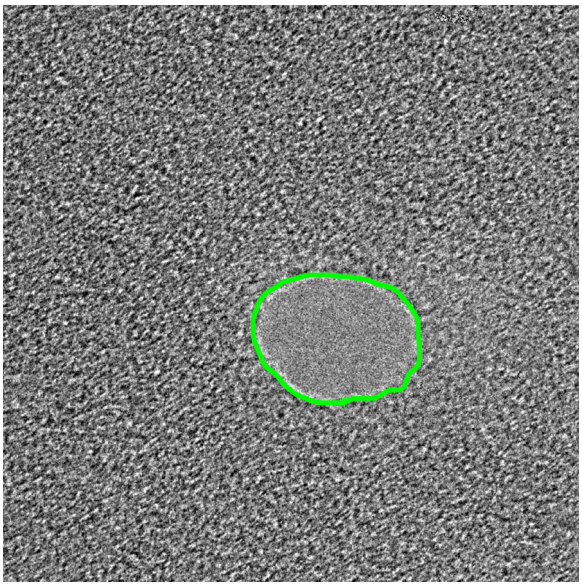
IEEE Photonics Journal
An IEEE Photonics Society Publication

Solid-State Nanopore Recognition and Measurement Using Shannon Entropy

Volume 3, Number 3, June 2011

Ty R. Wojcik, Student Member, IEEE
Diego Krapf, Member, IEEE



DOI: 10.1109/JPHOT.2011.2129503
1943-0655/$26.00 ©2011 IEEE

IEEE photonics SOCIETY

IEEE

Solid-State Nanopore Recognition and Measurement Using Shannon Entropy

Ty R. Wojcik,[1] ***Student Member, IEEE*****, and Diego Krapf,**[1,2] ***Member, IEEE***

[1]Department of Electrical and Computer Engineering, Colorado State University, Fort Collins, CO 80523 USA
[2]School of Biomedical Engineering, Colorado State University, Fort Collins, CO 80523 USA

DOI: 10.1109/JPHOT.2011.2129503
1943-0655/$26.00 ©2011 IEEE

Manuscript received February 14, 2011; revised March 10, 2011; accepted March 10, 2011. Date of publication March 17, 2011; date of current version May 6, 2011. Corresponding author: D. Krapf (e-mail: krapf@engr.colostate.edu).

Abstract: Solid-state nanopores are structures that can be fabricated using the electron beam of a transmission electron microscope (TEM). Nanopores can be used to electrically detect individual DNA molecules, and they have the potential to be applied to genomic sequencing. Current nanofabrication methods are manual and time-consuming, and thus, they do not allow for the fabrication of large-scale nanopore arrays. One of the requirements in the development of an efficient fabrication process is the automated recognition and measurement of nanopore dimension in real time. Here, we present a method for nanopore edge detection that uses Shannon entropy to identify nanopores in electron microscopy images. The nanostructure edges are determined by applying a Shannon entropy filter. This entropy image is segmented, and the texture-defined edges are determined. The diameters of nanopores as small as 3 nm are directly measured.

Index Terms: DNA sensing, microscopy, nanohole arrays, nanostructures.

1. Introduction

Solid-state nanopores are structures defined in silicon-based membranes that are typically fabricated using a focused electron beam in a transmission electron microscope (TEM) [1]. Nanopores can be used to electrically detect individual DNA molecules [2]–[4], and they possess the potential to be used for rapid genomic sequencing [5], [6]. In applications where a large amount of data is required, such as genomic sequencing, device parallelization is strongly desirable, and thus, nanopore arrays have been proposed [7]. Nanopores also show promise in the integration into microfluidic and optoelectronic devices [8], [9]. Nevertheless, any efficient industrial application would need the implementation of an automated process. Current nanofabrication methods are manual and time-consuming and therefore they do not allow for the fabrication of large-scale nanopore arrays. Furthermore, in the utilization of nanopores for nanobioscience it is imperative to work with devices of a given predetermined size in order to obtain reliable and reproducible data. One of the requirements in the development of an automated fabrication process is the availability of an electron microscope image processing tool that can recognize and measure nanopore dimensions. Unfortunately, image segmentation using pixel intensities alone does not yield satisfactory results due to the poor intensity contrast that characterizes nanopore images.

