CONTAMINATED PARTICLE DETECTION USING IMAGE PROCESSING

CHATCHAI MAHUNTADECHAKUL

A THESIS SUBMITTED IN PARTIAL FULFILLMENT

OF THE REQUIREMENT FOR THE DEGREE OF

MASTER OF ENGINEERING IN DATA STORAGE TECHNOLOGY

ENGLISH PROGRAM

INTERNATIONAL COLLEGE

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

2012

KMITL-2012-IC-M-005-009

หัวข้อวิทยานิพนธ์	การตรวจวัดเม็ดสิ่งสกปรกโดยใช้เทคนิคการประมวลผลภาพ
นักศึกษา	นายฉัตรชัย มหันตเดชากุล
รหัสประจำตัว	52600607
ปริญญา	วิศวกรรมศาสตรมหาบัณฑิต
ภาควิชา	เทคโนโลยีการบันทึกข้อมูล
พ.ศ.	2555
อาจารย์ที่ปรึกษาวิทยานิพนธ์	ดร.ชัยวัฒน์ หนูทอง

บทคัดย่อ

ปัญหาการทำงานผิดพลาดเนื่องจากมีสิ่งสกปรกภายในผลิตภัณฑ์นั้น มีความสำคัญอย่างมากในอุตสาหกรรมฮาร์ดดิสก์ไดร์ฟ งานวิจัยนี้ นำเสนอเทคนิคการประมวลผลภาพในการหาจำนวน และขนาดของเม็ดสิ่งสกปรก รวมทั้งศึกษาความเป็นไปได้ในการใช้กระบวนนี้เพื่อแยกและระบุชนิดของเม็ดสิ่งสกปรก สำหรับการประมวลผลภาพนั้น มีวิธีการที่สำคัญ คือ Otsu และ Kmeans โดยส่วนแรกของกระบวนการ คือ การแยกวัตถุออกจากพื้นหลังของภาพโดยใช้วิธีการ Otsu จากนั้นส่วนที่สองจะทำการแยกชนิดของวัตถุด้วยวิธีการ Kmeans และหาเฉลี่ยของแสงสีแดง, เขียว และน้ำเงิน เพื่อระบุชนิดของวัสดุในส่วนของการแสดงผลโดยอุปกรณ์ที่ใช้ในการถ่ายภาพ คือ กล้องที่มีกำลังขยาย 200 เท่าซึ่งให้ความละเอียดของภาพที่กว้างและยาว 0.557 ไมครอนต่อพิกเซล และเม็ดสิ่งสกปรกที่ใช้ศึกษามีอยู่ 3 ชนิดด้วยกัน คือ ทอง, สเตนเลสสตีล และพลาสติกจากบรรจุภัณฑ์ของหัวอ่านฮาร์ดดิสก์ไดร์ฟ วัสดุเหล่านี้จะถูกนำไปหาค่าเฉลี่ยของแสงสีแดง, เขียว และน้ำเงิน เพื่อกำหนดลักษณะเฉพาะของวัสดุในการประมวลผลภาพผลที่ได้จากการศึกษาพบว่าการตรวจวัดขนาดของเม็ดสิ่งสกปรกสามารถใช้ตรวจวัดวัตถุที่มีขนาดตั้งแต่ 25 ไมครอนขึ้นไป โดยมีความผิดพลาดประมาณ 1.96% การตรวจวัดจำนวนของเม็ดสิ่งสกปรกบนตัวอย่างที่มีทองและพลาสติกมีความผิดพลาดประมาณ 3% สำหรับตัวอย่างที่มีทอง และ สเตนเลสสตีล, และตัวอย่างที่มีสเตนเลสสตีลและพลาสติกพบความผิดพลาดในการตรวจวัดประมาณ 13% และ 12% ตามลำดับ สำหรับการระบุชนิดของเม็ดสิ่งสกปรกด้วยวิธีการประมวลผลภาพ ในงานวิจัยนี้ ให้ผลที่ถูกต้องกับตัวอย่างซึ่งมีเม็ดสิ่งสกปรกที่เป็นทอง และพลาสติกจากบรรจุภัณฑ์ของหัวอ่านฮาร์ดดิสก์ไดร์ฟเท่านั้น

Thesis	Contaminated Particle Detection using Image Processing
Student	Mr. Chatchai Mahuntadechakul
Student ID.	52600607
Degree	Master of Engineering
Program	Data Storage Technology
Year	2012
Thesis Advisor	Dr. Chaiwat Nuthong

ABSTRACT

The particulate failure is one of the most important issues in Hard Disk Drive manufacturing. This thesis presents contaminated particle detection algorithm in order to measure particle's quantity and sizes; including feasibility study of material identification. This algorithm contains two main sections; Otsu technique to extract particle from the image background and Kmeans to classify particle into multi clusters. After that the RGB information of each cluster is used to evaluate material's type. Many experiments using Optical Image Analysis (OIA) with 200X magnification with intentionally doped gold, stainless steel (SST) and tray powder as contaminated particles have been done. By using RGB baseline analysis the material identification criterion can be set. This proposed method shows that it can detect particle of size 25 micron and above. By detection particle's quantity the average errorness has found to be about 3%, 13% and 12% for the sample which are gold with tray, gold with SST and SST with tray, respectively. Finally, this algorithm shows the correct display for the sample of gold with tray only. However, the proposed method can be developed further in the future for more accurate material identification.

Acknowledgements

We would like to extend our profound thanks to Thailand Development Engineering and Material Science Laboratory Department at Seagate Technology (Thailand) Ltd., and College of Data Storage Innovation, King Mongkut's Institute of Technology Ladkrabang for all supports through out this research. We also would like to thank Dr. Chaiwat Nuthong, my advisor for his kind support, especially on technical consultance.

Finally, I would like to thank my family and my friend for all love.

Chatchai Mahuntadechakul

Table of Contents

Table of Contents (Cont)

List of Tables

List of Figures

List of Figures (Cont)

CHAPTER 1

Introduction

1.1 Background and Problems

Cleanliness control is required for Hard Disk Drive (HDD) component and HDD processes since contaminated particles could cause a serious HDD failure, i.e. particulate failure. Hence it is very important to able to detect the contaminated particle beforehand. There are many techniques, which are used for particle inspection and analysis, such as Liquid Particle Counter (LPC), Air Borne Particle Counter (APC), Scanning Electron Microscopy (SEM) and Optical Image Analysis(OIA). Among these, Optical Image Analysis (OIA) is the most interesting technique, since it has the lowest operating cost and the shortest operating time. The method has already been used for screen media lubricant contaminated on slider, particle mapping, particle distribution model of HDD process, monitoring cleaning process and other functions which are related to cleanliness.

Optimal Image Analysis (OIA) is operated the result of particle's quantity by using human eye. In particle inspection process, if the sample has a large amount of particles, the human error and operating time are limitation of this measurement. Moreover, it is not able to measure particle's size and also identify the type of material.

Figure 1.1 Particle measurement

1.2 Objectives

The main objective of this research is to develop on algorithm for OIA measurement in order to reduce operating time and provide information of particle's quantity and material. The algorithm is integrated with automatic optimal threshold finding to extract image of particles from background and use multi cluster technique to classify material of that particular particles. Finally, the information of an amount of particle, particle's size and expected material are evaluated and displayed.

1.3 Scope of Research

1.3.1 Develop an algorithm on MATLAB for operating with OIA measurement.

1.3.2 Use OIA measurement with the developed algorithm to specify sizes and total amount of contaminated particle.

1.3.2 Study feasibility to identify material of particle by using OIA measurement with the developed algorithm.

This thesis is organized as follows. Chapter 2 describes literature review of previous research articles which presented image segmentation techniques and also image optimal thresholding techniques. Chapter 3 describes image processing techniques which are used in the algorithm. Chapter 4 explains the algorithm to detect contaminated particle. Finally, Chapter 5 presents the algorithm's results. It also leads to conclusion of the research as well.

CHAPTER 2
Literature Review

Many research articles present inventions and applications of image segmentation algorithm. It has been evaluated in many works for object detection. Moreover, some works present applications of the algorithm to find object's size and quantity. In this chapter, the relative works are reviewed and studied feasibility to apply the concept with contaminated particle detection algorithm.

The first research article presents the successful results to count soybean grainsby using image processing with MATLAB. This paper "Grain Counting Method Base on Image Processing"[1] was published by Zhao Ping and Li Yongkui. The key learning is image processing algorithm which is operated on the input image from digital camera. The algorithm starts from convert RGB input image to gray scale image. Then the second step, the median filtering is applied to remove the noise and also improve the edge quality of image object. The third step, the optimal threshold value is calculated by Otsu algorithm to separate the pixel of object and background, and then the image information is converted from gray scale to binary image by respect to the selected threshold. This process is very essential for the accurate results. The forth step, the binary image from the previous step is operated with mathematical morphology to eliminate the error on object boundary. Finally, the object is counted and reported the quantity. This paper shows the successful results and gives inspiration to study about image thresholding techniques. The work flow of soybean grains detection is shown in Figure 2.1, the image results of each step and the results of grain counting and error are shown in Figure 2.2 and Figure 2.3, respectively.

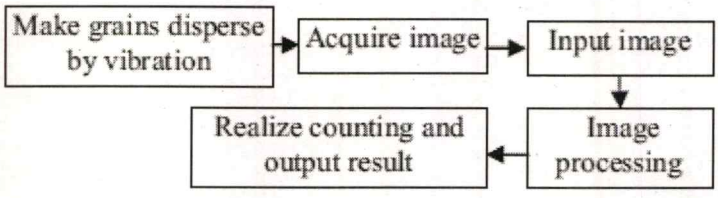


Figure 2.1 Work flow of soybean grains detection [1]

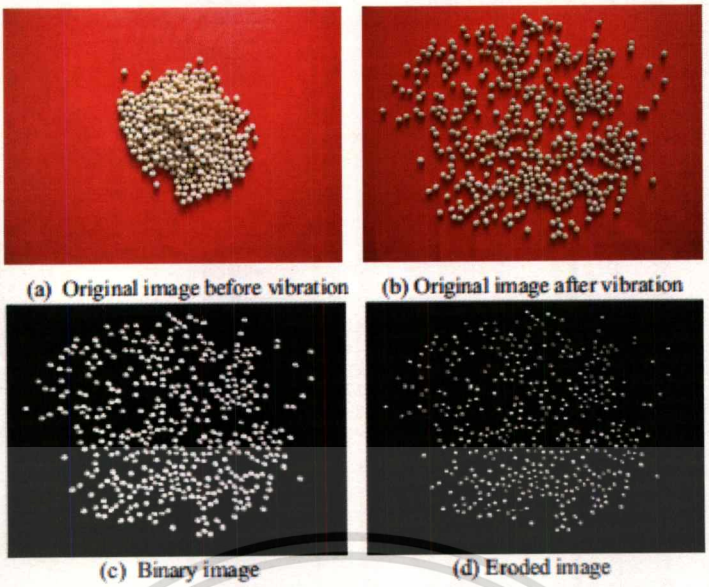


Figure 2.2 Image results of each processing step [1]

Table 2.1 Grain counting results and error [1]

Specimen	*Actual Number*	*Measured Number*	*Relative Error (%)*
Soybean	433	433	0.00
Soybean	341	341	0.00
Cow gram	278	278	0.00
Cow gram	206	206	0.00
Mung bean	493	494	0.29
Mung bean	329	329	0.00

The next research article is "A Review on Global Binarization Algorithms for Degraded Document Images"[2] by MaythapolnunAthimethphat. This article introduces the proposal and equation of 7 image thresholding algorithms. The main conclusion is not only one algorithm can be solved all problems in image segmentation process. However, the proper algorithm can be selected to perform better result in the work.

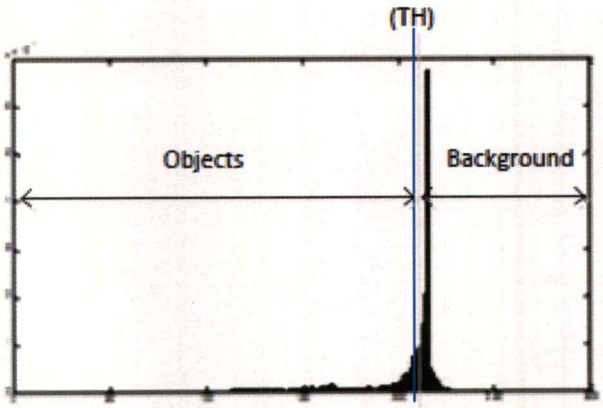


Figure 2.3 The image histogram with optimal threshold

- Ridler and Calvard Algorithm [8]: Ridler and Calvard, they developed the algorithm to optimize the threshold value between object and background of bimodal image histogram. The algorithm uses an iterative clustering by starting the initial mean of image intensity, then iterates the value until meet the criteria.

