View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Repository@USM

# DEFECT DETECTION AND CLASSIFICATION OF SILICON SOLAR WAFER FEATURING NIR IMAGING AND IMPROVED NIBLACK SEGMENTATION

**ZEINAB MAHDAVIPOUR**

**UNIVERSITI SAINS MALAYSIA**

**2016**

# DEFECT DETECTION AND CLASSIFICATION OF SILICON SOLAR WAFER FEATURING NIR IMAGING AND IMPROVED NIBLACK SEGMENTATION

**by**

**ZEINAB MAHDAVIPOUR**

**Thesis submitted in fulfillment of the requirements**
**for the degree of**
**Doctor of Philosophy**

**April 2016**

# ACKNOWLEDGEMENT

I wish to express my profound gratitude to God Almighty for His protection. I would like to appreciate my supervisor, Prof. Mohd Zaid Abdullah who has inspired and accompanied me throughout this study. His constant guidance, subjective criticism, continuous support, and great inspiration throughout the duration of my study are most valued. It is a great opportunity and honour to have worked under his supervision.

I would like to extend my heartiest appreciation to the Institute of Postgraduate Studies (IPS), Universiti Sains Malaysia for their contribution to this research through the Collaborative Research in Engineering, Science and Technology (CREST) grant304/PELECT/6050264/C121. I would also like to thank our industrial partner TT-Vision Technologies for partly sponsoring this project.

Special thanks go to the School Electrical and Electronic Engineering, USM for providing the necessary facilities and equipment. My sincere thanks go to all administrative and technical staff in the School, particularly Pn. Normala, En. Rahmat, Pn. Khalijah, En. Nor Azhar, En. Amir Hamid, and En. Naim. They offered instant assistance whenever needed. I would like to thank my colleagues and friends who provided encouragement and help.

Last but not least, I express my deepest respect and appreciation to my beloved husband and parents for their kindness, support, patience, and encouragement throughout my study. I will always be indebted to them. I sincerely thank you all for your understanding, trust, encouragement, and faith during the years of my study. Without their encouragement and love, this work could not have been completed.

