig. 3. Our new method did not suffer from this problem and was stable with a fixed window of 35×35 pixels.

It is clear that for cases where the spots in the image do not overlap, CCA performs well. However, as it can be seen in Figure 3, wherever the density and therefore the probability of having overlaps increase, CCA fails to estimate density and scale well. In those cases, TVscale is more robust to the layout of spots on the image. CCA also fails for minimum radius and density, since the image becomes too sparse to reliably find a suitable threshold for segmentation. In this experiment, the average computation time for quantifying the spots in each image was 0.015 sec for CCA and 0.529 sec for TVScale method. Most of that time, 0.510 sec, was spent on computing the scale image using TV flow with $T_{start} = 20$ and $T_{stop} = 300$.

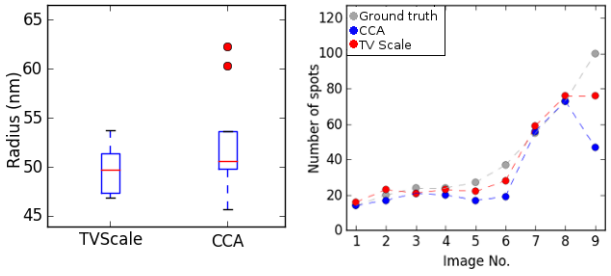

Fig. 5. Estimated radius and spot number from applying our method and connected component analysis to images of TetraSpeck microspheres. The radius of microspheres is known to be $50nm$, and they have been counted manually to obtain a ground truth.

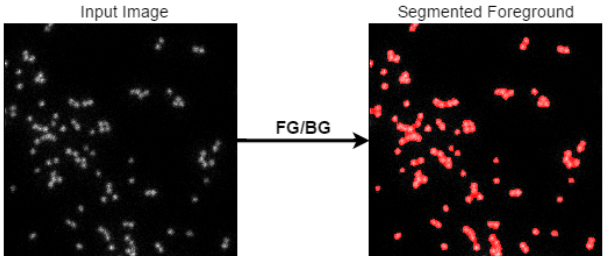

Fig. 6. In this example, connected component analysis fails to correctly quantify image structures, since spots are too close to each other. This image is listed as image 9 in Fig. 5.

2.6.3. Experiments on TetraSpeck Microspheres

We also tested our method on high resolution images of TetraSpeck™ microspheres. These are manufactured with a fixed diameter for the calibration of microscopes and are thus well-suited for quantitative evaluation of our method. As you can see in Fig. 5, our approach better estimates the radius, which is equal to $50nm$. The two outliers for the CCA case are the results from images where the spots are very close to each other, see Fig. 6 as an example. Our method outperforms CCA also for computing the number of spots.

2.6.4. Experiments on Plasma Membrane Protein Clusters

Fig. 7 shows a set of STED images from variants of the plasma membrane protein syntaxin 1A. Syntaxin 1A is known to organize in clusters [1]. An increase of the syntaxin 1A amount in the plasma membrane does not lead to an increase of its cluster size but rather to higher cluster densities [12]. Therefore the question arises if the aggregation level is regulated by an accumulation of charges or rather by steric hindrance of large protein domains. To address this issue, we are interested in measurements of cluster size and density.

Fig. 7 shows a scatter plot of a set of example STED images after using the TVScale method for computing their average spot scale and density. It illustrates the intended use of our method. In contrast to approaches that have focused on segmenting individual spots [2, 3, 4, 5], our primary goal is to provide fast, but robust global summary measures that quickly allow the analyst to find trends and outliers in image sets from high-throughput microscopy. This should allow them to identify smaller subsets of images of particular interest, to which more sophisticated segmentation-based approaches or manual segmentation can subsequently be applied.