The lack of intensity contrast in nanopore electron microscopy images is due to the intrinsic interaction between the electron beam and the membrane. The material of choice for nanopore fabrication is silicon nitride or a combination of silicon nitride and silicon oxide thin layers. When the membrane is thinner than 50 nm, a focused electron beam of a TEM operated at 250 keV is able to

efficiently remove material from the membrane. Then, the same beam is slightly defocused to reduce the beam intensity and the nanopore size is changed by means of surface tension [10]–[12]. Modifying the size is done in a controlled manner while the nanopore dimensions are monitored by real-time imaging. However, only a small fraction of the electrons arriving at the membrane are involved in collisions with the membrane material while most electrons continue their path unaffected. Furthermore, most of the electrons that undergo inelastic collisions still arrive at the detector, albeit with lower energy [13]. Therefore, the density of collected electrons going through the nanopore does not differ much from the electrons that are transmitted through the membrane. As a consequence, the rendered electron microscope image lacks any useful brightness contrast between the nanopore region and the surrounding material.

Traditional edge detection methods use gradient or Laplacian operators to detect edges. These methods are ineffective for nanopore recognition because they are based on intensity contrast. In this paper, we present a method for nanopore edge detection that uses Shannon entropy instead of intensity-based segmentation methods. Shannon entropy, which is a concept derived from information theory, is a parameter that quantifies the amount of information available in an image. Shannon entropy was previously used in many different image processing applications [14]–[16], and it is a mature edge detection/segmentation approach to texture-defined objects. Here, information-defined edges are determined by computing the local Shannon entropy in a small window that is scanned across the image. This entropy image is then contrast-enhanced and smoothed with a Gaussian filter to eliminate noise artifacts. The image obtained is segmented via automatic thresholding and the edges are determined. We show that nanopores with diameters as small as 3 nm can be directly recognized and measured using this protocol.

2. Methods

2.1. Shannon Entropy

Shannon entropy quantifies the information available in an image. It is a measure of the uncertainty of a set of random variables and is defined in terms of the probabilistic behavior of a source of information [17]. This property states that the entropy is minimal in homogenous areas and high in areas containing broad pixel distributions. The amount of self-information of an event is inversely proportional to the probability of the event. A random event that occurs with probability $p(X)$ contains $I(X)$ units of information defined by $I(X) = -\log_n p(X)$, which denotes the self-information of X. If $p(X) = 1$, then $I(X) = 0$, because there is no uncertainty of event X happening. The base of the logarithm, n, determines the unit which is used to measure the information. Thus, if we are using bits, the base of our algorithm is 2. If the probability mass function is presented as the self-information of X, then the Shannon entropy of X is the expected value of the self-information

$$H(X) = -\sum_i p_i \log_n(p_i). \tag{1}$$

The Shannon entropy of a 2-D image is calculated from (1), where p contains the gray-level histogram of the image (i.e., p_i are the counts of level i).

2.2. Algorithm Outline

The main part of our algorithm was based on replacing the value of each pixel by the entropy of a window surrounding the pixel. The Shannon entropy is computed in a scanning region of interest. We chose to use a 19×19 pixel window as the region of interest and an entropy image is formed where the calculated entropy of the window gives the intensity of its center pixel. A histogram transformation of the intensity levels was performed so that the entropy levels obtained were expanded to fill the 8-bit range, i.e., between 0 and 255. Fig. 1(a) and (b) show a representative nanopore image and its entropy image. In order to avoid the detection of artifacts in the image, the noise generated from the entropy calculation was eliminated by smoothing the contrast-enhanced entropy image with a Gaussian filter. The smoothed image is shown in Fig. 1(c). The problem of the