The object and background distribution probability are p_{obj} and p_{bg},

$$p_{obj} = \sum_{i=0}^{TH} p_i = \omega(TH) \tag{2.1}$$

$$p_{bg} = \sum_{i=TH+1}^{L} p_i = 1 - \omega(TH) \tag{2.2}$$

where TH is the image threshold, p_i is the intensity of pixel in image, $\omega(TH)$is the object's weight in the image, and $L$is the maximum threshold.

The mean value of object and background intensity are$\mu_{obj}$and μ_{bg},

$$\mu_{obj} = \sum_{i=0}^{TH} \frac{ip_i}{p_{obj}} = \frac{\mu(TH)}{\omega(TH)} \tag{2.3}$$

$$\mu_{bg} = \sum_{i=TH+1}^{L} \frac{ip_i}{p_{bg}} = \frac{\mu(L) - \mu(TH)}{1 - \omega(TH)} \tag{2.4}$$

The optimal threshold (TH_{opt}) is

$$TH_{opt} = argmin\left\{\frac{\mu_{bg}(TH) + \mu_{obj}(TH)}{2}\right\} \quad (2.5)$$

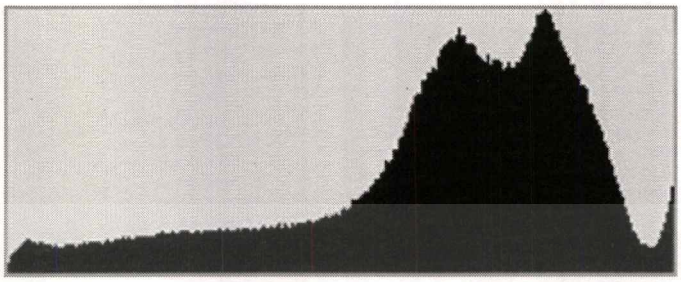


Figure 2.4 Bimodal image histogram [2]

- Kapur Algorithm [2]: This algorithm considers the object and background distributions by threshold level. It maximizes the entropy summation of two areas by optimizing a single threshold value.

The object and background distribution probability are

$$p_{obj} = \sum_{i=0}^{TH} p_i \quad (2.6)$$

$$p_{bg} = \sum_{i=TH+1}^{L} p_i = 1 - P_{obj} \quad (2.7)$$

The object and background distribution are$H(0,TH)$and $H(TH,L)$,

$$H(0,TH) = -\sum_{i=1}^{TH} \frac{P_i}{P_{obj}} ln \frac{P_i}{P_{obj}} \quad (2.9)$$

$$H(TH,L) = -\sum_{i=TH+1}^{L} \frac{P_i}{P_{bg}} ln \frac{P_i}{P_{bg}} \quad (2.10)$$

The optimal threshold value is

$$TH_{optimal} = H(0, TH) + H(TH, L) \tag{2.11}$$

- Fan Algorithm [2]: Fan presented the fast entropic technique by minimizing the complexity of computation.

The object and background distribution probability are

$$P_{obj}(TH) = \sum_{i=0}^{TH} p_i \tag{2.12}$$

$$P_{obj}(TH) = \sum_{i=0}^{TH} p_i \tag{2.13}$$

The object and background distribution are

$$H_{obj}(TH) = -\sum_{i=1}^{TH} \frac{P_i}{P_{obj}(TH)} In \frac{P_i}{P_{obj}(TH)} \tag{2.14}$$

$$H_{bg}(TH) = -\sum_{i=TH+1}^{L} \frac{P_i}{P_{bg}(TH)} In \frac{P_i}{P_{bg}(TH)} \tag{2.15}$$

The optimal threshold is

$$TH_{optimal} = Argmax\{H_{obj}(TH) + H_{bg}(TH)\} \tag{2.16}$$

- Portes de Albuquerque Algorithm [2]: Portes de Albuquerque developed this algorithm in 2004 by customizing the non-extensive entropy.

The object and background distribution probability are

$$P_{obj} = \sum_{i=0}^{TH} p_i \tag{2.17}$$

$$P_{bg} = \sum_{i=TH+1}^{L} p_i = 1 - P_{obj}(TH) \tag{2.18}$$

The object and background distribution are

$$H_{obj}(TH) = \frac{1 - \sum_{i=1}^{TH} \left(\frac{P_i}{P_{obj}}\right)^q}{q-1} \tag{2.19}$$

$$H_{bg}(TH) = \frac{1 - \sum_{i=TH+1}^{TH} \left(\frac{P_i}{P_{bg}}\right)^q}{q-1} \tag{2.20}$$

The optimal threshold value is

$$TH_{optimal} = Argmax\left\{\begin{matrix} H_{obj}(TH) + H_{bg}(TH) \\ +(1-q)\left(H_{obj}(TH)H_{bg}(TH)\right) \end{matrix}\right\} \tag{2.21}$$

- Xiao Algorithm [2]: Xiao presented the improvement method from the Kapur algorithm. He applied the gray-level spatial correlation (GLSD) histogram to develop the entropic algorithm.

The object and background distribution probability are

$$P_{obj} = \sum_{k=0}^{TH} \sum_{m=1}^{N\times N} P(k,m) \tag{2.22}$$

$$P_{bg} = \sum_{k=TH+1}^{255} \sum_{m=1}^{N\times N} P(k,m) \tag{2.23}$$

where $k$is the member of grey level in small windows which contain $N \times N$pixels, and $m$is the number of pixels in small windows as well.

The object and background distribution are

$$H_{obj}(TH,N) = -\sum_{k=0}^{TH} \sum_{m=1}^{N\times N} \frac{P(k,m)}{P_{obj}} In\left(\frac{P(k,m)}{P_{obj}}\right) \omega(m,N) \tag{2.24}$$

$$H_{bg}(TH,N) = -\sum_{k=TH+1}^{255} \sum_{m=1}^{N\times N} \frac{P(k,m)}{P_{bg}} In\left(\frac{P(k,m)}{P_{bg}}\right) \omega(m,N) \tag{2.25}$$

The optimal threshold value is

$$TH_{optimal} = ArgMax\{f_1(TH, N)\} \tag{2.26}$$

- Otsu Algorithm [6]: This algorithm is very simple and powerful to find the optimal threshold value. It has been selected to use in this work. Equation and the calculation example are shown in Chapter3.

- Kittler and Illingworth Algorithm [13]: Kittler and Illingworth, they presented the method to minimize the probability of classification error by evaluating the density of gaussian function.

The object and background distribution probability are

$$P_{obj} = \sum_{i=0}^{TH} p_i = \omega(TH) \tag{2.27}$$

$$P_{bg} = \sum_{i=TH+1}^{L} p_i = 1 - P_{obj}(TH) \tag{2.28}$$

The object and background mean are

$$\mu_{obj} = \sum_{i=0}^{TH} \frac{iP_i}{P_{obj}} \tag{2.29}$$

$$\mu_{bg} = \sum_{i=TH+1}^{L} \frac{iP_i}{P_{bg}} \tag{2.30}$$

The object and background variance areσ_{obj}^2 and σ_{bg}^2 ,

$$\sigma_{obj}^2 = \sum_{i=0}^{TH} \frac{(i - \mu_{obj})^2 P_i}{P_{obj}} \tag{2.31}$$

$$\sigma_{bg}^2 = \sum_{i=TH+1}^{L} \frac{(i - \mu_{bg})^2 P_i}{P_{bg}} \tag{2.32}$$

The continuous likelihood variablesare

$$p(x|\mu_{obj},\sigma_{obj}) = \frac{1}{\sqrt{2\pi\sigma_{obj}}} e^{-\frac{(x-\mu_{obj})^2}{2\sigma_{obj}^2}} \quad (2.33)$$

$$p(x|\mu_{bg},\sigma_{bg}) = \frac{1}{\sqrt{2\pi\sigma_{bg}}} e^{-\frac{(x-\mu_{bg})^2}{2\sigma_{bg}^2}} \quad (2.34)$$

The optimal threshold value is

$$TH_{opt} = Argmin[1 + 2[p_{obj} ln\sigma_{obj}(TH) + p_{bg} ln\sigma_{bg}(TH)] - 2[p_{obj} ln\, p_{obj} + p_{bg} ln\, p_{bg}] \quad (2.35)$$

This article gives many image processing eqations and also describes advantage and disadvantage for implementation of 7 image segmetation algorithms. This information is very useful to be applied to contaminated particle detection algorithm.

The next research article "Adaptive Image Thresholding for Real-Time Particle Mornitoring"[3] by Keivan Torabi, Seed Sayad and Stephen Thomas Balke in 2006. The authors studied and developed image thresholding teahnique to detect particle's quantity and also its sizes in the real time monitoring system. They usesd CCD camera hardware with resolution 5 μm^2 in their work. They classified the image thresholding techniques into 6 groups which are histogram shape-based methods, clustering-based methods, entropy-based methods, object attribute-based methods, spatial methods and local or adaptive thesholding methods. They also gave comments on the basic histogram thresholding which separates the object threshold and background threshold into two peaks that, this teahnique is not cappable to detect the particle in case of it has very small quantity. However, this issue can be solved by deviding image area into many squares area then process the minimum threshold value between object and background peaks. Finally, the results of each squares are combined. However, this technique is not good to support the realtime operation due to its cost is very expensive and also noise issue finding.

Finally, the authors introduced MaxMin thresholding to improve the sensitive error in particle's size measurement, and to reduce the background nosie also.

$$T_{MaxMin} = Max_{j=0\ to\ k}[Min_{i=1\ to\ n}(A_{i,j})] \quad (2.36)$$

Where T_{MaxMin} is the MaxMin thresholding value which calculated from the maxmimum value of the minimum area of particle detected or $A_{i,j}$.

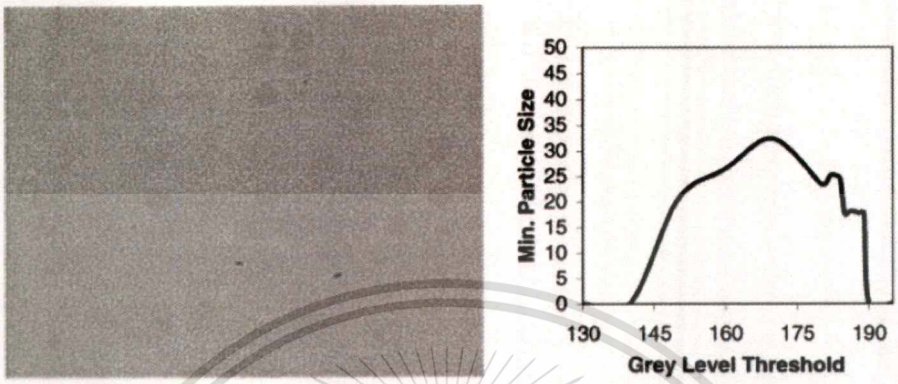


Figure 2.5 Example of a small contaminant particle [3]

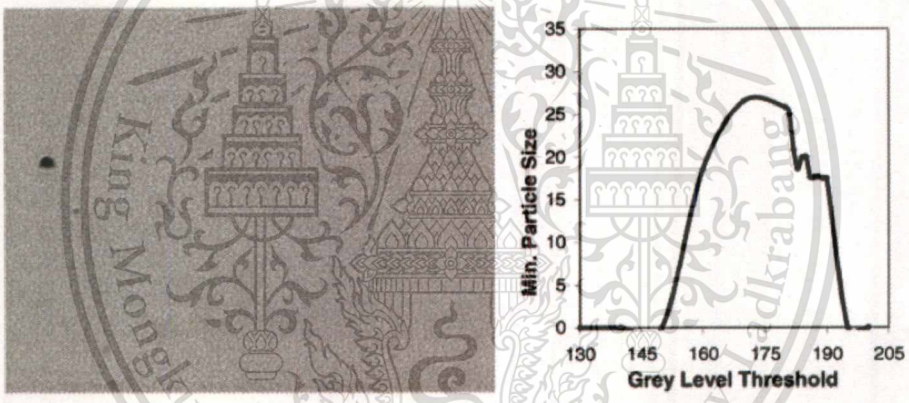


Figure 2.6 Example of a large contaminant particle [3]

The next article "Two-Stage Image Segmentation by Adaptive Thresholding and Gradient Watershed"[4] was presented by Ming Hong Pi and Hong Zhang in 2005. This article presents the image segmentation technique to measure size of oil-sand. The authors studied both watershed image segmentation and image threshold segmentation algorithm. They found that these techniques can be used to subtract oil-sand over the image background, but it have difference segmented weakness. Figure 2.7 shows the original image of oil-sand on conveyer belt and Figure2.8shows the image results of watershed image segmentation. The majority issue of watershed is over-segmentation results. In other hand Figure 2.9 shows binary image after segmentation by using image thresholding algorithm (Otsu's algorithm). However, it shows an error result due to some of oil-sand is not detected.

