Morphological Rank-Space Segmentation of Clumped Filaments in Fluorescence Microscopy Imagery

Yongjian Yu[1] and Jue Wang[2]

[1]Axon Connected, LLC, Earlysville, VA, USA
[2]Department of Mathematics, Union College, Schenectady, NY, USA
Corresponding author: wangj@union.edu

Abstract

We present a morphological rank-space segmentation method for automatically detecting and segmenting connected filamentous structures with application to unsupervised analysis of microscopic vaginal flora samples. This approach begins with a binarization process with adaptivity to local noise and object orientation. A novel morphological rank-space is proposed to decompose the connected filaments through morphological thinning, axial linking, labeling, and selection, resulting in a discrete rank-space representation. A ribbon delineation method is devised to reconstruct the boundaries of filaments through the rank-space. Our approach has been successfully applied to detect and segment the pseudohyphae for diagnosis of fungus Candida vaginitis.

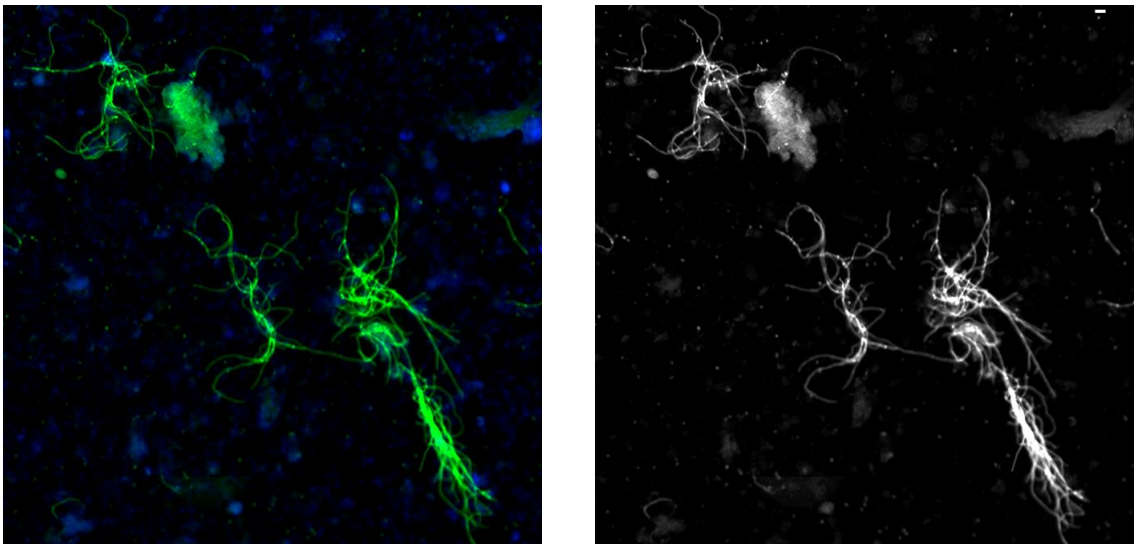
Keywords: Segmentation, morphology, rank-space, fluorescent microscope, filament, pseudohyphae

1. Introduction

Candida albicans is a pleiomorphic fungus and is the most common cause of vaginal fungal infections. Approximately 75% of women have at least one Candida vaginitis (CV) during their lifetimes while nearly half have at least two [4]. Candida overgrowth appears as mats of yeasts measuring 3 to 5 μm in diameter intermingled with pseudohyphae [6], that consist of invasive filaments comprising chains of conjoined, elongated yeast cells. Pseudohyphal growth is regarded as a defined developmental state and important fungal pathogen for CV [12]. The multiband fluorescent microscope is used to provide multi-band digital scans of sample microorganisms that have been pre-processed with immunofluorescent labeling or staining techniques. The fluorescence of the specific dyed microorganisms under the excitation of lights of various wavelengths better reveals the structures of cells of interest. Sample yeast cells and pseudohyphae are shown in Figure 1.

Manual finding and quantification of the pseudohyphae for CV screening and diagnosis is prohibitively time consuming. It is practically desired to develop automated techniques for segmenting the pseudohyphae from microscopic vaginal flora images towards providing visual evidence of fungal pathogens.

Unsupervised segmentation and quantification of the pseudohyphal structures is essential to perform automated high-throughput analysis, in order to compute, for instance, the distributions of lengths, widths, curvatures, budding scars, etc. The challenge for such unsupervised methods attributes to the fact that the pseudohyphae vary in shapes and sizes, form clumps, and the images often contain fiber contamination.



(a) (b)

Figure 1. A representative fluorescent image of yeast and pseudohyphae. (a) 470 nm green and 395 nm blue fluorophores; the former specifies the yeast and pseudohyphal cells and the latter indicates the nuclei (DNA). (b) The green channel.

