การแก้ไขภาพจากกล้องดิจิทัลโดยใช้ฟังก์ชันแก้ความโค้งแบบสเมือนพหุนาม
และฟังก์ชันลบแสงเงาแบบพหุนามลอการิทึม

นาย ดำรง สุตาภิรัตน์

DIGITAL-CAMERA CAPTURED IMAGE RESTORATION BY USING SPLINE-BASED DEWARPING

FUNCTION AND DESHADING FUNCTION BASED ON POLYNOMIAL-LOGARITHM

Mr. Dumrong Sutapirat

A Thesis Submitted in Partial Fulfillment of the Requirements

for the Degree of Master of Science Program in Imaging Technology

Department of Imaging and Printing Technology

Faculty of Science

Chulalongkorn University

Academic Year 2008

Thesis Title DIGITAL-CAMERA CAPTURED IMAGE RESTORATION BY USING

SPLINE-BASED DEWARPING FUNCTION AND DESHADING

FUNCTION BASED ON POLYNOMIAL-LOGARITHM

By Mr. Dumrong Sutapirat
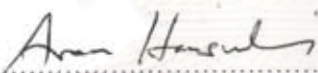
Field of Study Imaging Technology

Advisor Kingkarn Sookhanaphibarn, Ph.D.

Co-advisor Professor Chidchanok Lursinsap, Ph.D.

Accepted by the Faculty of Science, Chulalongkorn University in Partial
Fulfillment of the Requirements for the Master 's Degree
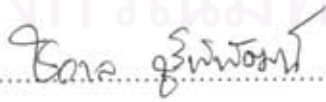
.... Dean of the Faculty of Science

(Professor Supod Hannongbua, Ph.D.)

THESIS COMMITTEE

.... Chairman

(Associate Professor Aran Hansuebsai, Ph.D.)

.... Advisor

(Kingkarn Sookhanaphibarn, Ph.D.)

.... Co-Advisor

(Professor Chidchanok Lursinsap, Ph.D.)

.... Examiner

(Assistance Professor Chawan Koopipad, Ph.D.)

.... External Examiner

(Associate Professor Prasong Praneetpolgrang, Ph.D.)

คำรง สุดาภิรัตน์ : การแก้ไขภาพจากกล้องดิจิทัลโดยใช้ฟังก์ชันแก้ความโค้งแบบเสมือนพหุ
นาม และฟังก์ชันลบแสงเงาแบบพหุนามลอการิทึม. (DIGITAL-CAMERA CAPTURED IMAGE
RESTORATION BY USING SPLINE-BASED DEWARPING FUNCTION AND DESHADING
FUNCTION BASED ON POLYNOMIAL-LOGARITHM) อ. ที่ปรึกษาวิทยานิพนธ์หลัก: อ.ดร.
กิ่งกาญจน์ สุขคณาภิบาล, อ.ที่ปรึกษาวิทยานิพนธ์ร่วม: ศ.ดร.ขัดชนก เหลือสินทรัพย์, 59 หน้า.

งานวิจัยนี้นำเสนอวิธีการใหม่ของการแก้ความโค้ง และลบเงาของภาพถ่ายเอกสารที่ได้
จากกล้องถ่ายภาพดิจิทัล ภาพถ่ายเอกสารดังกล่าวมีลักษณะ โค้งตามความหนาของหน้าหนังสือ
เทคนิคการแก้ความโค้งภาพถ่ายเอกสารในงานวิจัยนี้ทำการย้ายพิกัดของจุดภาพจากระบบพิกัด
โค้งไปยังระบบพิกัดตรงโดยใช้ตารางค้นหา สำหรับวิธีการลบเงาบนภาพถ่ายเอกสารจะใช้วิธี
ความคลาดเคลื่อนกำลังสองต่ำสุดในการลบเงาบนภาพถ่ายเอกสาร วิธีการในงานวิจัยนี้เป็น
อิสระต่อลักษณะเนื้อหาของเอกสาร, การจัดวางหน้า และลักษณะของตัวอักษร ในการทดลอง
ได้กำหนดภาพถ่ายเอกสารทดสอบซึ่งประกอบไปด้วย ลักษณะตัวอักษรประเภทต่างๆ และ
ขนาดของตัวอักษรที่ใช้ในตำราทั่วไป ความถูกต้องของวิธีที่ใช้ในงานวิจัยนี้ถูกนำมา
เปรียบเทียบกับฟังก์ชันแก้ไขความโค้งของโปรแกรมรู้จำตัวอักษร หรือ OCR ที่ชื่อ Abby
FineReader Professional 9.0 และ OmniPage Professtional 16 พบว่าวิธีการในการทดลองนี้ ให้
ประสิทธิภาพที่ดีกว่าฟังก์ชันแก้ความโค้งที่มีอยู่ในโปรแกรม Abby FineReader ถึง 49 % และดีกว่า
ฟังก์ชันแก้ความโค้งที่มีอยู่ในโปรแกรม OmniPage ถึง 10 % สำหรับความแม่นยำของคำ และ
ตัวอักษร
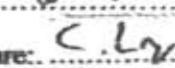
ภาควิชา วิทยาศาสตร์ทางภาพถ่ายและเทคโนโลยีการพิมพ์ ลายมือชื่อนิสิต...........................

สาขาวิชา เทคโนโลยีทางภาพ......................... ลายมือชื่ออ.ที่ปรึกษาวิทยานิพนธ์หลัก..................

ปีการศึกษา 2551......................... ลายมือชื่ออ.ที่ปรึกษาวิทยานิพนธ์ร่วม.....................

4972298523 : MAJOR IMAGING TECHNOLOGY

KEYWORDS: DOCUMENT IMAGE / DEWARPING FUNCTION / COORDINATE TRANSFORMATION / DIGITAL CAMERA / WARPED DOCUMENT IMAGE / DESHADING FUNCTION

DUMRONG SUTAPIRAT : DIGITAL-CAMERA CAPTURED IMAGE RESTORATION BY USING SPLINE-BASED DEWARPING FUNCTION AND DESHADING FUNCTION BASED ON POLYNOMIAL-LOGARITHM. ADVISOR : KINGKARN SOOKHANAPHIBARN, Ph.D., CO-ADVISOR : PROF. CHIDCHANOK LURSINSAP, Ph.D., 59 pp.

This paper proposes a novel dewarping and deshading technique for digital camera based document images, from which images are warped on the curve surface of the book bound volume. Our dewarping technique by using look-up table transforms pixels on the curve coordinate system into the planar coordinate system. In the deshading method, the shade on the image is removed by using Least Square Error method (LSE). Our technique is independent on document contents, page layout, and font styles. In the experiments, the test document images composed various font styles and sizes printed in common text books. The validity of our methods is compared with the program's curvature correction of Abby FineReader Professional 9.0 and Omnipage professional 16. Our technique outperforms the Abby's curvature correction up to 49 percent and the OmniPage's curvature correction up to 10 percent for word and character precisions.

Department: Photographic Science and Printing Technology   Student's Signature:................

Field of Study: Imaging Technology........................   Advisor's Signature:................

Academic Year: 2008........................   Co-advisor's Signature:................

ACKNOWLEDGEMENTS

CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

1.1 Statement of the problems

Document image analysis consists of three mainly processes that performs the overall interpretation of document images. First, the process of determining document structure may be viewed as guided by a model, explicit or implicit, of the class of documents of interest. The model describes the physical appearance and the relationships between the entities that make up the document such as text and graphic on the documents. Second, the image acquirement state by using scanners and cameras as mentioned in [3] is considerably degraded a degree of automation in document analysis systems. To alleviate the hardware degradation, several solutions regarding to image document restorations have been proposed such as the dewarping approaches to handle 2D images captured by a digital camera [1], [3], [4], [5], [6], [7], [10], [11], [13], [14] and ones obtained by a scanner [8], [9], [15], [16].

Third, digital optical character recognition (OCR) is a key role in document image analysis, which is now common in many offices. Nowadays, OCR has been an electronic tool to translate images of handwritten, typewritten or printed text into machine editable text. In academic researches, OCR is a field of pattern recognition, artificial intelligence and machine vision. Though academic research in the field continues, the focus on OCR has shifted to implementation of proven techniques. Junker *et al.* [2] researched on the evaluation of document analysis components by recall, precision, and accuracy based on the OCR commercial software. A common

question is how to evaluate the effectiveness of such complex document analysis systems involving rather distinct components.

The performance of an OCR system depends on the cumulative accuracy and speed of these steps. The steps include scanning, preprocessing, processing and post-processing. Skew detection is a part of preprocessing stage which affects the performance of an OCR system significantly if not detected and corrected properly. In our experiments, we chose the OCR software named *Abby FineReader Professional 9* for accuracy evaluation of our approaches.

In this research, we propose a novel dewarping function to transform pixels in the curve coordinate system into those in the planar coordinate system. Our dewarping function now derived from interpolation based on data collection from a digital studio kit. Before using the digital studio kit, the coordinates on the book surface are marked by using a graph sheet. Then we acquire the graph-sheet images by using a digital camera. In order to create dewarping function, we extract the pixel coordinates data of the graph-sheet images from the point being determined on the graph-sheet images. After obtaining the coordinates data, we interpolate data based on the graph-sheets coordinates data by using spline interpolation method and impute data in row by our row imputation algorithm, which can make the function complete. After that, the dewarping function being created is used to dewarp the warp document image. However, for better OCR's recognition rates accuracy, we impute the pixels in to a result image by our column imputation algorithm. To make our imaging result perfectly, we remove the undesirable shade appearing along the document surface due to the position of the light source and the book by using mathematic method.

1.2 Research problems

A warp document image is the type of image distortion called "perspective distortion". Perspective distortion can occur when the text plane is not parallel to the imaging plane. The effect is that the characters farther away look smaller and distorted.

Many novel approaches have been proposed for document image dewarping. These approaches can be grouped into two categories:

1. Page reconstruction using specialized hardware such as stereo camera and scanner etc [12], [13], [16].

2. Page reconstruction using a single camera spatially in an uncontrolled environment [6], [11], [14].

The first approaches require specialized hardware and based on 3-D shape reconstruction. These approaches are not flexibility and independent to the devices that is the approach based on stereo camera[12]. Moreover some approaches use the physical model of the device[16] or optical base model of the light source and the book surface [13] to constructed the 3-D shape modeling. For the second approaches, a document image are captured by using hand-held cameras in an uncontrolled environment, being more adaptable of acquiring document image. However, in these approaches, some of them use Text base line detection [6], [14] which is not independent to the content of the text, page layout and Font styles. Some approach require a physical setting of the equipment such as distance and the angle between camera's lens and page surface [11]. Thus, in this research, we propose the new dewarping document image approach in an uncontrolled environment which is

independent to the content, page layout, Font styles and physical setting of the equipments.

In this research we interest in one of another problem of using a single camera in an uncontrolled environment that is uneven lighting, which is one of the major problems of camera-based document image. It can occur whether the physical environment and uneven response from the devices especially when using artificial lights or to image reflective content. The effect is that the image has the shade appearing along the image plane. In our approach, we propose a method to deshaded the warp document image by using mathematical method that is least squares errors method to make our resulting image more completely.

1.3 Objectives

To develop a new dewarped and deshaded method to dewarp and deshade document images, which is acquired from digital camera.

1.4 Scope of study

1.4.1   The experimental equipments must be set in the condition as follows:

1.4.1.1 The focused distance of the cameras must be used at 45 mm. (equivalent to 90 mm.)

1.4.1.2 The book position must not be skew.

1.4.1.3 All of the equipments such as camera tripod, the book etc. must fix in the position while acquiring the document images.

1.4.1.4 Each of book pages must be in separate document images.

1.4.2   The image file used in our experiment must be set in the condition as follows:

      1.4.2.1 The file format must be JPEG with resolution of 1200 x 1600

        pixels.

      1.4.2.2 The document image must be 256 gray scales.

   1.4.3   The text on the documents should be black to help the OCR accuracy.

   1.4.4   The performance of our method is evaluated by OCR accuracy

      percentages using *Abby FineReader Professional 9* and *OmniPage*

      Professional 16.

1.5 Research procedures

   1.5.1   Study the related researches involving document image analysis.

   1.5.2   Set up the experimental equipments.

   1.5.3   Develop the dewarping and deshading method for the document

      images.

   1.5.4   Test against several document images test set by using *Abby*

      FineReader Professional 9 and *Omnipage Professionals 16.*

   1.5.5   Conclude thesis.

1.6 Conceptual framework

    In this research we propose the novel dewarping function based on transformation of pixels coordinates. Since we find that the pixel coordinates on the curve surface are obviously distort comparing with the same pixels on the planar image. Thus, to remove this image distortion, we create the dewarping function to transform the curve coordinate system into Cartesian coordinate system. The dewarping document image procedure is shown in Figure 1.1.

```
┌─────────────────────────────┐
│     Warp document image     │
└─────────────────────────────┘
              │
              ▼
┌─────────────────────────────┐
│      Image acquisition      │
└─────────────────────────────┘
              │
              ▼
┌─────────────────────────────┐
│      Dewarping Function     │
└─────────────────────────────┘
              │
              ▼
┌─────────────────────────────┐
│       Dewarping image       │
└─────────────────────────────┘
              │
              ▼
┌─────────────────────────────┐
│     Dewarp document image   │
└─────────────────────────────┘
```
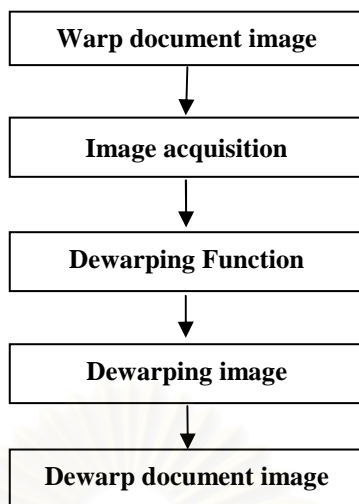
Figure 1.1 dewarping a document image procedure.

For the deshading approach, we suggest the deshading function base on shading behavior of the image. We offer the mathematical method to removing the shade along the document image plane by creating a deshading function based on least square errors technique. The deshading approach is shown in Figure 1.2.

```
┌─────────────────────────────┐
│     Dewarp document image   │
└─────────────────────────────┘
              │
              ▼
┌─────────────────────────────┐
│       Deshading function    │
└─────────────────────────────┘
              │
              ▼
┌─────────────────────────────┐
│        Deshading image      │
└─────────────────────────────┘
              │
              ▼
┌─────────────────────────────┐
│    Rectified document image │
└─────────────────────────────┘
```
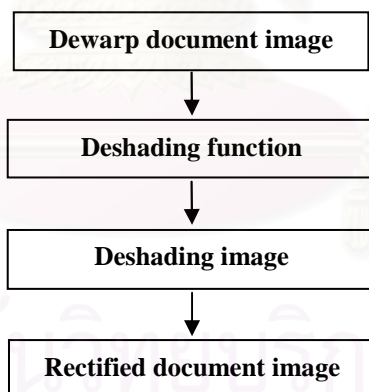
Figure 1.2 deshading a dewarp document image procedure.

