https://doi.org/10.2352/ISSN.2470-1173.2021.16.COLOR-341

Thin-film Nitrate Sensor Performance Prediction Based on Pre-processed Sensor Images[1]

Xihui Wang, Kerry Maize, Ye Mi, Ali Shakouri, George T.C. Chiu, Jan P. Allebach
Purdue University, West Lafayette, Indiana, United States

Abstract

Automating the assessment of sensor quality in the production of thin-film nitrate sensors can yield significant advantages. Currently, the inspection process is extremely time and labor intensive, requiring technicians to manually examine sensors from each batch to determine their performance. Not only is manually examining sensors costly, it also takes days to conclude the results. It is possible to utilize image based learning approach to entirely automate the quality assessment process by accurately predicting the performance of every sensor; this allows for instant performance analysis and rapid changes to the fabrication parameters.

The fabrication parameters will directly control the thickness of the ion-selective membrane (ISM) of the nitrate sensor. The thickness of the ISM directly affects the texture on the sensor's surface. Because of the reliable correlation between sensor performance and sensor surface texture, it allows us to use learning methods to predict sensor performance through images instead of direct measurements.

We propose a method to predict sensor quality using non-contact sensor images through a series of image processing techniques followed by machine and deep learning.

Introduction

The Scalable Manufacturing of Aware and Responsive Thin Films (SMART) [1] consortium concentrates on manufacturing affordable Internet of Things (IoT) sensors that can be produced at mass volumes and deployed over large areas. One of the major challenges is to efficiently and economically monitor the sensor quality during the fabrication process with a roll-to-roll (R2R) system in real-time. The study of nitrate sensor application has indicated the correlation between sensor performance and the non-uniform coating of the ISM determined by the process control parameters [2]. Such thickness variations of the ISM will create visible texture on the sensor surface. Giving those pieces of evidence, we developed the sensor performance prediction system based on non-contact sensor images shown in Figure 1.

The R2R sensor manufacturing system is used to fabricate our nitrate sensor by printing the electrode on a polyethylene terephthalate (PET) substrate and coating the electrode with the ISM and the silicon passivation layer as demonstrated in Figure 2.

The active region of the nitrate sensor is the electrode region coated with the ISM; and it is the region that draws the most attention. The non-contact sensor images fed into the prediction system are the sensor active region images captured using an Electro-Optical System (EOS)[2] camera with a microscope. During the sensor performance measurement procedure, the sensor active region will be immersed in the nitrate solution, while the non-coated electrode region will be connected with the NI[3] Module to record the sensor performance data. The idea that the non-uniform ISM coating will significantly impact sensor performance and simultaneously cause variations in sensor surface appearance provides mathematical confidence for the prediction system to associate the sensor performance data with the extracted texture features from sensor images.

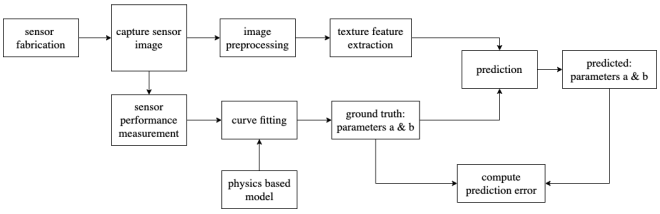


Figure 1. *Overview of Process Pipeline.*

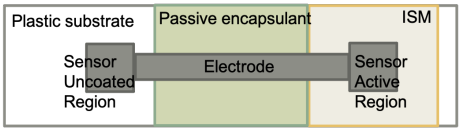


Figure 2. *Shape of the Nitrate Sensor.*

Prior research has shown that the image-based prediction system can predict the overall potentiometric response of the sensor given sensor active region images [3]. Both machine learning and deep learning approaches have been considered when designing the prediction system. The logarithmic function was proposed based on a physics model to represent the sensor performance. The local binary pattern (LBP)[4] visual descriptor and pre-trained convolutional neural network (CNN) were used to extract texture features from the sensor images. Manufacturing factors are also fused into the system along with image features.