The global thresholding methods are not appropriate for finding individual pseudohypha from multicellular clumps of pseudohyphae (colonies). The image decomposition approach through curvelet and wavelet transform is applied for segmenting actin filaments [2]. This method is inappropriate for extracting highly curved filaments due to poor computational efficiency as the so-called curvelet transform is actually the size-varying block-wise ridgelet transform through the scale space. Applying Hessain-based scale-space enhancement filters [5, 8] results in disconnections due to the low filter response along the bifurcation regions. These filters have been combined with tracking [3], bifurcation enhancement and suppression of non-tubular structures [15]. All enhancement filter methods require further decomposition and delineation. Other developments include fuzzy shape representation [1] and fuzzy connectedness [7]. An extensive review can be found in [10]. To date, not a single method is capable of successfully segmenting filamentous structures from every imaging modality and every organ.

In this paper, we explore innovative techniques using mathematical morphology and domain knowledge for automatically detecting and segmenting the pseudohyphae, or filamentous structures in general, from microscopic images. We present a paradigm composed of adaptive binarization, clue extraction, morphological rank-space transform, and rank-controlled ribbon contouring of filaments. The binarization is less sensitive to noise and object orientation. The new rank-space allows user specifications to be incorporated into the decomposition; a higher ranked filament is segmented with more integrity at places such as junctions. We demonstrate using clinical data that our system achieves visually agreeable results that enable the feasibility for providing assessment evidence for quick and accurate CV screening.

The rest of the paper is organized as follows. In section 2, we outline the framework for automatically detecting, segmenting, and quantifying pseudohyhae in the microscopic scans of vaginal flora samples. In section 3, we demonstrate the efficacy of our method using clinical data. Conclusions are drawn in section 4.

2. Methodology

Figure 2 depicts the overall structure of the process for segmenting connected filaments. The details of each block are described below.

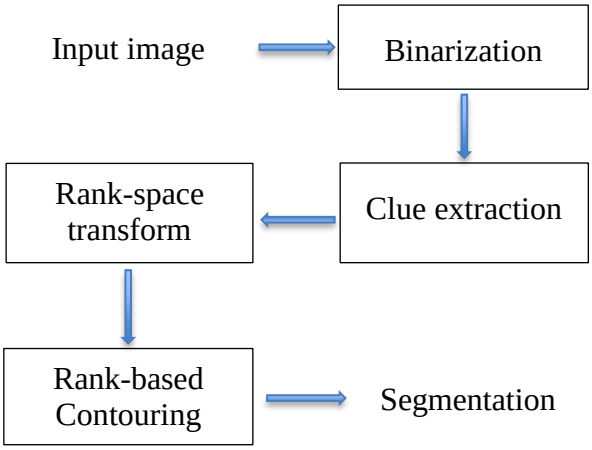


Figure 2. Flowchart of processing algorithms.

2.1 Binarization

The vaginal flora sample is labeled with the Caza Health® antibody package and illuminated with lights. A black and white digital camera captures and grabs the data. The image contains mainly the yeast, budding yeast, pseudohyphae and fibers (possibly from the swab used to collect the sample). The fluorescent microscopic scan must first undergo a series of pre-processing to correct image artifacts caused by channel cross-talk, LED non-uniform illumination, background auto-fluorescence, and so forth. The green channel (see Figure 1) captures most effectively the fluorescence emitted by the yeast and pseudohyphae.

The binarization process masks the interest areas of pseudohyphal colonies, resulting in a binary ridge image. This procedure adopts local adaptive thresholding algorithms to achieve the robustness to the local noise variation. The detection technique needs to be sensitive to the object orientation, as well as lowering the impact of interfering objects present in the test window upon the noise statistics. It is accomplished through the following steps. We slide M tilted line-segment windows over the image domain; at each point, we collect N reference samples per window to compute local statistics as a function of the window tilt angle. We then select the window that is most orthogonal to the pseudohyphal axis direction, and implement the adaptive threshold test, as described by

$$I - Z_{\perp} > C \cdot \left(1 + \left(\frac{Z_{\perp}}{Z_{max}}\right)^{r}\right), \tag{1}$$

where I is the green channel image with intensity normalized to [0, 1]. $Z_{\perp}$ corresponds to the test statistics obtained in the window that is most orthogonal to the pseudohyphal segment,

