DTU Library

Anomaly Detection in Float-Zone Crystal Growth of Silicon

Chen, Tingting; Tosello, Guido; Werner, Nico; Calaon, Matteo

Published in:
Procedia CIRP

Link to article, DOI:
10.1016/j.procir.2022.05.184

Publication date:
2022

Document Version
Publisher's PDF, also known as Version of record

Available online at www.sciencedirect.com

ScienceDirect

Procedia CIRP 107 (2022) 1515–1519

www.elsevier.com/locate/procedia

55th CIRP Conference on Manufacturing Systems

Anomaly Detection in Float-Zone Crystal Growth of Silicon

Tingting Chen[a], Guido Tosello[a], Nico Werner[b], Matteo Calaon[a]

[a]*Technical University of Denmark, Produktionstorvet, 2800 Kgs. Lyngby, Denmark*
[b]*Topsil GlobalWafers, Siliciumvej 1, 3600 Frederikssund, Denmark*

* Tingting Chen. Tel.: +45-5028-1068. *E-mail address:* tchen@mek.dtu.dk

Abstract

Float-Zone (FZ) crystal growth can manufacture silicon crystal with high purity and low oxygen concentration. However, due to the limited capability of FZ machines, oxidation contamination cannot be avoided, and an oxide layer sometimes would form on unmelted polycrystalline silicon surface. Oxide layers, so-called watermark, would negatively affect the crystal quality. Consequently, it is desirable to have an automated watermark detection in order to take corrective action in the early stages of the process. This paper proposed a complete framework for watermark detection on FZ images, including pre-processing, feature selection and classification. The results show it can act as a promising quality assurance tool for the FZ process.

Keywords: Type your keywords here, separated by semicolons ;

1. Introduction

With the evolution of electronics and optoelectronics, the demand for ultra-high-quality silicon wafer is continuously rising. As the first step in wafer fabrication, crystal growth of silicon crystal is naturally required to meet higher requirement on quality, cost and productivity as well. There are generally two approaches for the single crystal growth, including the Czochralski (CZ) and the Float-zone (FZ) methods[1– 2]. At present, CZ method is predominantly used in the production of single crystals due to high production efficiency and capability of producing large monocrystalline silicon from the melt in quartz crucible. However, silicon produced by CZ method does not fit the needs for some important electronic devices due to the oxygen impurities from the crucible[3]. Compared with CZ method, FZ can manufacture silicon crystal with higher purity and higher resistivity since it is crucible-free method, thus avoiding the contamination from crucible[4].

FZ has the potential of producing ultra-high-quality crystal with low impurities in particular low oxygen concentration. However, due to limited capability of FZ machines and residual impurities present within the machine, oxide layer sometimes would form on unmelted polycrystalline silicon surface, leaving an appearance of a shadow, so-called watermark. Watermark as oxygen contamination would negatively affect crystal quality. Consequently, it is desirable to implement an automatic detection for watermark in order to take corrective reaction in the early stage.

In this paper, the cone phase of the FZ process is selected to investigate the watermark development as it represents the process stage with the highest possibility of watermark formation. The camera system, which is integrated in the machine to measure geometrical quantities for FZ control system, is utilized to capture the image of polycrystalline surface when crystal grows into the specific diameter during cone phase. Figure 1 shows the comparison of normal image and abnormal image with watermark. We can see that in the normal process, the polycrystalline silicon surface has both homogeneous intensity and texture, with a smooth transition from darkness on top side to brightness on the melt area. When it comes to an abnormal image instead, an intensive gradient can be observed on the edge of the watermark.

Since normal images have similar characteristics, the watermark identification on polycrystalline surface can be turned into the anomaly detection on images by discriminating abnormal images from normal images. In general, the current

10.1016/j.procir.2022.05.184

research on anomaly detection on images can be categorized into reconstruction-based method[5-6] and classification-based method)[8-13]. In reconstruction-based method, only normal components of data are learned and reconstructed, then reconstruction error against original image can be used to indicate an anomaly score[7]. However, since it only learns from normal dataset, it is sensitive to samples which are out of training dataset and cannot guarantee the best discrimination ability[7].