# TABLE OF CONTENTS

Acknowledgement ………………………………………………………………. ii

Table of Contents……………………………………………………………… iii

List of Tables………………………………………………………………….. vi

List of Figures………………………………………………………………… vii

List of abbreviations…………………………………………………………. xi

List of Symbols……………………………………………..……………… xiii

Abstrak………………………………………………………………………… xv

Abstract……………………………………………………………………….. xvii

## CHAPTER 1- INTRODUCTION

1.1 Motivation ………………………………………………………... 1

1.2 Problem Statement ……………………………………………………… 4

1.3 Scope and Research Objectives ……………………………………..... 8

1.4 Thesis Outline …………………………………………………….... 9

## CHAPTER 2 - REVIEW OF LITERATURE

2.1 Introduction …………………………………………………………..... 10

2.2 Crack and Defect Detection Algorithms ……………………………... 10

2.3 Image Thresholding …………………..……………………………………. 16

2.3.1 Niblack Thresholding……………………………………………… 17

2.3.2 Segmentation Methods……..……………………………………….. 25

2.3.2.1 Sobel Technique……..…………………………………… 25

2.3.2.2 Canny Technique……..…………………………………….. 26

2.3.2.3 Otsu Technique……..……………………………………. 27

2.3.2.4 Anisotropic Diffusion Technique……..…………………. 28

2.4 Feature Extraction …………………………………………………….. 32

2.5 Machine Learning ………………………………………………………... 40

2.5.1 Support Vector Machines ……………………………………….. 43

2.5.2 Multi-class SVM ……………………………………………..……... 47

2.5.2.1 One-Versus-All (OVA) Method ……..…………………… 48

2.5.3 Unsupervised Classification……………………………………... 49

2.6 Summary ........................................................ 54

## CHAPTER 3 - METHODOLOGY

3.1 Introduction ........................................................ 55
3.2 Defect Inspection ........................................................ 56
  3.2.1 Sample Dataset ........................................................ 57
  3.2.2 Optical Setup........................................................ 57
3.3 Image Segmentation ........................................................ 66
  3.3.1 Adaptive Thresholding Methods ........................................................ 67
    3.3.1.1 Niblack's Segmentation ........................................................ 68
    3.3.1.2 Proposed Method 1........................................................ 68
    3.3.1.3 Proposed Method 2........................................................ 70
    3.3.1.4 Filtering ........................................................ 72
    3.3.1.5 Experimental Procedure ........................................................ 73
  3.3.2 Performance Evaluation ........................................................ 73
    3.3.2.1 Confusion Matrix ........................................................ 74
    3.3.2.2 Sensitivity and Specificity ........................................................ 75
    3.3.2.3 Pratt's Figure of Merit ........................................................ 77
3.4 Defect Feature ........................................................ 77
3.5 Classification ........................................................ 79
  3.5.1 Defect Classification ........................................................ 80
  3.5.2 Experimental Procedure ........................................................ 81
  3.5.3 Accuracy Measures of Clustering ........................................................ 85
    3.5.3.1 Rand Index ........................................................ 85
    3.5.3.2 Silhouette Index ........................................................ 85
3.6 Summary ........................................................ 86

## CHAPTER 4 - EXPERIMENTAL RESULTS

4.1 Introduction ........................................................ 87
4.2 Image Processing........................................................ 87
  4.2.1 Threshold Results 1 ........................................................ 88
    4.2.1.1 Quantitative Evaluation ........................................................ 96

4.2.2 Threshold Results 2………………………….…………………. 104
4.2.2.1 Quantitative Evaluation……………..……………………. 111
4.3 Classifiers Analysis …………………………………………………… 118
4.4 Summary ………………………………………………………….... 128

## CHAPTER 5 - CONCLUSION AND FUTURE WORK

5.1 Conclusion and Research Contributions ………………………………... 129
5.2 Limitation of the Techniques …………………………………………… 132
5.3 Suggestions for Future Work …………………………………………… 133

REFERENCES …………………………………………….…………….. 135

APPENDICES

LIST OF PUBLICATIONS

# LIST OF TABLES

**Page**

Table 3.1 The image datasets 57

Table 3.2 An instance of a confusion matrix with incorrectly and correctly classified categories for 10 samples, here, “⊞” signifies correctly classified and “⊟”signifies incorrectly classified 74

Table 3.3 Typical confusion matrix for a binary classifier or detection 75

Table 4.1 The *z* and *k* values computed from Figure 4.1 90

Table 4.2 The edge evaluation results averaged using 100 samples comparing several thresholding techniques with the proposed method 103

Table 4.3 The quantitative evaluation results calculated from (a) polycrystalline and (b) monocrystalline solar wafer samples. In each case the quality measures are averaged from 50 samples 113

Table 4.4 Total number of good and imperfect samples 123

Table 4.5 The evaluation result of *k*-mean clustering 123

Table 4.6 The SVM classification results of Dataset 2 128

# LIST OF FIGURES

**Page**

Figure 1.1 Examples of polycrystalline solar wafer images: (a) is an intact or good sample, (b) is a micro crack sample; (c) a defective sample that includes fingerprint, and (d) is defected by a Stain. In (b) and (d), the locations of micro-crack and stain are indicated by arrows 3

Figure 2.1 Binarization result: (a) Original image, (b) Niblack's result (Chiou *et al.*, 2012) 14

Figure 2.2 Binarization result: (a) Original image, (b) Niblack's result (Feng & Tan, 2004) 18

Figure 2.3 Binarization result: (a) Original image, (b) Sauvola & Pietikainen's result (Feng & Tan, 2004) 19

Figure 2.4 Binarization result: (a) Original image, (b) Wolf and Jolion's result (Feng & Tan, 2004) 20

Figure 2.5 Binarization result: (a) Original image, (b) Feng and Tan's result (Feng & Tan, 2004) 21

Figure 2.6 Binarization result: (a) Original image, (b) NICK's result (Bataineh *et al.*, 2011) 22

Figure 2.7 Binarization result: (a) Original image, (b) Bataineh *et al's* result (Bataineh *et al.*, 2011) 23