This paper will continue to expand the image-based prediction system by focusing on preprocessing the sensor active region images to achieve better accuracy on the predicted sensor performance curve. A template matching [5] method is implemented to segment the sensor active region from the non-contact image in the image data preparation step. The contrast limited adaptive histogram equalization (CLAHE) [6] technique is applied to enhance texture contrast in the sensor active region images. The Gaussian

[1]Research supported by the SMART films consortium (https://www.purdue.edu/discoverypark/birck/research/smart-films.php).

[2]Electro-Optical System Inc, Phoenixville, PA 19460.
[3]NI (formerly National Instruments), Austin, TX 78759-3504)

pyramid method [7] is investigated as a multiscale approach to extract texture features from sensor images.

Dataset Preparation

Sensor active region images and their ground truth data are required for the image-based prediction system. Before we train the prediction model, we need to generate them separately.

Image Data Preparation

To ensure the consistency of the experiment, we use the same equipment and following the same procedure to capture the non-contact sensor images as discussed in previous work [3]. As mentioned before, the sensor active region's texture appearance is related to the varying sensor performance. Therefore, we need to crop the sensor active region out of the original non-contact sensor image to avoid distraction to our prediction system. With the increasing amount of sensors fabricated under varying settings, separating the sensor active region from its background could be challenging. In this case, we proposed a more efficient and stable way to segment the sensor active region using the template matching method [5] as shown in Figure 3.

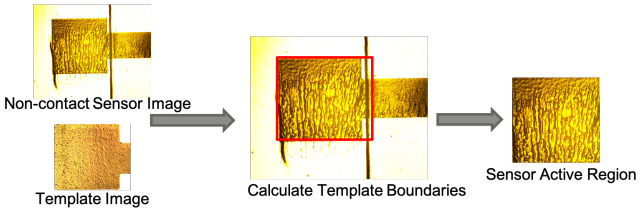


Figure 3. *Segment Sensor Active Region Using Template Matching Method.*

The template matching technique is intended for inspecting the source image and locating the area that best matches the object presented in the template image by minimizing the mean-squared error or maximizing the area correlation. In our case, the template matching algorithm gives us the best performance when we transfer the colored sensor image into the R-channel grayscale image and use the correlation coefficient method [8].

Ground Truth Data Preparation

We expect our image-based prediction system to predict the overall potentiometric response of the sensor. Therefore, the ground truth data should be the parameters that represent the entire sensor performance data. The physics-based model provides a logarithmic function representing the sensor performance signal, which simplifies the ground truth data into two parameters. Such parameters are the named performance parameters.

Figure 4.(a) shows the experimental setup for measuring the nitrate sensor performance [9]. The working electrode (WE) potential depends on the nitrate ion concentration, and the ISM ensures that only the nitrate ions impact the WE potential. The reference electrode (RE) provides a stable reference electrochemical potential via the solid electrolyte coating. The sensor performance data is the potential difference between the WE and the RE.

Figure 4.(b) gives an example of the sensor performance curve for one sensor set measured in a 0.001 molar nitrate solutions for 22 hours. After around 4.5 hours, the potentiometric

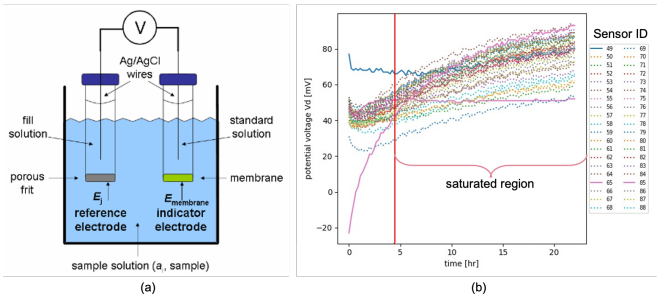


Figure 4. *(a) Sensor Voltage Measurement Experiment Setup; (b) Example of Measured Sensor Performance Signal.*

response will achieve its saturated phase; and this is the phase where we apply the physics-based model. It is worth mentioning that the solid line signals are the outliers caused by experimental error and will be eliminated when training the prediction system.

The physics-based model suggests that the change of potential voltage over time is a logarithmic growth [3]. Therefore, we can fit the saturated region of the sensor performance curve to the following equation. The parameters a and b are the performance parameters that represent the sensor performance curve after saturation.

$$V_{fit}(t) = a \cdot log(t) + b \tag{1}$$

The procedure to fit our saturated sensor performance curve to the logarithmic function and generate the ground truth parameters is shown in Figure 5.