In classification-based method, classification is made based on the features either handcrafted or automatically extracted. Handcrafted features are manually designed based on prior knowledge in the aspects of gradient[8], texture[9] and so on. However the description ability and representative ability of these handcrafted features are still limited when it comes to more complicated images[10]. Therefore, automatic feature extraction method is the preferable methodology in use. For instance, linear feature extraction methods like principle component analysis (PCA)[11]have been investigated for learning discriminative features. Nevertheless, these linear transformation techniques still limit the capability of capturing abstract features from images.

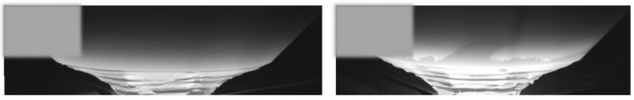

Fig. 1. (a) Normal process with homogeneous intensity on polycrystalline surface. (b) abnormal process with watermarks on polycrystalline surface.

Another powerful automatic feature extraction tool is Convolutional Neural Networks (CNN) [12], which is designed for extracting high-level features from input data with multiple array form, like image. Though CNN can perform both feature extraction and classification with high accuracy, the prediction of CNN from the last fully connected layer, which is less efficient[13] and cannot fully grasp the information of extracted features[14]. Therefore, some research combine CNN with other classifiers, for instance the combination of CNN and SVM[15], and the combination of CNN and XGBoost[13-14], in which CNN only acts as feature extractor. Some researches focus on improving the quality of inputs for CNN by replacing origin images with handcrafted features, for instance Local Binary Pattern(LBP)[16].

Motivated by extended CNN mentioned, in this paper, hybrid features combining the deep features and handcrafted features and hybrid classifiers are investigated. Our primary contributions of this work can be found in the following:

• A complete framework for anomaly detection on the floating zone crystal growth process, including prepro- cessing, feature selection and classification

• A novel classification approach based on XGBoost with the input of fusion features combining the deep features derived by CNN from Local Binary Pattern (LBP) and handcrafted features.

2. Methodology

The proposed framework of anomaly detection for float- zone crystal growth process can be seen in Figure 2. There are three main steps involved in the proposed approach: pre-processing, feature extraction and classification.

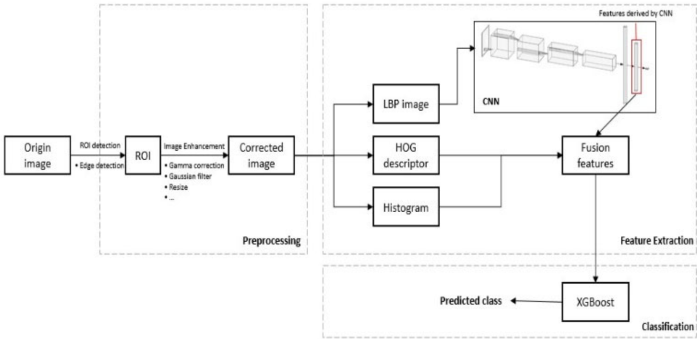

Fig. 2. Framework of anomaly detection in FZ process

2.1. Pre-processing

Since there are many disturbances in original images, for instance, varying background and varying size and intensity of polycrystalline silicon surface, which may affect the classification performance, pre-processing would be of great significance. In this case, pre-processing procedures include Region of interest (ROI) detection, image enhancement and resizing.

Since our images are from historical dataset collected for the past years, there are some text information on the left side. Therefore, we mainly focus on right side of images. In this case, since watermark normally happens around the melt edge on the polycrystalline silicon, ROI can be set based on area around the melt edge.

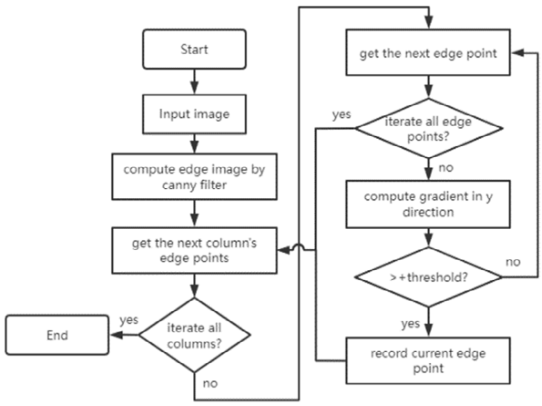

Fig. 3. The procedure of detecting melt edge.