1.7 Significance of the research

1.7.1 A dewarping function which is used for dewarped a warp document image.

1.7.2 A deshading function which is used for deshaded a document image.

1.7.3 Increase the OCR recognition rates.

1.7.4 Obtaining the document image that is available to the OCR.

1.7.5 Create a new way to research in document image analysis.

<p style="text-align: center;">**CHAPTER II**</p>

<p style="text-align: center;">**THEORY AND RELATED LITERATURE**</p>

2.1 Theory

2.1.1 Digital image

Digital image is an image being stored in digital form. It may be converted from analog data into numerical data by the image processor unit of imaging devices. The image can be defined as two dimensional function, $f(x, y)$, where x, y are spatial coordinates and f are intensity or gray levels of an image at these coordinates. The coordinates values and intensity values of the image are both discrete quantities. In general, we use integer values for those discrete quantities for notational clarity and convenience. The notation of the m x n digital image can be written as a matrix form as shown in Equation (1)

$$f(x, y) = \begin{bmatrix} f(0,0) & f(0,1) & \cdots & f(0, n-1) \\ f(1,0) & f(1,1) & \cdots & f(1, n-1) \\ \vdots & \vdots & & \vdots \\ f(1, n-1) & f(m-1,1) & \cdots & f(m-1, n-1) \end{bmatrix} \tag{1}$$

In this matrix, each element can be called pixel, image element, picture element or pel.

The intensity of any pixels x, y may be defined as a discrete gray level L, which canbe multiple values. Suppose that, if gray levels of the image have only 2 levels, the intensity values of the image will be 0 or 1 sometimes called binary image. However, if gray levels of the image have 8 levels, the image will be 0 to 255 or 256 of intensity values. Thus, it can be considered that the number of gray levels is an integer power of 2 as shown in Equation (2), where k is a bit number of the image:

$$L = 2^k \tag{2}$$

Discrete gray levels are integers values in the interval $[0, L-1]$ so that luminance of the image is depend on the gray levels of any pixel intensity values. From the previous information, it can be said that digital image composes of many pixels containing with several discrete intensity values and discrete quantities of coordinates as shown in Figure 2.1

Figure 2.1 Digital image containing with various pixels intensities.

2.1.2 Dewarping Function

Dewarping function is a coordinate transform $T()$ function, from the curved coordinate system given by $f(x, y)$, to a Cartesian coordinate system, $f(x', y')$ as shown in Figure 2.2.
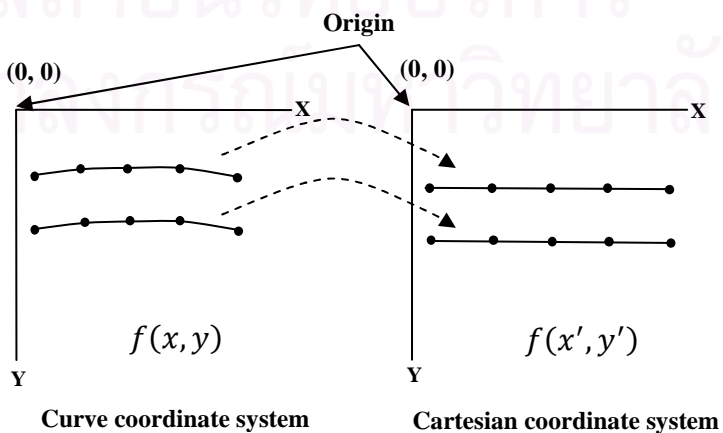
Figure 2.2 Transformation between curved coordinate system to a Cartesian coordinate system

The curved coordinate system [4], [10], [11] is defined as a set of pixel coordinates residing in the book-middle-page surface. The Equation is shown in Equation (3).

$$f(x', y') = T(f(x, y)) \qquad (3)$$

In our work, we create the function based on coordinates data of the images by using the spline interpolation method to interpolate the values between coordinates data. The method will be explained on the next section.

2.1.3 Spline Interpolation

Spline interpolation, a method of numerical analysis, is a form of interpolation where its interpolant is a type of piecewise polynomial called spline. It is used to represent the data values with function passing through the points called control points. The points are given by $(x_0, y_0), (x_1, y_1), ..., (x_n, y_n)$ from a given function $y=f(x)$. Suppose that, the function $f(x)$ is continuous functions that may be represented the $n+1$ data values with $f(x)$ passing through the $n+1$ points. Then we can find the value of y at any other value of x called interpolation.

Spline interpolation is preferred to use over another polynomial interpolation method because of its small interpolation error. While another interpolation method obtain higher order degrees of polynomial that may get oscillatory behavior in some cases (see Figure 2.3), spline interpolation use only low degrees of polynomial which is not more than degrees 3.
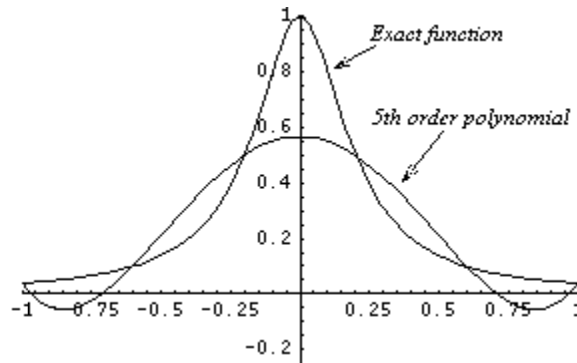
Figure 2.3 Oscillatory behavior due to using high polynomial degree.

Thus, spline interpolation uses information from more data points, but at the same time keeping the function true to the data behavior. However, for the fitting, smoothing and flexibility of the spline depend on its degrees of polynomial. The most common splines interpolations used are:

1) Linear spline interpolation

2) Quadratic spline interpolation

3) Cubic spline interpolation

2.1.3.1 Linear spline interpolation

The basic type of spline interpolation by using straight line passing through the point as shown in Figure 2.4. Given $(x_0, y_0), (x_1, y_1), \dots, (x_n, y_n)$, fit linear splines to the data. The linear splines are given by Equation (4), (5), (6). In the Equation we denoted the point as $y_i = f(x_i)$ and for the line as $f_i(x)$
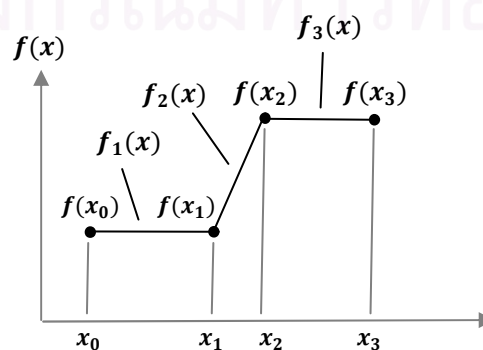
Figure 2.4 Linear spline interpolation sample

$$f_1(x) = f(x_1) + m_1(x - x_0) \quad : x_0 \le x \le x_1 \tag{4}$$

$$f_2(x) = f(x_2) + m_2(x - x_1) \quad : x_1 \le x \le x_2 \tag{5}$$

$$f_3(x) = f(x_2) + m_3(x - x_2) \quad : x_2 \le x \le x_3 \tag{6}$$

For $\quad m_i = \dfrac{f(x_{i+1}) - f(x_i)}{x_{i+1} - x_i} : i = 1, 2, 3$

m is the slope of linear function connected at the point.

Linear spline use data from the two consecutive data points. At the interior points of the data, the slope changes abruptly.

2.1.3.2 Quadratic spline interpolation

To making the spline function more fitting as possible, we will obtain the polynomial degrees 2 for *n+1* point of data. As shown in Figure 2.5, the line becomes more obviously smooth especially at the interior points.
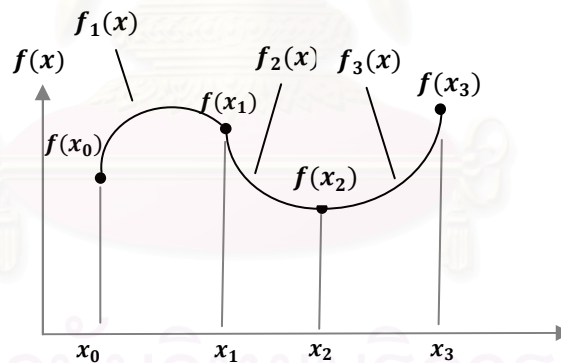
Figure 2.5 Quadratic spline interpolation sample

The quadratic function is shown in Equation (7), (8) and (9)

$$f_1(x) = a_1 x^2 + b_1 x + c_1 \quad : x_0 \le x \le x_1 \tag{7}$$

$$f_2(x) = a_2 x^2 + b_2 x + c_2 \quad : x_1 \le x \le x_2 \tag{8}$$

$$f_3(x) = a_3 x^2 + b_3 x + c_3 \quad : x_2 \le x \le x_3 \tag{9}$$

To finding the coefficient of the function, we defined the condition of the quadratic spline interpolation for *n+1* data using with *n* function and unknown coefficient with *3n* data:

1. The function is similar at the connection point or control point so that:

   at $f(x_1)$

   $$f_1(x_1) = a_1 x_1^2 + b_1 x_1 + c_1 = f(x_1)$$

   $$f_2(x_1) = a_2 x_1^2 + b_2 x_1 + c_2 = f(x_1)$$

   at $f(x_2)$

   $$f_2(x_2) = a_2 x_2^2 + b_2 x_2 + c_2 = f(x_2)$$

   $$f_3(x_2) = a_3 x_2^2 + b_3 x_2 + c_3 = f(x_2)$$

2. The first and the last function must pass through the first and the last points respectively:

   at $f(x_0)$

   $$f_1(x_0) = a_1 x_0^2 + b_1 x_0 + c_1 = f(x_0)$$

   at $f(x_3)$

   $$f_3(x_3) = a_3 x_3^2 + b_3 x_3 + c_3 = f(x_3)$$

3. First order derivative at the connection points must similar:

   at $f(x_1)$

   $$f_1'(x_1) = f_2'(x_1)$$

   $$2a_1 x_1 + b_1 = 2a_2 x_1 + b_2$$

   at $f(x_2)$

   $$f_2'(x_2) = f_3'(x_2)$$

   $$2a_2 x_2 + b_2 = 2a_3 x_2 + b_3$$

4. Second derivative at the first point must be zero value:

   $$f_1''(x_0) = a_1 = 0$$

2.1.3.3 Cubic spline interpolation

Cubic spline interpolation is the polynomial degrees of 3 as shown in Figure 2.6, which is the maximum degrees of its polynomial form for more flexibility, fitting and smoothing spline curve.
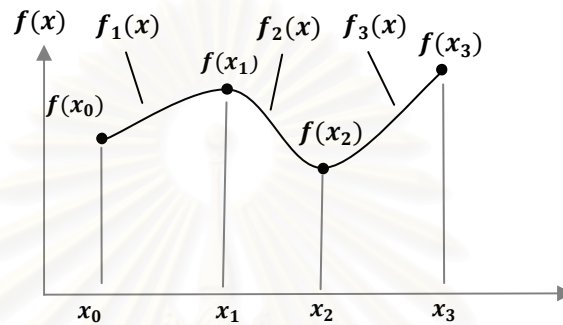
Figure 2.6 Cubic spline interpolation sample

The basic form of polynomial function is shown in Equation (10), (11) and (12):

$$f_1(x) = a_1 x^3 + b_1 x^2 + c_1 x + d_1 \quad : x_0 \leq x \leq x_1 \tag{10}$$

$$f_2(x) = a_2 x^3 + b_2 x^2 + c_2 x + d_2 \quad : x_1 \leq x \leq x_2 \tag{11}$$

$$f_3(x) = a_3 x^3 + b_3 x^2 + c_3 x + d_3 \quad : x_2 \leq x \leq x_3 \tag{12}$$

As similar to Quadratic spline that is data $n+1$, we portion the data in to n portions which make the $4n$ coefficient data. This method use $4n$ condition functions to finding the coefficient values. For the conditions of the function are:

1. The functions are similar at the connection points or control points so that:

at $f(x_1)$

$$f_1(x_1) = a_1 x_1^3 + b_1 x_1^2 + c_1 x_1 + d_1 = f(x_1)$$

$$f_2(x_1) = a_2 x_1^3 + b_2 x_1^2 + c_2 x_1 + d_2 = f(x_1)$$

at $f(x_2)$

$$f_2(x_2) = a_2 x_2^3 + b_2 x_2^2 + c_2 x_2 + d_2 = f(x_2)$$

$$f_3(x_2) = a_3 x_2^3 + b_3 x_2^2 + c_3 x_2 + d_3 = f(x_2)$$

2. The first and the last function must pass through the first and the last point respectively:

at $f(x_0)$

$$f_1(x_0) = a_1 x_0^3 + b_1 x_0^2 + c_2 x_0 + d_1 = f(x_0)$$

at $f(x_3)$

$$f_3(x_3) = a_3 x_3^3 + b_3 x_3^2 + c_3 x_3 + d_3 = f(x_3)$$

3. First order derivative at the connection point must similar:

at $f(x_1)$

$$f_1'(x_1) = f_2'(x_2)$$

$$3a_1 x_1^2 + 2x_1 b_1 + c_1 = 3a_1 x_2^2 + 2b_2 x_1 + c_2$$

at $f(x_2)$

$$f_2'(x_2) = f_3'(x_2)$$

$$3a_2 x_2^2 + 2x_2 b_2 + c_2 = 3a_3 x_2^2 + 2b_3 x_2 + c_3$$

4. Second order derivative at the connection point must similar:

at $f(x_1)$

$$f_1''(x_1) = f_2''(x_1)$$

$$6a_1 x_1 + 2b_1 = 6a_1 x_2 + 2b_2$$

at $f(x_2)$

$$f_2''(x_2) = f_3''(x_2)$$

$$6a_2 x_2 + 2b_2 = 6a_2 x_3 + 2b_3$$

5. Second order derivative at the first and the last points must be zero values:

at $f(x_0)$

$$f_1''(x_0) = 6a_1 x_0 + 2b_1 = 0$$

at $f(x_3)$

$$f_3''(x_3) = 6a_3 x_3 + 2b_3 = 0$$

2.1.4 Least Square Approximation

The objective of data approximation method is to find a simple representation of the data by a function of the best-fitting curve from the given set of points. We choose the form of function based on overall behavior of the data. The method of least squares assumes that the best-fit curve of a given function is the curve that has the minimal sum of the deviations squared (least square error) from a given set of data. Suppose that the data points are $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$. The fitting curve $f(x)$ has the deviation from each data point ε:

$$\varepsilon_i = Y_i - f(x_i)$$

Thus, the best fitting curve has the property that:

$$R = \varepsilon_1^2 + \varepsilon_2^2 + \cdots + \varepsilon_n^2 = \sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} (Y_i - f(x_i))^2$$