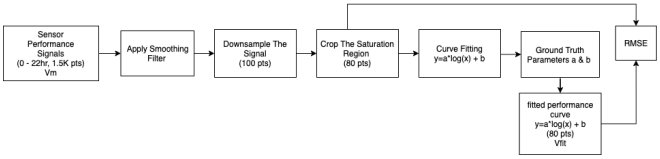


Figure 5. *Ground Truth Data Preparation Procedure.*

We denote the measured sensor performance signal as V_m and apply a smoothing filter on V_m. The smoothing filter we use here is the 5th order Savitzky-Golay filter [10]; and the filter window length is equal to 100 data points. We then downsample the smoothed signal from around 1.5k data points to 100 data points and denote the downsampled signal as V_d. The saturated region of V_d is the last 80 data points. After that, we use the Levenberg-Marquardt algorithm [11] to find the best fitting logarithmic curve for the saturated region of V_d. The fitted logarithmic curve is denoted as V_{fit}.

We use the root-mean-square error (RMSE) to evaluate the accuracy of the fitted curve. Figure 6 illustrates the difference between the original measured sensor performance signal V_m, the downsampled and smoothed signal V_d, and the fitted curve V_{fit} in the saturated region.

$$RMSE_{CF}(mV) = \sqrt{\frac{1}{N}\sum_{x}\left(V_{fit}(x) - V_d(x)\right)^2} \tag{2}$$

$$RMSE_{CF}(\%) = \sqrt{\frac{1}{N}\sum_{x}\left(\frac{V_{fit}(x) - V_d(x)}{V_d(x)}\right)^2} \times 100\% \tag{3}$$

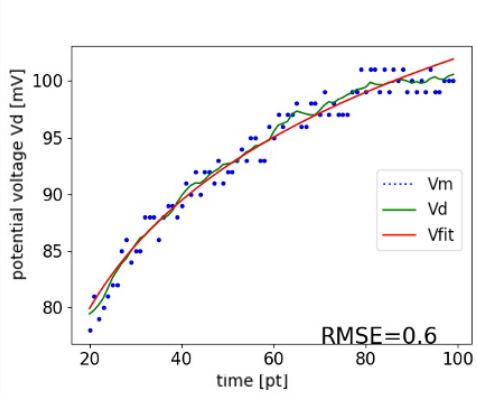


Figure 6. *Example of Fitted Logarithmic Curve: V_m is the measured performance signal, V_d is the downsampled & smoothed signal, V_{fit} is the fitted logarithmic curve.*

Here, CF stands for curve fitting process, and N is the total number of time points.

The dataset we generated contains 108 sensors. The performance data of those sensors was measured in a 0.001 molar nitrate solution for 22 hours. The average RMSE for the curve fitting process is around 1.2980 mV or 1.5231%. The result indicates that the idea of using two performance parameters as ground truth data to represent the saturated region of the sensor's potentiometric response is reliable.

Image Preprocessing

The connection between the texture feature of the sensor active region image and sensor performance data is the cornerstone for the image-based prediction system. Therefore, we propose an approach using the contrast limited adaptive histogram equalization (CLAHE) [6] method to improve the visibility level of the texture features of the active region image.

CLAHE is a variant of adaptive histogram equalization (AHE) [12], which improves local contrast, enhances the edges in each region of the image, and prevents overamplification of the noise in the meantime. The RGB color space of the active region sensor image is nonlinear since gamma correction is applied when capturing the sensor image. Hence we need to degamma correct the image and then apply CLAHE on the linear color space. Experiments show that the CLAHE works best on the L* channel. Two parameters are required for the CLAHE method. *ClipLimit* sets the threshold for contrast limiting, and *tileGridSize* represents the number of tiles in the row and column. Here, we set the *ClipLimit* to be 3 and *tileGridSize* to be 8×8. After the enhancement, we apply gamma correction to the enhanced image for display.

Figure 7 demonstrates the changes of the sensor active region image during the preprocessing step. The input image is the leftmost one, and the output preprocessed image is the rightmost one. As shown in the example, the texture becomes more noticeable in the preprocessed image.