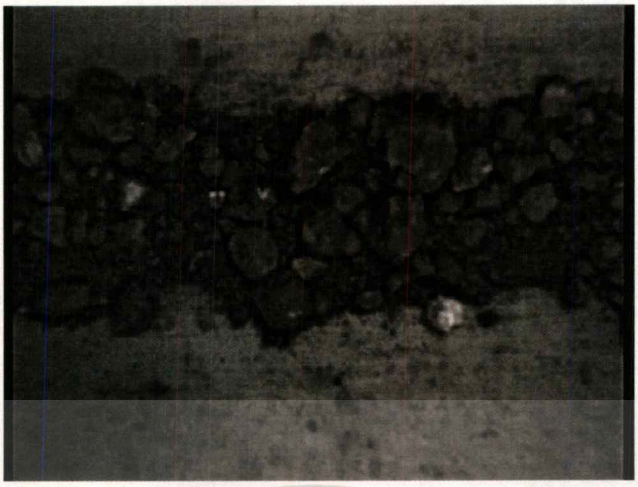


Figure 2.7 Original image of oil-sand on conveyor belt [4]

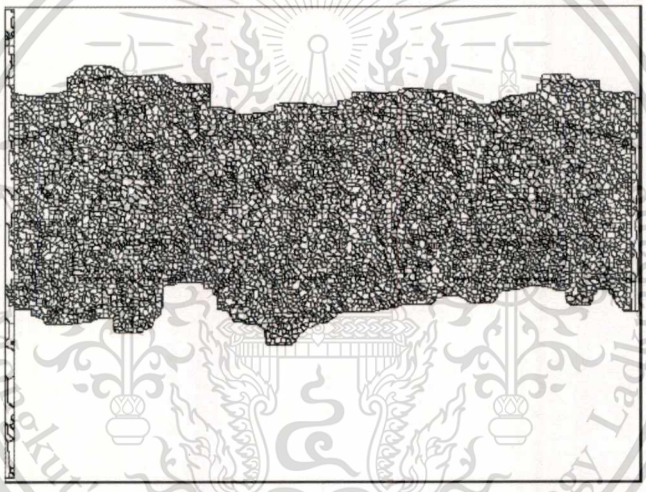


Figure 2.8 Image result of watershed image segmentation [4]

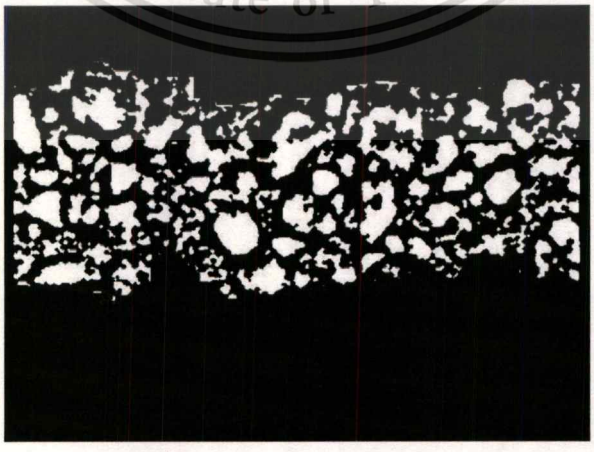


Figure 2.9 Binary image from Otsu thresholding [4]

To recover an error from both image segmentation algorithms, the authors proposed two-stage image segmentation algorithm to improve over-segmentation and also object identification error. The algorithm combines the advantage of watershed image segmentation (Procedure2 in Figure 2.10) to identify image object and image thresholding (Procedure1 in Figure 2.10) to eliminate the error from over-segmentation results.

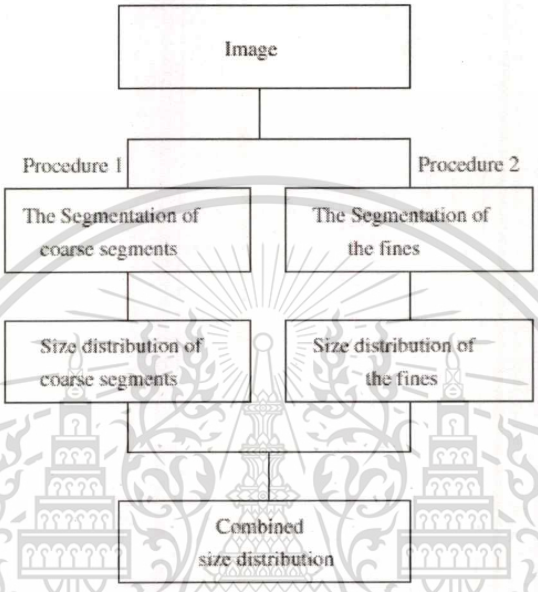


Figure 2.10 Flow chart of two stage image segmentation algorithm

Image processing techniques without thershold consideration are reviewed also, one pass bounddary detection and image edge density algorithm are found in the article "Fast Measuring Particle Size by using the Information of Particle Boundary and Shape"[5] by Weixing Wang. Several edge detection techniques were discussed and evaluated in this article. There are comparison results on 7 edge detection techniques which are Soble, Robert, Laplacian, Prewitt, Canny with high threshold, Canny with low threshold and new egde detection algorithm. Among these, the new edge detection algorithm gives the best results and also performes the lowest error. To achive on accurate result this algorithm uses four directions instead two directions to detect the valley as shown in Figure 2.11. Another advantage it can reduce some noise and illumination variations error from the original image. Finally, egde density is used to calculate the estimate size of the moving particle on coveyor belt.

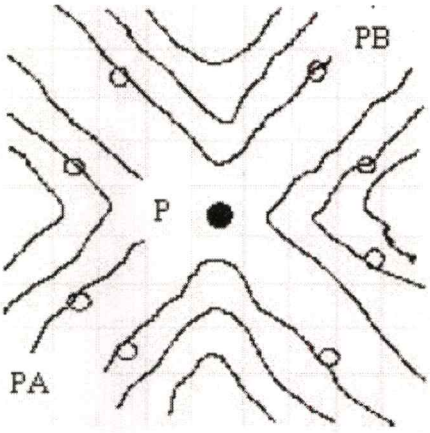


Figure 2.11 A grey value landscape [5]

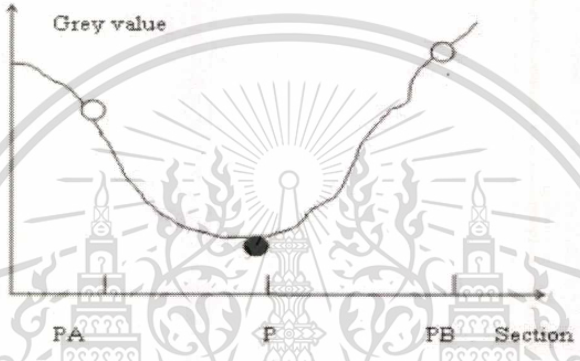


Figure 2.12 The optimal point between PA and PA [5]

CHAPTER 3

Image Processing Techniques

In the previous chapter, many research articles present the effectiveness of image segmentation algorithms in order to specify and evaluate the results of object's quantity and size. However, objective of the contaminated particle detection algorithm has been set not only to measure particle's size and quantity but also able to identify the type of material as well. Hence this chapter describes the techniques which are proposed to specify and classify contaminated particle in this work.

The design phase, work flow is divided into 5 main sections which are input RGB image, RGB to grey scale conversion, image thresholding, Kmeans and object detection as shown in Figure 3.1.

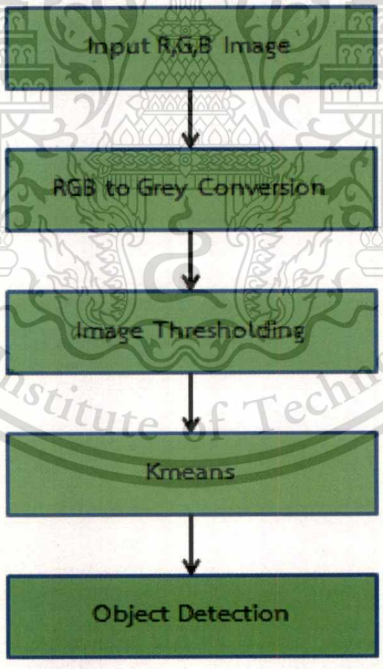


Figure 3.1 Structure of contaminated particle detection algorithm

3.1 RGB to Grey Scale Conversion

RGB image from OIA is acquired to MATLAB program. RGB image is composed of red, green and blue image information with matrix m×n×3 bytes. To process image thresholding, Formula 3.1 is used to convert RGB to grey scale image. The matrix I which has m×n bytes is the representative of grey scale information. It has only the brightness information.

$$I = \alpha \cdot R + \beta \cdot G + X \cdot B \tag{3.1}$$

Where α, β and X are the weight of R,G,B components respectively.

3.2 Image Threshold Optimization

Image thresholding is the crucial process to subtract particle over the image background. Otsu [6] is the selected technique to specify total particle in this work. It was invented by Nobuyuki Otsu in 1979. This technique is simple to use but effectiveness to find the optimum threshold between object and background. This threshold is used to separate object and background in process binary image conversion.

The object and background distribution probability are p_{obj} and p_{bg},

$$P_{obj} = \sum_{i=0}^{TH} P_i = W(TH) \tag{3.2}$$

$$P_{bg} = \sum_{i=TH+1}^{L} P_i = 1 - W(TH) \tag{3.3}$$

$$P_{obj} + P_{bg} = 1 \tag{3.4}$$

$$\mu_{obj} = \sum_{i=0}^{TH} \frac{iP_i}{P_{obj}} = \frac{\mu(TH)}{W(TH)} \tag{3.5}$$

$$\mu_{bg} = \sum_{i=TH+1}^{L} \frac{iP_i}{P_{bg}} = \frac{\mu(L) - \mu(TH)}{1 - W(TH)} \tag{3.6}$$

$$\mu_{Total} = \mu(L) = P_{obj}\mu_{obj} + P_{bg}\mu_{bg} = \sum_{i=0}^{L} iP_i \tag{3.7}$$

Where $\mu_{obj}$and μ_{bg} are mean of object threshold and background threshold, respectively. And σ_{obj}^2,σ_{bg}^2, σ_W^2 and σ_B^2 are variance of object, variance of background, variance within class and variance between class, respectively. Finally, the optimal threshold is evaluated by maximizing variance between class and minimizing variance within class.

$$\sigma_{obj}^2 = \sum_{i=0}^{TH} \frac{(i-\mu_{obj})^2 P_i}{P_{obj}} \tag{3.8}$$

$$\sigma_{bg}^2 = \sum_{i=TH+1}^{L} \frac{(i-\mu_{bg})^2 P_i}{P_{bg}} \tag{3.9}$$

$$\sigma_{Total}^2 = \sum_{i=o}^{L} (i-\mu(L))^2 P_i = \sigma_W^2 + \sigma_B^2 \tag{3.10}$$

$$\sigma_W^2 = W_{obj}\sigma_{obj}^2 + W_{bg}\sigma_{bg}^2 \tag{3.11}$$

$$\sigma_B^2 = W_{obj}W_{bg}(\mu_{bg} - \mu_{obj})^2 \tag{3.12}$$

Example to calculate the optimal threshold by Otsu's algorithm is shown as follows. The initial state threshold T=3 is selected, then the image is separated into initial background and initial object for calculation.

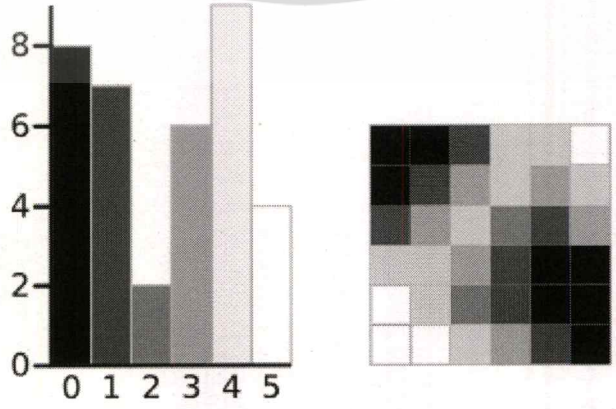


Figure 3.2 Simple6×6 image [23]

The initial state of background is shown in Figure 3.3.

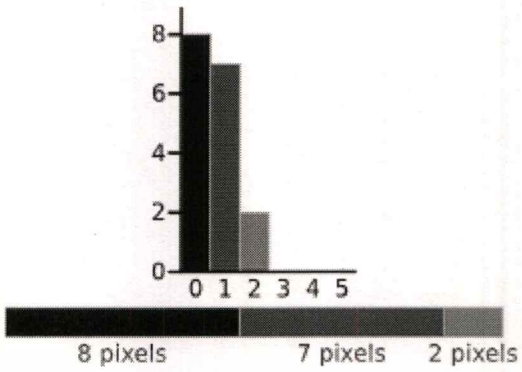


Figure 3.3 Threshold histogram of background [23]

Background mean is

$$\mu_{bg} = \frac{(0\times8)+(1\times7)+(2\times2)}{17} = 0.6471.$$

Background variance is

$$\sigma_{bg}^2 = \frac{((0-0.6471)^2\times8)+((1-0.6471)^2\times7)+((2-0.6471)^2\times2)}{17} = 0.4637.$$

Background weight is

$$W_{bg} = \frac{8+7+2}{36} = 0.4722.$$

The initial state of object is shown in Figure 3.4.