Figure 2.8 Binarization result: (a) Original image, (b) Chiu *et al* 's result (Chiu *et al.*, 2012) 25

Figure 2.9 x-direction ($G_x$) and y-direction ($G_y$) for Sobel convolution masks 25

Figure 2.10 Binarization result: (a) Original image, (b) Canny's result (Canny, 1986) 27

Figure 2.11 Binarization result: (a) Original, (b) Otsu 28

Figure 2.12 Binarization result: (a) Original image, (b) Tsai *et al.*'s result (Tsai *et al.*, 2010) 31

Figure 2.13 Binarization result: (a) Original image, (b) Anwar and Abdullah's result (Anwar & Abdullah, 2014) 32

Figure 2.14 Example of a contour illustrated by elliptic Fourier descriptors 38

Figure 2.15 SVM utilizes hyperplane margin to divide negative from positive classes 45

Figure 2.16 Example of *k*-mean algorithm: (a) input data set, (b) three seed points selected as cluster centers, (c) result after second iteration, (d) result after third iteration, (c & d) intermediate iterations updating cluster labels and their centers, and (e) final clustering by *k*-mean 51

Figure 3.1 Flow chart summarising the defect detection algorithm 56

Figure 3.2 Transmittance of light source with different wavelengths through a solar wafer (Ko *et al*., 2013) 59

Figure 3.3 (a) The schematic of backlight illumination, (b) the actual backlight illumination set-up 60

Figure 3.4 (a) The schematic of front-light illumination, (b) the actual front-light illumination system 61

Figure 3.5 The overall hardware set-up of the solar wafer inspector showing the important components 62

Figure 3.6 Examples of polycrystalline solar wafer images produced by the NIR system in Figure 3.5, (a) intact or good sample, (b) defected samples due to (i) micro-crack, (ii) fingerprint, and (iii) stain or glove marks. Dotted circles in (b) show the locations of the defects. 63

Figure 3.7 Examples of polycrystalline and monocrystalline solar wafer images as Dataset 2, (a) and (b) are defective polycrystalline and mono-crystalline samples respectively. The circles indicate the locations of the defect 63

Figure 3.8 (a) stain (b) micro-crack (c) fingerprint as defective samples captured by using front-light and backlight illuminator 64

Figure 3.9 Flowchart of the proposed image thresholding algorithm 67

Figure 3.10 Examples of polycrystalline solar wafer images produced by the NIR system in Figure. 3.6 as dataset 1, (a) is intact or good sample, (b) is micro crack sample; (c) fingerprint, and (d) is defected by stain or glove marks, the circle indicates the location of stain 78

Figure 3.11 Block diagram of defect detection and classification 81

Figure 3.12 Flowchart of *k*-mean clustering 84

Figure 4.1 Examples of thresholding results (a): Original images which are displayed as: (i) micro crack, (ii) defective images including micro cracks and other invisible defects such as scratches, and (iii) finger print defect. (b) Segmented images without *z* and *k* parameters, and (c) proposed segmented incorporating *z* and *k* parameters 89

Figure 4.2 Defect detection of NIR images comparing the proposed technique with Sobel (1970), Otsu (1979), Canny (1986), and anisotropic diffusion (Tsai et al., 2010) thresholdings: (a) Sobel, (b) Otsu, (c) Canny, (d) anisotropic diffusion, and (e) the proposed technique 91

Figure 4.3 Segmentation results comparing the proposed and the original Niblack formula including some of its popular variants. (a) input images (b) ground truth images corresponding to images in (a), (c) Niblack (1985), (d) Sauvola and Peitaikanen (2000), (e) Feng and Tan (2004), (f) Wolf and Jolion (2004), (g) Bataineh *et al.* (2011), (h) Chiu *et al.* (2012), and (i) proposed method 94

Figure 4.4 The average accuracy evaluation results 98

Figure 4.5 The average FOM evaluation results 99

Figure 4.6 The average DSC evaluation results 100

Figure 4.7 The average SEN evaluation results 101

Figure 4.8 The average FNR and FPR evaluation results 102

Figure 4.9 (a) The original images (b) the results without adaptive *k*, *z* and *x* parameters (c) the results of the proposed Method 2 104

