Investigation on Image Processing Techniques for Diagnosing Paddy Diseases

Nunik Noviana Kurniawati [#1], Siti Norul Huda Sheikh Abdullah [#2], Salwani Abdullah [#3], Saad Abdullah [*4]

[#] *Center for Artificial Intelligence Technology,*
Faculty of Information Science and Technology,
National University of Malaysia, 43600 UKM, Bangi, Selangor, Malaysia
[1]nunovia@yahoo.com
[2]mimi@ftsm.ukm.my, [3]salwani@ftsm.ukm.my

[*]*MARDI Alor Setar Station,*
PO BOX 105, Jalan Kuala Kedah,
05710Alor Setar, Kedah, Malaysia
[4]asaad@mardi.my

Abstract—**The main objective of this research is to develop a prototype system for diagnosing paddy diseases, which are Blast Disease (BD), Brown-Spot Disease (BSD), and Narrow Brown-Spot Disease (NBSD). This paper concentrates on extracting paddy features through off-line image. The methodology involves image acquisition, converting the RGB images into a binary image using automatic thresholding based on local entropy threshold and Otsu method. A morphological algorithm is used to remove noises by using region filling technique. Then, the image characteristics consisting of lesion type, boundary colour, spot colour, and broken paddy leaf colour are extracted from paddy leaf images. Consequently, by employing production rule technique, the paddy diseases are recognized about 94.7 percent of accuracy rates. This prototype has a very great potential to be further improved in the future.**

Keywords— **local entropy threshold, feature extraction, colour segmentation, paddy leaf diseases, production rule.**

I. Introduction

Paddy plantation is one of the most crucial agriculture activities in agricultural countries. Paddy is also one of the cereal crops and staple food to many people in the world including Malaysia as well as Asian countries. Paddy plantation is still threatened by many factors that make paddy-rice production become slow and less productive. It leads to hardly achieve the target of 10 tones per hectare for the national needs. One of the main factors is paddy disease.

A disease is an abnormal condition that injures the plant or causes it to function improperly. Diseases are readily recognized by their symptoms - associated visible changes in the plant. There are a lot of paddy diseases types, but this research focuses on three paddy diseases that have the same symptoms, which are Blast Disease (BD), Brown Spot Disease (BSD), and Narrow Brown Spot Disease (NBSD).

BD is caused by the fungus *Pyricularia grisea*. This fungal disease is estimated to cause production losses of US$55 million each year in South and Southeast Asia. BD is a deadly disease of rice plants, and it has widespread distribution in more than 80 countries [1]. The losses are even higher in East Asia and other more temperate rice growing regions around the world [2].

Small oval to circular brown spots are the first symptoms of BSD caused by the fungus *Bipolaris oryzae* (*Helminthosporium oryzae*). Most conspicuous symptoms of the disease occur on leaves and glumes of maturing plants.

NBSD is caused by the fungus *Cercospora janseana.* The disease varies in severity from year to year and usually becomes more severe as the rice approaches to maturity, causing premature death of leaves and leaf sheaths, premature ripening of kernels and lodging of plants. It decreases the market value of the grains because it causes grain discoloration and chalkiness, and reduces the milling recovery.

Traditionally, paddy farmers determine the type of disease as manual. The disease is visually detected by experienced growers who have the ability to detect subtle changes in plant colour or slight droop curls of plant leaves. Although paddy farmers are being trained by agricultural experts to recognize paddy diseases in order to ensure that early prevention or treatment is taken, but errors might occur in order to determine the type of diseases and control. However, this method is laborious and time consuming, and it is impossible to accurately estimate the infected areas and severity in large-scale farming [3]. The advent of computer technology offers great potential to automate the process.

Image processing and computer vision technology are very beneficial to the agricultural industry. They are more potential and more important to many areas in agricultural technology. Cunha [4] used recognition technique to analyse the pathological stress conditions and characterization of the fruits or plant leafs. Runtz and Dave [5,6] applied image processing technique for classification and identifying of the plant species.

This study aims to develop a prototype system to automatically and correctly detect and classify the paddy diseases with BSR, BBS, and PBR using image processing technique as an alternative or supplemental to the traditional manual method.