There are two types of least square problems:

2.1.4.1 Linear least square:

The simple form of linear function is:

$$y = ax + b$$

We have to find the values of two parameters a and b that produced the minimum errors of R:

$$R = \sum_{i=1}^{n} \epsilon_i^2 = \sum_{i=1}^{n} (Y_i - y_i)^2$$

$$= \sum_{i=1}^{n} (Y_i - f(x_i))^2$$

$$= \sum_{i=1}^{n} (Y_i - ax_i - b)^2$$

Thus, R is the function of both a and b denoted by $R(a, b)$. Then, We can find the minimum values of R by the conditions:

$$\frac{\partial R}{\partial a} = 0$$

$$\frac{\partial R}{\partial b} = 0$$

From the partial differentiate we have :

$$\frac{\partial R}{\partial a} = \frac{\partial}{\partial a}\left(\sum_{i=1}^{n}(Y_i - ax_i - b)^2\right) = 2\sum_{i=1}^{n}(Y_i - ax_i - b)(-x_i) = 0$$

$$\frac{\partial R}{\partial b} = \frac{\partial}{\partial b}\left(\sum_{i=1}^{n}(Y_i - ax_i - b)^2\right) = 2\sum_{i=1}^{n}(Y_i - ax_i - b)(-1) = 0$$

Thus that :

$$a\sum_{i=1}^{n}x_i^2 + b\sum_{i=1}^{n}x_i = \sum_{i=1}^{n}x_i Y_i$$

$$a\sum_{i=1}^{n}x_i + bn = \sum_{i=1}^{n}Y_i$$

2.1.4.2 Non-linear least square :

Given a function $f(x)$ of a variable x tabulated at n values $y_i = f(x_1), \ldots, y_n = f(x_n)$ assume the function is of known analytic form depending on m parameters $f(x; c_1, \ldots, c_n)$ and consider the over determined set of m Equations :

$$y_1 = f(x_1; c_1, \ldots, c_m)$$

$$y_m = f(x_n; c_1, \ldots, c_m)$$

Suppose we have the polynomial function:

$$Y(x_i) = c_1 + c_2 x_i + c_3 x_i^3 + \cdots + c_{m+1} x_i^m$$

We need to solve these Equations to obtain their values $c_1, \ldots, c_n$. The errors

from fitting curve for each $i = 1, \ldots, m$ is :

$$R = \sum_{i=1}^{n} \epsilon_i^2 = \sum_{i=1}^{n} (Y_i - c_1 - c_2 x_i - c_3 x_i^3 - \cdots - c_{m+1} x_i^m)^2$$

The minimum R values under the condition :

$$\frac{\partial R}{\partial c_j} = 0, \quad j = 1, \ldots, m + 1$$

Now we have:

$$\frac{\partial R}{\partial c_1} = 2 \sum_{i=1}^{n} (Y_i - c_1 - c_2 x_i - c_3 x_i^3 - \cdots - c_{m+1} x_i^m)(-1) = 0$$

$$\frac{\partial R}{\partial c_2} = 2 \sum_{i=1}^{n} (Y_i - c_1 - c_2 x_i - c_3 x_i^3 - \cdots - c_{m+1} x_i^m)(-x_i) = 0$$

$$\vdots \qquad\qquad\qquad \vdots$$

$$\frac{\partial R}{\partial c_{m+1}} = 2 \sum_{i=1}^{n} (Y_i - c_1 - c_2 x_i - c_3 x_i^3 - \cdots - c_{m+1} x_i^m)(-x_i^m) = 0$$

Reforming the Equations :

$$nc_1 + \sum_{i=1}^{n} x_i c_2 + \sum_{i=1}^{n} x_i^2 c_3 +, \ldots, + \sum_{i=1}^{n} x_i^m c_{m+1} = \sum_{i=1}^{n} Y_i$$

$$\sum_{i=1}^{n} x_i c_1 + \sum_{i=1}^{n} x_i^2 c_2 + \sum_{i=1}^{n} x_i^3 c_3 +, \ldots, + \sum_{i=1}^{n} x_i^{m+1} c_{m+1} = \sum_{i=1}^{n} Y_i x_i$$

$$\vdots \qquad\qquad\qquad \vdots$$

$$\sum_{i=1}^{n} x_i^m c_1 + \sum_{i=1}^{n} x_i^{m+1} c_2 + \sum_{i=1}^{n} x_i^{m+2} c_3 +, \ldots, + \sum_{i=1}^{n} x_i^{2m} c_{m+1} = \sum_{i=1}^{n} Y_i x_i^m$$

We can rewrite it in a matrix form :

$$\begin{bmatrix} \sum_{i=1}^{n} x_i^0 & \sum_{i=1}^{n} x_i^1 & \sum_{i=1}^{n} x_i^2 & \sum_{i=1}^{n} x_i^m \\ \sum_{i=1}^{n} x_i^1 & \sum_{i=1}^{n} x_i^2 & \sum_{i=1}^{n} x_i^3 & \sum_{i=1}^{n} x_i^{m+1} \\ \vdots & \vdots & \vdots & \vdots \\ \sum_{i=1}^{n} x_i^m & \sum_{i=1}^{n} x_i^{m+1} & \sum_{i=1}^{n} x_i^{m+2} & \sum_{i=1}^{n} x_i^{2m} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ \vdots \\ c_{m+1} \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} Y_i \, x_i^0 \\ \sum_{i=1}^{n} Y_i \, x_i^1 \\ \sum_{i=1}^{n} Y_i \, x_i^2 \\ \vdots \\ \sum_{i=1}^{n} Y_i \, x_i^m \end{bmatrix}$$

2.2 Related literature

Many researcher have been proposed various techniques for document image dewarping approaches. These approaches can be classified into two board categories following to the acquisition of images:

1. 3-D shape reconstruction of the page using specialized hardware like stereo-cameras[12], structured light source[13], or laser scanners[16].

2. Reconstruction of the page using single camera in an uncontrolled environment[6], [11], [14]

2.2.1 3-D Shape reconstruction of the page

These approaches are based on 3-D shape reconstruction. One of the major drawbacks of the approaches requiring specialized hardware is that they limit the flexibility of capturing documents with cameras or scanners.

2.2.1.1 Yamashita, A. Kawarago *et al.* [12] proposed the 3-D shape reconstruction approaches by using stereo-cameras. The images are reconstructed in 3-D shape executed from the result of the stereo-

cameras. A flat surface is recovered from the curved or folded surface of the 3-D shape.

2.2.1.2 Zhang, L. and Tan, C.L. [13] proposed the Shape-From-Shading (SFS) techniques to reconstruction the 3-D shape by considered perspective projection under oblique light source. Then, after reconstruct the shape, the images are flattened the warp surface using a physically-based deformation model.

2.2.1.3 Zhang, L. and Zhang, Z. *et al.* [16] proposed the dewaping method for scanner-based document image. In this approach they offer the shape reconstruction base on shading information in a scanned document image using SFS. According to SFS, it is characterized by 1) a proximal and moving light source, 2) Lambertian reflection 3) nonuniform alberdo distribution 4) document skew to build the practical models for reconstruct the 3-D shape of the book surface. Then they restore the scanned document image base on deshading and dewarping model.

2.2.2 Reconstruction of the page using single camera in an controlled environment

These approaches, having caught more attention recently, are based on a single camera in an uncontrolled environment. Our method is base on these approaches.

2.2.2.1 Wu, M. and Li, R. *et al.* [11]proposed a book dewarping model which flattens a curved book page to its original flat rectangle shape. This model handle warped book images taken from top of the book surface. The angle between camera lens and the book surface are

realized and take it in to an account in the transform model. Then, when they get the model, they map each pixel in warped region to a pixel in its original rectangle region.

2.2.2.2 Zhang, L. and Tan, C.L. [14]proposed an image restoration technique that corrects various warping distortions such as rolling curl, binding curl and fold distortions in camera-based image. Their method use a Spline Interpolation techniques based on ruled surface model constructed from text lines in the 2D document image.

2.2.2.3 Schneider, D.C. and Block M. *et al.* [6] proposed using local orientation features to interpolate a vector field from which a warping mesh is derived. Then, they correct the image by approximating the non-linear distortion with multiple linear projections.

CHAPTER III

RESEARCH METHODOLOGY

Our experimental process is classified as shown in Figure 3.1

Document

Image acquisition

Warp document image

Dewarping Function
- Recording the image coordinates
- Cubic spline interpolation
  - Row imputation

Dewarping image
- Column imputation

Dewarp document image

Deshading function

Deshading image
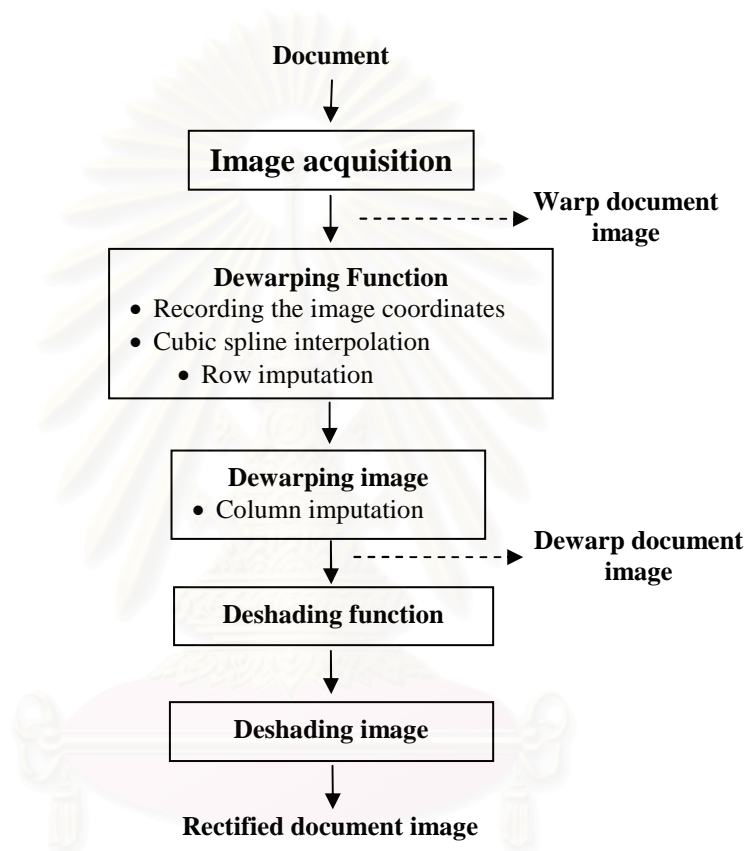
Rectified document image

Figure 3.1 Experimental process

3.1 Image acquisition

To capture document images by a digital camera, we use the equipments as shown in Figure 3.2.

Digital Camera

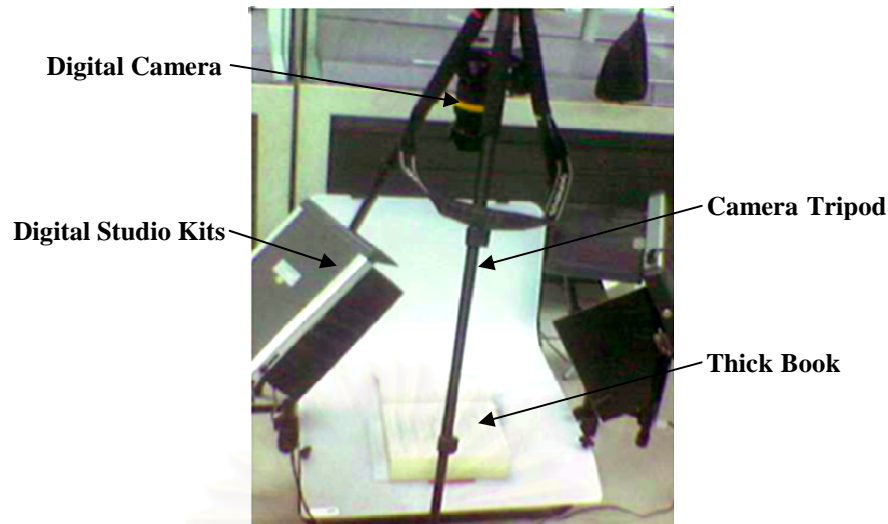Digital Studio Kits

Camera Tripod

Thick Book

**Figure 3.2 The experimental setting consisting of digital camera,
digital studio kits, camera tripod and thick book**

The following specification of our experimental equipments are described
below:

1) Our digital camera model is Olympus E-1 with 5 million effective

resolutions with resolution of 1200 x1600.

2) The focusing distance, we set at 45 mm, which is equivalent to 90 mm due

to the crop factor of the digital camera or the size of camera's sensor.

3) For the thick book used in our experiment is 600-pages published by Wrox

in uncoated paper.

According to the experimental setting, the book and the camera must be in fixed
position unless the function cannot dewarp the warp document images correctly.

3.2 Dewarping Function

In this section, we will describe the process of creating our purpose a

dewarping function to dewarping a warp document image. In order to creating the

function, firstly, we have to recording coordinates data from our two types of graph-

sheets images by measuring coordinates data from a graphic program, described in Section 3.2.1. Next, after getting the data from the graph-sheets, we interpolate the given pair of coordinates by cubic spline interpolation method and impute the line in to the gaps by our algorithm, described in Section 3.2.2. After creating the function, we will take the function in to dewarping image process.

3.2.1 Recording the image coordinates

According to the graph-sheet images, there are two different images as follows:

1) The image of graph-sheet laid on a planar surface is called a planar graph-sheet image.