Figure 4.10 Examples of thresholding results (a) original images, (i-ii) mono-crystalline, (iii-iv) polycrystalline solar wafers, (b) proposed segmented incorporating $z$, $x$ and $k$ parameters, (c) Sobel (1970), (d) Otsu (1979), (e ) Canny (1986), (f) Diffusion (Tsai *et al*., 2010) 105

Figure 4.11 Examples of binarised results comparing the proposed and the original Niblack formula, including its recent variants (a) original images, (b) Niblack (1985), (c) Sauvola and Pietikainen (2000), (d) Bataineh *et al*. (2011), (e) Chiu *et al*. (2012), (f) Feng and Tan (2004), (g) Wolf and Jolion (2004), (h) proposed method, and (i) Ground truth images corresponding to images in (a). In all images, (i-ii) are mono-crystalline and (iii-iv) are polycrystalline solar wafers. 108

Figure 4.12 (a) defective images of Dataset 1, (b) results of proposed Method 2 (c) results of proposed Method 1 115

Figure 4.13 (a) the defective images of Dataset 2, (b) the results of proposed Method 1 (c) the results of proposed Method 2 116

Figure 4.14 The normalized $\overline{|EFD|}$ of solar wafer images samples: (a) original images, (b) results after thresholding and filtering, (c) results after connected component, (d) boundary results and (e) spectrum of normalized $\overline{|EFD|}$ 120

Figure 4.15 Elbow chart showing the variation of SSE with $k_c$ 122

Figure 4.16 Elbow chart showing the variation in the performance variance with $k_c$ 123

Figure 4.17 Distribution of clustering results of 2000 samples Dataset 1, comprising of different types of Dataset 1 samples 124

Figure 4.18 Scatter distribution of clustering results on different types of samples 124

Figure 4.19 The Silhouette result of clustering 125

Figure 4.20 The Rand index results of clustering 126

Figure 4.21 The accuracy results of classification 128

# LIST OF ABBREVIATIONS

| | |
|---|---|
| ACC | Accuracy |
| AFKM | Adaptive fuzzy k-means clustering |
| AEFD | Average of Elliptic Fourier descriptors |
| Am | Amplitude |
| A2DKM | Automated two-dimensional k-Means |
| ANN | Artificial neural network |
| BCC | Bandwidth Cluster Center |
| CCD | Charge-coupled device |
| DA | Discriminant analysis |
| DAGSVM | Directed acyclic graph SVM |
| DSC | Dice similarity coefficient |
| DWT | Discrete wavelet transform |
| EFDs | Elliptic Fourier descriptors |
| EL | Electroluminescence |
| EM | Elbow method |
| FA | Factor analysis |
| FNR | False negative rate |
| FOM | Pratt’s figure of merit |
| FPR | False Positive Rate |
| FTA | Focal plane array |
| FCM | Fuzzy c-means |
| GK | Gustafson–kessel |
| HRBF | Hybrid radial basis function |
| IR-LED | Infrared Light Emitting Diode |
| L | Level |
| LR | Linear regression |
| MATLAB | Matrix laboratory |
| MDRL | Manifold discriminant regression Learning |
| NIR | Near Infrared |
| OVA | One-versus-all |

| | |
|---|---|
| PL | Photo luminescence |
| RBF | Gaussian radial basis function |
| RHT | Radiant heat thermography |
| RI | Rand index |
| RMDRL | Robust manifold discriminant regression learning |
| RUV | Resonance ultrasonic vibration |
| SAM | Scanning Acoustic Microscopy |
| SEN | Sensitivity |
| SI | Silhouette index |
| SMO | Sequential minimum optimization |
| SPE | Specificity |
| SSE | Sum of square errors |
| SVM | Support vector machine |
| SVs | Support vectors |
| TN | True negative |
| TNR | True negative rate |
| TP | True positive |
| TPR | True positive rate |
| WCSS | Within-cluster sum of squares |