978-0-7695-3879-2/09 $26.00 © 2009 IEEE
DOI 10.1109/SoCPaR.2009.62

The rest of the paper is organized as follows. In section II, the methodology used is introduced that includes 5 steps i.e. data acquisition, image enhancement, image segmentation and feature extraction, and the last step is imaging classification. Computational results are discussed in Section III. Section IV gives our conclusion and future work.

II. METHODOLOGY

The methodology for diagnosing paddy diseases can be simplified as Fig. 1. This process involves several tasks, such as image acquisition and collection, image segmentation and pre-processing, shape feature extraction and colour feature extraction, and paddy diseases classification based on lesion type, boundary colour, spot colour, and broken paddy leaf colour.

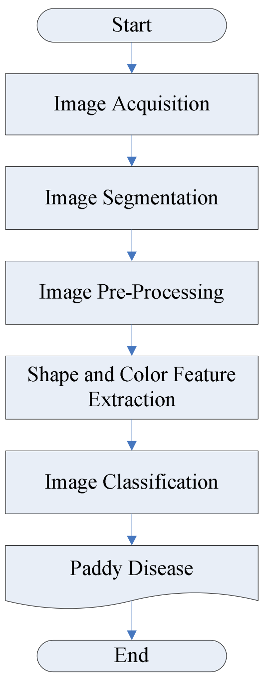
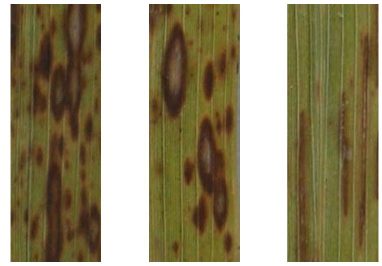


Fig. 1 Methodology Process.

A. Image Acquisition

In this process, it is a preparation process to obtain paddy leaf images. The RGB colour images of paddy leaf are captured using a Canon PowerShot G2 digital camera, with pixel resolution 768x1024. The digitized images are about 225 KB size each.

Those images are cropped into a smaller image with dimension of 109 x 310 pixels. There have collected about 94 data samples. It consists of three types of paddy diseases as shown in Fig. 2. Images are stored in BMP format. The prototype uses Matlab image processing library.

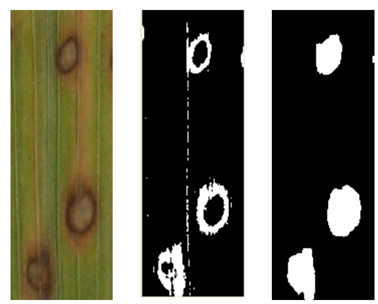


(a) (b) (c)

Fig. 2 Three Types of Paddy Diseases: (a) BD; (b) BSD; (c) NBSD

B. Image Segmentation and Pre-Processing

The segmentation and pre-processing task are the initial stage before the image is used for the next process. The main objective of this process is to obtain the binary image with less noise or noise free. In order to achieve high accuracy, an appropriate silhouette should be obtained. The RGB image (Fig. 3(a)) is converted into a binary image using threshold method, as shown in Fig. 3(b). In this paper, we used local entropy threshold methods of Eliza and Chang [7] and Otsu method.



(a) (b) (c)

Fig. 3 (a) RGB Image; (b) Binary Image with Noise; (c) Noise free Binary Image.

Entropy is widely used for measuring of local information content or uncertainty and the information content in a probability distribution [8,9]. An occurrence matrix is generate from the input image in accordance with probability distribution needed for entropy measures. The local entropy is defined as follow:

$$H = -\sum_{i=1}^{n} p_i \log p_i \ , \tag{1}$$

$$p_i = \frac{a_i}{a_1 + a_2 + \ldots + a_n} \ , \tag{2}$$

where H is local entropy value, p_i are the probability distribution, and $a_1, a_2, \ldots, a_n$ are the brightness levels in the windows located at the central pixel,

Note that the name 'local entropy' does not refer to a locally generated threshold, but is the name originally given to the method.

The Otsu method is based on selecting the lowest point between two classes of the histogram by considering the between-class variance [10].

Camera flash can act as noise and affects the image quality as shown in Fig. 3(b). Hence, median filter and morphological operators are applied to remove unnecessary spots by using a region filling technique. As a result, a noise free binary image is produced as shown in Fig. 3(c).