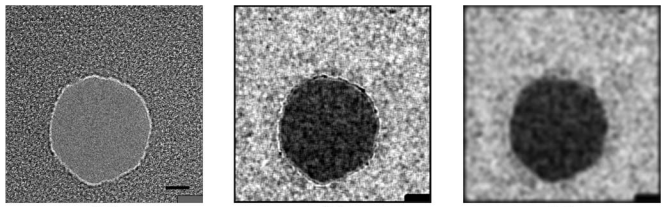


Fig. 1. Shannon entropy of a nanopore TEM image. (a) Original nanopore, 1024 × 1024 pixels. The black scale bar on the bottom right is 5 nm. (b) Shannon entropy image processed with a 19 × 19 pixel window and normalized to fill the 8 bit (0 to 255) range. (c) Entropy image smoothed with a Gaussian filter.

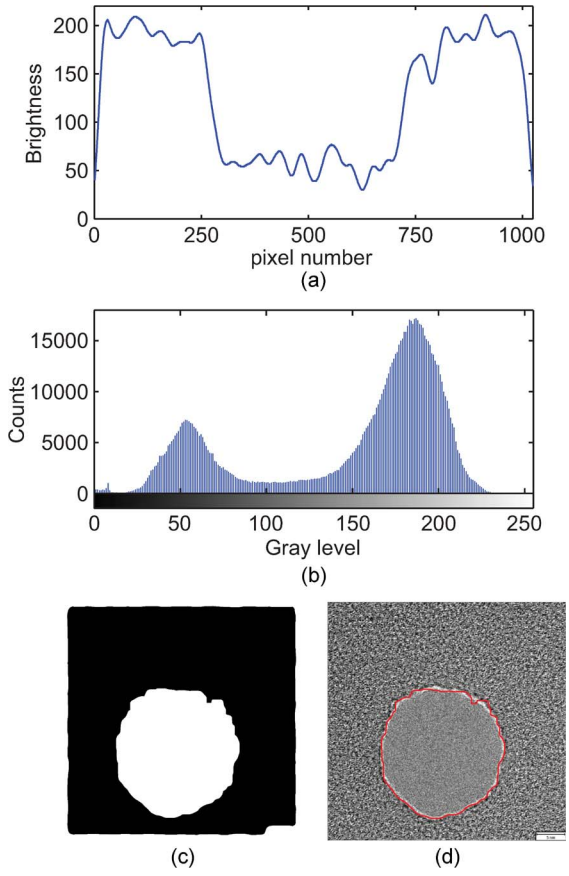


Fig. 2. Analysis of Shannon entropy image. (a) Brightness values of a cross section along the nanopore, showing the intensity of the pixels in the 512th row. (b) Histogram of the whole entropy image. (c) Binary image obtained by automatic thresholding of the entropy image. (d) Boundary of the binary image superimposed on the original nanopore.

logarithm-of-zero problem was addressed by adding 10^{-4} to all p_i values, which directly yields $\lim_{p\to 0} p\log(p) = 0$.

The resulting entropy image [see Fig. 1(c)] presented a strong contrast between the nanopore and the surrounding nitride membrane. Fig. 2(a) shows a cross section along the center of the nanopore. The gray-level histogram of the entropy image exhibited two well-resolved peaks [see Fig. 2(b)] that identified the region inside the nanopore and the nitride membrane. This image was segmented by automatic thresholding [18]. The threshold was set by finding the lowest value in the valley between the two peaks. More involved thresholding algorithms could be used here [19],

