



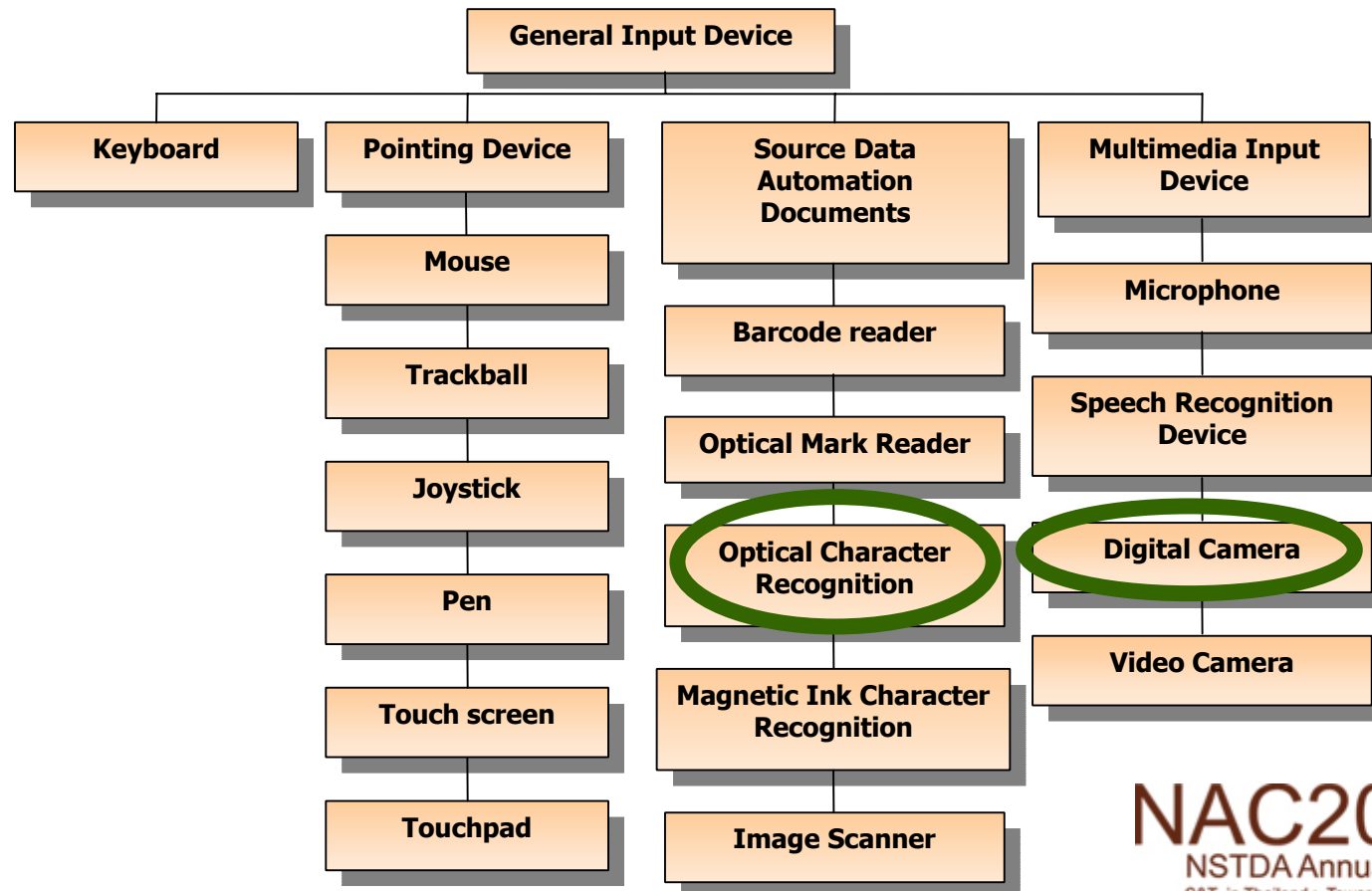
การประมวลผลส่วนหน้าสำหรับภาพจากกล้องเพื่อการรู้จำตัวอักษร