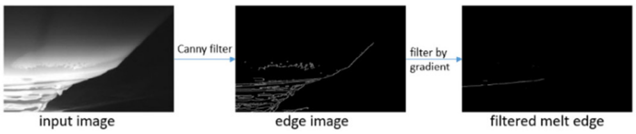

Fig. 4. Filtered melt edge

Melt edge is characterized by a sharp transition of pixel intensity. Beyond the melt edge, it is the surface of polycrystalline silicon with smooth intensity transition from darkness to brightness. Under the melt edge, there is solid residual of polycrystalline silicon containing multiples solid edges. To locate melt edge, canny filter is first applied to detect edges in the image. Considering watermark edges may also be detected by canny filter, detected edge points are filtered by predefined gradient magnitude and gradient direction in y direction. In this case, the wanted gradient direction is positive y direction(from darkness to brightness) because watermark edge points normally have positive y direction and they need to be filtered out. The procedure of detecting melt edge can be seen in Figure 3. Figure 4 presents the filtered melt edge.

Then ROI, is defined by cropping the image starting from the top centre point and ending with the point on the melt edge with fixed width.

Considering some light watermark with relatively high intensity is difficult to differentiate with the emitting polycrystalline silicon surface, gamma correction[17] is used to enhance image contrast. Gamma correction is a non-linear operation on pixel intensity, which aims to control the brightness of the image to suit human perception. The equation for gamma correction is formulated as follows:

$$I_{corrected} = I_{max} * \left(\frac{I}{I_{max}}\right)^{\gamma} \quad (1)$$

where I is the pixel intensity and Imax is the maximum intensity of the image, here the maximum value is 255. γ is a fixed value for controlling the contrast. When γ is higher than 1, the corrected image will become darker and pixels with low intensity will be compressed to lower degree. When γ is lower than 1, the corrected image will become brighter and pixels with low intensity will be enhanced to higher degree. In this case, we would like watermark can be emphasized with lower intensity, therefore γ is chosen to be higher than 1.

Then Gaussian filtering is applied to smooth the image and reduce the image noises before further feature extraction. Lastly the pre-processed image would be resized to the same shape for CNN inputs. Figure 5 shows an example of pre-processing images.

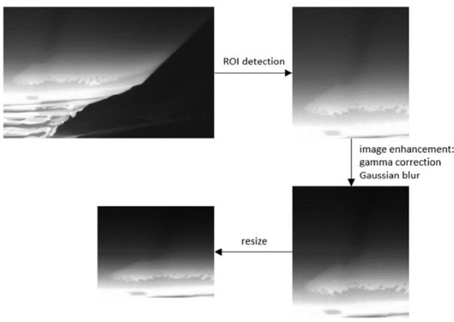

Fig. 5. An example of pre-processing image.

2.2. Feature extraction

The performance of classifiers is greatly dependent on the extracted features. Therefore features should be selected carefully. In this case, watermark normally has the appearance of black spots or shadow, which affects smooth texture with low gradient change shown in the origin image. Hence, our primary focus would be on gradient and texture. In this work, gray histogram, HOG descriptor, and LBP are utilized and extracted from the pre-processed image.

Gray histogram is the gray intensity distribution in the image, which is the invariance to the translation and rotation[10]. However, on the other hand, it leaves out the spatial information in the image, since it is unable to extract the correlation between pixels.

Histogram of oriented gradient(HOG) descriptor[18], is the feature by counting occurrences of gradient orientation in local area of the image. HOG descriptor presents the robustness to illumination changes and variance in translation and rotation.

LBP proposed by Ojala[19], is a typical type of texture descriptor, which can extract local texture information with high computational efficiency. LBP is computed at each pixel by comparing the intensity value with the neighbourhood and multiplying with the weights.