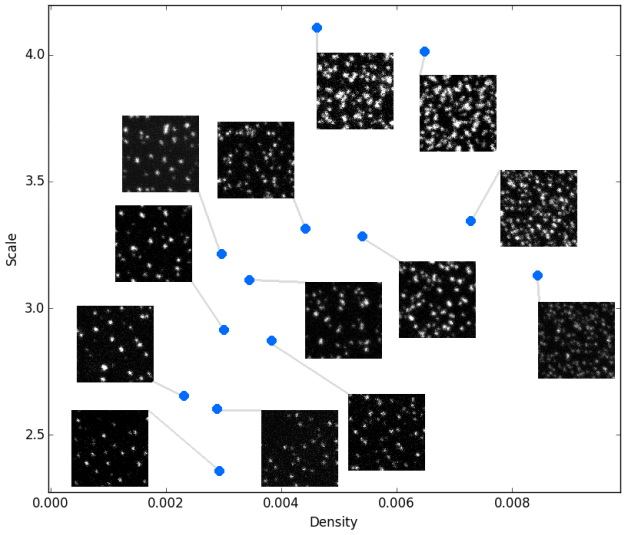

Fig. 7. A set of high resolution STED images of protein clusters, represented by small patches, laid out according to their average scale and spot density, as computed by our method.

In the future, we would like to compare our method with the very recent work of Basset et al. [13] , which is similar to ours in its aim to provide a global estimate of scale and density.

3. CONCLUSION

We present a novel scale measure that is derived from TV flow and allows for per-pixel estimation of the sizes of spots in biological images. We combine this approach with a fully automated foreground / background segmentation and calibrate it to produce estimates of spot density and radius in physical units of length. On simulated data and a physical phantom, we demonstrate that our method is more robust to noise, image artifacts, and segmentation errors than standard Connected Component Analysis. We also demonstrate that it produces plausible results on real-world images from STED microscopy. We believe that our method is well-suited for automated analysis of high-throughput microscopy, where manual counting and measuring are infeasible.

4. REFERENCES

[1] Jochen J Sieber et al., "Anatomy and dynamics of a supramolecular membrane protein cluster," *Science*, vol. 317, no. 5841, pp. 1072–1076, 2007.

[2] Jean-Christophe Olivo-Marin, "Extraction of spots in biological images using multiscale products," *Pattern recognition*, vol. 35, no. 9, pp. 1989–1996, 2002.

[3] Sonal Kothari, Qaiser Chaudry, and May D Wang, "Automated cell counting and cluster segmentation using concavity detection and ellipse fitting techniques," in *Biomedical Imaging: From Nano to Macro, 2009. ISBI'09. IEEE International Symposium on*. IEEE, 2009, pp. 795–798.

[4] Marcin Smereka and Ignacy Duleba, "Circular object detection using a modified hough transform," *International Journal of Applied Mathematics and Computer Science*, vol. 18, no. 1, pp. 85–91, 2008.

[5] Elena Bernardis et al., "Pop out many small structures from a very large microscopic image," *Medical Image Analysis*, vol. 15, no. 5, pp. 690–707, 2011.

[6] Fuensanta Andreu et al., "Minimizing total variation flow," *Differential and integral equations*, vol. 14, no. 3, pp. 321–360, 2001.

[7] Joachim Weickert, BM ter Haar Romeny, Max Viergever, et al., "Efficient and reliable schemes for nonlinear diffusion filtering," *Image Processing, IEEE Transactions on*, vol. 7, no. 3, pp. 398–410, 1998.

[8] Thomas Brox, *From Pixels to Regions: Partial Differential Equations in Image Analysis*, Ph.D. thesis, Saarland University, 2005.

[9] Thomas Brox and Joachim Weickert, "A TV flow based local scale estimate and its application to texture discrimination," *Journal of Visual Communication and Image Representation*, vol. 17, no. 5, pp. 1053–1073, 2006.

[10] David Moroni Strong, *Adaptive Total Variation Minimizing Image Restoration*, Ph.D. thesis, University of California at Los Angeles, 1997.

[11] Agus Zainal Arifin and Akira Asano, "Image segmentation by histogram thresholding using hierarchical cluster analysis," *Pattern Recognition Letters*, vol. 27, no. 13, pp. 1515–1521, 2006.

[12] Jochen J Sieber, Katrin I Willig, Rainer Heintzmann, Stefan W Hell, and Thorsten Lang, "The snare motif is essential for the formation of syntaxin clusters in the plasma membrane," *Biophysical journal*, vol. 90, no. 8, pp. 2843–2851, 2006.

[13] Antoine Basset, Jerome Boulanger, Jean Salamero, Patrick Bouthemy, and Charles Kervrann, "Adaptive spot detection with optimal scale selection in fluorescence microscopy images," *Image Processing, IEEE Transactions on*, vol. 24, no. 11, pp. 4512–4527, 2015.