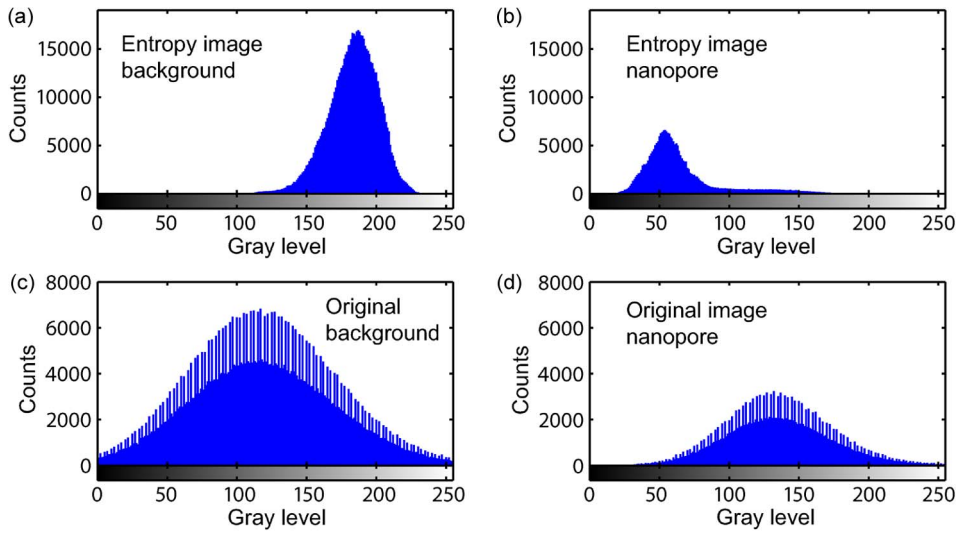


Fig. 3. Nanopore and background histograms. (a) Entropy histogram of the background. (b) Entropy histogram of the nanopore. (c) Gray-level histogram of the background in the original image. (d) Gray-level histogram of the nanopore in the original image.

but given the simplicity of the obtained histogram, we did not evaluate using other thresholding techniques. This threshold is used to generate a binary image of the nanopore [see Fig. 2(c)] by assigning all pixels below the threshold value to the background and pixels above the threshold value to the object of interest.

The boundary pixels were found using the built-in MATLAB function *bwboundaries* [20]. This function returned an array of cells, with each cell corresponding to an object in the binary image. We defined an object as an identifiable portion of the image having at least 1000 connected pixels. A cell contained the (x_i, y_i) pixel locations that define the edge of the object. The nanopore boundary superimposed on the original nanopore image is shown in Fig. 2(d). The boundary pixel locations were fitted to a circle equation $x^2 + y^2 - 2x_c x - 2y_c y + C = 0$ using a least squares algorithm.

2.3. Algorithm Implementation

We used 64-bit MATLAB 7.11 running on Windows 7 for implementing the nanopore recognition algorithm. We run this algorithm in a commercial desktop computer with an Intel Core i7 processor and 12-GB DDR3 RAM. The whole MATLAB code is available in the supplemental information.

We selected an entropy window of 19 $\times$ 19 pixels in the algorithm implementation by trial and error. The optimal window size is determined by the magnification and camera pixel size; thus, it is instrument dependent. We find that a 19 $\times$ 19 pixel window provides an excellent starting point in a broad range of images. The Gaussian kernel that we implement to smooth the entropy image has a standard deviation 10. The sensitivity of the outcome of our algorithm to changes in this parameter is minor.

3. Results and Discussion

The performance of the entropy-based segmentation algorithm is appreciated in Fig. 3. Fig. 3(a) and (b) show the separate histograms of the background and the nanopore, respectively, of the entropy image. The joint histogram is shown in Fig. 2(b). On the other hand, the gray-level histograms of the nanopore and the background of the original (segmented) image show very similar gray level [see Fig. 3(c) and (d)]. The radical improvement in the bimodality of the entropy histogram justifies the proposed transformation.