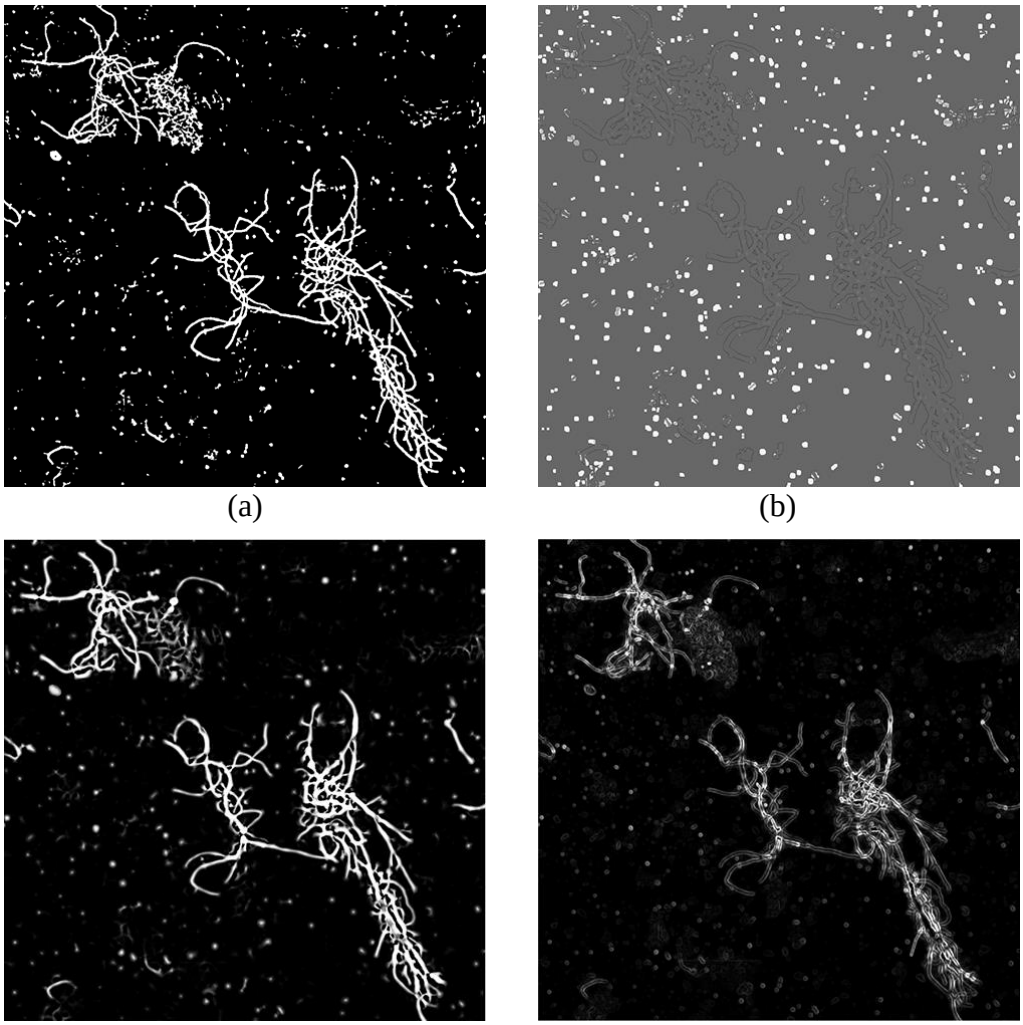
$$Z_{\perp} := \min_{1 \le k \le M} Z_k,$$

$$Z_{max} = \max_{1 \le k \le M} Z_k;$$

Z_k is the mean intensity in the N-pixel line segment window W_k centered at position (i, j) in the image domain with a tilt angle of $\frac{k-1}{M}\pi$ from the x-axis, *i.e.*,

$$Z_k = \text{mean}\{I_{ij} \in W_k\}.$$

The r in Eq. (1) is a positive real factor; C is a threshold that is proportional to the noise standard deviation. It is set to maintain an acceptable filament detection probability for low SNR with a consistent false detection rate. The orientation adaptivity manifests through the multiplier of C. For elongated structures, $Z_{max} > Z_{\perp}$, and the right-hand side of Eq. (1) is approximately equal to C. For round objects, $Z_{max} = Z_{\perp}$, and the right-hand side is doubled C so that the round or close to round ones are not being detected as pseudohyphae. In this study, we choose the parameters experimentally: M = 8~16, N = 15, C = 0.01~0.05, and r = 2. The resulting binary ridge mask is shown in Figure 3(a).



(a) (b) (c) (d)

Figure 3. (a) Binary ridge mask, (b) Spotness map, (c) Ridgeness map, (d) Gradient magnitude of Figure 1(b).