# LIST OF SYMBOLS

| | |
|---|---|
| $a_n, b_n, c_n, d_n$ | Coefficients to extract features |
| $b$ | Bias term |
| $B$ | Magnitude average of EFD threshold image |
| $C$ | Regularization parameter |
| $\boldsymbol{c_c}$ | Centre, centroid |
| $dc$ | Diffusion coefficient |
| $d^*$ | Degree of polynomial kernel |
| $d, d_{ij}, d_e$ | Euclidean distance |
| $div$ | Divergent operator |
| $f, f_i, f_{ij}, \hat{f}_k$ | Classification function |
| $G$ | Gray-level of an image |
| $I$ | Input image |
| $\bar{I}$ | Binarised input image |
| $i, j, m, n$ | Index variable |
| $K$ | Kernel function |
| $K^*$ | Edge stopping threshold |
| $k$ | Normalisation factor |
| $k_c$ | Number of cluster |
| $l$ | Number of samples |
| $l_i$ | Number of samples in class *i* |
| $\mu$ | Mean |
| $\mu_g$ | Mean value of the global image pixels |
| $\mu_i$ | Mean of points belonging to each cluster |
| $\mu_w$ | Mean value of the pixels inbinarizationwindow |
| $M$ | Minimum gray value of image |
| $max_{level}$ | Maximum gray level value |
| $\boldsymbol{N}$ | Number of Fourier harmonics |
| $N_c$ | Number of dataset in cluster |
| $N_D$ | Numbers of detected edge pixels |
| $N_I$ | Numbers of ideal edge pixels |
| $N_p$ | Neighboured pixels |

| | |
|---|---|
| $N$ | Total number of pixels in an image |
| $p_i$ | Pixel value |
| $R$ | Maximum gray-value standard deviation |
| $Rs$ | Dynamic range of standard deviation |
| $\sigma$ | Standard deviation |
| $\sigma_{adaptive}$ | Adaptive standard deviation |
| $\sigma_{max}$ | Maximum standard deviation |
| $\sigma_{min}$ | Minimum standard deviation |
| $\sigma_w$ | Window standard deviation |
| $Sm_{ij}$ | Separability measure between class *i* and class *j* |
| $T$, $T_H$ , $T_L$ | Threshold |
| $\boldsymbol{T}$ | Total contour length |
| $t$ | Iteration |
| $u$, $v$, | Index variable |
| $\boldsymbol{W}$ | Weight vector |
| $\boldsymbol{X}$ , $\boldsymbol{X}_i, \boldsymbol{X}_n$ | Input feature vector for each sample |
| $\overline{\boldsymbol{X}}_i$ | Vector of group mean |
| $y_i$, $y_n^{ij}$ | Class label (+1 or -1) |
| $z, x, k$ | Proposed defined parameters |
| $\nabla$ | Gradient |
| $\delta$ | Sigma term of the Gaussian kernel function |

# PENGESANAN KECACATAN DAN PENGELASAN WAFER SOLAR SILIKON DENGAN MENGGUNAKAN KAEDAH PENGIMEJAN NIR DAN SEGMENTASI NIBLACK YANG DITAMBAHBAIK