OCR Pre-Processor for Image captured from Digital Camera

ณัฐ ปิยะปราโมทย์

โรงเรียนสารสิทธิ์พิทยาลัย

General Input Devices



Document Image

1. บทนำ

ปรากฏการณ์สำคัญอย่างหนึ่งที่จะเกิดขึ้นได้ทั่วไปในประเทศต่าง ๆ ทั้งในอดีตและปัจจุบัน คือ ในระยะที่เศรษฐกิจของประเทศได้ขยายตัวมากขึ้น ความสำคัญของสาขาเกษตรในการจ้างงานและรายได้ของประเทศจะลดลง ปรากฏการณ์ดังกล่าวจะสังเกตเห็นได้ชัดเจนจากผลิตภัณฑ์ของเหมืองแร่และรายได้จากสาขาเกษตรที่ลดลง เมื่อรายได้เฉลี่ยต่อหัวของประชากรเริ่มสูงขึ้น เมื่อเศรษฐกิจขยายตัวมากขึ้นและรายได้เฉลี่ยต่อหัวสูงขึ้นถึงระดับหนึ่ง ทั้งผลิตภัณฑ์และจำนวนของกำลังแรงงานเกษตรจะลดลงด้วย บางประเทศเกิดปัญหาการขาดแคลนแรงงานในสาขาเกษตรได้ ทั้งนี้เพราะแรงงานสาขาเกษตรส่วนหนึ่งได้โยกย้ายไปทำงานนอกสาขาเกษตร เมื่อโครงสร้างเศรษฐกิจของประเทศเปลี่ยนแปลงไปภายในสาขาเกษตรเองก็มีการเปลี่ยนแปลง โครงสร้างการผลิตและการจ้างงานด้วย การผลิตบางชนิดอาจขยายตัวสูง แต่บางชนิดอาจหดตัว ซึ่งมีผลต่อการเปลี่ยนแปลงของโครงสร้างการจ้างงานในสาขาเกษตร ลักษณะและอัตราการเปลี่ยนแปลงของกำลังคนและการจ้างงานในสาขาเกษตรย่อมขึ้นอยู่กับความแตกต่างของปัจจัยทางเศรษฐกิจและสังคมของแต่ละ