2.2 Spotness and ridgeness features

We adopt a newly devised *spotness* feature [14] and *ridgeness* feature for providing segmentation criteria for the next section. They are illustrated in Figure 3 (b) and (c). The spotness appears bright on compact structures such as yeast while showing dark boundaries on elongated structures such as pseudohyphae. Thus, the spotness feature is tailored to discriminating between yeast and pseudohyphae. The spotness image is a rotational morphological top-hat transform of an input image I, defined as

$$p = I - \max\{\gamma_{i,B}(I)\}_{i=0}^{K-1}, \tag{2}$$

where $\gamma_{i,B}(I) := R_{-\theta_i}\left(R_{\theta_i}(I) \circ B\right)$; R_{θ_i} is the rotation w.r.t. the image center by an angle $\theta_i := i\frac{\pi}{K}$; K is the number of angles. The operator $\circ$ denotes opening by a horizontal line structure element B of l-pixel long and 1-pixel wide. The second term in Eq. (2) suppresses spot structures of size less than l while preserving the interiors and augmenting the borders of filamentous structures of dimension longer than l, regardless of orientations. In our study, we set K = 36 and l = 20.

The ridgeness feature is based on the multi-resolution enhancement filter response [5]. For an input image I, the Hessian matrix H is defined as

$$H_{ij}(x,\sigma) = \sigma^2 I(x) * \frac{\partial^2}{\partial x_i \partial x_j} G_\sigma(x),\ i,j = 1,2, \tag{3}$$

where $G_\sigma(x)$ is the 2D Gaussian of standard deviation σ and * denotes convolution. The eigenvalues of H are sorted, $|\lambda_1| \leq |\lambda_2|$. The larger intensity variation takes place along the second eigenvector, corresponding to the direction across the intensity boundary. To enhance the elongated more than the spherical structures, a ridgeness indicator function is computed as

$$R = \max_{\sigma \in D} e^{-\frac{(\lambda_2/\lambda_1)^2}{2\alpha^2}} \left(1 - e^{-\frac{\sqrt{\lambda_1^2+\lambda_2^2}}{2\beta^2}}\right), \tag{4}$$

where α and β are sensitivity parameters. The set D is preset according to the expected range of filament width. R lies between 0 and 1. We set α = 2, β = 8, and D = {1,2,3,4}.

2.3 Segmentation of connected filaments

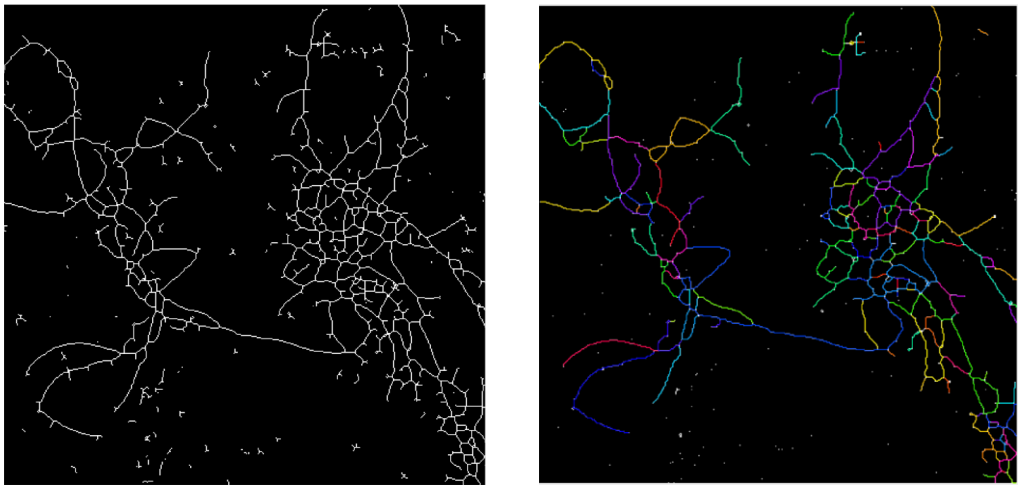
The pseudohyphae tend to clump together in the developed state of Candida infection. In order to assess the infection by taking the geometrical and other measurements of individual pseudohypha, it is necessary to split and segment the clumped cells. Our segmentation algorithm comprises two steps: analysis and synthesis. In the first step, the binary mask of ridges is decomposed into distinct, thin ridge markers of descending rank-scale orders; each marker represents a disjoint cell axis (or central line). The synthesis of the rank-scale space restores the filamentous shapes while preventing the merging of the boundary contours of split ridge markers.

The analysis procedure is described as follows. We first extract appropriate pseudohypha-like ridge markers by morphological thinning and pruning of the binarized yeast mask map, and then link the 1-pixel wide ridge elements to form chains of connected ridges. Finally, a morphological rank-space representation of pseudohypha-like filaments is generated by conditioned, scale-rank prioritized reconstruction and filtering based on significance metrics and thresholds derived from clinical relevance.

The thinning algorithm removes pixels on the boundaries of connected components without breaking them apart. This operation is repeated until the image is stable, followed by morphological pruning that removes the end points of lines without removing small objects completely.