Texture Feature Extraction

As mentioned previously, the non-uniform coating of ISM during the sensor fabrication process causes visual differences in

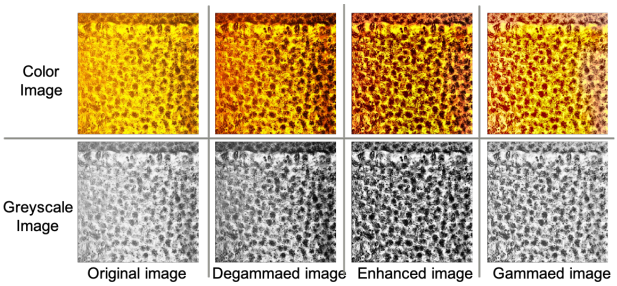


Figure 7. *Example of Sensor Image During The Preprocessing Procedure: top row is the color image; bottom row is the L* channel monochrome image.*

the sensor active region image. It is necessary to extract meaningful features from the active region sensor image that describe the texture properties. This section will focus on the local binary patterns (LBP) [4] method, and the combinational method of LBP and the Gaussian pyramid [7] method.

LBP, a powerful texture operator, plays a vital role in the study of pattern classification in computer vision. Various methods have been developed since the default method [4] of LBP was first proposed. This paper will focus on the application of the uniform method [13] and the nri_uniform (Non-rotation invariant uniform) method [14] of LBP. The uniform method of LBP is grayscale and rotation invariant for uniform patterns, while the nri_uniform method is only grayscale invariant. The pattern is called uniform if the binary array contains at most two bitwise transitions from 0 to 1, or vice versa.

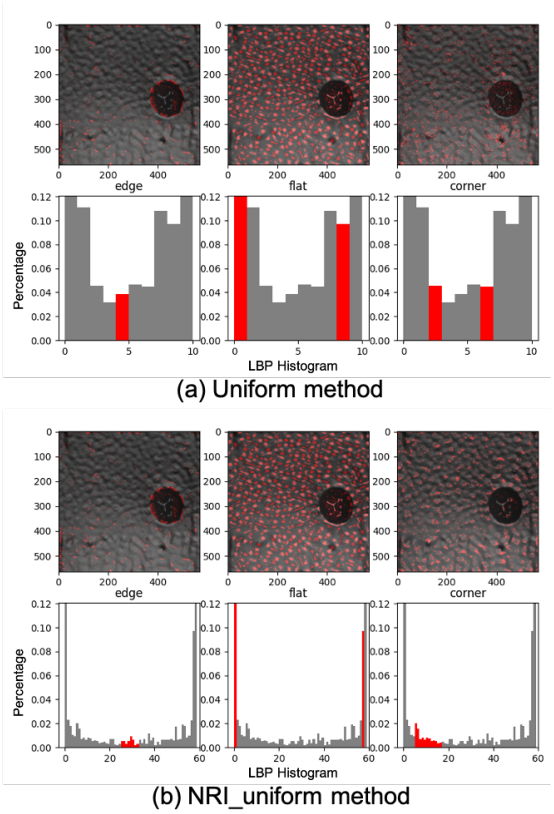


Figure 8. *Example of LBP Histogram with $P = 8$ and $R = 15$.*

Two parameters are essential for generating the LBP of an image. P represents the number of circularly symmetric neighbor points, and R defines the radius of the neighbor circle around the target pixel. With the same parameter settings, the generated LBP histograms are entirely different for the uniform method and the nri_uniform method as is shown in Figure 8.

The Gaussian pyramid method [7] is often used as a multi-scale image processing technique. The idea is to apply a Gaussian filter on the image and then downsample the image, so that the resolution for each layer will be 1/4 of that of the previous layer. In our case, the Gaussian pyramid contains three layers (layer0, layer1, and layer2). And we denote the original sensor active region image as layer0. The image size for each layer will be 555 $\times$ 555 pixels, 278 $\times$ 278 pixels, and 139 $\times$ 139 pixels, respectively.

The combinational method applies the LBP method on each layer of the Gaussian pyramid to extract texture features over different scales. Figure 9 shows an example of applying the LBP method on the Gaussian pyramid with the same parameter setting for the LBP at each scale.