C. Feature Extraction Using Texture Analysis

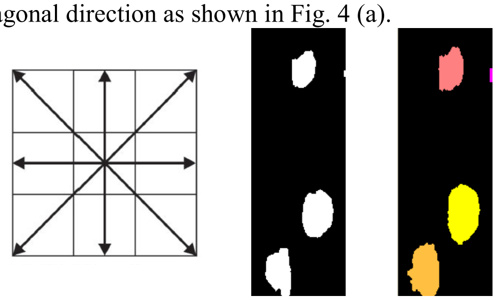
As the image of paddy disease shown in Fig. 2, several of disease varieties had different lesion shape and lesion colour, hence the shape and colour feature of disease spot were extracted for study.

The image analysis focused on the shape feature extraction and colour based segmentation.

1) Shape Feature Extraction

Shape is an important parameter of an image. People often understand and distinguish an object by its shape. General descriptors such as number of the object, width and length of the object are important characteristics to describe its shape. Those characteristics are used to extract feature the lesion. Blob Analysis is used in this research to get number of the object for labelled regions in a noise free binary image.

Object is a lesion on the paddy leaf. Normally, there are more than one object on paddy leaf image. Using 8-connected neighbourhood technique, counting the object is determined to get characteristic of the shape (Fig. 4(c)). This means that a pair of adjoining pixels is part of the same object only if they are both on and are connected along the horizontal, vertical, and diagonal direction as shown in Fig. 4 (a).



(a) (b) (c)

Fig. 4. (a) 8-Connected neighbourhood pixels to count the number of the object; (b) Noise free binary image; (c) Counted number of the object.

The width and height of the object are common features to determine the type of the lesion. A simple approach of measuring the width and height of the object is to count the number of object pixels. Suppose that I is a noise free binary image as shown in Fig. 4 (b) then $I(x,y) = 0$ for the object pixels and $I(x,y) = 1$ for the background pixels. Every object in an image is analysed to get height and width of the object (Fig. 5). Hence the type of the lesion can be determined by agricultural experts through width and height of the object.

100 data image samples consist of 20 images of each lesion type are analysed to obtain the value of width and height of the lesion. The type of the lesion (Fig. 6) has been determined as the following rules:

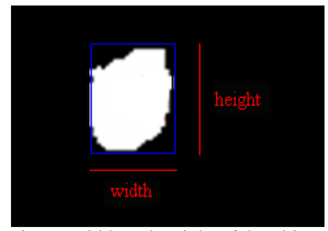


Fig. 5. Width and Height of the object.

a) Spot $\rightarrow$ $width \leq 30 \cap height \leq 35$
b) Round $\rightarrow$ $((height : width) \leq 2.5) \cap (width \leq 45)$
c) Oval $\rightarrow$ $((height : width) \leq 4.0) \cap (width \leq 45)$
d) Taper $\rightarrow$ $((height : width) > 2.0) \cap (width \leq 18)$
e) Spindle $\rightarrow$ $((width : height > 4.0) \cap (width > 15))$ or $((width : height \leq 4.0) \cap (width > 45))$

Based on experiment the constant values are set.

(a) (b) (c) (d) (e)

Fig. 6. Type of the Lesion: (a) Spot; (b) Round; (c) Oval; (d) Taper; (e) Spindle.

2) Colour Feature Extraction

Colour always plays a most important role in image processing and an important sign in recognizing different classes. Digital image processing produces quantitative colour measurements that are very helpful when investigating the lesion for early diagnosis [11].

The pixel in a colour image is commonly represented in the RGB space, in which the colour at each pixel is represented as a triplet (R,G,B), where R, G and B are respectively the red, green, and blue value from a colour image capturing device. Other colour spaces like the HSI and CIE colour model are also used in many other segmentation methods where their benefits and limitations are analysed and reported [12,13,14,15].

Generally, the colour difference is evaluated using the distance between two colour points in a colour space. The most common distance is Euclidean distance. Our proposed technique is based on the CIELab colour space, which is a uniform chromaticity colour space to get boundary colour, spot colour and broken leaf colour. It is known that Euclidean distance of two colours is proportional to the difference that human visual system perceived in the CIELab colour space [16].

Every pixel of an image is determined its specific colour according to the bank colour. Using Blob Analysis, boundary points and centroid point of each object on the original image (RGB image) is identified at early stage. This process produces 9 points on different locations, such as top-left, top-right, right-top, right-bottom, bottom-right, bottom-left, left-bottom, left-top, and centroid as shown in Fig. 7.