The linking algorithm takes the thinned binary mask of ridges and ridge direction angle image (obtained from multi-resolution enhancement filter), and outputs a list of labeled ridge chains, as shown in Figure 4. The ridge-linking algorithm is outlined below:

1) Start a row-wise raster scanning of the binary ridge map from the upper left corner, *i.e.*, the image origin.
2) Follow the ridge with 8-connectivity to the end. Search through its neighbors within a preset distance for the points that are most compatible with respect to the ridge direction.
3) A chain of ridges either terminates at a joint or branch to grow if its continuity is stronger than that of other chains connected to the same joint.
4) Flip the tracked pixels, and track the rest of the ridge segment, if any, to the end of the ridge chain.
5) Finally make sure that the starting point of each ridge chain is closer to the image origin than its end.



(a) (b)

Figure 4. (a) Zoomed-in map of thinned ridges representing the axial lines of the filamentous pseudohyphae. (b) Segmented ridge central lines with random color-labeling.

The morphological rank-space filament clump decomposition depends on the priority criteria that are devised specifically according to application. We elaborate in this paper the morphological multi-rank space with focus on segmenting pseudohyphae. This paradigm is applicable to general curved filament segmentation tasks.

Three metrics are measured for each ridge marker: the length, average fluorescence intensity, and average ridgeness. The prioritizing algorithm performs as follows:

1) Each ridge chain receives three separate scores $\{t_i\}_{i=1,2,3}$, respectively, according to its index in the descending order of length, average fluorescent intensity, and average ridgeness.
2) A total score $t = \sum_{i=1}^{3} t_i$ is assigned to each ridge chain. The ridge chains are then sorted again in the ascending order based on their total scores.
3) The first P ridge chains are selected from the final sorted list. The number P depends on a predefined maximum number of pseudohyphae.

With these metrics, the long, bright ridges (or filaments) receive a higher priority than short, dark, non-filamentous ones (such as edges). We define the rank of a ridge to be its index of the descending total scores. In the rank-space, the originally crowded ridge markers are split

and redistributed into multiple rank subspaces. The markers with higher priority reside sparsely in the high-rank subspace.

In the synthesis process, the exterior boundaries of the disjoint filamentous structures are reconstructed in a bottom-up manner through the rank-space. Starting from ridge marker(s) with the highest rank, we delineate the boundary contours of filaments that have the same rank *s*, then append them to the output sequence indexed with ranks $(1, 2, \dots, s-1)$. During this process, contours do not merge. The algorithm is provided below.

1) Compute the gradient magnitude $g(x)$ of the green channel image. An example is shown in Figure 3(d).
2) Remove 1 pixel at the joint end of each line marker of rank *s* resulting in a trimmed ridge marker. Then compute a rank-based binary influence zone $B_s(x)$, which is given by the watershed transform of the distance map of the union of trimmed ridge markers at rank *s* and the already-segmented filament ribbons from rank 1 to $s-1$.
3) At rank *s*, loop through each point along the curve marker, and search in the direction perpendicular to the curve, within the influence zone $B_s(x)$, to locate the maximum gradient magnitude (boundary) points on both sides of the curve marker. The boundary contour of the marker is described by a marker curve $\{x_{s,i}; i = 1, \dots, l_s\}$, where $x_{s,i}$ represents the row and column index of the i^{th} ridge point along the central line of rank *s*; l_s is the length of that marker. The variable radius is given by

$$\hat{r}_{s,i} = \min\left\{\max_{r_i}\{g(x_i^+)\}, \max_{r_i}\{g(x_i^-)\}\right\}, \tag{5}$$

where $x_i^{\pm} = x_i \pm r_i \frac{(x_i - x_{i-1})^{\perp}}{|x_i - x_{i-1}|}$; the symbol $\perp$ denotes transverse of a vector. For simplicity the subscript *s* is omitted. The geometry is illustrated in Figure 5.
4) Append $\{(x_{s,i}, \hat{r}_{s,i}); i = 1, \dots, l_s\}$ to the previously reconstructed contours.
5) Compute the average width and minimum spotness for each segmented filament. The combination of length, average fluorescence intensity, average ridgeness, average width, and minimum spotness fully characterizes each segmented pseudohyphal suspect.
6) Finally, identify the pseudohyphal filaments by means of length thresholding or width thresholding.

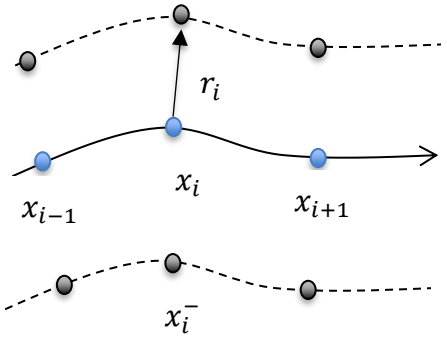


Figure 5. Geometry of filament segmentation.