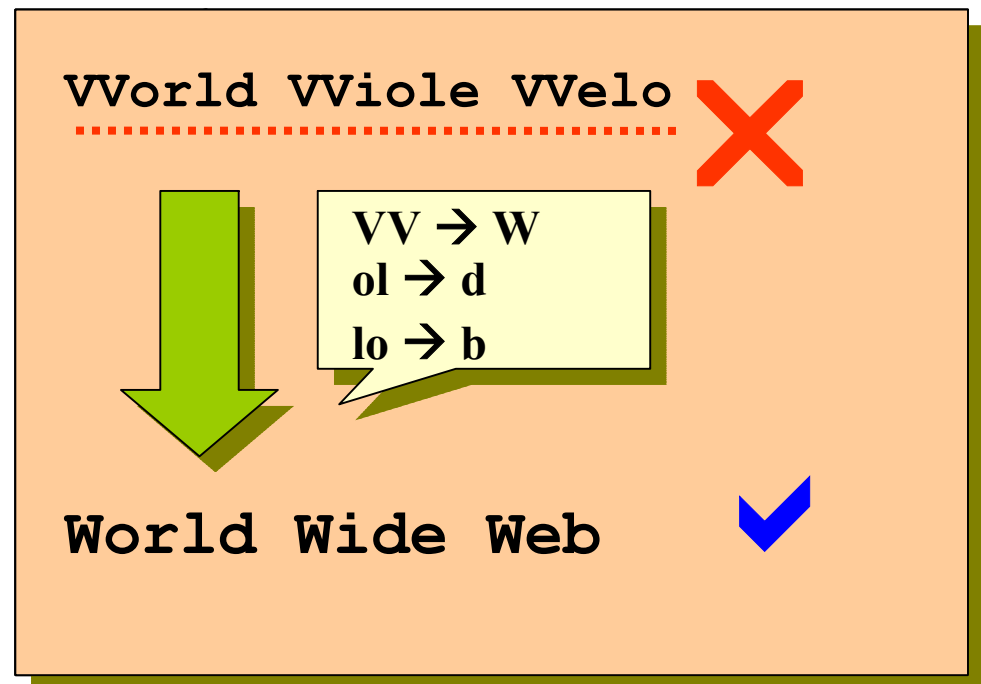
From Scanner



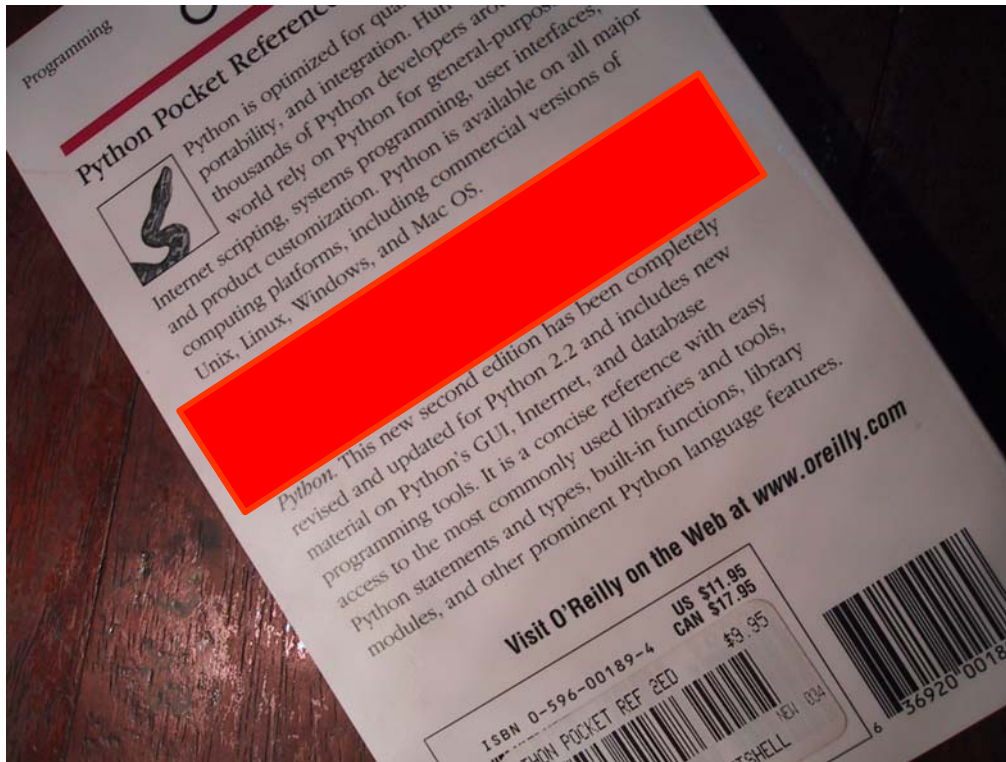
From Digital Camera

Optical Character Recognition

- **Pre-Processing**
- OCR Processing