2) The image of graph-sheet laid on a curve surface is called a curve graph-sheet image.
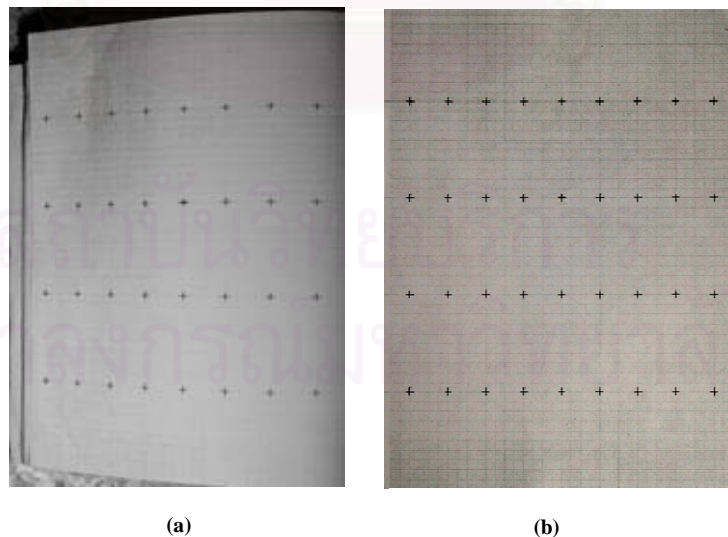
The two graph-sheets images are shown in Figure 3.3.

(a)   (b)

Figure 3.3 The two Graph-sheets in the experiment.
(a) Curve graph-sheet (b) Planar graph-sheet

To creating a dewarping function, at the beginning, we must find out the position of coordinates data from two graph-sheet images captured by our digital camera with resolution of 1600 x 1200 pixels. We determine the position of pixel coordinates *(X, Y)* (see Figure 3.4) by using Adobe Photoshop CS2 software at the defining point (along the dash lines) on the images for 18 x 24 size of data, as shown in Table 3.1 .
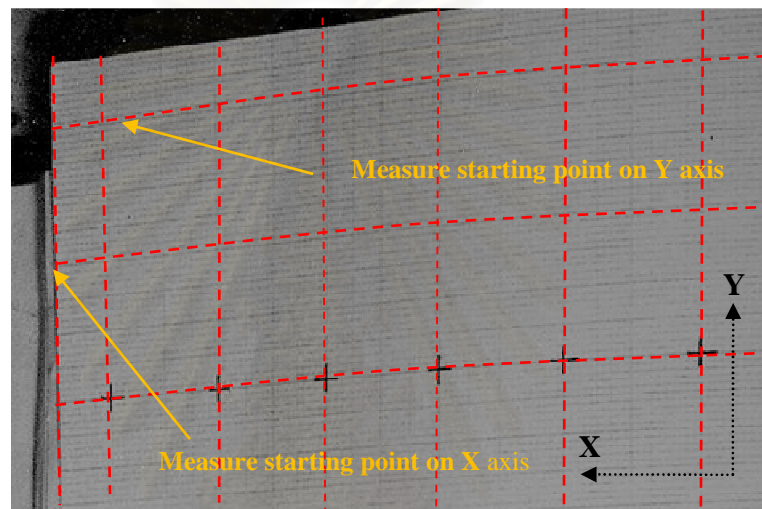
Figure 3.4 The determining measure point on the cross point of dash line

From Table 3.1 each line of coordinates data can be written as a matrix W of coordinates as shown in Equation 3.1. In our experiment, there are 18 lines each of which consists of 24 *(x, y)* coordinates.

$$W = \begin{bmatrix} x_{11}y_{11} & x_{12}y_{12} & \cdots & x_{1n}y_{1n} \\ x_{21}y_{21} & x_{22}y_{22} & \cdots & x_{2n}y_{2n} \\ \vdots & \vdots & & \vdots \\ x_{m1}y_{m1} & x_{m2}y_{m2} & \cdots & x_{mn}y_{mn} \end{bmatrix} \qquad (3.1)$$

Where $m = 18$ and $n = 12$

Table 3.1 Coordinate data sample from the graph-sheets images

Line 1		Line2		Line3		Line4	
x1	**y1**	**x2**	**y2**	**x3**	**y3**	**x4**	**y4**
288	182	288	292	288	401	288	512
331	177	331	288	331	399	331	510
374	171	374	283	374	395	374	507
418	166	418	279	418	391	418	504
464	161	464	274	464	387	464	500
508	155	508	269	508	383	508	497
552	149	552	264	552	380	552	495
599	143	599	259	599	375	599	492
648	137	648	254	648	371	648	489
698	132	698	250	698	368	698	486
753	127	753	246	753	364	753	484
809	123	809	242	809	362	809	482
867	119	867	239	867	359	867	480
928	117	928	237	928	358	928	479
988	117	988	237	988	358	988	478
1049	117	1049	237	1049	358	1049	479
1104	121	1104	240	1104	360	1104	481
1138	130	1138	247	1138	366	1138	486

3.2.2 Cubic spline Interpolation

From the previous section, the sizes of both graph-sheets image are
1600x1200. Obviously, however, the coordinates data recorded by the software have
only 18 x 24 size of coordinates matrix. Thus, we must create the function having the
size closely to graph-sheets images. Firstly, we interpolate the data from the
coordinates matrix by using cubic spline interpolation method, see Figure 3.5.
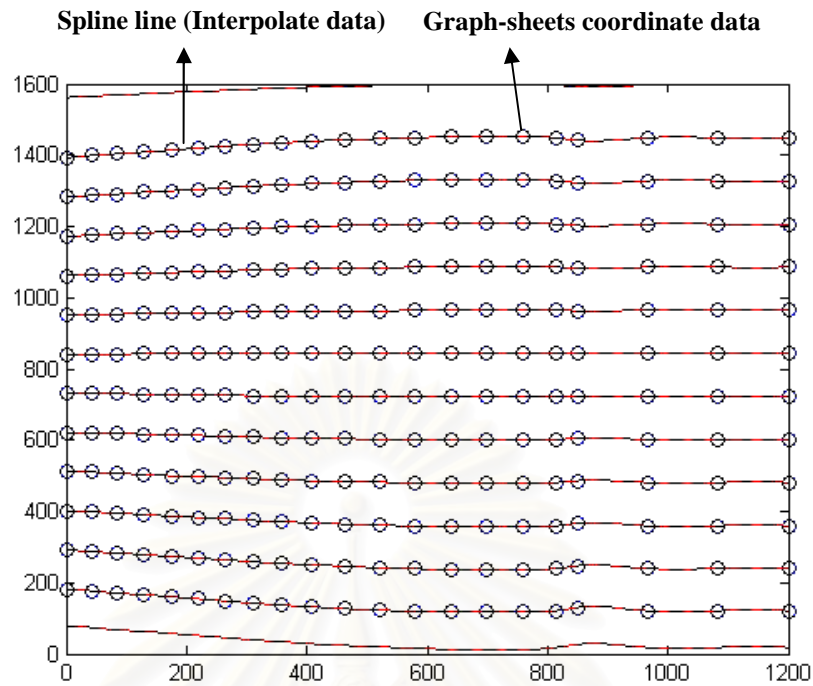
Spline line (Interpolate data) **Graph-sheets coordinate data**

Figure 3.5 List of spline lines,

Dot markers assisted to generate splines along the book surface.

From Figure 3.5, the graph-sheets coordinates data are substituted with dot markers and solid lines are the result from the spline interpolation. However, the gaps appear on the graph between the lines (see Figure 3.6), cannot make our function perfectly.
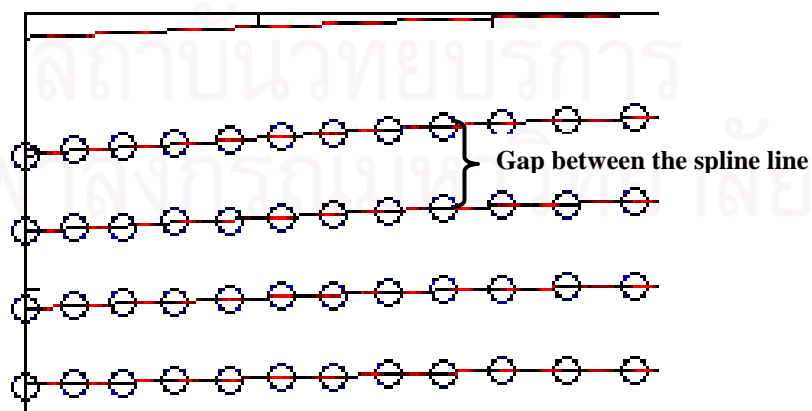

Figure 3.6 The gaps appearing between the spline lines.

Thus, we have to generate data in row between each of the splines lines by our algorithm as depicted in Figure 3.7 called row imputation algorithm. Our algorithm helped us to ensure that all type of data – recording coordinate data, interpolated data and imputation data – are spread along the image surface and the function is a good representation of the document surface.

GR = Array of Gap distance between interpolated data of Planar graph-sheet.

GC = Array of Gap distance between interpolated data of Curve graph-sheet.

For $i = 1$: number of GR array

$$n = \frac{GC(i)}{\text{number of period}};$$

$$prop = \frac{GC(i)}{GR(i)} \times n;$$

$$new\_period = n \times prop;$$

up = Spline line below;

$down$ = Spline line above;

For $j = 1$: periods number

$$sub\_periods = new\_period \times j;$$

$$newline = up + sub\_periods;$$

if $newline <= down$

store in LUT;

end if

end for

end for

Figure 3.7 Row imputation algorithm

According to the algorithm, the purpose of our algorithm is to impute the lines in row of our function. The process begin with reading the Gap data of both graph-sheets images, which are the average size of gaps between each of the spline lines (upper and below) along the image plane, as shown in Figure 3.6 previously. The average gaps data are denoted as *GR* and *GC* being the arrays of double. Next, we divide each values of *GC* in to periods by periods number 12. Each size of the periods is denoted as *n*. After that, the *new_periods* is made by creating a proportion: $\frac{GC}{GR} \times \boldsymbol{n.}$ The *new_periods* variable just having been created previously are appropriate to imputing the line being parallel to the spline lines in row. Finally, we impute the data by adding the upper spline line *up* with *sub_period* = *new_periods x j* in row, where *j* is a For Loop from 1 to periods number. The result of our algorithm is shown in Figure 3.8.
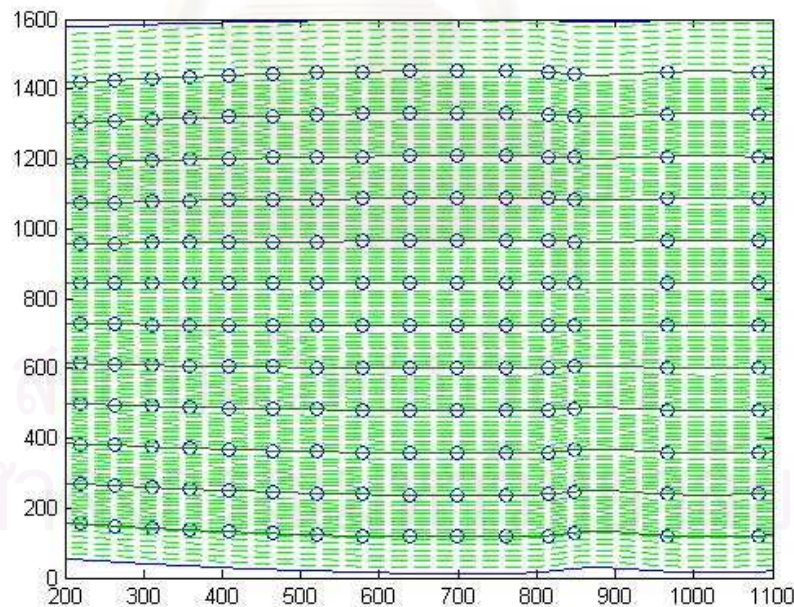

Figure 3.8 Result of rows imputation algorithm, dash lines being the imputation data.

After impute the lines, the gaps still appears but the imputing lines are quite parallel to the spline line (see Figure 3.9). Then we impute another data in to the gaps linearly. Finally, we have the dewarping function as well as look-up table (LUT) with the size of 1585 x 1085 corresponded to original curve graph-sheets image size to transform the curve coordination system in to Cartesian coordinate system as shown in Figure 3.10.
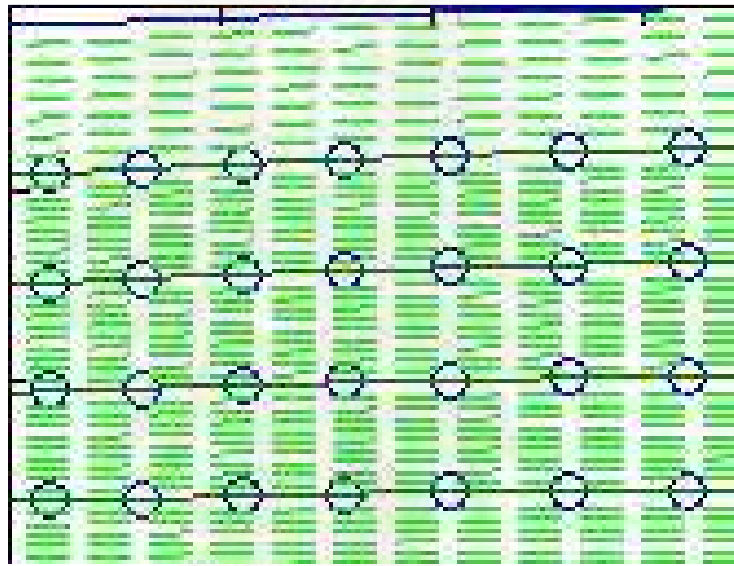

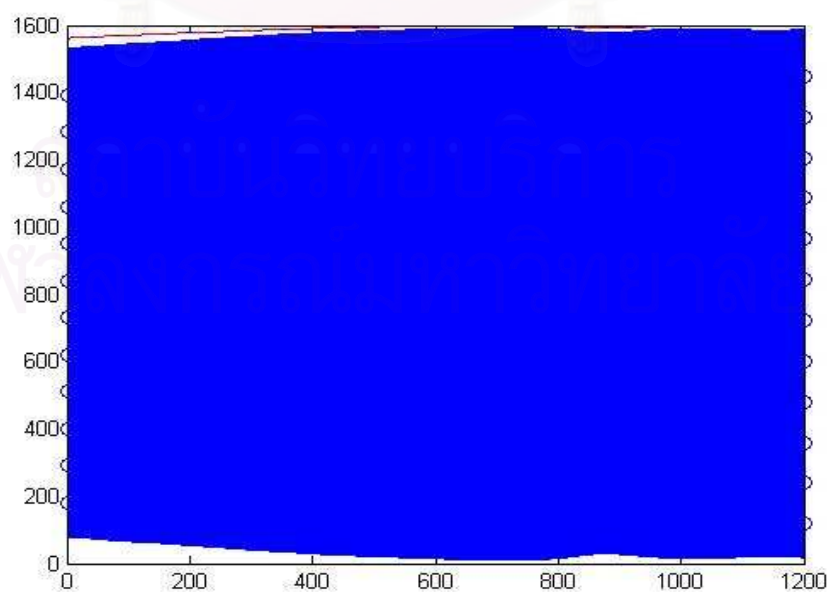**Figure 3.9 The gap still appear after imputing the data.**

Figure 3.10 Final lookup-table of curve graph-sheet.