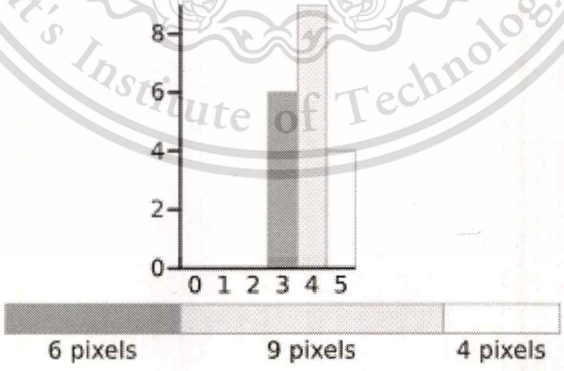


Figure 3.4 Threshold histogram of object [23]

Object mean is

$\mu_{obj} = \frac{(3\times6)+(4\times9)+(5\times4)}{19} = 3.8947.$

Object variance is

$\sigma^2_{obj} = \frac{((3-3.8947)^2\times6)+((4-3.8947)^2\times9+((5-3.8947)^2\times4)}{19} = 0.5152.$

Object weight is

$W_{obj} = \frac{6+9+4}{36} = 0.5278.$

Variance within class and variance between class are

$\sigma^2_W = (0.4722 \times 0.4637) + (0.5278 \times 0.5152) = 0.4909.$

$\sigma^2_B = (0.4722 \times 0.5278)\ (0.6471\text{-}3.8947(0.6471\text{-}3.8947)^2) = 2.6287.$

The summary of variance within class and variance between class from threshold T=0 to T=5 is shown in Table 3.1. Threshold T=3 is considered to be the optimal value between object and background because it has minimum variance within class and maximum variance between class.

Table 3.1 Variance within class and variance between class from threshold T=0 to T=5.

Threshold	T=0	T=1	T=2	T=3	T=4	T=5
Within Class Variance	$\sigma^2_W = 3.1196$	$\sigma^2_W = 1.5268$	$\sigma^2_W = 0.5561$	$\sigma^2_W = 0.4909$	$\sigma^2_W = 0.9779$	$\sigma^2_W = 2.2491$
Between Class Variance	$\sigma^2_B = 0$	$\sigma^2_B = 1.5928$	$\sigma^2_B = 2.5635$	$\sigma^2_B = 2.6287$	$\sigma^2_B = 2.1417$	$\sigma^2_B = 0.8705$

3.3 Morphology Transformation

Although Otsu is a powerful technique to segment the image object but some pixel in the boundary between object and background is not good handled enough. To recover this issue, morphology technique is applied. This technique is also generally used in many proposes such as, image pre-processing, enhancing object structure, image segmentation and quantitative description of object. In this work, Morphological Opening [7] is applied to binary image from thresholding process in order to improve quality of object's boundary. To understand Morphological Opening, Morphological Dilation and Erosion [7] are needed to understand as well.

3.3.1 Morphological Dilation

Morphological Dilation combines vector information of two images for possible vector additions. It is represented with Equation 3.13, where X and B are sets of image vector, and p is image pixel. This method is used for filling the small holes and gulfs in objects and also increasing the object size. It is not good for the work that original size is required.

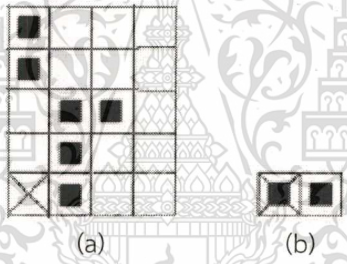
$$X \oplus B = \{p \in \varepsilon^2 : p = x + b, x \in X \text{ and } b \in B\} \tag{3.13}$$

Example,

X = {(1,0), (1,1), (1,2), (2,2), (0,3), (0,4)}

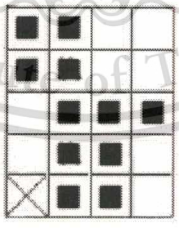
B = {(0,0), (1,0)}

$X \oplus B$= {(1,0), (1.1), (1,2), (2,2), (0,3), (0,4), (2,0), (2,1), (2,2), (3,2), (1,3), (1,4)}



(a) (b)

Figure 3.5 (a) Set X, (b) Set B[7]



(c)

Figure 3.5 (c) Set $X \oplus B$[7]

3.3.2 Morphological Erosion

Morphological Erosion also uses vector information of the images similar to Morphological Dilation but the calculation procedure is difference. This technique, the subtraction vectors are considered instead of the possible addition vectors.

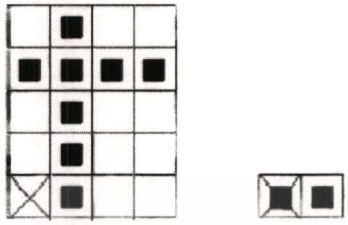
$$X \ominus B = \{p \in \varepsilon^2 : p = x + b \in X \text{ for every } b \in B\} \tag{3.14}$$

Example,

X= {(1,0), (1,1), (1,2), (0,3), (1,3), (2,3),(3,3),(1,4)}

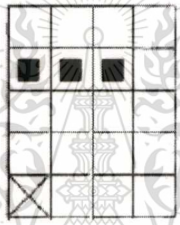
B = {(0,0), (1,0)}

$X \ominus B$= {(0,3), (1,3), (3,3)}



(a) (b)

Figure 3.6 (a) Set X, (b) Set B[6]



(c)

Figure 3.6 (c) Set$X \ominus B$[7]

3.3.3 Morphological Opening

Morphological Opening is the method to resolve variation of object's size from Morphological Dilation and Erosion process. This method is composed of 2 operations which are erosion then following with dilation. The result of Equation 3.15 is Morphological Opening on image X by element$B$.

$$X \circ B = (X \ominus B) \oplus B \tag{3.15}$$

3.4 Kmeans Image Clustering

Kmeans image clustering [16] [17] is the technique to classify image object into multiple classes and. In this work, it is proposed to separate particle into k cluster by following 4 main steps. The data points around centroids are minimized the sum-of-squares criterion

$$V = \sum_{i=1}^{k} \sum_{x_j \in S_i} (x_j - \mu_i)^2 \tag{3.16}$$

where x_j is a vector representing the *jth*data point and μ_i is the geometric centroid of the data points in S_i.

Step 1 : The number of cluster (k) is assigned.

Step 2 : The object data is assigned into each clusters base on the number of cluster and its distance from the centroid. In this step, the centroid value is initialized by using random value from the intensity distribution.

Step 3 :The centroid of each cluster is recomputed by respecting to the centroid intensity ($c^{(i)}$).

$$(c^{(i)}) = arg\ min_j \left\| x^{(i)} - \mu_j \right\|^2 \tag{3.17}$$

Step 4 : The new centroid value is iterated until meet the criterion.

$$\mu_i = \frac{\sum_{i=1}^{m} 1\{c_{(i)} = j\} x^{(i)}}{\sum_{i=1}^{m} 1\{c_{(i)} = j\}} \tag{3.18}$$

The advantages of Otsu, Morphological Opening and Kmeans Image Clustering in this chapter are leaded to develop contaminated particle detection algorithm in Chapter 4.

CHAPTER 4

Contaminated Particle Detection

Contaminated particle detection algorithm is developed to measure particle's quantity and size. Moreover, the feasibility to use OIA for material identification is studies by using this algorithm also. To design process flow in the algorithm, many image processing techniques are reviewed in Chapter 2 and 3. Finally, Otsu, Kmeans and Morphological Opening are the main techniques which are applied to the algorithm as shown in Figure 4.1.

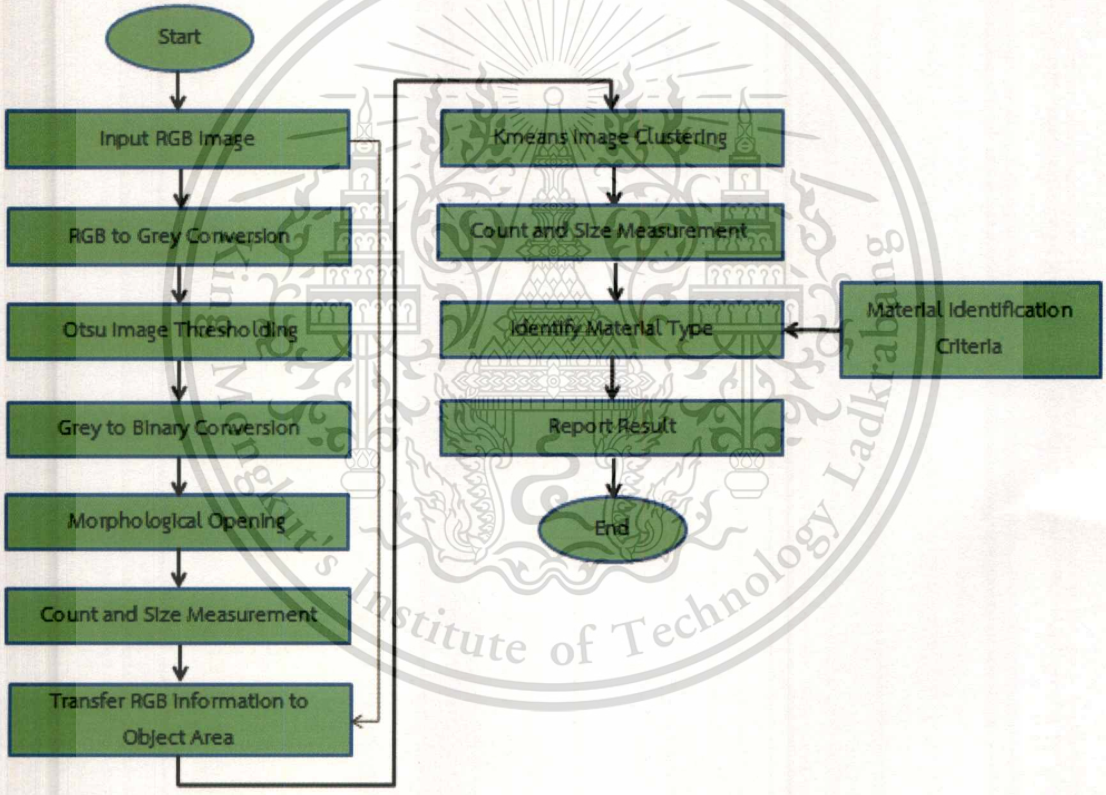


Figure 4.1 Flow chart of contaminated particle detection algorithm

4.1 Input RGB Image

The prepared samples are captured the image by optical microscope (OIA). The magnification of OIA is set at 200X with resolution 0.557 × 0.557 microns per pixel. The original RGB image is shown in Figure 4.2.

Figure 4.2 Original RGB image

4.2 RGB to Grey Conversion

RGB image from OIA is converted to grey scale image as show in Figure 4.3. The grey scale information is prepared for Otsu image thresholding process.

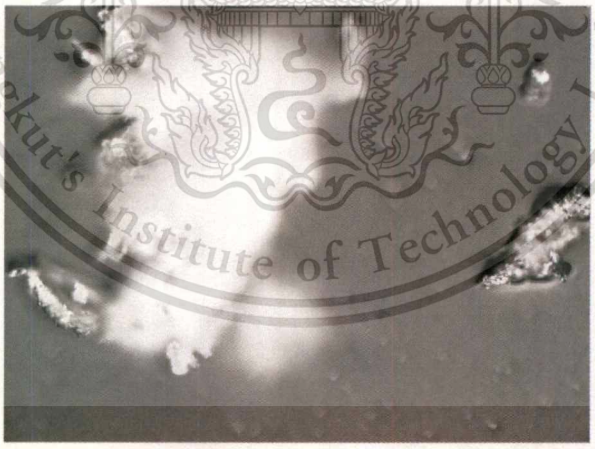


Figure 4.3 Grey scale image

4.3 Otsu Image Thresholding

Otsu technique is applied to evaluate the optimal threshold between object and background distribution. Figure 4.4 shows histogram of image intensity. The red line or threshold 186 is the optimal value which is used to separate grey level into 2 classes for object and background.