- Post-Processing

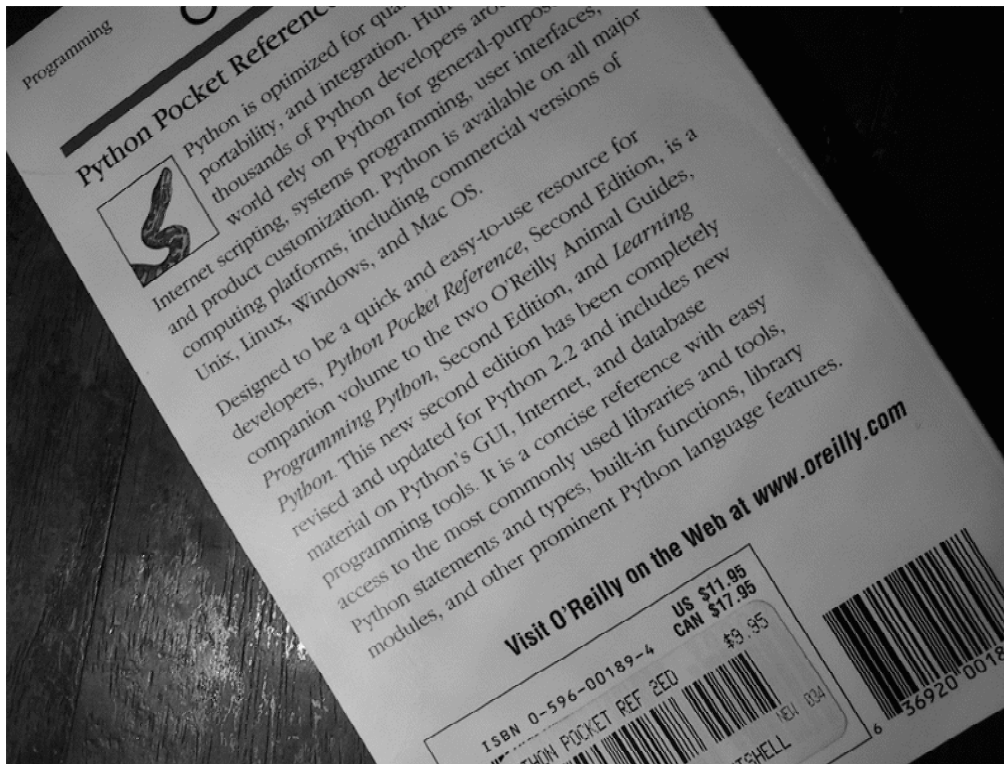


Text Localization



- Polygon Area

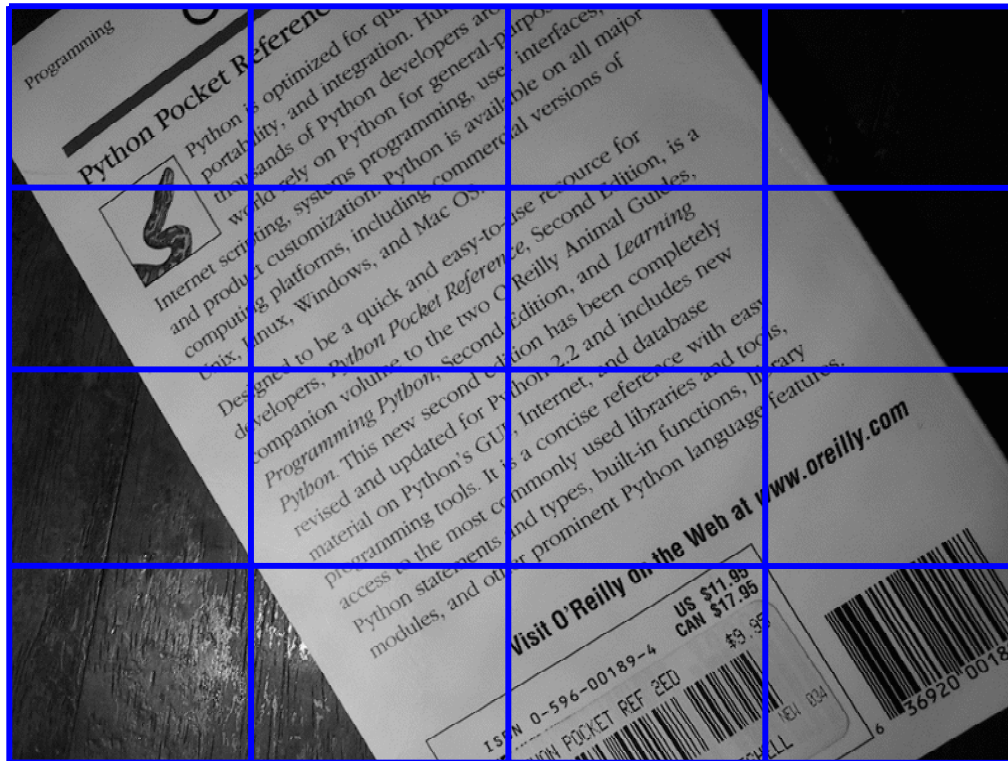
Image Conversion (RGB to Grayscale)