After the feature extraction, all features would be standardized to zero mean and unit variance using z-score. The normalized LBP image would be direct input for the CNN. CNN is trained to be feature extractor. After training phase, the in-features of output layer would be treated as deep features for further classification.

By merging deep features and gray histogram as well as HOG features into one vector, fusion features are established, which would be the input for next stage classification.

2.3. Classification

Even though compared with original data, the complexity of features is already reduced, there are still challenges for conventional single classifier to meet the requirements of identifying relevant features from noisy data, high speed and interpretability of results[20]. Confronted with this problem, we turned to ensemble learning based on decision trees to obtain better prediction. On one hand, decision trees can achieve high computational efficiency and produce interpretable model while performing internal feature selection. On the other hand, ensemble methods combines the outputs of many weak classifiers to produce a powerful ‘committee’[20], helping mitigate inaccuracy problems. In this case, Extreme Gradient Boosting (XGBoost)[21], a powerful gradient boosting decision tree methodology, is utilized. In this work, a binary XGBoost classifier is established for the fusion of features.

3. Experiments

The proposed methodology was evaluated on the FZ process. The images were captured in the cone phase when the crystal diameter reached the same predefined diameter. There are 425 normal images and 517 abnormal images recorded respectively, 70% of which would be used for training and the rest for testing. Then images would start going through pre-processing procedure in the sequence of ROI detection and image enhancement and end up with resized images with a resolution of (200,256). The predefined gradient threshold for obtaining melt edge was 80 and γ of 1.5 for gamma correction was used.

Then fusion features were obtained by applying proposed feature extraction method as shown in Figure 2. The number of orientation bins and the size of a cell and the number of cells in a block used in HOG descriptor are 8, 20 and 2 respectively. The radius and number of neighbors of LBP descriptor are 1 and 8 respectively. The architecture of CNN used for extracting deep features can be seen in Table 1. Afterwards, all features would be fused into one vector and standardized to zero mean and unit variance with Z-score method. Then XGBoost was utilized to classify features.

Table 1. Architecture of CNN used.

Layer	Output shape	Kernel, stride, padding
Input	(200,256,1)	
Conv2D	(200,256,16)	5,1,same
Batch Normalization	(200,256,16)	
ReLU	(200,256,16)	
Maxpooling	(100,128,16)	
Conv2D	(100,128,32)	5,1,same
Batch Normalization	(100,128,32)	
ReLU	(100,128,32)	
Maxpooling	(50,64,32)	
Conv2D	(50,64,64)	5,1,same
Batch Normalization	(50,64,64)	
ReLU	(50,64,64)	
Maxpooling	(25,32,64)	
Conv2D	(25,32,128)	5,1,same
Batch Normalization	(25,32,128)	
ReLU	(25,32,128)	
Maxpooling	(12,16,128)	
Flatten	(24576,)	
Dense	(512,)	
Dropout	(512,)	p=0.2
Dense	(1,)	

For comparison, several baseline models also conducted experiments. To investigate the effectiveness of LBP images for CNN, the same CNN model was trained and tested with pre-processed original images and LBP images, which were called IMG+CNN and LBP+CNN in this paper. Besides, to prove the efficiency of the handcrafted features used in the proposed method, we remove them in order and test them on the same test dataset. Furthermore, we also tested the fusion features according to Section 2.2 by SVM with radial basis function.

In this paper, four metrics including accuracy, precision, recall and F1-score are used for the evaluation, which are defined in the Equations 2-5. Accuracy is the ratio of correctly predicted samples among all data samples. However, when it comes to imbalanced dataset, Accuracy would be unreliable. Precision is the fraction of correctly predicted positives among all predicted positives. Recall indicates the sensitivity of the model to the anomaly. High Precision relates to low false alarm rate, while high Recall may lead to high false alarm. F1 Score is the harmonic mean of Precision and Recall, which can be a balanced indicator of performance of a model. Considering skewness of our dataset distribution is slight, we mainly focus on Accuracy and F1 Score. Each model was trained repeatedly for 10 times, and average and standard deviation of metrics were recorded.