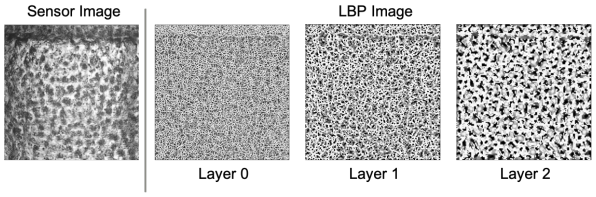


Figure 9. *Example of LBP Method Applied with Gaussian Pyramid: $P = 24$ and $R = 3$ for LBP in each layer.*

Prediction Models

The prediction system is constructed to predict sensor performance based on the active region images. As mentioned, the performance parameters a and b can represent the potentiometric response in the saturated region. Following previous work [3], the support vector regression (SVR) [15] model and a CNN-based (Convolutional Neural Network) regression model are selected to be the prediction models. The system takes the generated performance parameters and the texture features extracted from the sensor image as input during the training process for the SVR model. To test the accuracy of the prediction model, the system takes the feature vector as input and outputs the predicted performance parameters during the testing process. The structure of the SVR prediction system is shown in Figure 10.

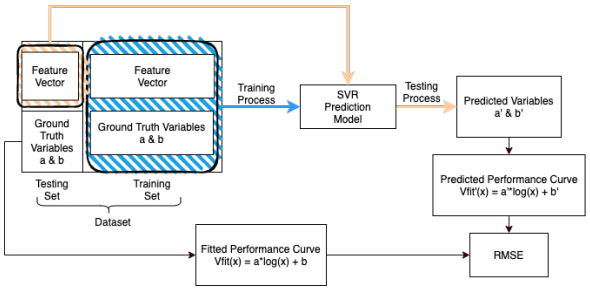


Figure 10. *Overview of The SVR Prediction System.*

The SVR model will find an appropriate hyperplane in higher dimensions to fit the input data by setting the proper hyper-parameters. The radial basis function (RBF) [16] kernel is used in the SVR model because of the non-linear relationship between the feature vector and the performance parameters.

For the CNN-based regression model, the input for the system is the sensor image instead of texture features. The structure of the CNN prediction system is shown in Figure 11. The deeper network can learn more complex features from the image in the convolutional layers, but gradients would become infinitely large or zero and fail the training if the network contains too many layers. The residual network provides an idea to overcome the vanishing gradient problem by using skip connections. Hence, the architecture of ResNet-34 [17] is selected for the CNN-based regression model. Two modifications are made here to fit our quest. The number of neurons in the fully connected output layer is adjusted to be two. The loss function is replaced by the $L2$ loss between the predicted performance curve and the fitted performance curve as shown in Figure 11.

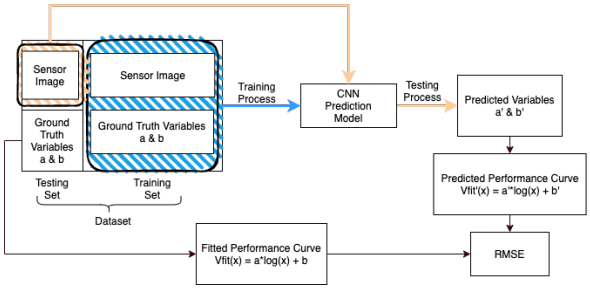


Figure 11. *Overview of The CNN Prediction System.*

Experiment Results

The dataset used in this experiment contains 108 sensors, as mentioned in the dataset preparation section. To get a reliable estimate of the system performance, we follow the 5-fold cross validation procedure to train and evaluate our system. The number of sensors in each fold is 22, 22, 22, 21, 21. The system will train on four folds and evaluate the remaining one fold each time. When all folds have been evaluated exactly once, we take the average performance across all five folds as the system performance.

The prediction models with which we have experimented are shown in Table 1.