3. Results

We evaluate the morphological rank-space segmentation method using two sets of images. In the first test, images of size 1024 by 1024 pixels were acquired using a Nikon inverted

microscope, charge-coupled device camera and 4X objective lens. The proposed method is applied to a selection of the images. The magnified results are shown in Figure 6.

We compare our method to the marker-controlled watershed transform [9] and morphological multiscale decomposition (MSD) [11, 13]. Our approach outperforms the other methods in terms of preserving the integrality of long filaments and providing visually agreeable splitting of clumped cells. An average ribbon width is computed and adopted per pseudohypha display for easy visualization in comparison. The actual ribbon size distribution can be used for pseudohyphae quantification. The 15-scale MSD as applied to the binary ridge detection breaks the long continuous filament into multiple segments. The watershed transform fails with the same markers as used in our method.

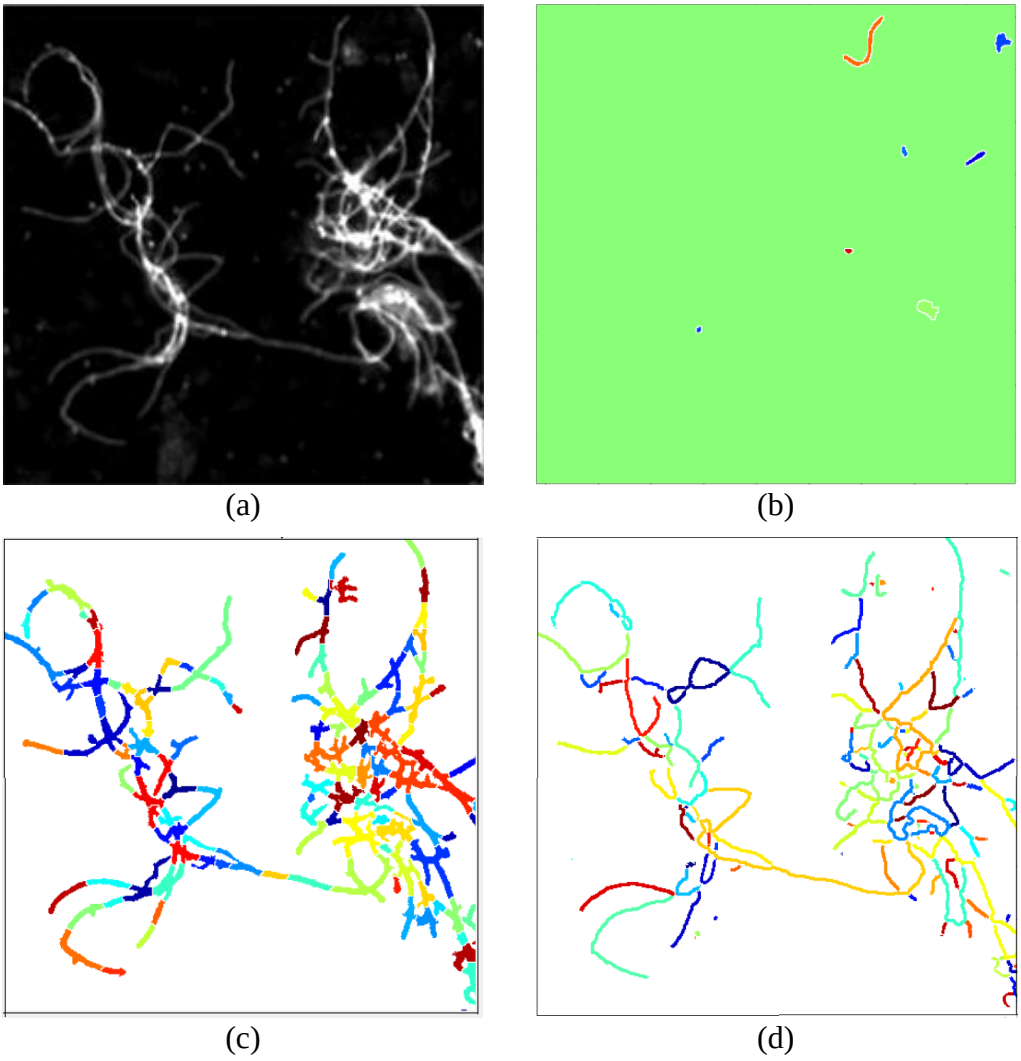
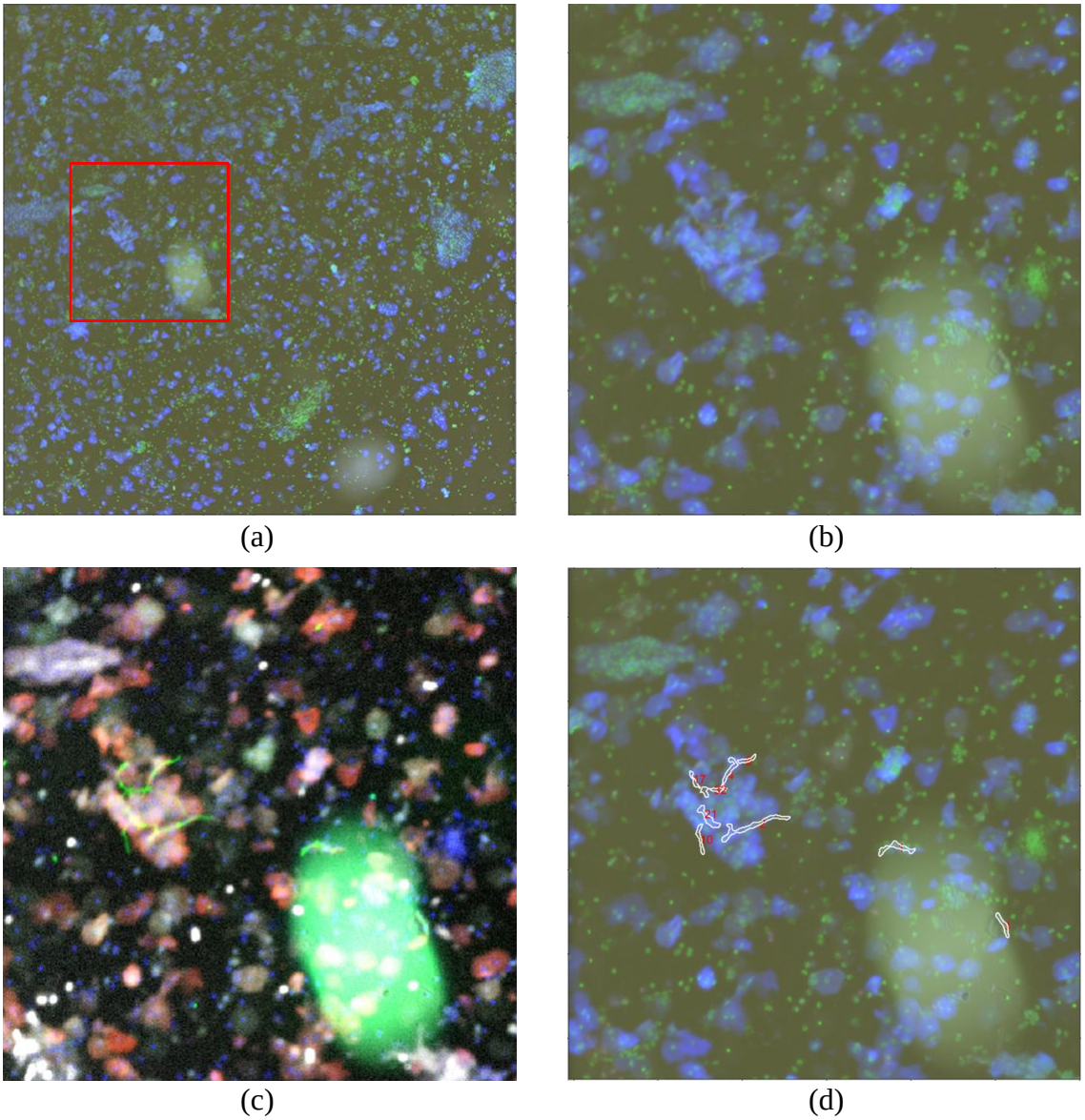