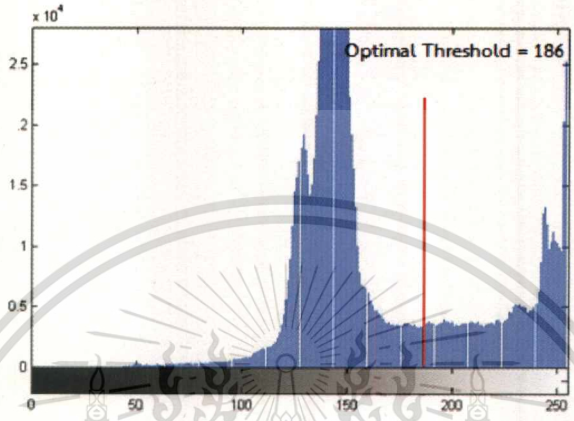


Figure 4.4 Image histogram with optimal threshold value

4.4 Grey to Binary Conversion

The optimal threshold value from the previous process is applied to segment contaminated particle and image background into black and white pixel as binary image in Figure 4.5.

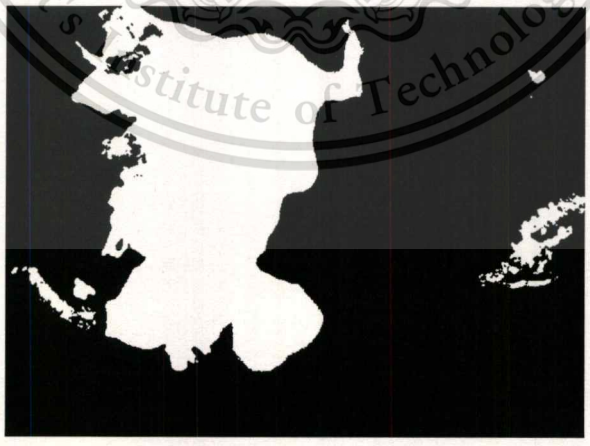


Figure 4.5 Binary image

4.5 Morphological Opening and Fill Hole

Morphological Opening is applied to improve quality of object boundary. The lost pixels are resolved by filling all holes inside the object boundary. The result of binary image after Morphological Opening and fill hole process is shown in Figure 4.6.

Figure 4.6 Image after Morphological Opening and fill hole

4.6 RGB Information to Object Transformation

To classify image object or contaminated particle into multi clusters, RGB information is needed. Hence RGB information from input image process is mapped and replaced over the pixel which is identified as particle. The result is shown in Figure 4.7.

Figure 4.7 Image after transferring RGB information to object area

4.7 Kmeans Image Clustering

Kneans image clustering is used to divide image object into multi clusters. Particles are classified into cluster 1 and cluster 2 as shown in Figure 4.8 and Figure 4.9 respectively.

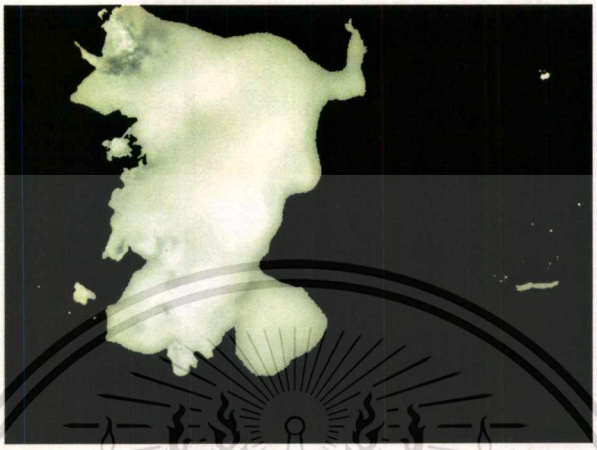


Figure 4.8 Cluster1 from Kmeans image clustering process

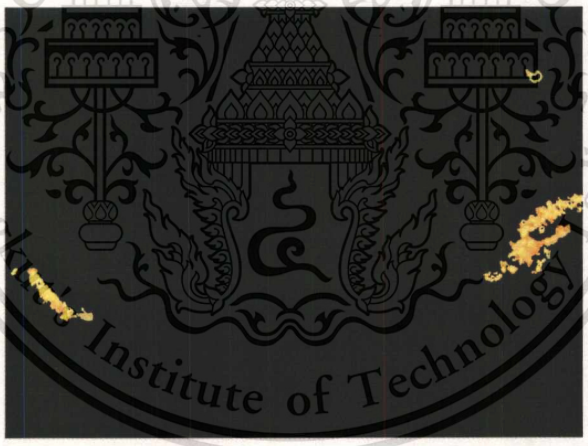


Figure 4.9 Cluster2 from Kmeans image clustering process

4.8 Object Counting and Sizing

Particle's quantity and size are evaluated and displayed for image thresholding result and image clustering result.

4.9 Material Identification

A number of material of particle, i.e. gold, stainless steel (SST) and tray materials are set in the experiments for identification propose. The characteristics if these materials are investigated in order to be identification criteria in the algorithm. In sample preparation process, gold stainless steel and tray material powder are doped on carbon tape as shown in Figure 4.10. Image of the samples are capture on OIA 200X to analyze RGB baseline by using MATLAB. The results of material baseline are shown in Table 4.1.

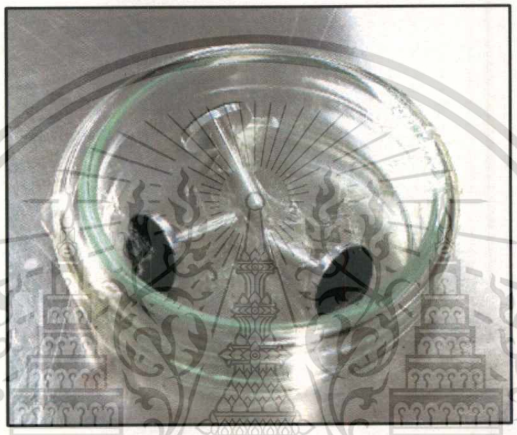


Figure 4.10 Samples of gold, SST and tray powder on carbon tape

The results of RGB baseline of gold, stainless steel and tray material in Table 4.1, it shows that the ranges in B Chanel of these three materials are difference. Hence, the ranges of RGB baseline are set as criteria to identify material on the algorithm. Moreover, the hypothesis tests by using 2-Sample *t*are used to investigate the difference of mean of RGB baseline also.

Table 4.1 RGB baseline of material

Material	Sample Size	R				G				B			
		Mean	Median	StDev	Range	Mean	Median	StDev	Range	Mean	Median	StDev	Range
Gold	9	235.65	234.96	9.71	219-248	205.98	210.5	12.39	182-218	105.89	107.95	8.97	91-119
Stainless steel	21	227.83	230.15	10.97	194-243	241.14	243.13	6.41	220-250	246.75	248.03	3.79	234-252
Tray	20	210.53	213.46	14.53	181-235	228.65	233.86	12.31	202-247	202.07	207.02	17.89	165-233

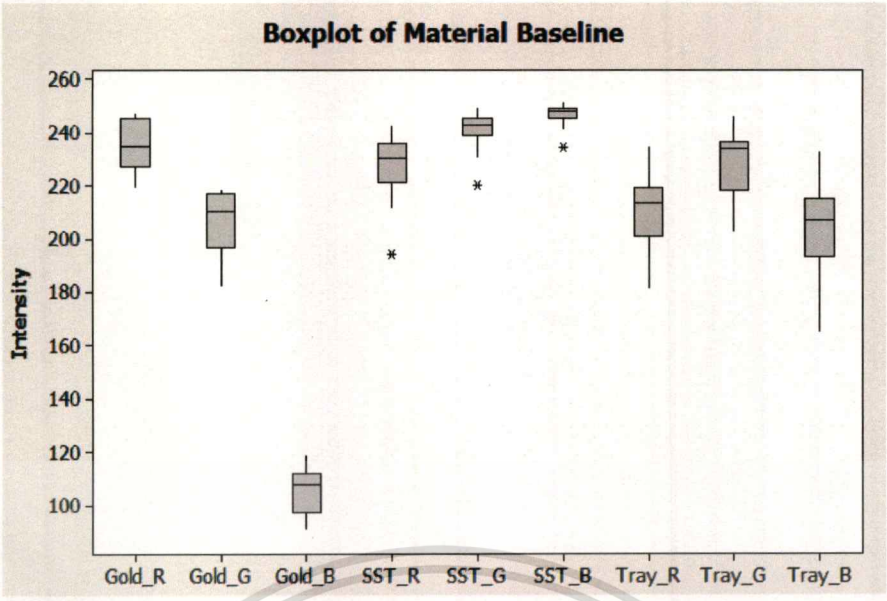


Figure 4.11 Boxplot of gold, SST and tray material

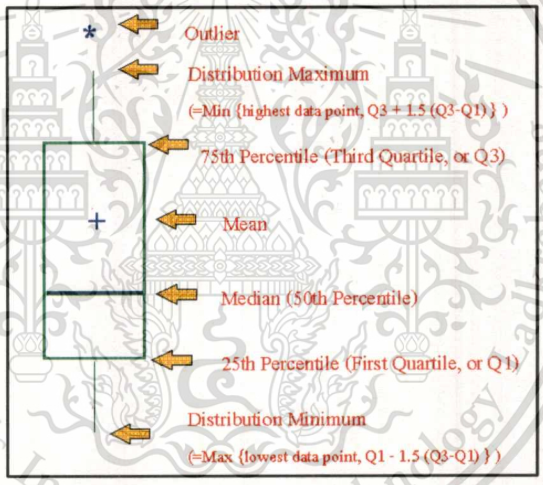


Figure 4.12 Boxplot structure [24]

Figure 4.11displays boxplot of the data in Table 4.1 in order to compare the different RGB information of gold, stainless steel and tray material.

- R and B Channel, boxplot presents that these materials have some related data.
- G Chanel, boxplot presents that these materials have unique characteristics.

The structure of boxplot is shown in Figure 4.12, the data of 3 quartiles or 25^{th} - 75^{th} percentile is shown in the box, and the 0 - 25^{th} percentile and 75^{th} - 100^{th} percentile of data are represented by bottom and top line respectively, and the

median line in the box is represented the 50th percentile of data. Lastly, the data outlier is represented by star.

4.9.1 Hypothesis Test

To consider the difference of RGB means of three materials, hypothesis tests using 2 - Sample *t* are investigated. The hypothesis is set mean 1 = mean 2. The conclusion is true, if mean 1 - mean 2 = 0. If not, the difference of mean is significantly larger or smaller than 0, the conclusion is means different. According to sample size is limited. The power of tests are investigated the risk to accept error or β in Figure 4.14.

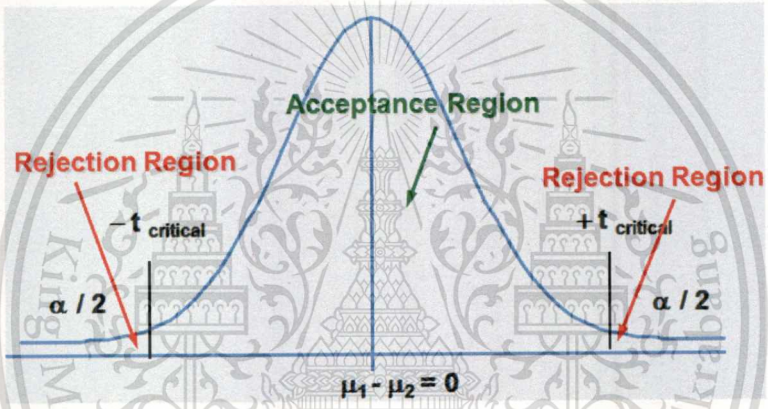


Figure 4.13 Normal distribution with acceptance region and rejection region

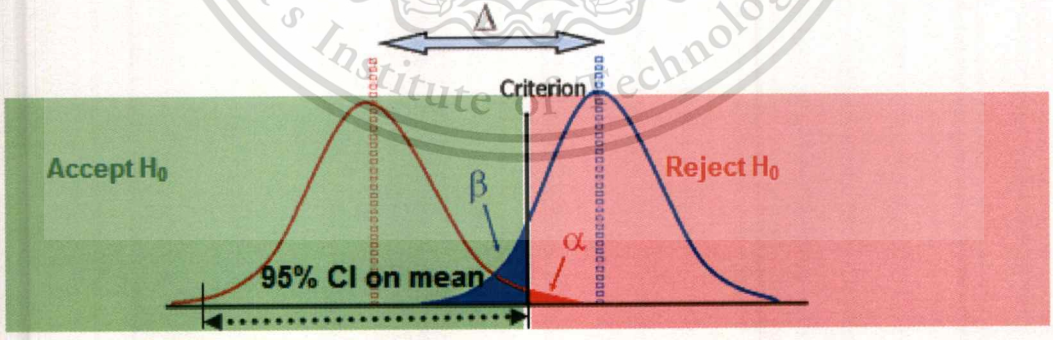


Figure 4.14 Risk to accept error (β) in hypothesis test

In experiments there are 3 materials with RGB data. The comparisons are set into 3 cases with 9 hypotheses. These hypotheses are tested by 2-Sample *t* with alpha error (α) 0.05.The summary of hypothesis tests is showed in Table 4.2. The results show that only case of R Chanel between gold VS stainless steel has p-Value

greater than 0.05 and the power of test is only 50%. Hence this case can't reject the hypothesis.