Table 2. Confusion matrix.

	Predict Normal	Predict Normal
Actual Normal	True Negative(TN)	False Positive(FP)
Actual Abnormal	False Negative(FN)	True Positive(TN)

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (2)$$

$$Precision = \frac{TP}{TP + FP} \quad (3)$$

$$Recall = \frac{TP}{TP + FN} \quad (4)$$

$$\mathrm{F1-Score} = \frac{2 * Precision * Recall}{Precision + Recall} \quad (5)$$

Where TP, TN, FP, FN are true positive, true negative, false positive, false negative respectively, which can be seen in confusion matrix shown in Table 2.

Finally, the detailed results of different approaches can be seen in Table 3. As seen, the proposed method with fusion features and XGBoost classifier outperforms all baseline models in each evaluation metrics.

Besides, we can easily see that using extracted LBP image as the input for CNN can improve the performance compared with using the original pre-processed image, which indicates that watermark highlighted from LBP descriptor can help accelerate CNN to extract deep features with relatively small dataset. When replacing the CNN classifier with XGBoost on deep LBP features (DLBP-XGB), an increase on all metrics can be observed as expected. In the tests of the handcrafted features, it is worth mentioning that histogram features seem to have larger contribution to the proposed method according to the results of DLBP+HOG+XGB.

The performance of fusion features with SVM is lower than the proposed method. However, obvious increase in all metrics compared with CNN methods can still be observed, which verifies the effectiveness of fusion features.

Table 3. Performance of different approaches.

Model	Accuracy	Precision	Recall	F1 Score
IMG+CNN	90.63±2.3	92.89±3.1	92.89±2.92	91.75±1.99
LBP+CNN	93.25±1.69	94.34±4.98	93.74±4.28	93.85±1..42
DLBP+XGB	95.09±0.8	95.38±0.9	95.68±1.12	95.52±0.74
DLBP+HOG+ XGB	95.27±0.5	95.81±1.1	95.55±0.93	95.67±0.45
DLBP+HIST+ XGB	95.19±0.76	95.56±1.05	95.68±0.76	95.62±0.68
DLBP+HOG+ HIST+SVM	94.38±0.68	94.92±1.30	94.84±0.41	94.87±0.58
DLBP+HOG+ HIST+XGB (proposed)	**95.41±0.61**	**96.18±1.12**	**95.42±1.11**	**95.79±0.56**

Note: IMG is pre-processed image, DLBP means deep features from LBP image captured by CNN, XGB is XGBoost Classifier.

4. Conclusion

In this paper, a hybrid model with fusion features is proposed for anomaly detection of FZ process. This method utilizes CNN as deep feature extractor based on LBP images, and then derived deep features are fused with handcrafted features including HOG descriptor and histogram. Finally, XGBoost is applied for binary classification. Overall, the experimental results show that the proposed method obtains

95.41% accuracy and 95.79% F1-score on detecting watermarks for floating zone crystal growth process of silicon.

Nevertheless, despite the good classification performance shown in the proposed approach, there is still some room for improvements in the future. Firstly, in this case, only one image is captured for one run of the process and ROI is used for filtering out noise background, which can lead to high potential to miss detection when watermark lies on the back side or out of the ROI. This problem can be handled by online classification as polycrystalline silicon would rotate during the process. Hence, in the future, anomaly detection on videos would be needed. Finally, the watermark developed during the process is in fact a symptom of the problem, therefore further studies related to watermark formation will be performed by linking image detection of watermark with other data sources generated by machine during production.

Acknowledgements

This research work was undertaken in the context of DIGI-MAN4.0 project ("DIGItal MANufacturing Technologies for Zero-defect Industry 4.0 Production", https://www.digiman40.mek.dtu.dk/). DIGIMAN4.0 is a European Training Network supported by Horizon 2020, the EU Framework Programme for Research and Innovation (Project ID: 814225)

References

[1] W. Zulehner, Historical overview of silicon crystal pulling development, Materials Science and Engineering B: Solid-State Materials for Advanced Technology 73 (1) (2000) 7–15.