3.3 Dewarping Image

To dewarp a given image, we must convert an original document image or warp document image into the dewarp image by the obtained function from previous section. Our initial result image is shown in Figure 3.11 (a).

(a)

(b)

(c)
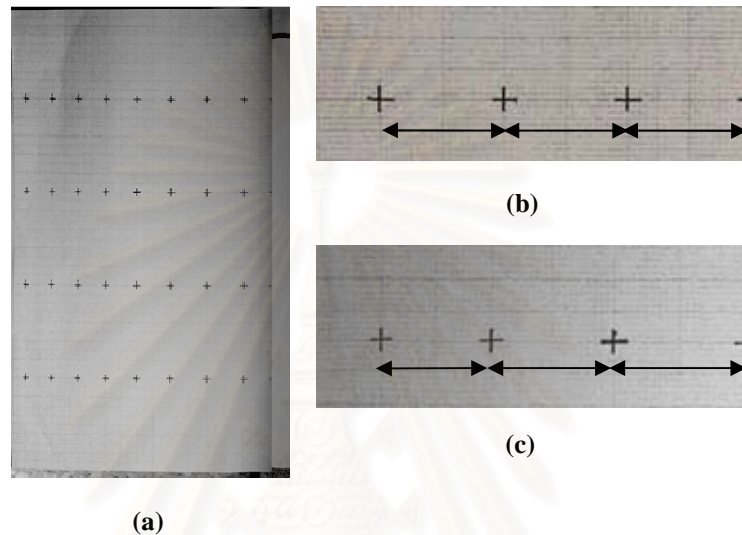
Figure 3.11 Dewarping image using the Dewarping function
(a) Our initial result after dewarping
(b) Showing distance of planar graph-sheet image in X axis
(c) Showing distance of our initial result image in X axis

From the result image, it shows that the proportion along the column of our result image is not appropriate (see Figure 3.11 (b) comparing with Figure 3.11 (c)) due to the curvature of the book surface. Thus, we create the other process to expand the image along the column by our pixel imputation algorithm as shown in Figure 3.12.

DX = read distance data between coordinates
point that is recorded previously

$$Prop = \frac{max\,(DX)}{max(DX)-DX} \; ;$$

For i = 1 : number of *DX*

    Old = x(i);

    New = x(i);

      While Old < x(i+1)

        Impute = prop(i) × L;

        Loop = Prop(i);

          For j = 1 : Loop

          imNew(New) = imOld(Old);

          New ++; Old++; Count++;

           If Count = Impute

            For k = 1: row

             ImNew(k,New) = mean(Imold(row-1:row+1,old-1:old);

             New++;

            End For

           End If

          End For

        L++;

      End While

End For

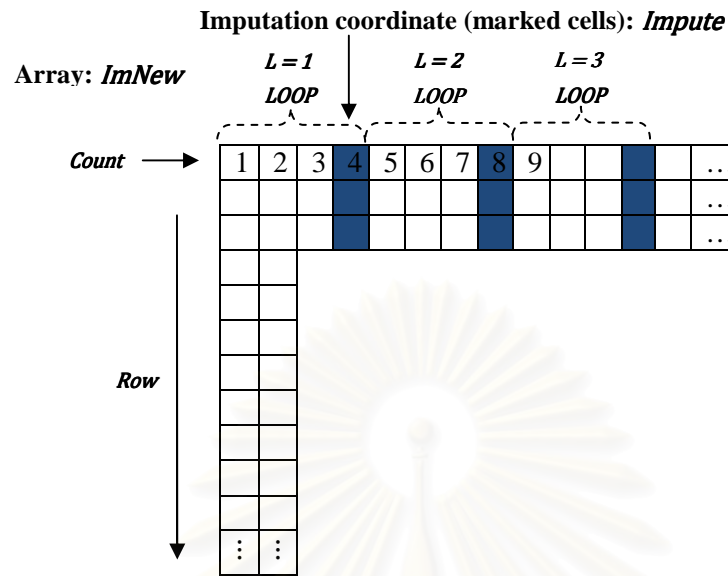**Figure 3.12 Column imputation algorithm**

Figure 3.13 Method to selecting the position to impute the pixels.

According to our algorithm, the process starts with creating distance data in curve graph-sheets column denoted as array of *DX*. Then, we make the array of proportion *Prop* helping to define the period of pixel imputation, computing from $\frac{max(DX)}{max(DX)-DX}$. Next, we define the 2 variables *New and Old* to keep the coordinate values *x* (of *x, y* graph-sheets image coordinates) of input and output images respectively. After that, we look for the purpose pixel coordinates to impute the pixels called imputation coordinates for *imnew* array (see Figure 3.13) by variables: *Impute*, increased following the variable *L* from *Prop(i) * L* with the condition that is *Old < x(i+1).* However, if that coordinate is not the imputation coordinate, the process will copy the pixel from *Imold* image array to *Imnew* image array – *Imold* belongs to the dewarp image and *Imnew* belongs to the final pixel imputation image. In the final process, this process is a pixel average before impute the pixel values in to the imputation coordinates by averaging the pixels of the *Imold* image at the determine

coordinates (*Row, Old*) into *Imnew* image (see Figure 3.14). Finally, the final image after impute is shown in Figure 3.15 (C).
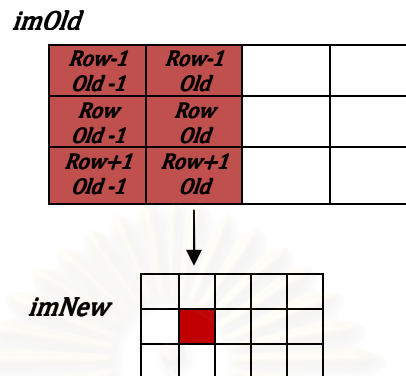

**Figure 3.14 Average of six pixels of *imOld* imputing
in one pixels of *imNew***

(a)　　　　　　　　(b)　　　　　　　　(c)
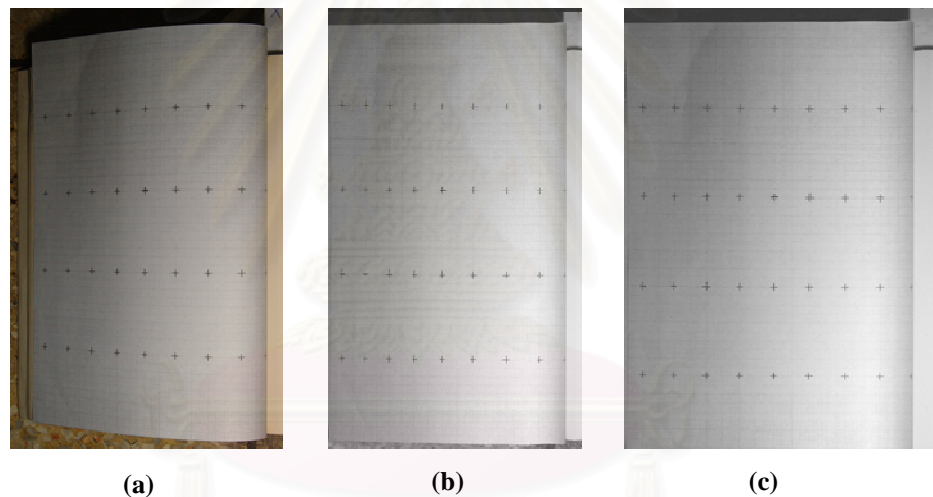
**Figure 3.15 Comparing image results (a) Original image (b) Dewarp
image (c) Final dewarp image after pixel imputation process**

3.4 Deshading Function

After we dewarp the document images, on the dewarp image, there is a shade appearing on our dewarp image due to position of the light source and warping of the book surface or sometime called uneven lighting effect. According to the shade, the shade emerges in 2 forms – shadow and light shade (see Figure 3.16). The light shade is happened because this area is adjacent to the light source; in contrast, the shadow

shade is taken place due to its being farther from the light source. In the shadow area, the image may lose their detail due to the shade being dark.
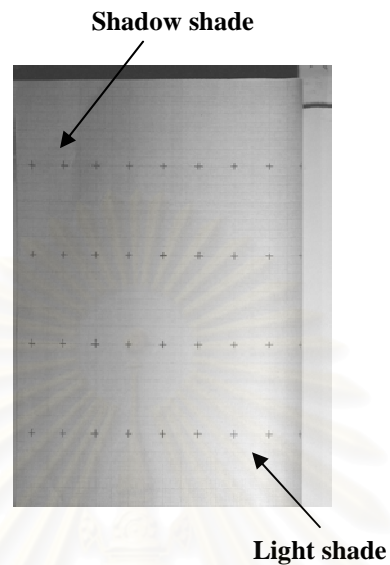

Figure 3.16 The shade appearing on the warp document image in two forms as Shadow shade and light shade respectively.

From this problem, we have to deshading the warp document image by using least square errors method (LSE). According to the LSE, it is a method used to find the function of the best-fitting curve from the given set of points. In our works, we use the LSE to get the function of overall intensities of the images and fitting to the purpose given set points. These mean that we have the two sub-process to deshading our dewarp document images.

In order to get the function of image's shade, we have to recoord the overall intensities of the images with the 2 shades on it by making the gray-scale image as shown in Figure 3.17. The gray-scale has 0-100% of 21 gray-levels increasing 5 % to each scale. The image is captured by our digital camera and dewarping completely by our method.
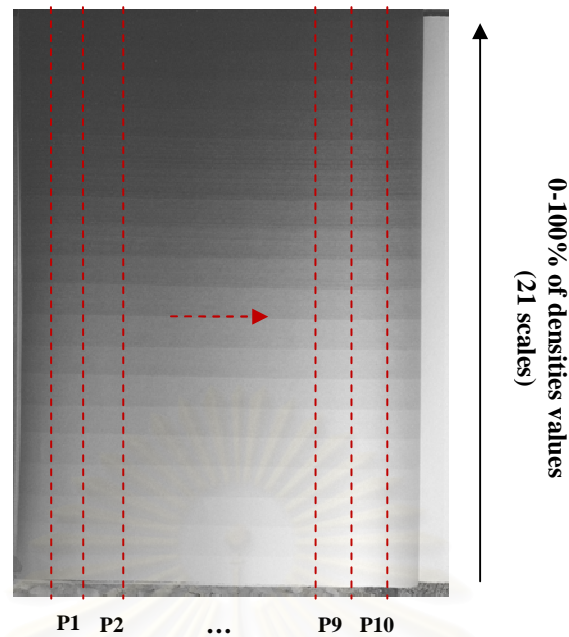
**Figure 3.17 The Gray-scale image with densities from 0 - 100% (21
grey-scales levels) increasing 5 % to each scale and divide it in to
10 periods in vertical line (see dash lines).**

After preparing the gray-scale image, we have to collect the intensities data by
divide it in to 10 periods denoted as P following in vertical line (see Figure 3.17).
Thus, we have 10 x 21 set of intensities data as shown in Table 3.2.

From the intensities data, the maximum data are in period 9. It means that this
period has the maximum intensities values. Thus, we plot all periods with the
maximum periods data as shown in Figure 3.18. The graph shows that the data is
likely to be a logarithmic form.

Table 3.2 Sample of intensities data from 10 x 21 values.

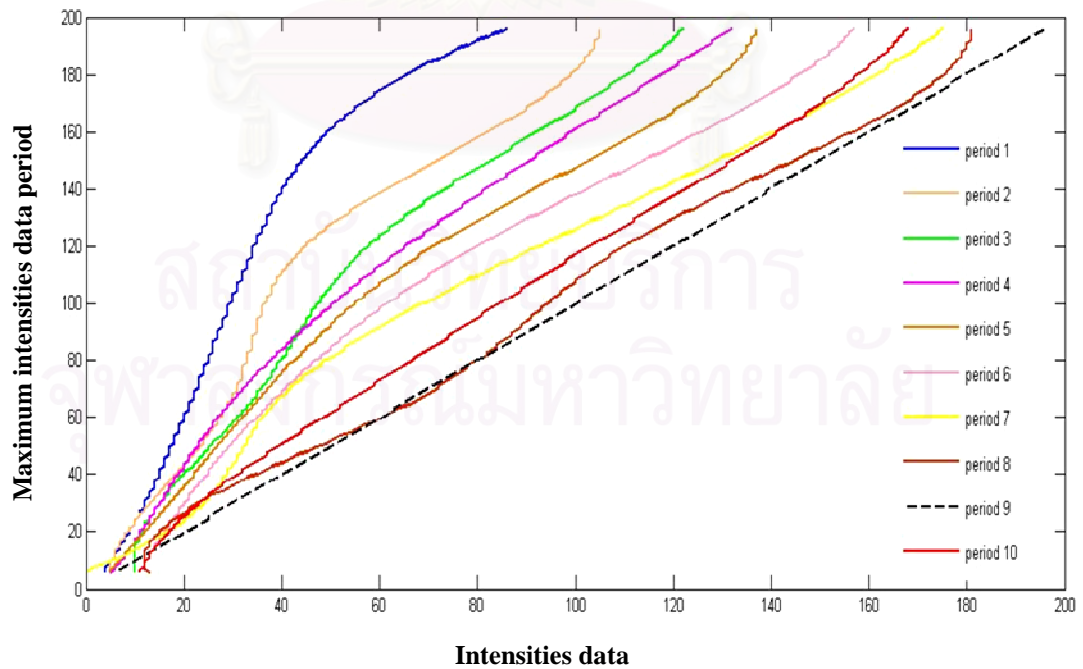| P1	P2	P3	P4		P9
7	11	9	11	...	17
8	13	13	11		19
11	15	13	16		36
13	16	14	18		42
14	17	21	25	...	46
14	21	21	26		37
19	31	31	30		60
24	39	34	40		66
24	42	41	41		69
29	42	47	53	...	92
31	52	61	58		111
40	64	70	69		122
42	69	78	78		139
49	77	89	79		146
52	81	99	96	...	165
60	85	102	102		165
68	93	103	113		180
74	103	105	118		188
84	104	113	125		192
87	108	118	131		195
102	110	120	136	...	197

Figure 3.18 The graph between 9 periods and maximum period.