Table 4.2 Summary of hypothesis tests

Case	Hypothesis	α	p-Value	Power of Test (1-β)
Gold vs Stainless steel(SST)	(R Chanel) Mean Gold = Mean SST	0.05	0.069	50%
	(G Chanel) Mean Gold = Mean SST	0.05	0	99%
	(B Chanel) Mean Gold = Mean SST	0.05	0	> 99%
Gold vs Tray	(R Chanel) Mean Gold = Mean Tray	0.05	0	> 99%
	(G Chanel) Mean Gold = Mean Tray	0.05	0	98%
	(B Chanel) Mean Gold = Mean Tray	0.05	0	> 99%
Stainless steel(SST) vs Tray	(R Chanel) Mean SST = Mean Tray	0.05	0	99%
	(G Chanel) Mean SST = Mean Tray	0.05	0	97%
	(B Chanel) Mean SST = Mean Tray	0.05	0	> 99%

Case # 1 Comparison between gold and stainless steel material

- (R Chanel) Equal test between gold and stainless steel: 2-Sample *t* shows not significant different mean with p-Value = 0.069. But the power of test refer to the sample size in Table 4.1 is only 50%. Boxplot also presents that many data are related between two materials.

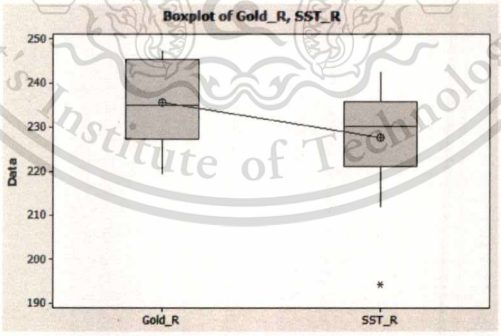


Figure 4.15 Boxplot of R Chanel of gold VS stainless steel (SST)

- (G Chanel) Equal test between gold and stainless steel: 2-Sample *t* shows significant different mean with p-Value = 0. Power of test refer to the sample size in Table 4.1 is 99%. Boxplot presents that the data are not related between two materials.

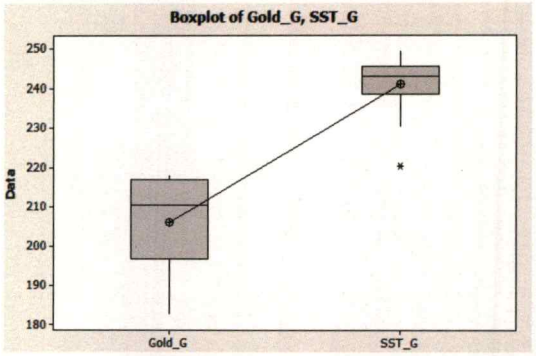


Figure 4.16 Boxplot of G Chanel of gold VS stainless steel (SST)

- (B Chanel) Equal test between gold and stainless steel: 2-Sample *t* shows significant different mean with p-Value = 0. Power of test refer to the sample size in Table 4.1 is greater than 99%. Boxplot presents that the data are not related between two materials.

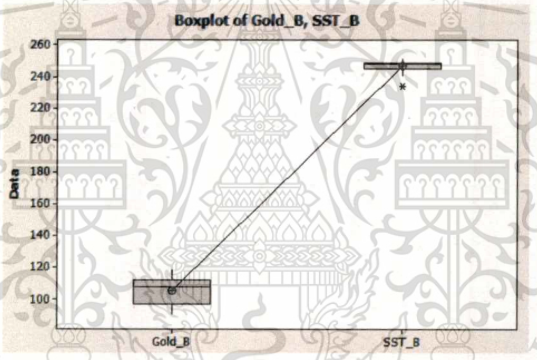


Figure 4.17 Boxplot of B Chanel of gold VS stainless steel (SST)

Case#2 Comparison between gold and tray material

- (R Chanel) Equal test between gold and tray: 2-Sample *t* shows significant different mean with p-Value = 0. Power of test refer to the sample size in Table 4.1 is greater than 99%. But boxplot presents that some data is related between two materials.

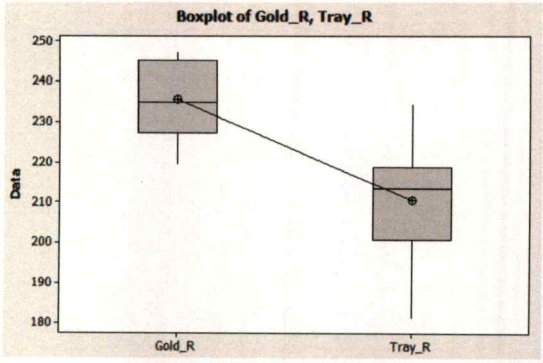


Figure 4.18 Boxplot of R Chanel of gold VS tray

- (G Chanel) Equal test between gold and tray: 2-Sample *t* shows significant different mean with p-Value = 0. Power of test refer to the sample size in Table 4.1 is 98%. But boxplot presents that some data is related between two materials.

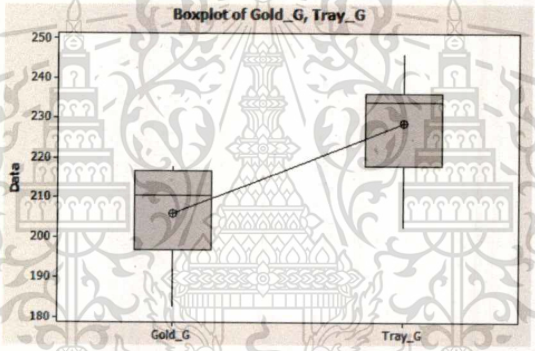


Figure 4.19 Boxplot of G Chanel of gold VS tray

- (B Chanel) Equal test between gold and tray: 2-Sample *t* shows significant different mean with p-Value = 0. Power of test refer to the sample size in Table 4.1 is greater than 99%. Boxplot presents that the data are not related between two materials.

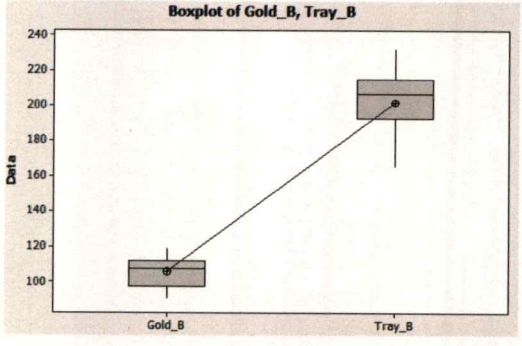


Figure 4.20 Boxplot of B Chanel of gold VS tray

Case#3 Comparison between stainless steel and tray material

- (R Chanel) Equal test between SST and tray: 2-Sample *t* shows significant different mean with p-Value = 0. Power of test refer to the sample size in Table 4.1 is 99%. But boxplot presents that some data is related between two materials.

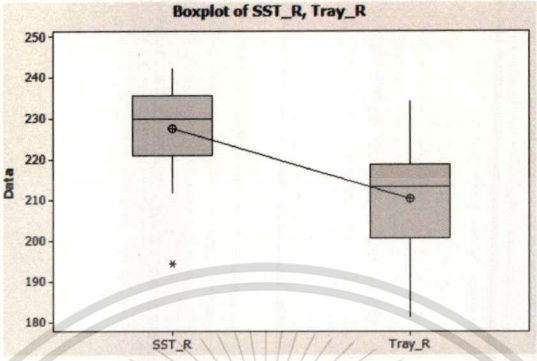


Figure 4.21 Boxplot of R Chanel of SST VS tray

- (G Chanel) Equal test between SST and tray: 2-Sample *t* shows significant different mean with p-Value = 0. Power of test refer to the sample size in Table 4.1 is 97%. But boxplot presents that some data is related between two materials.

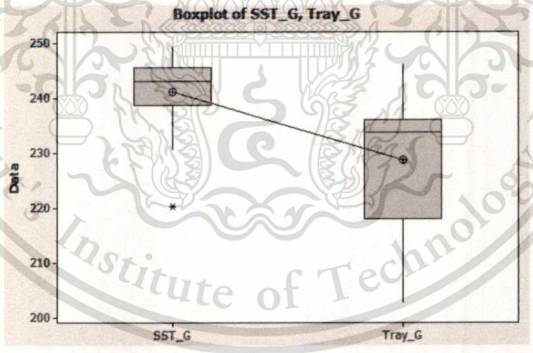


Figure 4.22 Boxplot of G Chanel of SST VS tray

- (B Chanel) Equal test between SST and tray: 2-Sample t shows significant different mean with p-Value = 0. Power of test refer to the sample size in Table 4.1 is greater than 99%. Boxplot also presents that the data are not related between two materials.

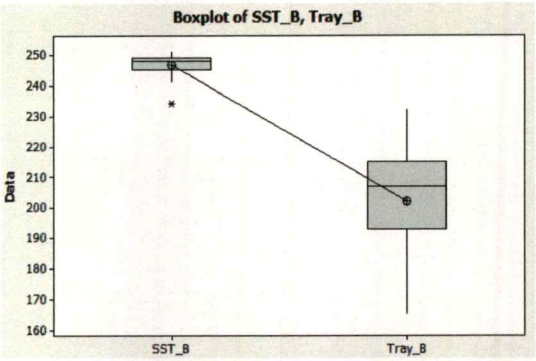


Figure 4.23 Boxplot of B Chanel of SST VS tray

4.10 Report Result

Finally, there are 3 results in display. Figure 4.24shows the display result of total particle detection from image thresholding process. The results of Kmeans image clustering are showed in Figure 4.25 and Figure 4.26 for cluster 1 and cluster 2, respectively.

- Figure 4.24 presents that there are total 3 particles detected from image thresholding process. Two particles have size smaller than 15000 pixels and another particle has size about 72000 pixels.

- Figure 4.25 presents that there are 2 expected gold particles. The particle sizes are about 5200 and 13000 pixels.

- Figure 4.26 presents that there are 39 expected tray particles. These particles have many sizes. However, the biggest size is about 2300 pixels.

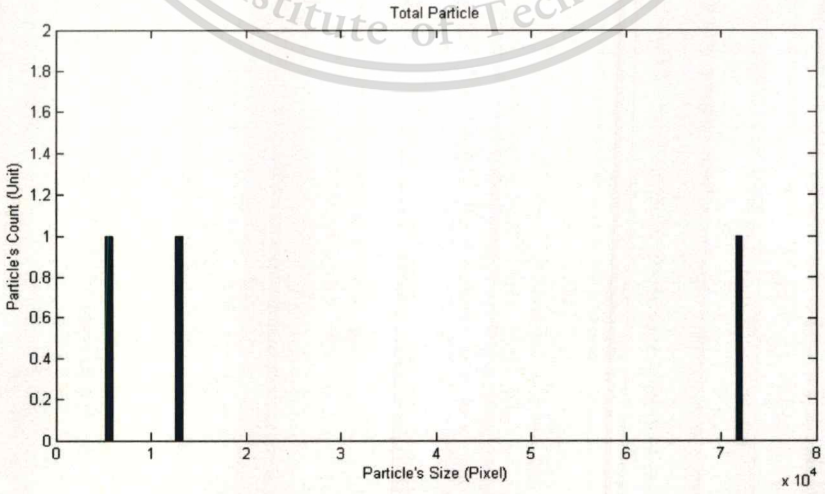


Figure 4.24 Display of image thresholding result

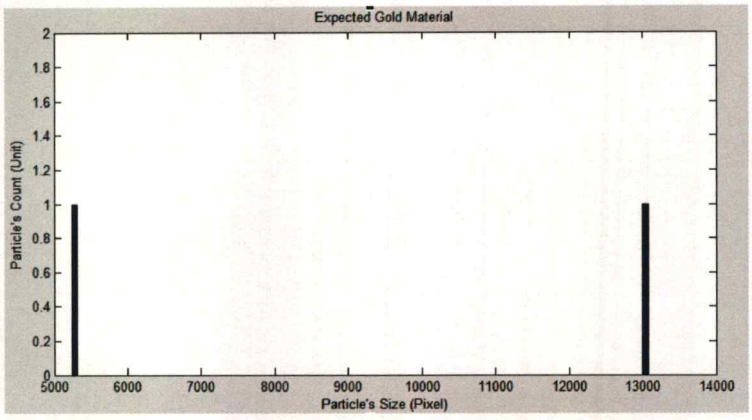


Figure 4.25 Display result of cluster 1

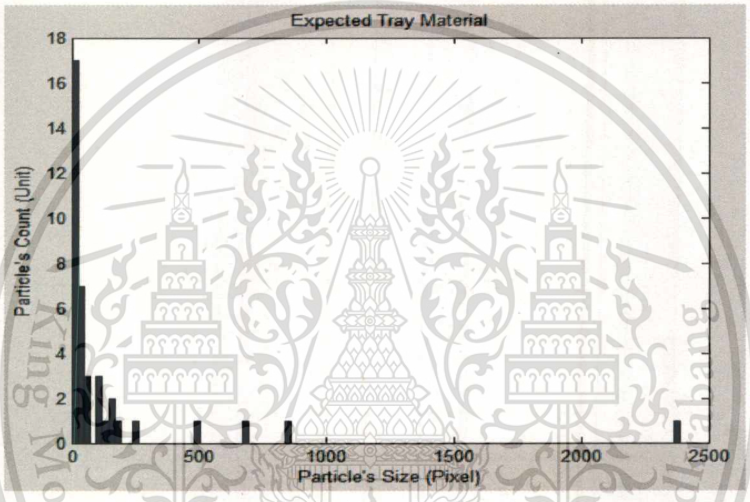


Figure 4.26 Display result of cluster 2

CHAPTER 5

Results and Conclusion

5.1 Results and Discussion

The experiments are set for finding particle's size, quantity and type of material. These results are used in particle analysis, particle risk assessment and particle mapping in order to screen particle on incoming components, identify particle generation in HDD process or particle removal in cleaning process. The device used in these experiments is OIA; to be precise the magnification used for OIA is 200X which provides pixel resolution at 0.557 × 0.557 microns.