## ABSTRAK

Menghasilkan tenaga yang boleh diperbaharui berkuantiti tinggi memerlukan kecekapan yang tinggi dalam fabrikasi produk wafer silikon, yang juga merupakan komponen asas panel solar. Oleh yang demikian, pemeriksaan kualiti yang tinggi untuk wafer solar semasa proses pengeluaran sangat penting. Dalam tesis ini, sistem pengesanan kecacatan yang cekap dan automatik menggunakan strategi pengelasan dan kelompok termaju telah dicadangkan. Dalam kajian ini, satu skema mesin penglihatan untuk mengesan keretakan mikro dan kecacatan-kecacatan yang lain dalam pembuatan polihabluran dan mono kristal wafer solar dicadangkan dan dibangunkan. Pemeriksaan retak mikro sangat mencabar kerana kecacatan ini sangat kecil dan tidak boleh dilihat dengan mata kasar. Kewujudan struktur heterogenus yang lain dalam wafer solar seperti bahan-bahan kasar dan kawasan gelap menjadikan pemeriksaan lebih mencabar. Dalam tesis ini, sebuah inspektor retak mikro yang mengandungi pencahayaan inframerah yang dekat dan algoritma segmentasi Niblack yang diperbaharui telah dicadangkan. Keputusan emperikal dan visual menunjukkan ketepatan dan prestasi yang lebih baik dari segi angka merit Pratt dan kaedah penilaian yang lain berbanding dengan formula pengambangan Niblack yang sedia ada. Keputusan angka merit (FOM), ketepatan (ACC), pekali kesamaan dadu (DSC) dan sensitiviti yang masing-masingnya sentiasa lebih tinggi daripada 0.871, 99.35 %, 99.68 %, dan 99.75 % bagi imej-imej dalam kajian ini. Sementara itu, satu set deskriptor bersepadanan dengan penerangan ciri-ciri bentuk Fourier eliptik, diekstrak bagi setiap kecacatan yang telah dikesan, dan dinilai bagi setiap kluster bagi tujuan pengelompokan dan pengelasan. Pengelasan menggabungkan analisis ciri keamatan kecacatan, penggunaan tanpa pengawasan kelompok purata-*k* dan pelbagai kelas algoritma SVM. Kaedah-kaedah ini telah digunakan untuk pengesanan, pengelompokan dan klasifikasi imej wafer solar polihabluran, bersepadanan dengan kecacatan seperti keretakan mikro, kekotoran, dan cap jari. Keputusan kajian menunjukkan bahawa kaedah purata-*k* dan penklasifikasi SVM mampu mengelompok dengan tepat kecacatan-kecacatan tersebut dengan ketepatan, indeks Rand, dan Bayang indeks dengan nilai purata masing-masing sebanyak 99.8 %, 99.788 %, dan 98.43 %.

# DEFECT DETECTION AND CLASSIFICATION OF SILICON SOLAR WAFER FEATURING NIR IMAGING AND IMPROVED NIBLACK SEGMENTATION

## ABSTRACT

Producing a high yield of renewable energy requires a high efficiency in product fabrication of silicon wafers, which is the basic building component of solar panels. For this reason, the high quality inspection of solar wafers during the procedures of production is very important. In this thesis, an automatic and efficient defect detection system, utilising advanced classification and clustering strategies are proposed. In this study a machine vision scheme for detecting micro-cracks and other defects in polycrystalline and monocrystalline solar wafer manufacturing is proposed and developed. Micro-crack inspection is very challenging, because this type of defect is very small and completely invisible to the naked eye. The presence of other heterogeneous structures in solar wafers like grainy materials and dark regions further complicates the problem. In this study an efficient micro-crack inspector comprising near infrared illumination and an improved Niblack segmentation algorithm is proposed. Empirical and visual results demonstrate that the proposed solutions are competitive when compared to existing Niblack thresholding formulas and other standard methods, and achieve better precision and performance in terms of Pratt's figure of merit and other evaluation methods. Result in a figure of merit (FOM), accuracy (ACC), dice similarity coefficient (DSC), and sensitivity were consistently higher than 0.871, 99.35 %, 99.68 %, and 99.75 %, respectively, for all images tested in this study. Meanwhile, a set of descriptors corresponding to Elliptic Fourier Features shape description is extracted for each defect and is evaluated for each cluster to use for clustering and classification part. The classification combines the analysis of defect intensity features, the application of unsupervised *k*-mean clustering and multi-class SVM algorithms. The methods have been applied for detecting, clustering and classification polycrystalline solar wafer images, corresponding to defects such as micro cracks, stain, and fingerprints. Results indicate that the *k*-mean and SVM classifier can accurately cluster the defects with accuracy, Rand index, and Silhouette index averaging at 99.8 %, 99.788 %, and 98.43 %, respectively.