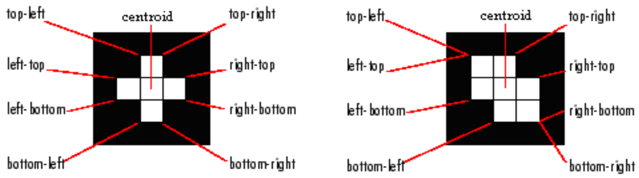


Fig. 7. Eight Locations of the boundary points and centroid point.

Generally the image is composed of RGB colour components. Thus, we must convert RGB colour components to CIELab colour component. To do this, we first converted RGB to CIEXYZ using the Rec. 709 HDTV primaries and D_{65} white point [17] as following equations:

$$\begin{bmatrix} X_i \\ Y_i \\ Z_i \end{bmatrix} = \begin{bmatrix} .4124 & .3576 & .1805 \\ .2126 & .7152 & .0722 \\ .0193 & .1192 & .9505 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}, \tag{3}$$

where $i = 1,2,\dots n$.

The CIELAB equation is then applied.

$$\begin{aligned} L^* &= 116 f\left(\frac{Y}{Y_o}\right) - 16, \\ a^* &= 500\left[f\left(\frac{X}{X_o}\right) - f\left(\frac{Y}{Y_o}\right)\right], \\ b^* &= 200\left[f\left(\frac{Y}{Y_o}\right) - f\left(\frac{Z}{Z_o}\right)\right], \end{aligned} \tag{4}$$

where

$$\begin{aligned} f(q) &= \sqrt[3]{q} \quad, && q > 0.008856 \\ f(q) &= 7.787q + \frac{16}{116} \quad, && q \le 0.008856 \end{aligned} \tag{5}$$

X_0, Y_0, and Z_0 are the tristimulus values of the illuminant D_{65}.

The colour difference ΔE^*_{ab} between two colours (colour pixel and colour marker), in terms of L^*, a^*, b^* is given by the Euclidean metric:

$$\Delta E^*_{ab} = \left[(\Delta L^*)^2 + (\Delta a^*)^2 + (\Delta b^*)^2\right]^{\frac{1}{2}} \tag{6}$$

The smallest distance (ΔE^*_{ab}) represents the pixel most closely match to the colour marker. For example, if the distance between a pixel and the red colour marker is the smallest, then the pixel is labelled as a red pixel.

D. *Image Classification*

Based on above characteristics, such as lesion type, boundary colour, spot colour, and broken paddy leaf colour, paddy diseases is recognized using production rule method with forward-chaining method. The production rules have been developed through serial interviews with agricultural expert.

III. Result and Analysis

There are two applied methods to determine the threshold value: local entropy threshold and Otsu threshold, in the diagnosis system. These methods are used in the segmentation phase in order to gain the optimum result for diagnosing paddy diseases.

Ninety-four image samples including 14 images of BD, 33 images of BSD and 47 images of NBSD have been tested using three threshold methods and compared with the expert supervision. The recognition accuracy rates of two methods are 94.7% and 61.2% for local entropy threshold and Otsu threshold respectively as represented in Table 1.

TABLE I
Recognition Accuracy Rates for Diagnosis Paddy Disease Based on Variable, Fixed and Otsu Threshold.

Threshold type	Correct	Incorrect	Accuracy rate (% Correct)
Local Entropy	89	5	94.7%
Otsu	58	36	61.2%

Since the intensity value is unique for each image, the Otsu method disables to perform segmentation task accurately. The intensity value is highly correlated to time and distance of the captured images. Furthermore, image capturing is closely related to the light source [18]. In this research, uncontrolled light source is used. Therefore, the captured image may suffer illumination problem and influence the intensity value. Consequently, it causes difficulty to normalize perfectly the viewing surface of the source image.

TABLE 2
Error Rates for Local Entropy and Otsu Thresholds in Diagnosing Paddy Disease Based on Image Processing Phases.