- ITU-R 601-2
Luma transform

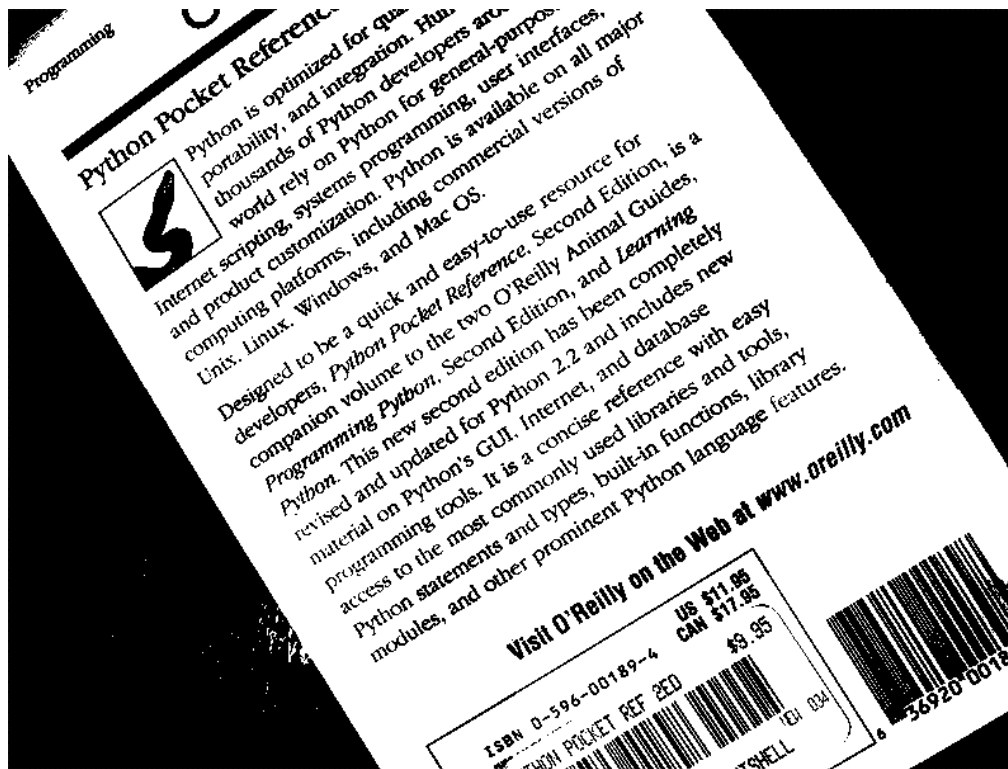
$$L(x, y) = \frac{299 \times R(x, y)}{1000} + \frac{587 \times G(x, y)}{1000} + \frac{114 \times B(x, y)}{1000}$$

Image Thresholding



- Adaptive Threshold
- K-mean Clustering

Image Thresholding



- Adaptive Threshold
- K-mean Clustering

Plane Recovery

- The first (leftmost) argument in a class method function is commonly named `self`.