Thus, we choose an initial form of function as a logarithmic function forms as shown in Equation 3.2.

$$y = a + b \ln x \qquad (3.2)$$

the coefficients can be found from least squares fitting as Equation (3.3) and (3.4)

$$b = \frac{n \sum_{i=1}^{n}(y_i \ln x_i) - \sum_{i=1}^{n} y_i \sum_{i=1}^{n} \ln x_i}{n \sum_{i=1}^{n}(\ln x_i)^2 - (\sum_{i=1}^{n} \ln x_i)^2} \qquad (3.3)$$

$$a = \frac{\sum_{i=1}^{n} y_i - b \sum_{i=1}^{n}(\ln x_i)}{n} \qquad (3.4)$$

When we find the coefficients data a and b, we plot the graph to see the relation between the two coefficients as show in Figure 3.19.
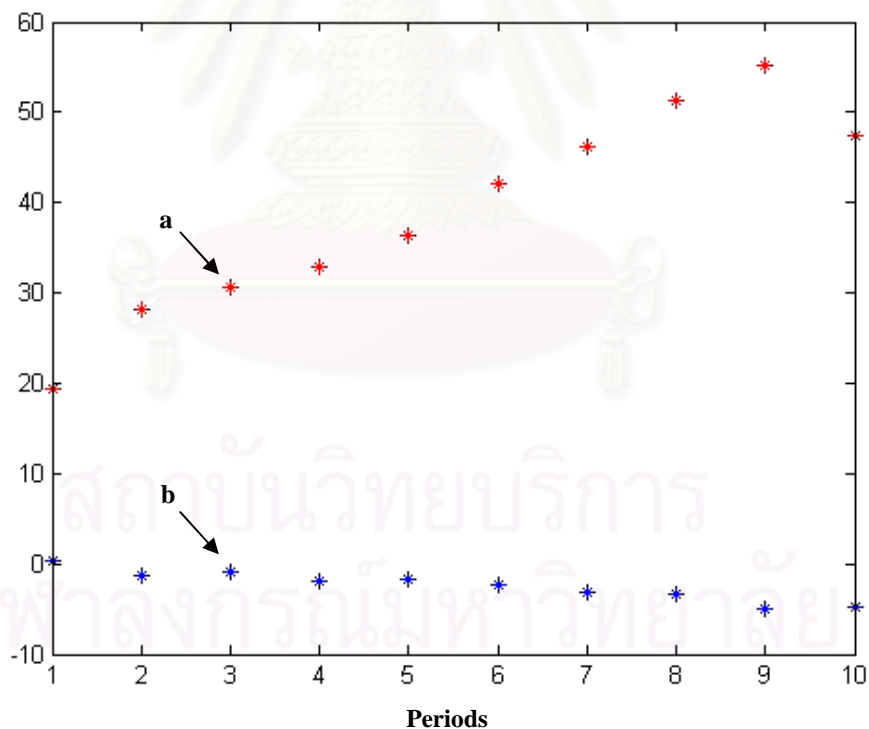

Figure 3.19 The two coefficients data from intensities data base on logarithmic function.

From the graph, a and b coefficients values are likely to be a polynomial degree 2 and

linear form respectively. Thus, we define the form of 2 coefficients data as Equation (3.5) and (3.6)

$$a = c_1 z^2 + c_2 z + c_3 \tag{3.5}$$

$$b = c_4 z + c_5 \tag{3.6}$$

Rewrite Equation (3.2), (3.5) and (3.6) by substituting $x = x_1$ and $z = x_2$ as follows:

$$f(x_1, x_2) = a_1 x_2^2 + a_2 x_2 + a_3 x_2 ln x_1 + a_4 ln x_1 + a_5 \tag{3.7}$$

Where variables x_2 and x_1 belong to number of 9 periods (1-9 excluding maximum periods data) and intensities data period respectively. From Equation (3.7), there are 5 coefficients to solve, to find their values, we create matrix form base on LSE method as Equation (3.8)

$$
\begin{bmatrix}
\sum_{i=1}^{n} x_{2,i}^4 & \sum_{i=1}^{n} x_{2,i}^2 & \sum_{i=1}^{n} x_{2,i}^3 ln x_{1,i} & \sum_{i=1}^{n} x_{2,i}^2 ln x_{1,i} & \sum_{i=1}^{n} x_{2,i}^2 \\
\sum_{i=1}^{n} x_{2,i}^3 & \sum_{i=1}^{n} x_{2,i}^3 & \sum_{i=1}^{n} x_{2,i}^2 ln x_{1,i} & \sum_{i=1}^{n} x_{2,i} ln x_{1,i} & \sum_{i=1}^{n} x_{2,i} \\
\sum_{i=1}^{n} x_{2,i}^3 ln x_{1,i} & \sum_{i=1}^{n} x_{2,i}^2 ln x_{1,i} & \sum_{i=1}^{n} x_{2,i}^2 (ln x_{1,i})^2 & \sum_{i=1}^{n} x_{2,i} (ln x_{1,i})^2 & \sum_{i=1}^{n} x_{2,i} ln x_{1,i} \\
\sum_{i=1}^{n} x_{2,i}^2 ln x_{1,i} & \sum_{i=1}^{n} x_{2i} ln x_{1i} & \sum_{i=1}^{n} x_{2i} (ln x_{1i})^2 & \sum_{i=1}^{n} (ln x_{1,i})^2 & \sum_{i=1}^{n} ln x_{1,i} \\
\sum_{i=1}^{n} x_{2,i}^2 & \sum_{i=1}^{n} x_{2,i} & \sum_{i=1}^{n} x_{2,i} ln x_{1,i} & \sum_{i=1}^{n} ln x_{1,i} & n
\end{bmatrix}
\begin{bmatrix}
a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5
\end{bmatrix}
=
\begin{bmatrix}
\sum_{i=1}^{n} x_{2,i}^2 y_i \\
\sum_{i=1}^{n} x_{2,i} y_i \\
\sum_{i=1}^{n} (x_{2,i} ln x_{1,i}) y_i \\
\sum_{i=1}^{n} (ln x_{1,i}) y_i \\
\sum_{i=1}^{n} y_i
\end{bmatrix}
$$

$$\tag{3.8}$$

Where y is maximum intensities data period (period 9).

3.5 Deshading Image

After we obtain the coefficients values $a_1 \dots a_5$ from Equation (3.8), we

represent them in to the Equation (3.7). Then, we can get new intensities of the dewarp document image with no shade along its plane. The result is shown in Figure 3.20 (b).
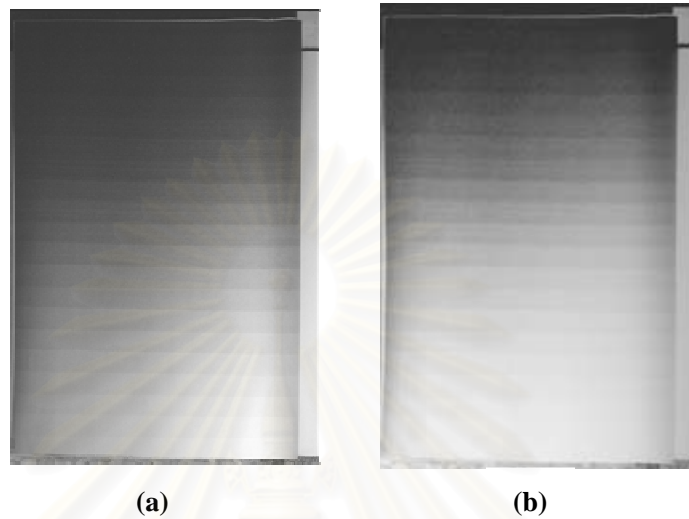

(a)   (b)

Figure 3.20 The rectified image from the deshading image process (a) dewarp image (b) rectified image

EXPERIMENTAL RESULTS

From previous chapter, we introduce a novel dewarping technique for camera-based document   images. To implement the proposed approach in MATLAB scientific software package, the program was run on MACBOOK Core2DUO 2.0 with a gigabyte memory. Our techniques tested against the following data sets :

(1) Text documents with varying font sizes between 12 – 18 pt.

(2) Text documents with eight combination of varying font styles: bold letters, italic letters, or/and underlined letters. The sample is shown in Figure 4.1

(3) One-page documents with pictures.
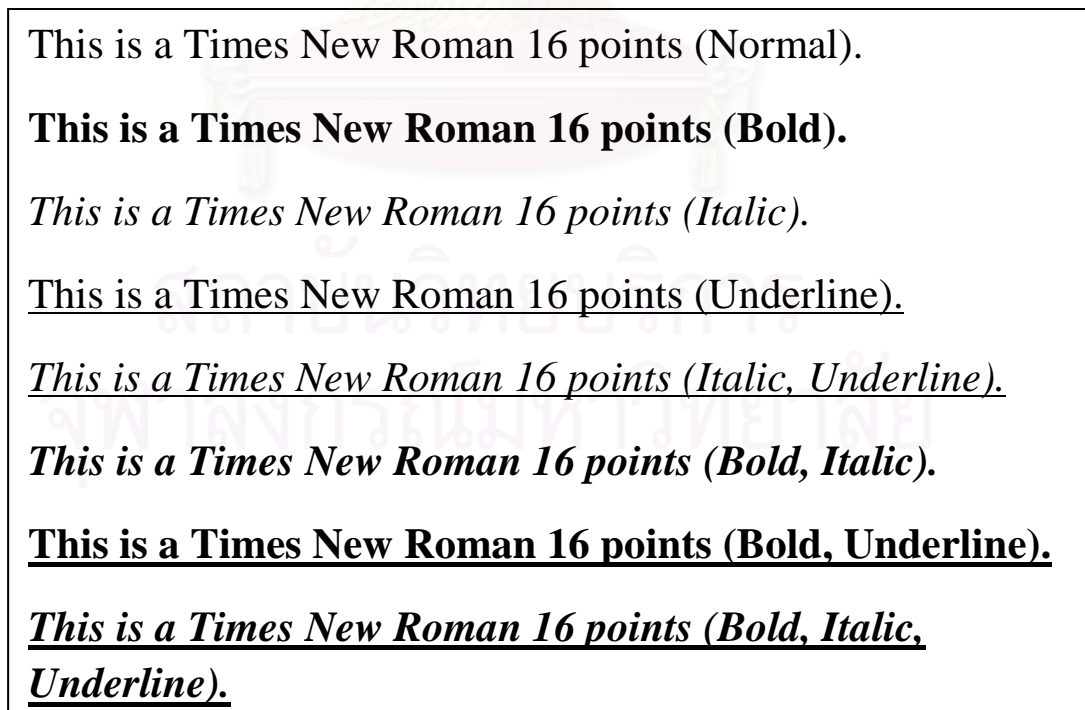
(4) Page documents with varying book bound volumes.

This is a Times New Roman 16 points (Normal).

This is a Times New Roman 16 points (Bold).

This is a Times New Roman 16 points (Italic).

This is a Times New Roman 16 points (Underline).

This is a Times New Roman 16 points (Italic, Underline).

This is a Times New Roman 16 points (Bold, Italic).

This is a Times New Roman 16 points (Bold, Underline).

This is a Times New Roman 16 points (Bold, Italic, Underline).

Figure 4.1 Text-page document image with the various font styles at 16 pt.
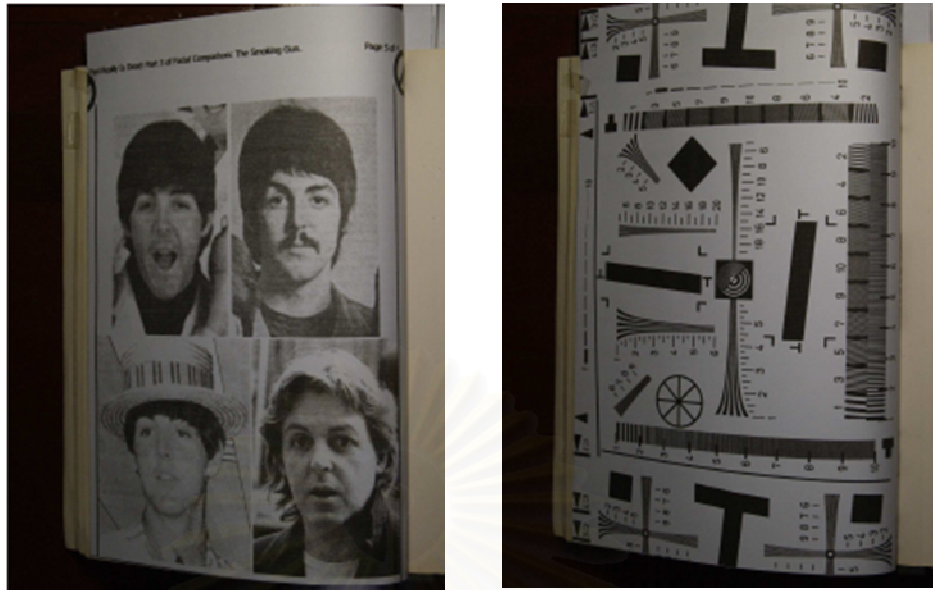
Figure 4.2 One-page Document images with picture.