5.1.1 Particle's Quantity and Material Identification

The size of particle to be focused in the experiments is 25 microns and above. The samples used for experiments are carbon tapes which are intentionally doped with gold, stainless steel and tray particle in clean room environment. Human eyes-based measurement with OIA 200X results would be set as references since it is a present method to screen particle defected of some incoming component, and all the samples are controlled the material of particle as well. The experiments are set into 3 cases which are gold and tray material, gold and stainless steel material, stainless steel and tray material.

Case#1 gold and tray material

The amount of particles which are counted manually by human eyes-based are set as reference. The algorithm automatically preforms the display of total particle's quantity and size by Otsu, and the display of particle's material with its quantity and size by two stages of Otsu and Kmeans. The results in Table 5.1 show that the Otsu stage gives satisfy results. The average difference is about 3% comparing with human eye based measurement. In addition, it takes faster operating time about 60%. For material identification results, the algorithm can be classified both gold and tray particle. However, total particle's quantity after material identification shows relatively high difference.

To consider the increasing of particle'squantity in material identification process, image cluster1 (Figure 5.2(a)) and image cluster2 (Figure 5.3(b)) show that the

objects are split into many components. Hence particle's quantity and size after material identification can't be used due to relatively high error.

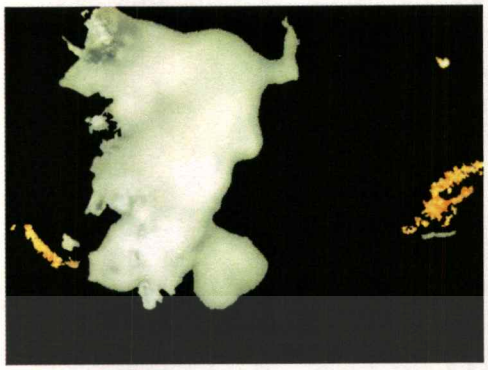


Figure 5.1 Objects before Kmeans classification (Case#1)

(a) (b)

Figure 5.2 (a) Kmeans Cluster 1 (Case#1)
(b) Kmeans Cluster 2 (Case#1)

Table 5.1 Results of Case#1 gold and tray material

Sample	Reference: OIA 200X (Human Eye)	OIA 200X (Otsu)		OIA 200X (Otsu+Kmeans)					
				Cluster 1		Cluster2		Total	%Difference
	Measured (counts)	Measured (counts)	%Difference	Material	Measured (counts)	Material	Measured (counts)	(Cluster1+2)	
1	5	5	0%	Gold	2	Tray	12	14	180%
2	6.5	7	8%	Gold	16	Tray	73	89	1269%
3	5	5	0%	Gold	4	Tray	3	7	40%
4	5	5	0%	Gold	29	Tray	18	47	840%
5	3	3	0%	Gold	1	Tray	2	3	0%
6	7	8	14%	Gold	5	Tray	16	21	200%
7	3	3	0%	Gold	6	Tray	5	11	267%
8	4.5	5	11%	Gold	8	Tray	31	39	767%
9	2	2	0%	Gold	2	Tray	7	9	350%
10	8	8	0%	Gold	4	Tray	10	14	75%
Avg %Difference		3%		399%					
Operating time	3 minutes/Sample	1 minute/Sample		1 minute/Sample					

Remark OIA with Human Eye: Three times and two operators measured on each sample

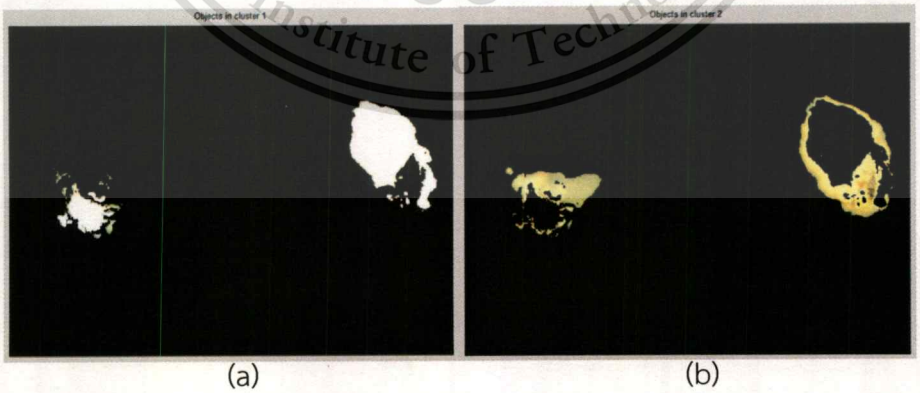
Table 5.1 shows the summary results of Case#1. Total particle's quantity from Otsu and Otsu + Kmeans process are different from human eye-based about 3% and 399%, respectively. For operating time, OIA with the algorithm takes only 1 minute/sample.

Case#2 gold and stainless steel material

The measurement method is set similar toCase#1. The summary results are shown in Table 5.2. It shows that total particle's quantity by OIA with Otsu and Otsu + Kmeans are different from human eyes-based about 13% and 1864%, respectively. For material identification results, this algorithm is not capable to classify gold with stainless steel on the sample.

Figure 5.4 shows that Kmeans image clustering is not capable to classify gold and stainless steel on the sample due to these materials have overlap intensity.

Figure 5.3 Objects before Kmeans classification (Case#2)



(a) (b)

Figure 5.4 (a) Kmeans Cluster 1 (Case#2)
(b) Kmeans Cluster 2 (Case#2)

Table 5.2 Result of Case#2 gold and SST material

Sample	Reference: OIA 200X (Human Eye)	OIA 200X (Otsu)		OIA 200X (Otsu+Kmeans)					
				Cluster 1		Cluster2		Total (Cluster1+2)	%Difference
	Measured (counts)	Measured (counts)	%Difference	Material	Measured (counts)	Material	Measured (counts)		
1	5	5	0%	Unknown	21	Unknown	28	49	880%
2	2	2	0%	Unknown	26	Unknown	34	60	2900%
3	8	8	0%	Unknown	12	Unknown	20	32	300%
4	4	4	0%	Unknown	22	Unknown	31	53	1225%
5	9	6	33%	Unknown	87	Unknown	29	116	1189%
6	2	3	50%	Unknown	36	Unknown	30	66	3200%
7	2	3	50%	Unknown	25	Unknown	14	39	1850%
8	3	3	0%	Unknown	43	Unknown	44	87	2800%
9	2	2	0%	Unknown	36	Unknown	28	64	3100%
10	5	5	0%	Unknown	28	Unknown	37	65	1200%
Avg %Difference		13%		1864%					
Operating time	3 minutes/Sample	1 minute/Sample		1 minute/Sample					

Remark OIA with Human Eye: Three times and two operators measured on each sample

Table 5.2 shows the summary results of Case#2. Total particle's quantity from Otsu and Otsu + Kmeans process are different from human eye-based about 13% and 1864%, respectively. For operating time, OIA with the algorithm takes only 1 minute/sample.

Case#3 stainless steel and tray material

The measurement method is set similar Case#1 and Case#2. The summary results are shown in Table 5.3. It shows that total particle's quantity by OIA with Otsu and Otsu + Kmeans are different from human eyes-based about 12% and 1743%, respectively. For material identification results, this algorithm is not capable to classify stainless steel with tray material on the sample.

Figure 5.6 shows that Kmeans image clustering is not capable to classify stainless steel and tray on the same sample due to these materials have overlap intensity.

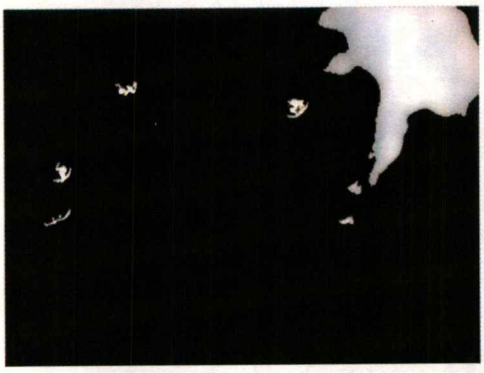


Figure 5.5 Objects before Kmeans classification (Case#3)

Figure 5.6 (a) Kmeans Cluster 1 (Case#3)
(b) Kmeans Cluster 2 (Case#3)

Table 5.3 Result of Case#3 stainless steel and tray material

Sample	Reference: OIA 200X (Human Eye)	OIA 200X (Otsu)		OIA 200X (Otsu+Kmeans)					
				Cluster 1		Cluster2		Total (Cluster1+2)	%Difference
	Measured (counts)	Measured (counts)	%Difference	Material	Measured (counts)	Material	Measured (counts)		
1	7	7	0%	Unnown	84	Tray	54	138	1871%
2	8.5	8	6%	Unnown	132	Tray	152	284	3241%
3	4	5	25%	Unnown	124	Tray	45	169	4125%
4	5	5	0%	Unnown	26	Tray	39	65	1200%
5	3	3	0%	Unnown	22	Tray	15	37	1133%
6	9	8	11%	Unnown	33	Tray	42	75	733%
7	2	3	50%	Unnown	18	Tray	14	32	1500%
8	7.5	7	7%	Unnown	42	Tray	32	74	887%
9	5	5	0%	Unnown	54	Tray	22	76	1420%
10	5	4	20%	Unnown	29	Tray	42	71	1320%
Avg %Difference		12%		1743%					
Operating time	3 minutes	1 minute		1 minute					

Remark OIA with Human Eye: Three times and two operators measured on each sample

Table 5.3 shows the summary results of Case#3. Total particle's quantity from Otsu and Otsu + Kmeans process are different from human eye-based about 12% and 1743%, respectively. For operating time, OIA with the algorithm takes only 1 minute/sample.

5.1.2 Size Measurement

The experiment is set to study accuracy of size measurement in order to bring this information to use with particle detection work. The sizing standards which are 10, 25, 50 and 100 micronsare set in the experiment. These four standards are measured its size by using OIA 200X with the algorithm. The results in Table 5.4 show that the maximum error is 5.31% on standard size 10 microns. Moreover, this study is shown that OIA with the algorithm is capable to support the particle's size 25 microns and above as the setting of experiments in Section 5.1.1.

However, particle's size in the result displays of Section 5.1.1. are found some variation due to some lost pixel in image thresholding process and some process in algorithm to improve quality of an image object. Although the results of particle's size have some variation, but OIA with the algorithm gives faster operating time more than SEM measurement which is used to measure particle's size in HDD manufacturing. Hence OIA with the algorithm can be used to support some work in

HDD manufacturing but it is needed to optimize and develop by respect to the application's requirement.

Table 5.4 Sizing measurement results

Sample	Sizing Standard (Microns)	OIA 200X (Otsu)	
		Measured (microns)	%error
1	100	100.00	0.00%
2	50	49.50	1.00%
3	25	24.51	1.96%
4	10	9.47	5.31%

Remark Operating time is one minute per sample.

5.2 Conclusion

The algorithm is developed to measure particle's quantity, size and to identify type of material. Various image processing techniques are reviewed, i.e. Ridler and Calvard algorithm, Kapur algorithm, Fan algorithm, Max Min thresholding, Gradient Watershed, Otsu algorithm, Kmeans image clustering, Morphology transformation. Among these, Otsu algorithm is selected to segment total particle from the image background, Morphology transformation is used to improve border quality of particle detected, Kmeans image clustering is applied to classify particle into multi clusters in order to identify the material. Moreover, the characteristic of materials are investigated for identification feasibility as well. Gold, stainless steel and tray materials are set in the experiments due to these materials are found in many particle failures in HDD manufacturing. In addition, these materials also related with some assemble process.