Results of nanopore recognition and analysis are shown in Fig. 4. The nanopore images used in this algorithm were 8-bit TIFF images with a resolution of 1024 $\times$ 1024 pixels. The average time

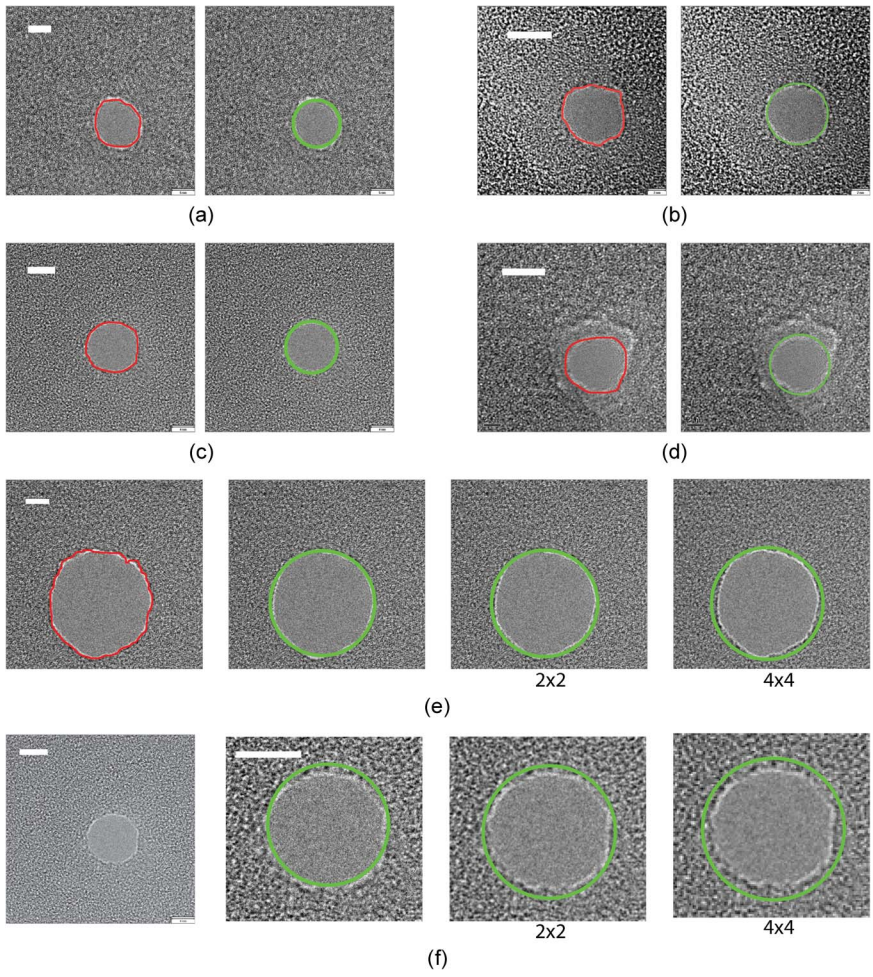


Fig. 4. Results of the nanopore edge finding and circle fitting algorithms. All white scale bars on the upper left corners are 5 nm. (a)–(d) Representative results showing four different nanopores. The images on the left show the boundary (red) superimposed on the nanopore and the images on the right the circle fitting (green) to the same boundary. (e) Same nanopore shown in Figs. 1 and 2. The image on the left shows the original nanopore with the boundary. The next three images show the circle fitting results when the unbinned image was processed, using the 2 × 2 binned image, and 4 × 4 binning. (f) Circle fitting results to a nanopore image that was cropped to 448 × 420 pixels and binned. The left image shows the original nanopore.

required to execute the algorithm on a nonmodified, full-size image was 130 s. We explored several options to drastically reduce this processing time with our current computational capabilities. The simplest way to improve run-time was to use smaller images by selecting a smaller region of interest and analyzing only that region. In a TEM, this can be very easily accomplished because the region where a nanopore is being sculpted is known *a priori*. The average time to implement our algorithm on a 256 × 256 pixel image was 7 seconds and on a 200 × 200 pixel image was only 5 s.