From Figure 4.1, total test images in first and second data sets were 3 combinations, each of which page layout was one column with typing English letters of Times New Roman. The test images size of 1200×1600 pixels were captured by a digital camera model of Olympus E-1. The performance of our method was evaluated by OCR accuracy percentages using *Abby FineReader Professional 9* and *OmniPage Professional 16* as widely used in several researches [2, 6]. According to *Abby FineReader and OmniPage Professional*, their functions of the curvature correction based on text-line detection and geometric detection respectively were used for accuracy comparisons [6]. Junker [2] proposed the standard measurements, which can prove that our algorithm has a better or worse effectiveness than another algorithm, such as recall, precision, and accuracy. Our accuracy evaluation was defined as follows:

1) $Character\ Precision = \dfrac{Number\ of\ characters\ correctly\ detected\ by\ OCR}{Total\ number\ of\ characters\ detected\ by\ OCR}$

2) $Word\ Precision = \dfrac{Number\ of\ words\ correctly\ detected\ by\ OCR}{Total\ number\ of\ words\ detected\ by\ OCR}$

3) $Character\ Recall = \dfrac{Number\ of\ characters\ correctly\ detected\ by\ OCR}{Total\ number\ of\ characters\ in\ the\ document}$

4) $Word\ Recall = \dfrac{Number\ of\ words\ correctly\ detected\ by\ OCR}{Total\ number\ of\ words\ in\ the\ document}$
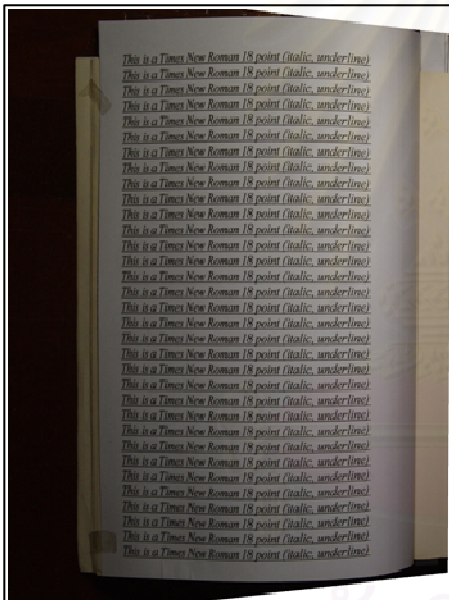
Using the FineReader's curvature correction engine, of which are deskewed image and straighten text line, it automatically detects the text component on the document ,and finally deskews and dewarps the document image. However, the OmniPage's curvature correction engine, 3D deskew, it automatically detects the geometric attribute of the image by building the mesh grids mapping on the image . It can dewarp the document image and allows users adjusting the grids manually to increase the accuracy. However, their algorithms could not applied to some document images as shown in Table 4.1 and Figure 4.3, which applicable and not applicable rectification were represented by A and NA, respectively.
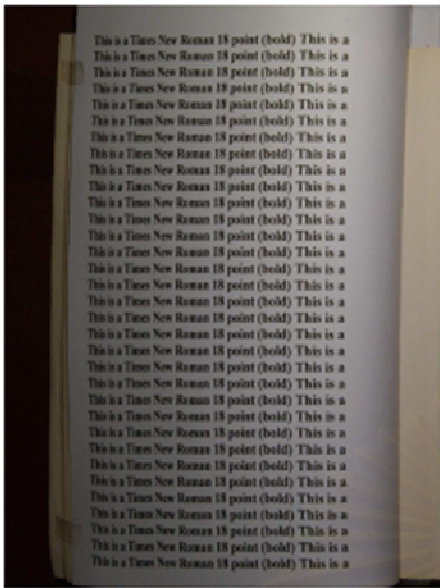
(a)

(b)

(c)

(d)

**Figure 4.3 Results from FineReader's and Omni's curvature correction engine.
(a) Valid document content from FineReader approach denoted as A (b) Invalid
document content from FineReader approach denoted as NA (c) Valid document
content from Omnipage approach denoted as A (d) Invalid document content
from Omnipage approach denoted as NA**

Table 4.1 Applicable or not applicable to 32 combinations of document images with various Font styles comparing between our technique (Our tech) and Abby's curvature correction.
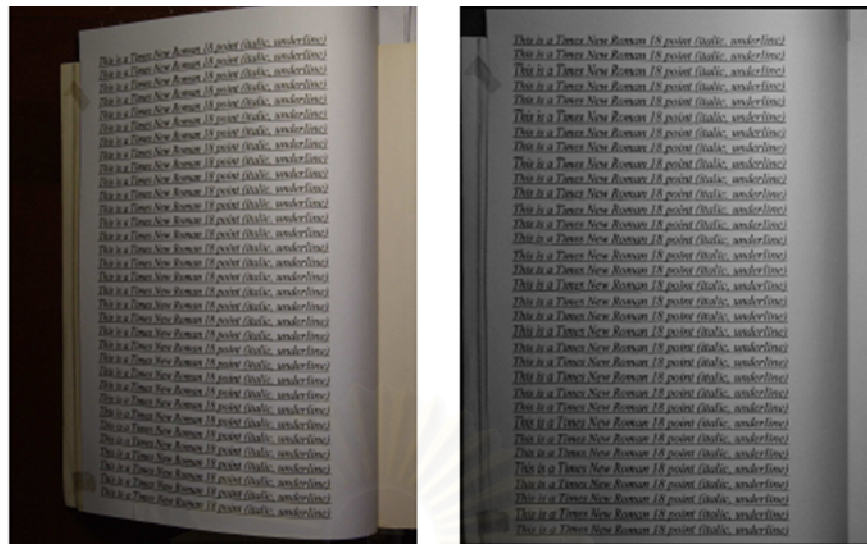**Note: Font code with # [NBIU] is an abbreviation of Font size, Normal, Bold, Italic and Underline**

Font	Our tech using Abby	Our tech using Omni	Abby's	Omni's
18n	A	A	A	A
18b	A	A	A	A
18i	A	A	NA	A
18u	A	A	NA	A
18bi	A	A	A	A
18bu	A	A	NA	A
18iu	A	A	NA	A
18biu	A	A	A	A
16n	A	A	A	A
16b	A	A	NA	A
16i	A	A	NA	A
16u	A	A	NA	A
16bi	A	A	A	A
16bu	A	A	A	A
16iu	A	A	NA	NA
16biu	A	A	A	A
14n	A	A	NA	A
14b	A	A	A	A
14i	A	A	A	A
14u	A	A	A	NA
14bi	A	A	A	A
14bu	A	A	A	A
14iu	A	A	NA	A
14biu	A	A	NA	NA
12n	A	A	A	A
12b	A	A	A	A
12i	A	A	A	A
12u	A	A	NA	A
12bi	A	A	A	A
12bu	A	A	NA	A
12iu	A	A	NA	A
12biu	A	A	NA	A

For the text-page document images, we test our text-page document data set by our method. Our method includes the 2 processes as dewarping and deshading. The result images by using dewarping function are shown in Figure 4.4.

(a)

(b)

Normal font-style

Bold font-style

Underline font-style

Italic font-style

Bold and underline font-style

Italic and underline font-style

Bold and Italic font-style

All font-style

(c)

Figure 4.4 Experiments on text-page document images with the following font styles : 18 pt Italic Underline Time New Roman (a) Original document Image (b) Dewarp document image (c) Eight various font style results.

After we get the dewarp image from our method, we test our method performance by OCR to convert the text being on our text-page document. The results of our method performance comparing with OCR's curvature correction are shown in Table 4.2 - 4.3 with excluding NA and including NA for Table 4.4 – 4.5.

Table 4.2 Average word and character recall rates (%) for varying Font sizes excluding NA image test

OCR	Font size	Average word recall		Average character recall	
		Our method	FineReader	Our method	FineReader
FineReader	18 pt.	99.43	95.76	99.90	98.82
	16 pt.	99.30	82.83	99.90	91.15
	14 pt.	98.76	82.27	99.42	87.91
	12 pt.	88.32	78.62	91.32	86.07
		Our method	OmniPage	Our method	OmniPage
OmniPage	18 pt.	99.49	96.45	99.80	97.52
	16 pt.	98.49	96.72	99.50	98.23
	14 pt.	97.00	97.40	98.74	97.50
	12 pt.	97.13	94.78	98.17	96.40

Table 4.3 Average word and character precision rates (%) for varying Font sizes excluding NA image test

OCR	Font size	Average word precision		Average character precision	
		Our method	FineReader	Our method	FineReader
FineReader	18 pt.	99.57	96.65	99.87	99.21
	16 pt.	99.55	88.53	99.85	94.21
	14 pt.	98.80	87.85	99.41	92.45
	12 pt.	89.25	85.07	93.83	90.21
		Our method	OmniPage	Our method	OmniPage
OmniPage	18 pt.	99.63	97.76	99.86	98.41
	16 pt.	98.89	97.54	99.68	98.67
	14 pt.	97.04	97.45	98.73	97.51
	12 pt.	97.40	94.97	98.76	97.30

Table 4.4 Average word and character recall rates (%) for varying Font sizes image test

OCR	Font size	Average word recall		Average character recall	
		Our method	FineReader	Our method	FineReader
FineReader	18 pt.	99.43	47.88	99.90	49.41
	16 pt.	99.3	41.42	99.90	45.58
	14 pt.	98.76	51.42	99.42	54.95
	12 pt.	88.32	39.31	91.32	43.28
		Our method	OmniPage	Our method	OmniPage
OmniPage	18 pt.	99.49	96.45	99.80	97.52
	16 pt.	98.49	84.63	99.50	85.95
	14 pt.	97.00	73.05	98.74	73.13
	12 pt.	97.13	94.78	98.17	96.40

Table 4.5 Average word and character precision rates (%) for varying Font sizes image test

OCR	Font size	Average word precision		Average character precision	
		Our method	FineReader	Our method	FineReader
FineReader	18 pt.	99.57	47.83	99.87	49.61
	16 pt.	99.55	44.27	99.85	47.10
	14 pt.	98.80	54.91	99.41	57.78
	12 pt.	89.25	42.53	93.83	45.06
		Our method	OmniPage	Our method	OmniPage
OmniPage	18 pt.	99.63	97.76	99.86	98.41
	16 pt.	98.89	85.34	99.68	86.33
	14 pt.	97.04	73.09	98.73	73.13
	12 pt.	97.40	94.97	98.76	97.30

From Table 4.2-4.3, our results improve 11.58 % and 6.65 % for word

and character recall, respectively. The results also yield the improvement of 7.73 %

and 4.22 % for word and character precision respectively. these percentage excludes

NA. In case of including NA shown in table 4.4-4.5, testing with FineReader, the

results improve 52.63 % and 48.15 % for words and character recall. The results also

yield the improvement of 49.51 % and 48.35 % for word and character precision. In

the same manner, testing with OmniPage, the results improve 10.8 % for word and

character recall and 10.45 % and 10.53 % for word and character precision. In the

summary, our method outperforms the FineReder's correction engine up to 48 – 52 % for words and character recall and 48 – 49 % for word and character precision respectively and the OmniPage's engine up to 10 - 11% for word and character of both recall and precision. Our results from both FineReader and OmniPage show that the 18pt. font size from all font styles yield the best accuracy. Then we also examined the eight 18pt. text-pages with varying book bounds volumes for finding the accuracy of our dewarping function to other bound volumes of book. We start at 10 to 40 sheets. The results are shown in Table 4.6 and 4.7. They show that the accuracy begin decrease obviously at 40 sheets.

Table 4.6 Average words and character recall rate (%) for
difference pages on a thick bound volumes

Page difference	Average recall (%) using FineReader		Average recall (%) using OmniPage	
	Word	Character	Word	Character
10	99.53	99.84	99.38	99.58
20	99.72	99.96	99.56	99.89
30	99.37	99.8	99.05	99.47
40	97.12	99.08	98.54	99.14

Table 4.7 Average words and character precision rate (%) for
difference pages on a thick bound volumes

Page difference	Average precision (%) using FineReader		Average precision (%) using OmniPage	
	Word	Character	Word	Character
10	99.60	99.82	99.70	99.90
20	99.82	99.94	99.63	99.92
30	99.41	99.86	99.48	99.75
40	97.87	99.14	99.06	99.51

According to the two One-page document images with picture, the results of our dewarping method are shown in Figure 4.5
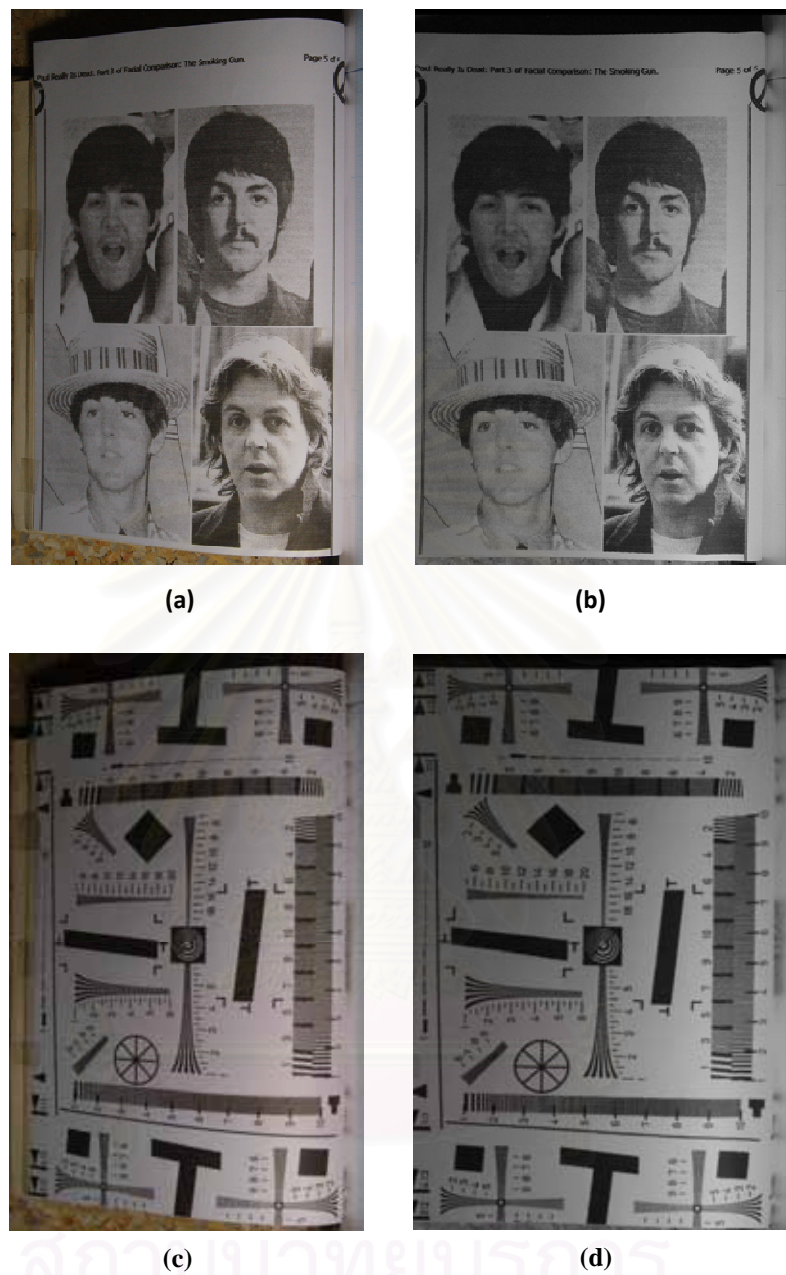
Figure 4.5 Experiment on a one-page document image with one picture (a) and (c) are warp image (b) and (d) are our result images.