[2] Müller, G., Métois, J.J. and Rudolph, P. eds., 2004. Crystal Growth-From fundamentals to technology. Elsevier. (2004) 239–270.

[3] A. Muiznieks, J. Virbulis, A. Ludge, H. Riemann, N. Werner,¨ Floating Zone Growth of Silicon, second edi Edition, Vol. 2, Elsevier B.V., 2015.

[4] N. Werner, Analysis and Automation of the Crucible-free Floating Zone (FZ) Growth of Silicon Crystals, 2014. [5] Strunk Jr W, White EB. The elements of style. 3rd ed. New York: Macmillan; 1979.

[5] Q. Wei, B. Shi, J. Y. Lo, L. Carin, Y. Ren, R. Hou, Anomaly detection for medical images based on a one-class classification (Feb 2018) (2018)57.

[6] F. Zhang, B. Du, L. Zhang, Saliency-guided unsupervised feature learning for scene classification, IEEE Transactions on Geoscience and Remote Sensing 53 (4) (2015) 2175–2184

[7] T. Ehret, A. Davy, J.-m. M. M. Delbracio, Image Anomalies : a Review and Synthesis of Detection Methods, Journal of Mathematical Imaging and Vision 61 (5) (2017) 710–743..

[8] Chandrakala, M. and Devi, P.D., 2021. Two-stage classifier for face recognition using hog features. Materials Today: Proceedings, 47, pp.5771-5775.

[9] Liu, P., Guo, J.M., Chamnongthai, K. and Prasetyo, H., 2017. Fusion of color histogram and LBP-based features for texture image retrieval and classification. Information Sciences, 390, pp.95-111.

[10] G. Cheng, J. Han, S. Member, X. Lu, S. Member, Remote Sensing Image Scene Classification : Benchmark and State of the Art, in: Proceedings of the IEEE, IEEE, 2001, pp. 1865–1883.

[11] M. Singh, G. Kaur, SAR Image Classification Using PCA and Texture, in: International Conference on Advances in Information Technology and Mobile Communication, Springer, Berlin, 2011, pp. 435–439.

[12]Y. LeCun, K. Kavukcuoglu, C. Farabet, Convolutional networks and applications in vision, ISCAS 2010 - 2010 IEEE International Symposium on Circuits and Systems: Nano-Bio Circuit Fabrics and Systems (2010) 253–256.

[13] E. Bugingo, D. Zhang, A New Hybrid Convolutional Neural Network and eXtreme Gradient Boosting Classifier for Recognizing Handwritten Ethiopian Characters, IEEE Access 8 (2020) 17804–17818.

[14] W. Jiao, X. Hao, C. Qin, The Image Classification Method with CNNXGBoost Model Based on Adaptive Particle Swarm Optimization, Information 12 (4) (2021) 156.

[15] X.-x. Niu, C. Y. Suen, A novel hybrid CNN – SVM classifier for recognizing handwritten digits, Pattern Recognition 45 (4) (2012) 1318–1325.

[16] Touahri, R., AzizI, N., Hammami, N.E., Aldwairi, M. and Benaida, F., 2019, April. Automated breast tumor diagnosis using local binary patterns (LBP) based on deep learning classification. In 2019 International Conference on Computer and Information Sciences (ICCIS) (pp. 1-5). IEEE.

[17] Huang, S.C., Cheng, F.C. and Chiu, Y.S., 2012. Efficient contrast enhancement using adaptive gamma correction with weighting distribution. IEEE transactions on image processing, 22(3), pp.1032-1041.

[18] W. T. Freeman, M. Roth, Orientation Histograms for Hand Gesture Recognition, in: International workshop on automatic face and gesture recognition, 1994, pp. 296–301.

[19] T. Ojala, A comparative study of texture measures with classification based on featured distributions, Pattern Recognition 29 (l) (1996) 51–59.

[20] J. Friedman, T. Hastie, R. Tibshirani, The elements of statistical learning, Vol. 1.

[21] T. Chen, C. Guestrin, XGBoost : A Scalable Tree Boosting System, in: Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining, 2016, pp. 785–794.