At the present, it takes relatively long time to get the result of particle's material due to SEM is only main measurement. SEM measurement has a long operating time and also expensive operating cost compare with OIA measurement. In other hand, OIA with human eyes-based measurement is implemented to screen contaminated particle on some incoming of component, such as operation to screen incoming of slider in HGA (Head Gimbal Assembly) process. It is manually operated by human operator. Hence, OIA with human eyes-based measurement can gives only

the result of particle's quantity and it takes long operating time to count a huge number of particles. Moreover, human error is unpredicted factor in measurement's result.

The experiment's results, it can be seen that OIA with the algorithm is capable to identify gold and tray particles. It also gives an acceptable error on total particle count as well. For information of the particle's size, although it has small error to measure standard sizing but in the application on real particles which are gold, stainless steel and tray, it is found a variation from image thresholding and some process which is used to improve quality of image object as well. The results also show that OIA with the algorithm is not capable to classify both gold with stainless steel particles and stainless steel with tray particles.

Lastly, this work presents the opportunity to develop and manage OIA equipment in HDD manufacturing for operating time, operating cost and human error reduction.

REFERENCES

Z. Ping, L. Yongkui. "Grain Counting Method Based on Image Processing" **Information Engineering and Computer Science**, 2009.

A. Maythapolnun. "A Review on Global Binarization Algorithms for Degraded Document Images" **AU Journal of Technology**, 2011. Pp. 188-195.

K. Torabi, S. Sayad, S.T. Balke. "Adaptive Image Thresholding for Real-Time Particle Monitoring" **International Journal of Imaging Systems and Technology**, Vol. 16, 2006. Pp. 9-14.

M.H. Pi, H. Zhang. "Two-Stage Image Segmentation by Adaptive Thresholding and Gradient Watershed" **Proceedings of The Second Canadian Conference on Computer and Robot Vision**, 2005.

W. Wang. "Fast Measuring Particle Size by Using the Information of Particle Boundary and Shape" **International Conference on Machine Learning and Computing**, 2006. Pp. 1086-1095.

N. Otsu. "A Threshold Selection Method from Gray-Level Histograms" **IEEE Transaction on Systems Man andCybernetics**, 1979. Pp. 62-66.

S. Milan, H. Vaclav, B. Roger. "Image Processing Analysis, and Machine Vision" 2008.

T.W. Ridler, E.S. Calvard. "Picture Thresholding Using an Iterative Selection Method" **IEEE Transactions on Systems Man and Cybernetics**, vol. 8, 1978. Pp. 629-632.

P.S. Liao,T.S. Chen, P.C. Chung. "A Fast Algorithm for Multilevel Thresholding" **Journal of Information Science and Engineering**, vol. 17, 2001. Pp. 713-727.

W.A. Yasnoff, J.K. Mui, J.W. Bacus "Error Measure for Scene Segmentation" **Pattern Recognition**, Vol. 9, 1977. Pp. 217-231.

P.K. Sahoo, S Soltani, A.K. Wong, Y.C. Chan. "A Survey of Thresholding Techniques. Computer Vision" **Graphics and Image Processing**, vol. 41, 1988. Pp. 233-260.

Y. Xiao, Z.G. Cao, T.X. Zhang. "Entropic Thresholding Based on Gray-Level Spatial Correlation Histogram" **Proc. 19th Int. Conf. On Pattern Recognition (ICPR 2008)**, Tampa, FL, USA, 2008.

J. Kittler, J. Illingworth. "Minimum Error Thresholding" **Pattern Recognition**, vol. 19, 1986. Pp. 41-47.

S.U. Lee, S.Y. Chung, R.H. Park. "A Comparative Performance Study of Several Global Thresholding Technique for Segmentation Computer Vision" **Graphics and Image Processing,** vol. 52, 1990. Pp. 171-190.

S. Mehmet, S. Bulent. "Survey Over Image Thresholding Technique and Quantitative Performance Evaluation" **Journal of Electronic Imaging,** 2004. Pp. 146-165.

S. Tatiraju, A. Mehta."**Image Segmentation Using K-means Clustering EM and Normalized Cuts**" [Online]. Available: www.ics.uci.edu.

M.C.J. Christ, R.M.S. Parvathi. "Segmentation of Medical Image Using K-Means Clustering and Marker Controlled Watershed Algorithm" **European Journal of Scientific Research,** vol. 71, Pp. 190-194.

N. Laitinen, O. Antikainen, J. Yliruusi. "Characterization of Particle Sizes in Bulk Pharmaceutical Solids Using Digital Image Information", **AAPS PharmSciTech,** 2003.

D.J. Graham, I. Reid, S.P. Rice. "Automated Sizing of Coarse-Grained Sediments: Image-Processing Procedures" **Mathematical Geology,** vol. 37, 2005.

B.N. Saha, N. Ray, H. Zhang. "Computing Oil Sand Particle Size Distribution by Snake-PCA Algorithm" **International Conference on Acoustics, Speech and Signal Processing,** 2008.

A.Z. Chitade, S.K. Katiyar. "Colour Based Image Segmentation Using K-Means Clustering" **International Journal of Engineering Science and Technology,** vol. 2, 2010.

A.S. Samma, R.A. Salam. "Adaptation of K-Means Algorithm for Image Segmentation" **World Academy of Science, Engineering and Technology 50,** 2009.

A. Greensted, "**Otsu Thresholding Explained**" [Online]. Available: www.labbookpages.co.uk/software/imgProc/otsuThreshold.html

C. Srikaewsiw, "**Statistics: Graphical Data Presentation**" [Online]. Available: www.sites.google.com/site/mystatistics01/chapter2/graphical-data-presentation

APPENDIX

Contaminated Particle Detection Algorithm

```
%%Clear all and clear screen
clear all
close all
clc

%%Input image and transfer to grey scale
A = imread('Image');
B = rgb2gray(A);
figure
imshow(A);
figure
imshow(B);

%%Image thresholding to define background and object
O = graythresh(B);              %% Otsu
C1 = im2bw(B,O);                %% Grey scale to binary conversion
figure
imshow(C1);
C2 = bwareaopen(C1,50);         %% Morphological Opening
Figure
imshow(C2);
C = imfill(C2,'holes');         %% Fill hole on object
Figure
imshow(C);

%%Evaluate particle's size and number
countO = bwconncomp(C,8);           %% Evaluate particle's number
countO.NumObjects
sizeO = regionprops(countO,'basic');  %% Evaluate particle's size
size_areasO = [sizeO.Area];
figure
hist(size_areasO,100),title('Total Particle')
```

%%Transfer RGB information from input image to the pixel which was identified to be particle

```
[r,c] = find(C==0);                    %% Identify background pixel
D = A;
ind = sub2ind(size(C),r,c);
D1 = D(:,:,1);
D2 = D(:,:,2);
D3 = D(:,:,3);
D1(ind) = 0;
D2(ind) = 0;
D3(ind) = 0;
D(:,:,1) = D1;
D(:,:,2) = D2;
D(:,:,3) = D3;
figure
imshow(D);
```

%%Kmeans image clustering

```
E = A;
cform = makecform('srgb2lab');
lab_E = applycform(E,cform);
%%
ab = double(lab_E(:,:,2:3));
nrows = size(ab,1);
ncols = size(ab,2);
ab = reshape(ab,nrows*ncols,2);
nColors = 2;                           %% Number of cluster
[cluster_idxcluster_center] = kmeans(ab,nColors,'distance','sqEuclidean','Replicates',2);
pixel_labels = reshape(cluster_idx,nrows,ncols);
imshow(pixel_labels,[]), title('image labeled');
segmented_images = cell(1,3);
rgb_label = repmat(pixel_labels,[1 1 3]);
for k = 1:nColors
color = D;
color(rgb_label~=k) = 0;
segmented_images{k} = color;
end
figure
```

```
imshow(segmented_images{1}), title('Objects in cluster 1');  %% Display cluster1
figure
imshow(segmented_images{2}), title('Objects in cluster 2');  %% Display cluster2

%%Evaluate particle's size and number in cluster1
AAA = segmented_images{1};
BBB = rgb2gray(AAA);
TTT = graythresh(BBB);
CCC = im2bw(BBB,TTT);
count1 = bwconncomp(CCC,8);        %%Evaluate particle's number in cluster1
count1.NumObjects
%% Sizing Kmeans_1
size1 = regionprops(count1,'basic');   %% Evaluate particle's size in cluster1
size_areas1 = [size1.Area];
figure,hist(size_areas1,100)

%%Evaluate particle's size and number in cluster2
AAAA = segmented_images{2};
BBBB = rgb2gray(AAAA);
TTTT = graythresh(BBBB);
CCCC = im2bw(BBBB,TTTT);
count2 = bwconncomp(CCCC,8);       %%Evaluate particle's number in cluster2
count2.NumObjects
size2 = regionprops(count2,'basic');   %%Evaluate particle's size in cluster2
size_areas2 = [size2.Area];
figure,hist(size_areas2,100)

%%Identify material in cluster1 and cluster2
MMM = AAA;
MMMR = MMM(:,:,1);
MMMG = MMM(:,:,2);
MMMB = MMM(:,:,3);
MMMVR = MMMR(:);
MMMVG = MMMG(:);
MMMVB = MMMB(:);
idxMMMR = find(MMMR~=0);
idxMMMG = find(MMMG~=0);
idxMMMB = find(MMMB~=0);
```

```
M1R = mean(MMMVR(idxMMMR));          %% Mean of R-Chanel in cluster1
M1G = mean(MMMVG(idxMMMG));          %% Mean of G-Chanel in cluster1
M1B = mean(MMMVB(idxMMMB));          %% Mean of B-Chanel in cluster1
MMMM = AAAA;
MMMMR = MMMM(:,:,1);
MMMMG = MMMM(:,:,2);
MMMMB = MMMM(:,:,3);
MMMMVR = MMMMR(:);
MMMMVG = MMMMG(:);
MMMMVB = MMMMB(:);
idxMMMMR = find(MMMMR~=0);
idxMMMMG = find(MMMMG~=0);
idxMMMMB = find(MMMMB~=0);
M2R = mean(MMMMVR(idxMMMMR));        %% Mean of R-Chanel in cluster2
M2G = mean(MMMMVG(idxMMMMG));        %% Mean of G-Chanel in cluster2
M2B = mean(MMMMVB(idxMMMMB));        %% Mean of B-Chanel in cluster2
SWC1 = 0;
if (M1R> 219) && (M1R<248)
if (M1B> 91) && (M1B<119)
if (M1G>182) && (M1G<218)
        SWC1 = 1;
end
end
end
if (M1R> 194) && (M1R<243)
if (M1B> 234) && (M1B<252)
if (M1G>220) && (M1G<250)
        SWC1 = 2;
end
end
end
if (M1R> 181) && (M1R<235)
if (M1B> 165) && (M1B<233)
if (M1G>202) && (M1G<247)
        SWC1 = 3;
end
end
end
```

```
SWC2 = 0
if (M2R> 219) && (M2R<248)
if (M2B> 91) && (M2B<119)
if (M2G>182) && (M2G<218)
        SWC2 = 1;
end
end
end
if (M2R> 194) && (M2R<243)
if (M2B> 234) && (M2B<252)
if (M2G>220) && (M2G<250)
        SWC2 =2;
end
end
end
if (M2R> 181) && (M2R<235)
if (M2B> 165) && (M2B<233)
if (M2G>202) && (M2G<247)
        SWC2 =3;
end
end
end

%%Display the identification results from cluster1
switch SWC1
case 1
figure
hist(size_areas1,100), title('Expected Gold Material')

case 2
figure
hist(size_areas1,100), title('Expected SST Material')
case 3
figure
hist(size_areas1,100), title('Expected Tray Material')
otherwise
figure
hist(size_areas1,100), title('Expected Unknown Material')
```

```
end

%%Display the identification results from cluster2
switch SWC2
case 1
figure
hist(size_areas2,100), title('Expected Gold Material')
case 2
figure
hist(size_areas2,100), title('Expected SST Material')
case 3
figure
hist(size_areas2,100), title('Expected Tray Material')
otherwise
figure
hist(size_areas2,100), title('Expected Unknown Material')
end
```

Author's Biography

Name-Surname :	Mr. Chatchai Mahuntadechakul
Date of Birth:	July 11'1982
Place of Birth:	Bangkok, Thailand
Education:	2005 Bachelor of Electrical Engineering, Srinakharinwirot University

Work Experience :

2012 - Current	Senior FTD Process Engineer, Seagate Technology (Thailand) Ltd.
2008 - 2012	HSA Process Development Engineer, Seagate Technology (Thailand) Ltd.
2005 - 2008	ESD & Contamination Process Engineer, Western Digital (Thailand) Ltd.

Publication:

ChatchaiMahuntadechakul, ChaiwatNuthong, "HDD Contaminated Particle Detection using Optimal Thresholding Image Analysis", DST-CON2011