Table 1. Prediction Methods Implemented in the Image-based Prediction System

Method	Description
M1	OI + LBP(uniform) + MF + SVR
M2	OI + LBP(nri_uniform) + MF + SVR
M3	EI + GP + LBP(uniform) + MF + SVR
M4	EI + GP + LBP(nri_uniform) + MF + SVR
M5	EI + Pre-trained CNN + MF + SVR
M6	OI + Trained CNN

To examine the effect of image preprocessing step, texture features will be extracted from the original active region image and the preprocessed sensor image. We denote the enhanced sen-

sor active region image as EI, and the original sensor image as OI.

In addition to the texture features, the manufacturing factors (MF) are added to the feature vector as input to the prediction system based on the SVR model. The manufacturing factors include the average measured sensor thickness data and three process control parameters, which are solid content, line speed, and flow rate. Each manufacturing factor is a floating-point number, and will be normalized to the range [0, 1].

The uniform method LBP generates a 10-element 1D feature array by setting $P = 8$ and $R = 3$. The feature array is then normalized such that the sum of the elements in the array will be one. The nri_uniform method LBP generates a 58-element 1D feature array under the same setting. The Gaussian pyramid (GP) contains three layers. Hence, applying the LBP method on each GP layer generates a 1D feature array three times longer.

Another approach is to extract the feature vector learned from the pre-trained CNN model. The same architecture ResNet-34 is used here, as we discussed in the previous section. The feature vector will be a 512-element 1D array that outputs from the last average pooling layer.

The RMSE is used to evaluate the accuracy of the predicted sensor performance for each fold. The average RMSE and the standard deviation of RMSE are used to estimate the performance of the image-based prediction system.

$$RMSE_{predict}(mV) = \sqrt{\frac{1}{N}\sum_{x}\left(V_{fit}(x) - V'_{fit}(x)\right)^2} \quad (4)$$

$$RMSE_{predict}(\%) = \sqrt{\frac{1}{N}\sum_{x}\left(\frac{V_{fit}(x) - V'_{fit}(x)}{V_{fit}(x)}\right)^2} \times 100\% \quad (5)$$

Table 2. Prediction Results (five-fold cross-validation)

Method	RMSE (mV)	RMSE (%)	StDev (mV)	StDev (%)
M1	6.00	8.24	1.31	2.59
M2	5.91	8.06	1.49	2.67
M3	**5.69**	**7.75**	**0.74**	**1.45**
M4	5.87	7.98	1.53	2.79
M5	5.81	8.12	1.62	3.01
M6	6.22	9.15	0.87	1.68

The accuracy and the robustness of the image-based prediction system can be described by the average RMSE and the standard deviation shown in Table 2. M1 and M5 are the prediction models used in the previous work [3]. The texture features extracted using different LBP methods alone and then fused with MF cannot make a noticeable difference by comparing M1 and M2. Texture features extracted using the combinational method with multiresolution features help improve the accuracy of the prediction system by comparing M1 with M3 and M2 with M4. The results show that the preprocessed sensor image and the Gaussian pyramid method improves the performance of the system. M3 achieves the best performance among all six models, which means applying LBP method on the Gaussian pyramid of the preprocessed sensor image will help improve the performance of the prediction system. The result also verifies the correctness of the physics-based model.

Conclusion

Image based learning methods are able to accurately predict sensor quality with RMSE of 7.75 - 9.15% depending on the process and model. To monitor the sensor quality during the fabrication process with a R2R system in real-time, the image-based prediction system is developed to accurately predict the potentiometric response of the nitrate sensor given preprocessed sensor active region images. A novel way of segmenting the active region from the non-contact sensor image is introduced to prepare the image dataset for the prediction system. The active region sensor images will be preprocessed before being fed into the prediction system to enhance the texture features that appear on the sensor surface. The physics-based model suggests a logarithmic relationship between time and the potentiometric response in the saturated phase, which helps us generate the ground truth dataset. The LBP descriptor, the Gaussian pyramid method, and the pre-trained CNN model are used to extract texture features from the preprocessed active region images. The feature vector, one of the inputs to train the SVR based prediction system, is generated by appending the extracted image feature with the normalized manufacturing factors. Both machine learning and deep learning approaches are able to achieve highly accurate predictions of sensor quality.

References

[1] SMART Films Consortium, Brick Nanotechnology Center, Purdue University, West Lafayette, IN. [Online]. Available: https://engineering.purdue.edu/SMART

[2] J. Hu, A. Stein, and P. Bühlmann, "Rational design of all-solid-state ion-selective electrodes and reference electrodes," TrAC Trends in Analytical Chemistry, vol. 76, pp. 102–114, 2016.