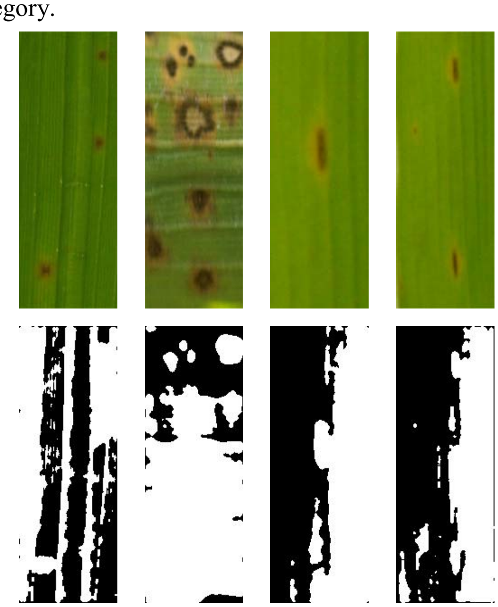
Image processing Phases	Error rates	
	Local Entropy	Otsu
Segmentation	-	61.1%
Feature extraction or Classification	100%	39.9%

The error rates for Local Entropy and Otsu threshold methods based on image processing phases are listed in Table 2. The error rates are categorized into two groups: segmentation and, feature extraction or classification. These errors are obtained and analysed in accordance to numbers of incorrect recognition (Table 1).

From Table 2, the Otsu Method which obtained the highest incorrect recognition number has the highest segmentation error rates about 61.1% while the balance (38.9%) is due to the feature extraction or classification errors. On the other hand, by using Local Entropy Method, it can perfectly segmenting the important objects, and the incorrect recognition number (Table 1) is totally (100%) due to feature extraction or classification errors.

The segmentation errors can occur due to uneven illuminants on each captured image. As a result, this uneven illuminants can cause large white areas to be formed. This problem occurs frequently in Otsu method as shown in Fig. 8. Otsu Method is calculated based on standard deviation and mean values calculations from histogram distributions of an image. The method does not make any assumption about the probability density functions. It only relies on their mean and variances. Therefore, the result might not be accurate. Apart from that, the correct maximum is not necessarily the global maximum [18].

Another causal error is determined in the feature extraction or classification process. Quite often it extracts inaccurately the type of the lesion. In general, the type of the lesion is computed based on width and height of the object. Then, these values are analysed using production rule method to obtain the shape of the lesion. However, mistaken in determining the width and height of the lesion can cause distance calculation imprecise. As a result, the lesions have different measures, even though they are supposed to be in the same object category.



(a)

(b)

Fig. 8 Sample images of error in the segmentation process by using Otsu methods; (a) RGB images; (b) Binary images

In conclusion, determination of threshold value is an important step in a segmentation process. Incorrect determination of threshold value can infer the result of segmentation task becoming inaccurate, and it leads to be erroneous in the classification phase.

The methodologies of this research as described in the previous sections are implemented in a user friendly interface as shown in Fig. 9. The user interface shows the paddy leaf image, grayscale image, binary image after applying the threshold value and a filter image after applying the median filtering process. The result of shape feature extraction and colour segmentation are also shown in Fig. 9.

IV. CONCLUSIONS AND FUTURE WORK

A system for diagnosing paddy diseases, including BD, BSD and NBSD mainly based on Matlab application has been developed in this study. The image processing techniques were used to establish the classification system. Four characteristics of lesion type, boundary colour, spot colour, and broken paddy leaf colour were tested for used to establish the classification system. The ratio of height and width of the lesion spot provided a unique shape characteristic for determining the type of the lesion.

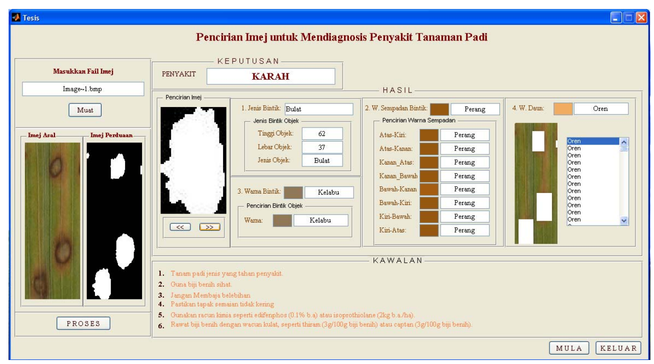


Fig. 9 User Interface of Prototype System

Two thresholding methods have been applied to get the best result in diagnosing ninety-four paddy leaf images. The best accuracy of two methods that used local entropy threshold is about 94.7%. Different intensity values and less prone to illumination, thus Otsu method is disabled to perform segmentation task accurately.