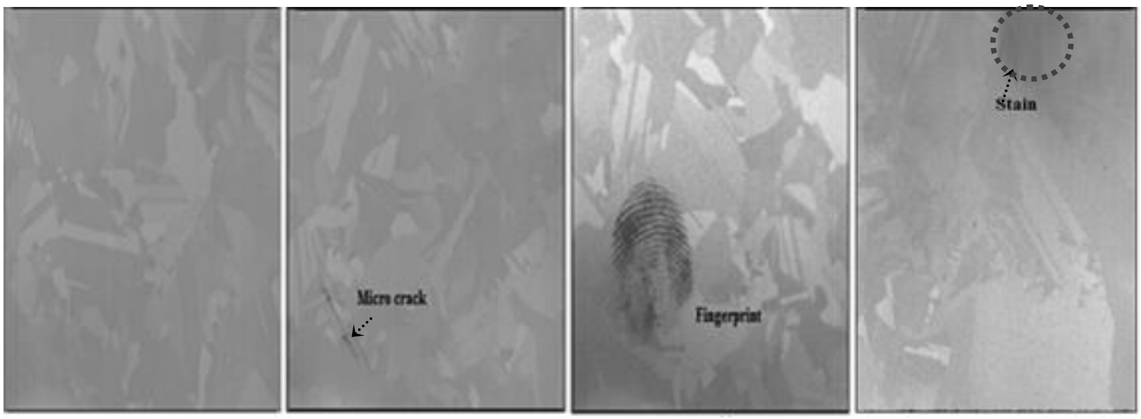
# CHAPTER 1

# INTRODUCTION

## 1.1 Motivation

The increasing demand for renewable energy has led to the growth in the production of solar cells and wafers. Naturally, there has also been an increase in silicon wafer production, which forms the basic building component of many solar panels. According to the statistics published by the Silicon Manufacturer Groups, the world-wide shipments of solar wafers achieved a record high of 2,587 millions square inches shipped in the second quarter of 2014 (Calif, 2014). Depending on the materials used in the manufacturing, solar wafers and cells can be divided into two major types. They are (i) monocrystalline wafers, which are utilized in the manufacture of integrated circuits and transistors, and (ii) polycrystalline silicon wafers, which are commonly utilized in solar power and semiconductor industries (Sparenberg, 2009). In industrial applications, polycrystalline wafers are the preferred material in the production of solar wafers due to lower manufacturing costs (Tsai *et al*., 2010; Belyaev *et al*., 2006). However, the imperfection of manufacturing processes has led to a substantial reduction in production yields. Around 5 %-10 % of the total numbers of wafers produced are defective, which in turn causes energy wastage due to increases production costs (Chiou *et al*., 2011; Rupnowski & Sopori, 2009). Thus, one of the most important procedures in the production of solar wafers is the inspection defects. Chief among these defects are micro-cracks, which contribute to stress fractures and thus equipment down time. The problem is very challenging because this type of defect is very small and completely invisible to the naked eye, which is formed inside the solar wafer and can only be visualised electronically or sensed mechanically. Depending on its size, the micro-cracks can be categorised into two groups. The first group comprises micro-cracks with sizes less than 30 $\mu m$, while the second group comprises those at bigger than 30 $\mu m$ in size (Chiou *et al*., 2011; Israil *et al*., 2013). Moreover, the presence of other heterogeneities in the solar wafer, such as grainy material or broken metal fingers, can cause the wafer images to be highly textured with a densely heterogeneous background when visualised electronically. The low contrast between intact and defective pixels further complicates the problem. Traditionally, the near infrared (NIR) spectrum has been used for the purpose of imaging, since this type of radiation is transparent to most of the materials which make-up solar wafers. Compared to other imaging techniques, NIR offers advantages in terms of high accuracy, good sensitivity and faster response time (Israil *et al*., 2013). However, NIR imaging requires very powerful and advanced image processing techniques, since the image that it produces usually contains many artefacts.

Examples of polycrystalline solar wafer images, which include an intact sample and several other samples that possess such imperfections as micro cracks, stains, and fingerprints, are shown in Figure 1.1. In reality there are other types of defects in solar wafers products but the common defects as micro crack, fingerprint and stain are shown in this figure. In order to solve this kind of multi-class problem found within photovoltaic industry, several methods have been used.



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

Figure 1.1: Examples of polycrystalline solar wafer imagesDataset 1: (a) is an intact or good sample, (b) is a micro crack sample; (c) a defective sample that includes fingerprint, and (d) is defected by a stain. In (b) and (d), the locations of micro-crack and stain are indicated by arrows.