[3] Q. Yang, K. Maize, X. Jin, H. Jiang, M. A. Alam, R. Rahimi, G. Chiu, A. Shakouri, and J. P. Allebach, "Predicting Response of Printed Potentiometric Nitrate Sensors Using Image based Machine Learning," Printing for Fabrication, 2020.

[4] T. Ojala, M. Pietikainen, and D. Harwood, "A comparative study of texture measures with classification based on featured distributions," Pattern Recognition, vol. 29, no. 1, pp. 51-59, 1996.

[5] R. Brunelli, Template matching techniques in computer vision: theory and practice. John Wiley & Sons, 2009.

[6] A. Reza, "Realization of the contrast limited adaptive histogram equalization (CLAHE) for real-time image enhancement," Journal of VLSI Signal Processing Systems for Signal, Image and Video Technology, vol. 38, no. 1, pp. 35-44, 2004.

[7] E. H. Adelson, C. H. Anderson, J. R. Bergen, P. J. Burt, J. M. Ogden, "Pyramid methods in image processing," RCA Engineer, vol. 29, no. 6, pp. 33-41, 1984.

[8] J. P. Lewis, "Fast Normalized Cross-Correlation," Vision Interface, pp. 120-123, 1995.

[9] S. Sedaghat, S. Jeong, A. Zareei, S. Peana, N. Glassmaker, and R. Rahimi, "Development of a nickel oxide/oxyhydroxide- modified printed carbon electrode as an all solid-state sensor for potentiometric phosphate detection," New Journal of Chemistry, vol. 43, no. 47, pp. 18619–18628, 2019.

[10] A. Savitzky, M. J. Golay, "Smoothing and differentiation of data by

simplified least squares procedures," Analytical Chemistry, vol. 36, no. 8, pp. 1627-1639, 1964.
[11] J. J. Moré, "The Levenberg-Marquardt algorithm: implementation and theory," Numerical Analysis. Springer, Berlin, Heidelberg, pp. 105-116, 1978
[12] R. Szeliski, Computer Vision: Algorithms and Applications. Springer Science & Business Media, 2010.
[13] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 7, pp. 971-987, 2002.
[14] T. Ahonen, A. Hadid, and M. Pietikäinen, "Face recognition with local binary patterns," European Conference on Computer Vision, Springer, Berlin, Heidelberg, 2004.
[15] A. J. Smola, and B. Schölkopf, "A tutorial on support vector regression," Statistics and Computing, vol. 14, no. 3, pp. 199-222, 2004.
[16] M. T. Musavi, W. Ahmed, K. H. Chan, K. B. Faris, and D. M. Hummels, "On the training of radial basis function classifiers." Neural Networks 5, no. 4, pp. 595-603, 1992.
[17] K. He, X. Zhang, S. Ren, J. Sun, "Deep residual learning for image recognition," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770-778, 2016.

Author Biography

Xihui Wang *received her B.S. (2016) and M.S. (2019) in Electrical Engineering from Purdue University and is currently a Ph.D. candidate in Purdue ECE. Her research focuses on image processing, computer vision, and machine learning.*

Kerry Maize *received his B.S. in Electrical Engineering and Computer Science from UC Berkeley in 2002. He came to UCSC in 2005 to pursue graduate study in the areas of quantum electronics and nanoscience. His current work is focused on optical coherence tomography and thermal device characterization. Prior to engineering, Kerry worked in journalism.*

Ye Mi *received his Ph.D. in Nuclear Engineering from Purdue University in 1998. He received both his M.S. and B.S. from Tsinghua University in China, majoring in Nuclear Physics and Nuclear Instrumentation respectively. In recent years, he has been focusing on R&D of industrial processing tomography. Currently, he is working on inline metrology of R2R systems.*

Ali Shakouri *is an electrical and computer engineering professor and director of the Birck Nanotechnology Center at Purdue University. He received his PhD from Caltech. His research focuses on quantum electronics, mutual interaction of heat, light, and electricity in nanomaterials and devices, lock-in imaging, and advanced image processing with applications to nanoscale thermal measurements and roll-to-roll process monitoring. He leads a team to manufacture low-cost smart internet of thing (IoT) devices and sensor network for applications in advanced manufacturing and agriculture.*