A different and attractive alternative to reduce the image size and thus the run-time is to perform pixel binning on the whole image. Algorithm results in nanopore images without binning and in 2 × 2 and 4 × 4 binned images are shown in Fig. 4(e) and (f). The binning factor represents the number of pixels that are combined to form each larger pixel [21]. The algorithm proves to yield accurate results and the obtained radius is within a 10% error. The image in Fig. 4(f) was cropped to a region of interest of 448 × 420 pixels and further binned to reduce run-time. The time to process these images was 23.6 s, 5.6 s, and 1.5 s for the case without binning, with 2 × 2 and 4 × 4 binning,

respectively. Nanopores ranging from 3 nm to 100 nm in diameter have been analyzed without changing any of the algorithm parameters. This provides evidence for the robustness of this method. However, adjusting the window size according to different nanopore size gives more accurate and faster results.

The membrane surrounding the nanopore is frequently thinned by the tails of the electron beam. This effect introduces a particular challenge in the recognition of nanopores as the thinned membrane presents a modified texture. Fig. 4(d) shows a nanopore with these characteristics. However, as observed in the figure, our method is robust enough to operate in this situation and accurately recognizes a nanopore in the presence of a thinned surrounding membrane.

4. Conclusion

We have successfully analyzed TEM images of nanopores using texture analysis methods. Traditional edge detections such as gradient or Laplacian operators could not be used for this application due to the lack of contrast of the nanopore. To overcome this problem, we used a Shannon entropy filter to automatically identify edge boundaries. The coordinates of edges can be quantified by a circle-fitting algorithm. This allows for a clear and efficient method to determine nanopore edges. The time to compute Shannon entropy filtering with current personal computers for 256×256 images is of the order of a few seconds. While it is possible to use this time resolution for automatic nanopore manufacturing, selecting a smaller region of interest where the nanopore is being processed can reduce the time requirements for this algorithm to the millisecond regime. Furthermore, it is expected that with stronger computing capability and multiprocessor computing implementation the time it will take to perform such operations in the near future will be greatly reduced. These results together with the provided MATLAB scripts can be readily implemented in a system to automatically fabricate solid state nanopores and nanopore arrays.

Acknowledgment

The authors thank Prof. Dr. H. W. Zandbergen and Dr. M.-Y. Wu for their help with nanopore fabrication, B. Simon for his help with the MATLAB scripts, and A. Weigel for carefully reading the manuscript.