Future task will concentrate in ANN to classify paddy disease.

ACKNOWLEDGMENT

This research is based on a research grant from Ministry of Science, Technology and Innovation, Malaysia entitled "Image Based Fuzzy Expert System for Paddy Diseases Diagnosis" (01-01-02-SF0381).

REFERENCES

[1] S.C. Scardaci et al., Rice Blast: A New Disease in California, *Agronomy Fact Sheat Series* 1997-2, Department of Agronomy and Range Science University of California, Davis

[2] Robert W. Herdt, Research Priorities for Rice Biotechnology, in Rice Biotechnology, G.S. Khush and G.H. Toenniessen (eds.), Alden Press Ltd., London, 1991, pp. 35-37.

[3] Kobayashi, T., Kanda, E., Kitada, K., Ishiguro, K., Torigoe, Y., 2001. Detection of rice panicle blast with multispectral radiometer and the

potential of using airborne multispectral scanners. *Phytopathology*, **91**(3):316-323. [doi:10.1094/PHYTO.2001.91.3.316].

[4] J.B. Cunha. 2003. Application of Image Processing Techniques in the Characterization of Plant Leafs. *Proc. IEEE Intl' Symposium on Industrial Electronics.*

[5] K.J. Runtz. 1991. Electronic Recognition of Plant Spesies for Machine Vision Sprayer Control Systems. Proc. IEEE Western Canada Conference on Computer, Power, and Communication System in a Rural Environment (WESCANEX'91), pp. 84-88.

[6] S.Dave and K. Runtz. 1995. Image Processing Methods for Identifying Species of Plants. Proc. IEEE Western Canada Conference on Computer, Power, and Communication System in a Rural Environment (WESCANEX'95), pp.403-408.

[7] Eliza Yingzi Du and Chein-I Chang. 2003. An Unsupervised Approach to Colour Video Thresholding. *IEEE Proc. International Conference on Multimedia and Expo ICME'03* 3:III-337-40.

[8] Alhouse, Mark L.G. & Chein-I Chang. 1995. Image Segmentation by Local Entropy Methods. *Proc. International Conference on Image Processing* 3:61-64.

[9] Ping Yang and Otman A. Basir. 2003. Adaptive Weighted Median Filter using Local Entropy for Ultrasonic Image Denoising. Proc. of the 3rd International Symposium on Image and Signal Processing and Analysis 799-803.

[10] A.K. Jain, 1989. Fundamental of Digital Image Processing. *NJ: Prentice Hall.*

[11] L. Lucchese and S.K. Mitra, 2001. Colour Image Segmentation: A State-of-the-Art Survey. *Proceeding of the Indian National Science Academy, Vol. 67A, No. 2, 2001*, pp. 207-221.

[12] H. Wang, 2001. Automatic Character Location and Segmentation in Colour Scene Images. *Proc. of the 11th International Conf. on Image Analysis and Processing*.

[13] S. Lefevre, L. Mercier, V. Tiberghien and N. Vincent, 2002. Multiresolution Colour Image Segmentation Applied to Background Extraction in Outdoor Images. *IS&T European Conf. on Colour in Graphics, Image and Vis., 2002*, pp.363-367.

[14] J. Liu and Y-H. Yang, 1994. Multiresolution Colour Image Segmentation. *IEEE Transactions on PAMI, Vol.16*, 1994.

[15] K. J. Yoon and I.S. Kweon. Colour Image Segmentation Considering the Human Sensitivity for Colour Pattern Variations. Photonics Boston, SPIE, 2001, pp. 269-278.

[16] A. Hanbury and J. Serra. Mathematical morphology in the L*a*b* colour space. Technical Report N-36/01/MM, Centre de Morphologie Mathématique, Ecole des Mines de Paris, 2001.

[17] H. Tian, S.K. Lam and T. Srikanthan, 2003. Implementing Otsu's Thresholding Process using Area-Time Efficient Logarithmic Approximation Unit. *Proc. Of ISCAS'03, IV-21-IV-24.*

[18] Petrou, Maria & Bosdogianni, Panagiota. *Image Processing the Fundamentals*. John Wiley & Sons, Ltd. 1999.