George T.C. Chiu *is a Professor with the School of Mechanical Engineering with courtesy appointmentsin the School of Electrical and Computer Engineering and the Department of Psychological Sciences at Purdue University. He worked on designing printers and multifunction devices with Hewlett-Packard, Palo Alto, CA, USA. From 2011 to 2014, he served as the Program Director for the Control Systems Program with the National Science Foundation while on leave from Purdue University, West Lafayette, IN, USA. Chiu is a fellow of ASME and a fellow of the Society for Imaging Science and Technology (IS&T) and IEEE senior member.*

Jan P. Allebach *is a Hewlett-Packard Distinguished Professor of electrical and computer engineering at Purdue University. He is a fellow of the National Academy of Inventors, IEEE, the Society for Imaging Science and Technology (IS&T), and SPIE. He was named the Electronic Imaging Scientist of the Year by IS&T and SPIE, and was named the Honorary Member of IS&T, the highest award that IS&T bestows. He has received the IEEE Daniel E. Noble Award, the IS&T/OSA Edwin Land Medal, the IS&T Johann Gutenberg Prize, and is a member of the National Academy of Engineering.*

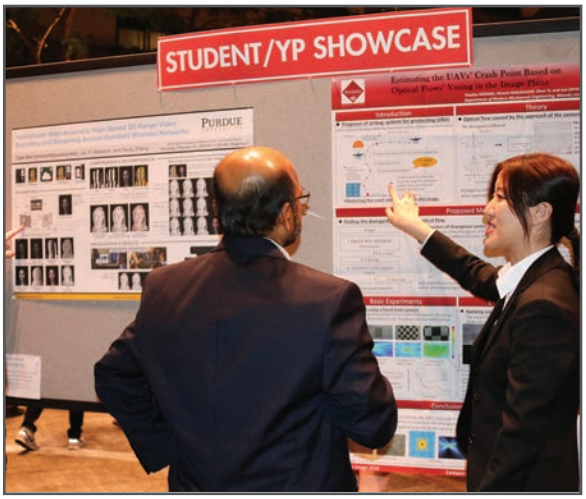
JOIN US AT THE NEXT EI!

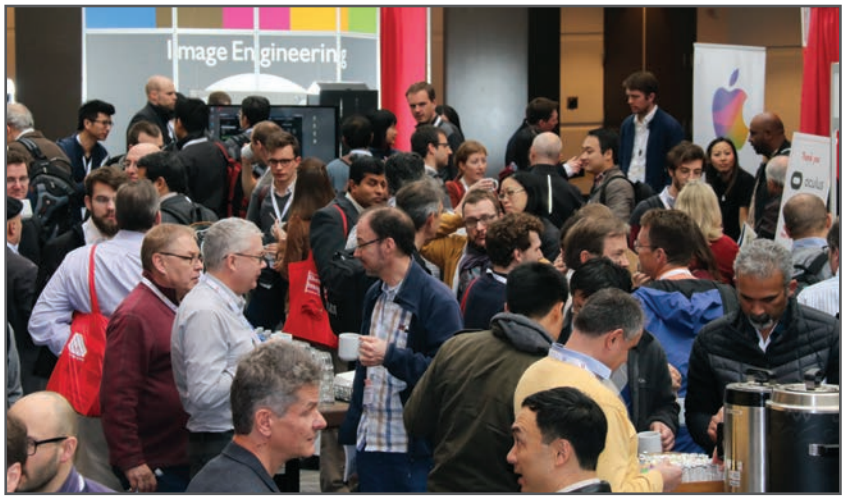
IS&T International Symposium on

Electronic Imaging

SCIENCE AND TECHNOLOGY

Imaging across applications . . . Where industry and academia meet!





• SHORT COURSES • EXHIBITS • DEMONSTRATION SESSION • PLENARY TALKS •
• INTERACTIVE PAPER SESSION • SPECIAL EVENTS • TECHNICAL SESSIONS •

www.electronicimaging.org

IS&T
imaging.org