From the resulting images, there is a shade along the document due to the position of the light source. According to the shade behavior, the image may be divided in two sections; the curved section located on the right hand is brighter than on the right hand and the left hand section is darker and some detail may hide in the shadow of the

image (see Figure 4.6). Thus, we deshade the images by our process called deshading image by our deshading function. The results of deshading are shown in Figure 4.7
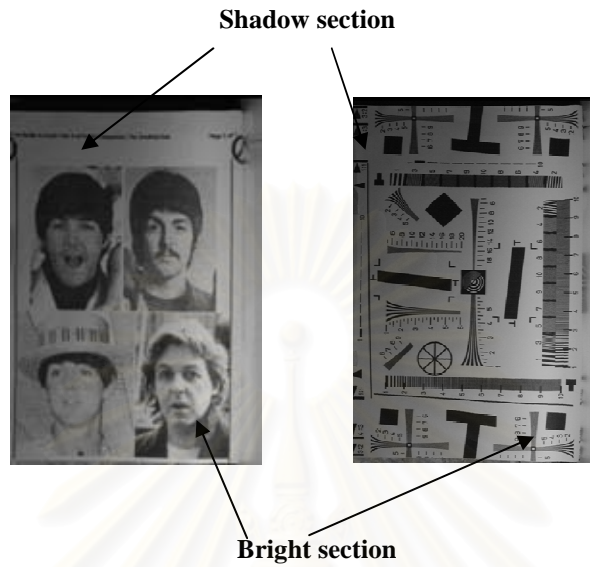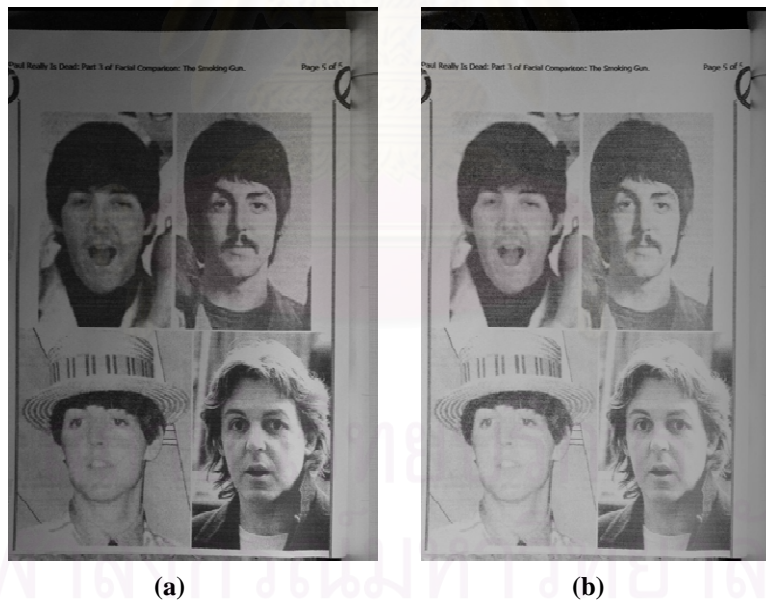
Shadow section

Bright section

Figure 4.6 The image being divided in two sections – Shadow section and Bright section

(a)   **(b)**
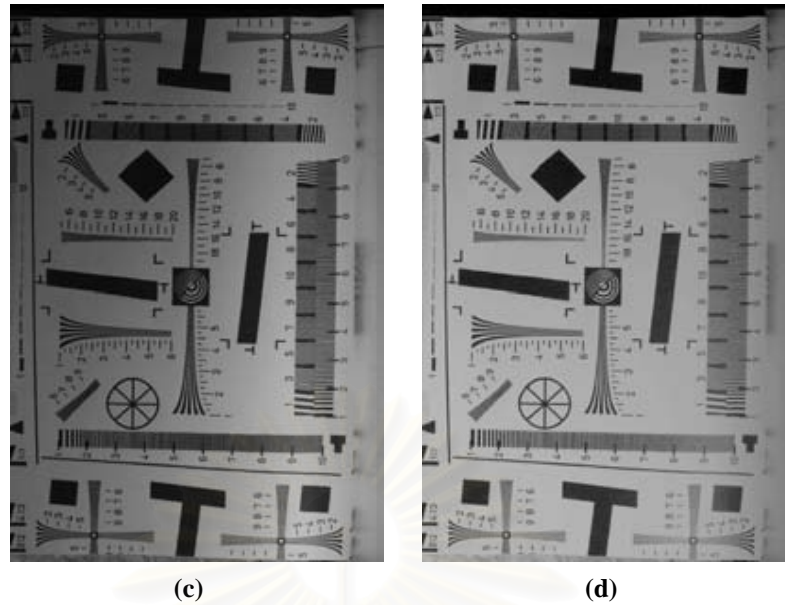
(c)　　　　　　　　　　**(d)**

**Figure 4.7 Experiment on a one-page dewarped document image with
one picture (a) and (c) are dewarped one-page document image with
one picture (b) and (d) are deshaded document image.**

The results show that the shadow on the image has been removed. However, the

image scale and brightness on the document images are preserved. The image detail in

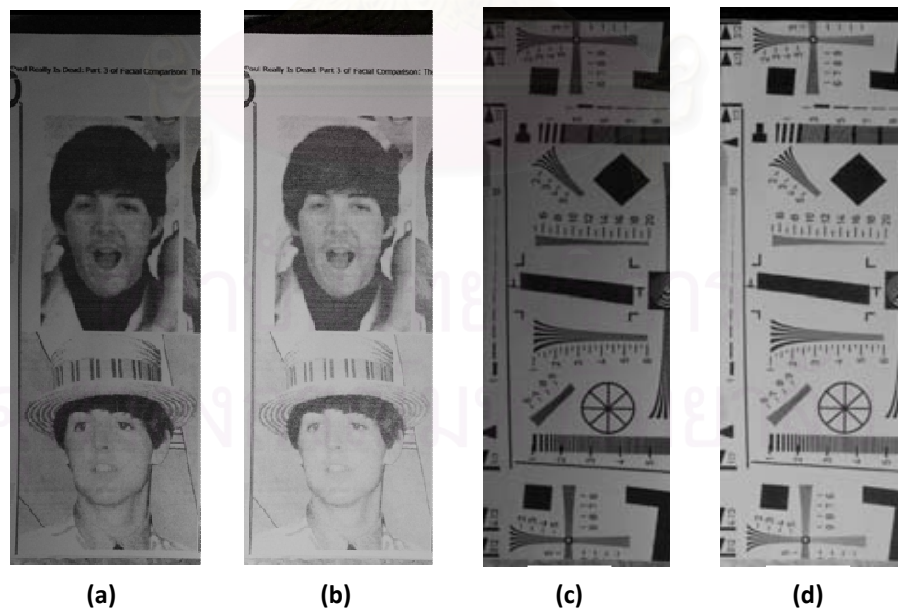the shadow section is clearer (see Figure 4.8).

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

**Figure 4.8 Rectified images (b) and (d) comparing with dewarped document
images (a) and (c)**

CHAPTER V

CONCLUSIONS AND DISCUSSIONS

5.1 Conclusions and Discussions

Finally, we evaluate the restoration results by comparing our estimated surface shape with the real shape as well as the OCR performance on original and restored document images. The results show that the geometric and photometric distortions are mostly removed and the OCR results are improved markedly. Our method outperforms the FineReader's curvature correction up to 48 and 49 percent for word and character precisions, respectively and 10-11% for the Omnipage's curvature correction. As a result, our approach was capable to rectify all documents while the other failed to rectify some document images. In the experimental results, the capability of our approach is limited to the font size of 12 pt., which yields the worst efficiency as a small size of document images. According to the OCR's curvature correction engine, the result show that OmniPage's engine has more efficiency than FineReader's engine due to it allows the user adjust the engine manually rather than only automatically warping the image as in FineReader. Our approach is independent to neither font styles such as bold, italic, and underline letters nor serif or non-serif letters. For the camera test chart image, the straight lines rotated through 45 degrees were not restored perfectly by using our approach as shown in Figure 5.1. Our approach does not cover the problem of text restoration near and at the spine of the book. For the deshading image, our deshading approach can remove a shade on the document image clearly and make the image brighter than the original. All of

document image components (text and graphic) are quite obvious, better seeing their details.
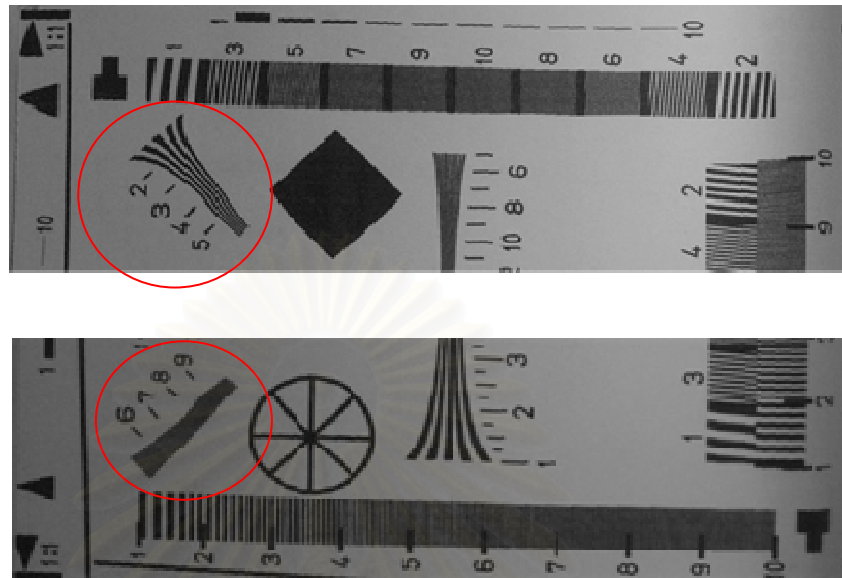

Figure 5.1 The straight line through 45 degree not being restored perfectly

5.2 Future works

In future works, the dewarping function should be adjusted especially in row imputation algorithm to solve the non straight line through 45 degree problem and in column imputation algorithm to find a new method that is better than average the pixel before impute in to an imputation coordinate. Next, we will increase Text quality for more accuracy of the OCR recognition.

According to the commercial version, if possible, the image coordinates will be preferred to being recorded automatically and the function as a look up table will be stored in database being various in book page sizes and book bound volumes. The commercial software should automatically detect the document skew, warped and shade taking place along the book surface. The equipment setting may be set as instant equipment such as camera fixing, which is better than a camera tripod, for more accuracy of the image result and reducing the errors as possible.

REFERENCE

1. Fu, B., Wu, M., Li, R., Li, W., Xu, Z., Yang, C., A Model-based Book Dewarping Method Using Text Line Detection, Second International Workshop on Camera-Based Document Analysis and Recognition, 22 Sep. 2007

2. Junker, M., Hoch, R., Dengle, A., On the Evaluation of Document Analysis Components by Recall Precision and Accuracy, International Conference on Document Analysis and Recognition, IEEE Proceedings on Vol.1, 1999, pp. 713 - 716

3. Kaufman, H., Tekalp, A.M., Survey of estimation techniques in image restoration, Control Systems Magazine 11, 1(1991): 16-24.

4. Liang, J., DeMenthon, D., Doermann, D., Flattening curved documents in images, Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, Volume 2, 20-25 June 2005, pp. 338 – 345

5. Ma, Y., Wang, C., Dai, R., Correcting Bound Document Images Based on Automatic and Robust Curved Text Lines Estimation, ICCPOL 2006, Proceedings. Lecture Notes in Computer Science 4285 Springer 2006, pp. 197-204.

6. Schneider, D.C., Block, M., Rojas, R., Robust Document Warping with Interpolated Vector Fields, Ninth International Conference on Document Analysis and Recognition, IEEE Proceedings on Vol.1, 2007, pp. 113 – 117

7. Tan, C.L, Zhang, L., Zhang, Z., Xia, T., Restoring Warped Document Images through 3D Shape Modeling, Pattern Analysis and Machine Intelligence, Transaction on, Vol. 28, No. 2, February 2006, pp. 195 – 208

8. Tan, C.L., Zhang, L., Zhang, Z., Xia, T., Restoring warped document images through 3D shape modeling, Pattern Analysis and Machine Intelligence, IEEE Transactions on Volume 28, Issue 2, Feb. 2006, pp. 195 - 208

9. Wada, T., Ukida, H., Matsuyama, T., Shape from shading with interreflections under proximal light source-3D shape reconstruction of unfolded book surface from a scanner image, Computer Vision, 1995. Proceedings., Fifth International Conference on, 20-23 June 1995, pp. 66 - 71

10. Wu, M., Li, R., Fu, B., Li, W., Xu, Z., A Model-based Book Dewarping Method to Handle 2D Images Captured by a Digital Camera, Ninth International Conference on Document Analysis and Recognition, IEEE Proceedings on Vol.1, 2007, pp. 158 - 167

11. Wu, M., Li, R., Fu, B., Li, W., Xu, Z., A Page Content Independent Book Dewarping Method to Handle 2D Images Captured by a Digital Camera, Image Analysis and Recognition, 4th International Conference, ICIAR 2007, Montreal, Canada, August 22-24, 2007, Proceedings. Lecture Notes in Computer Science 4633 Springer 2007, pp. 1242-1253

12. Yamashita, A. Kawarago, A. Kaneko, T., Miura, K. T., Shape Reconstruction and Image Restoration of Non-Flat Surfaces of Documents with a Stereo Vision System, Pattern Recognition, 2004. Proceedings., 17th International Conference on, 23-24 August 2004, pp. 482-485

13. Zhang, L., Tan, C.L., Warped Document Image Restoration Using Shapefrom-Shading and Physically-Based Modeling, Applications of Computer Vision, 2007. WACV '07. IEEE Workshop on Feb. 2007, pp. 29 - 29

14. Zhang, L., Tan, C.L., Warped image restoration with applications to digital libraries Document Analysis and Recognition, 2005. Proceedings. Eighth International Conference on 29 Aug.-1 Sept. 2005, Volume 1, pp. 192 - 196

15. Zhang, L., Zhang, Z., Tan, C.L., Xia, T., 3D geometric and optical modeling of warped document images from scanners, Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on Volume 1, 20-25 June 2005, pp. 337 - 342

16. Zhang, Z., Lim, C.L., Fan, L., Estimation of 3D shape of warped document surface for image restoration Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on, 23-26 Aug. 2004 Volume 1, 2004, pp. 486 - 489

VITA

Dumrong Sutapirat was born on June 24, 1982 in Pethchaboon, Thailand. He received his B.Sc. Degree in Apply Statistic from the Faculty of Science from King Mongkut Institute of Technology Ladkrabang 2004, and he has been a graduate student in the Imaging Technology Program, Graduate school, Chulalongkorn University since 2009.