In the recent years, there has been an increasing trend in the use of machine vision in the manufacturing sectors and industry. This includes methods used to supply imaging-based automatic inspection and analysis for applications such as automatic inspection, process control, and robot guidance in industry. Typical tasks of machine vision in the industrial vision inspection system include: image acquisition, image processing, feature extraction, and decision making (Malamas *et al*., 2003). The use of machine vision in industrial automation provides a better solution, as it helps to increase productivity and quality through consistent, accurate and fast inspection. However, due to the lack of image processing and artificial intelligence algorithms which are suitable andaccurate in solving the inspection tasks involved, the inspection and grading processes continue to be manual or semi-manual efforts (Anwar, 2014). Inevitably, the problem of detecting defects in solar wafer also exhibit similar circumstances. Conventionally, the solar wafers consist of invisible and visible defects. The main defects as the invisible defects are micro-cracks and the visible defects are stains, fingerprints. Because of the increasing of using solar cells and wafer applications, even the defects that do not directly link to reliability issues such as water mark and surface stain, fingerprint are detected and considered as fail or secondary grade of cells for the solar cell and wafer buyers. Those defects are visually inspected by operators. However, the inconsistent inspection results caused by human error make the fully automatic optical inspection solution become essential equipment for crystalline cell and wafer products (Chroma, 2015).

Therefore, there is a research prospective, specifically in the field of machine vision, to solve the problem of micro-cracks and the detection of other defects and classifications in solar wafers. Motivated by this need, this thesis presents the methods and techniques for detecting defects in the images of polycrystalline and monocrystalline solar wafers. This study integrates an image-processing and machine-learning platform toward an application in invisible and visible defect inspection and classification. It addresses image processing techniques based on an adaptive Niblack filter and its application in solar wafer images. Additionally, this work examines the EFD method for feature extraction. Furthermore, machine learning and classification based on unsupervised clustering is investigated. For the sake of completeness this work also examines the classification based on multiclass supervised support vector machines (SVM).

## 1.2 Problem Statement

Among the tests that need to be carried out on solar wafers includes the inspection of micro-cracks and other defects such as stains and fingerprint. In an effort to reduce the cost of production, manufacturers are making increasingly thinner solar wafers. Though cheaper to produce, such products suffer from a serious drawback, in that they are relatively more fragile and hence easily broken if not handled properly. Moreover, the thinner the wafer, the greater the chances of forming micro-cracks and other defects. On average, about 5 to 10 percent of solar wafers tend to break during production (Rupnowski & Sopori, 2009). It is important that these micro-cracks and other defects are detected and extracted as early as possible to minimise machine outages or other complications resulting from processing defected samples, especially during assembly or packaging.

There are various methods which can be used in the detection of micro-cracks and other defects. Among them are Radiant Heat Thermography (RHT) (Devitt *et al*., 1992; Pilla *et al*., 2002), eddy currents (Johnson & Esquivel, 2006; Zenzingera *et al*., 2007), dye inspection (Zenzingera *et al*., 2007), the ultrasonic method (Reber & Beller, 2003), the Scanning Acoustic Microscopy Method (SAM) (Knauss *et al*., 1995; Connor *et al*., 1998), Resonance Ultrasonic Vibration (RUV) (Dallas *et al*., 2008; Polupan, & Ostapenko, 2006), optical transmission (Ko *et al*., 2013; Abdelhamid *et al*., 2014), Photo Luminescence (PL) (Chiou *et al*., 2011; Trupke *et al*., 2006a; Trupke *et al*., 2006), Electro Luminescence (EL) (Takahashi *et al*., 2006; Dreckschmidt *et al*., 2007; Tsai *et al*., 2012; Anwar & Abdullah, 2014), and infrared thermography (Pilla *et al*., 2002).

Some of these methods, especially infrared thermography and RHT, are less popular because of their limited capability in distinguishing micro-cracks from other textures in a solar wafer image. Meanwhile, methods like dye mapping and RUV have limited use because they can potentially damage the sample during inspection.

In contrast, EL does not suffer from the same problems as mechanical methods, since it is a completely non-destructive inspection technique. However, this method