Figure 6. Method comparison. (a) Input image, (b) Marker based watershed, (c) 15-scale MSD, (d) Our method.

In the second test, we apply the method to clinical vaginal samples from Discovery Life Sciences (DLS), Inc. The samples are prepared using Axon Pac and scanned using nCyte microscope. The image size is 2048 by 2048 pixels. One of the test images is shown in Figure 7(a). In this fluorescent image, the pseudohyphae are highlighted in red, the nuclei (DNA) in green, the epithelial tissue cells in blue, and trich markers in white. Our analysis is performed

for the entire image. For detail viewing, magnified displays of the region in the red box are illustrated in Figures 7(b)-(d).

A color-enhanced view is shown in Figure 7(c). In this view, different color encoding allows the pseudohyphae to be highlighted in green so that the viewer can see the ground truth more clearly. The blue dots are nuclei (DNA); the red are epithelial tissue cells; the white blobs are trich markers. In Figure 7(d), the pseudohyphae detected and segmented by our method are highlighted as white contours and overlaid on the original image, along with their count numbers. This provides key information for CV diagnosis. A comparison of Figure 7(d) with 7(c) indicates that our algorithm achieves consistent performance.



(a) (b) (c) (d)

Figure 7. Test on DLS clinical vaginal sample scan. (a) 2048 by 2048 input image. (b) Magnified view of the red box in (a). The pseudohyphae are highlighted in red, nuclei (DNA) in green, epithelial cells in blue, and trich markers in white. (c) Color-enhanced view. The pseudohyphae are highlighted in green, nuclei (DNA) in blue, epithelial cells in red, and trich markers in white. (d) The pseudohyphae detected and segmented by our method are overlaid as white contours, along with their count numbers in red.

4. Conclusions

We have demonstrated a generic morphological rank-space decomposition technique for segmenting connected filamentous structures. This approach enables a user-specific interest priority and natural scales to be integrated in a novel segmentation paradigm. Our method is evaluated using clinical fluorescent microscopic images of vaginal samples. It is computationally efficient. The results are visually expected, consistent and promising. They provide clues for diagnosis and aid in quantitative Candida vaginitis screening.

References

[1] G. Agam, S.G. Armato and C. Wu, "Vessel tree reconstruction in thoracic CT scans with application to nodule detection", *IEEE Trans. Med. Imag.*, vol. 24, no. 4, pp. 486-499, 2005.

[2] M. Alisocha-Perez, C. Benadiba, K. Goossens, S. Kasas, G. Dietler, R. Willaert and H. Sahli, "A robust actin filaments image analysis framework", *PLOS Computational Biology*, 1005063, 2016.

[3] S. Aylward and E. Bullitt, "Initialization, noise, singularities, and scale in height ridge traversal for tubular object centerline extraction", *IEEE Tans. Med. Imag.*, vol. 21, no. 2, pp. 61-75, 2002.

[4] M.E. Egan and M.S. Lipsky, "Diagnosis of vaginitis", *Am. Fam. Physician*, vol. 62, no. 5, pp. 1095-104, 2000.

[5] A.F. Frangi, W.J. Niessen, K.L. Vincken and M.A. Viergever, "Multiscale vessel enhancement filtering", *Proc. Med. Image Comput. Assist. Interv.*, vol. 1496, pp. 130-137, 1998.

[6] J. Guarner and M.E. Brandt, "Histopathologic diagnosis of fungal infections in the 21st century", *Clin. Microbiol. Rev.*, vol. 24, no. 2, pp. 247-280, 2011.

[7] J.N. Kaftan, A.P. Kiraly, A. Bakai, M. Das, C.L. Novak and T. Aach, "Fuzzy pulmonary vessel segmentation in contrast enhanced CT data", *Proc. SPIE Medical Imaging 2008: Image Processing*, vol. 6914, 2008.

[8] K. Krissian, G. Malandain, N. Ayache, R. Vaillant and Y. Trousset, "Model based detection of tubular structures in 3D images", *Comp. Vis. Image Understand.*, vol. 80, no. 2, pp. 130-171, 2000.

[9] C. Rambabu and I. Chakrabarti, "An efficient immersion-based watershed transform and its prototype architecture", *J. Syst. Archit.*, vol. 53, no. 4, pp. 210-226, 2007.

[10] D.R. Rina *et. al.*, "Comparing algorithms for automated vessel segmentation in computed tomography scans of the lung: the VESSEL12 study", *Med. Image Anal.*, vol. 18, no. 7, pp. 1217-1232, 2014.

[11] O. Schmitt and M. Hasse, "Morphological multiscale decomposition of connected regions with emphasis on cell clusters", *Comput Vis. Image Underst.*, vol. 113, no. 2, pp. 188-210, 2009.

[12] V. Veses and N.A.R. Gow, "Pseudohypha budding patterns of Candida albicans", *Medical Mycology*, vol. 47, no. 3, pp. 268-275, 2009.

[13] F. Xing and L. Yang, "Robust nucleus/cell detection and segmentation in digital pathology and microscopy images: a comprehensive review", *IEEE Rev. Biomed. Eng.*, vol. 9, pp. 234-263, 2016.

[14] Y. Yu and J. Wang, "Automated enumeration and classification of bacteria in fluorescent microscopy imagery", *Proc. IEEE LSC*, 2018.

[15] C. Zhou, H.P. Chan, B, Sahiner, L.M. Hadjiiski, A. Chughtai, S. Patel, J. Wei, J. Ge, P.N. Cascade and E.A. Kazerooni, "Automated multiscale enhancement and segmentation of pulmonary vessels in CT pulmonary angiography images for CAD applications", *Med. Phys.*, vol. 34, no. 12, pp. 4567-4577, 2007.