References

[1] C. Dekker, "Solid-state nanopores," *Nat. Nanotechnol.*, vol. 2, no. 4, pp. 209–215, Apr. 2007.
[2] D. Fologea, J. Uplinger, B. Thomas, D. S. McNabb, and J. L. Li, "Slowing DNA translocation in a solid-state nanopore," *Nano Lett.*, vol. 5, no. 9, pp. 1734–1737, Sep. 2005.
[3] J. B. Heng, A. Aksimentiev, C. Ho, P. Marks, Y. V. Grinkova, S. Sligar, K. Schulten, and G. Timp, "The electromechanics of DNA in a synthetic nanopore," *Biophys. J.*, vol. 90, no. 3, pp. 1098–1106, Feb. 2006.
[4] J. B. Heng, A. Aksimentiev, C. Ho, P. Marks, Y. V. Grinkova, S. Sligar, K. Schulten, and G. Timp, "Stretching DNA using the electric field in a synthetic nanopore," *Nano Lett.*, vol. 5, no. 10, pp. 1883–1888, Oct. 2005.
[5] M. Rhee and M. A. Burns, "Nanopore sequencing technology: Nanopore preparations," *Trends Biotechnol.*, vol. 25, no. 4, pp. 174–181, Apr. 2007.
[6] M. Rhee and M. A. Burns, "Nanopore sequencing technology: Research trends and applications," *Trends Biotechnol.*, vol. 24, no. 12, pp. 580–586, Dec. 2006.
[7] M. J. Kim, M. Wanunu, D. C. Bell, and A. Meller, "Rapid fabrication of uniformly sized nanopores and nanopore arrays for parallel DNA analysis," *Adv. Mater.*, vol. 18, no. 23, pp. 3149–3153, Dec. 4, 2006.
[8] M. A. Burns, B. N. Johnson, S. N. Brahmasandra, K. Handique, J. R. Webster, M. Krishnan, T. S. Sammarco, P. M. Man, D. Jones, D. Heldsinger, C. H. Mastrangelo, and D. T. Burke, "An integrated nanoliter DNA analysis device," *Science*, vol. 282, no. 5388, pp. 484–487, Oct. 16, 1998.
[9] H. Craighead, "Future lab-on-a-chip technologies for interrogating individual molecules," *Nature*, vol. 442, no. 7101, pp. 387–393, Jul. 27, 2006.
[10] A. J. Storm, J. H. Chen, X. S. Ling, H. W. Zandbergen, and C. Dekker, "Fabrication of solid-state nanopores with single-nanometre precision," *Nat. Mater.*, vol. 2, no. 8, pp. 537–540, Aug. 2003.
[11] D. Krapf, M. Y. Wu, R. M. M. Smeets, H. W. Zandbergen, C. Dekker, and S. G. Lemay, "Fabrication and characterization of nanopore-based electrodes with radii down to 2 nm," *Nano Lett.*, vol. 6, no. 1, pp. 105–109, Jan. 2006.
[12] W. M. Zhang, Y. G. Wang, J. Li, J. M. Xue, H. Ji, Q. Ouyang, J. Xu, and Y. Zhang, "Controllable shrinking and shaping of silicon nitride nanopores under electron irradiation," *Appl. Phys. Lett.*, vol. 90, no. 16, p. 163102, Apr. 16, 2007.
[13] M. Y. Wu, D. Krapf, M. Zandbergen, H. Zandbergen, and P. E. Batson, "Formation of nanopores in aSiN/SiO_2 membrane with an electron beam," *Appl. Phys. Lett.*, vol. 87, no. 11, p. 113106, Sep. 12, 2005.

[14] A. S. Abutaleb, "Automatic thresholding of gray-level pictures using two-dimensional entropy," *Comput. Vis. Graph. Image Process.*, vol. 47, no. 1, pp. 22–32, Jul. 1989.
[15] J. N. Kapur, P. K. Sahoo, and A. K. C. Wong, "A new method for gray-level picture thresholding using the entropy of the histogram," *Comput. Vis. Graph. Image Process.*, vol. 29, no. 3, pp. 273–285, Mar. 1985.
[16] J. H. Lin, "Divergence measures based on the Shannon entropy," *IEEE Trans. Inf. Theory*, vol. 37, no. 1, pp. 145–151, Jan. 1991.
[17] A. P. S. Pharwaha and B. Singh, "Shannon and non-Shannon measures of entropy for statistical texture feature extraction in digitized mammograms," in *Proc. WCECS*, 2009, vol. I/II, pp. 1286–1291.
[18] J. C. Russ, *The Image Processing Handbook,* 5th ed. Boca Raton, FL: CRC, 2007.
[19] M. Sezgin and B. Sankur, "Survey over image thresholding techniques and quantitative performance evaluation," *J. Electron. Imaging*, vol. 13, no. 1, pp. 146–168, Jan. 2004.
[20] R. C. Gonzalez, R. E. Woods, and S. L. Eddins, *Digital Image Processing Using MATLAB*. Englewood Cliffs, NJ: Prentice-Hall, 2004.
[21] T. J. Fellers, K. M. Vogt, and M. W. Davidson, *Nikon MicroscopyU, CCD Signal-To-Noise Ratio*, 2000. [Online]. Available: http://www.microscopyu.com/tutorials/java/digitalimaging/signaltonoise/index.html