Specific Statements

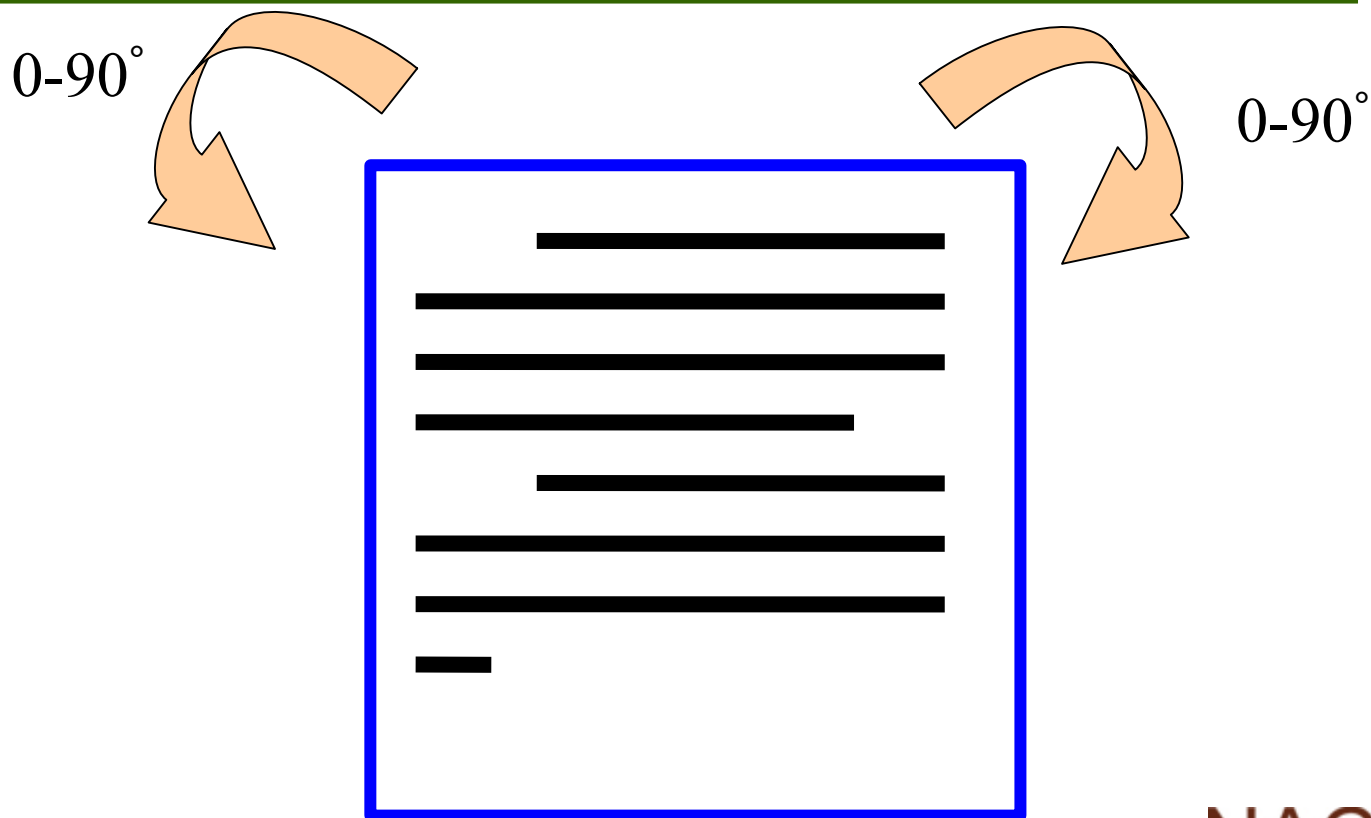
The following sections describe all Python statements. Each section lists the statement's syntax formats, followed by usage details. For compound statements, each appearance of a suite in a statement format stands for one or more other statements, possibly indented as a block under a header line. A suite must be indented under a header if it contains another compound statement (`if`, `while`, etc.); otherwise, it can appear on the same line as the statement header. The following are both valid constructs:

```
if x < 42:
    print x
while x: x = x - 1

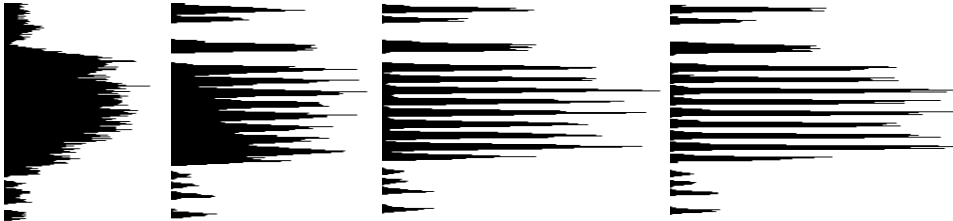
if x < 42: print x
```

- สร้าง กราฟแสดงความหนาแน่นของจุดสีดำ ในแต่ละแถว Pixel

Plane Recovery (cont.)



Plane Recovery (cont.)



- The first (optional) argument to a class method function is conventionally named `self`.

Specific Statements

The following sections describe all Python statements. Each section lists the statement's syntax forms, followed by usage details. For compound statements, each appearance of a suite in a statement format stands for one or more other statements, possibly indented as a block under a header line. A suite must be indented under a header if it contains another compound statement (`if`, `while`, etc.); otherwise, it can appear on the same line as the statement header. The following are both valid constructs:

```
if x < 42:  
    print x  
if x < 42: print x
```

Plane Recovery (cont.)

- Method A

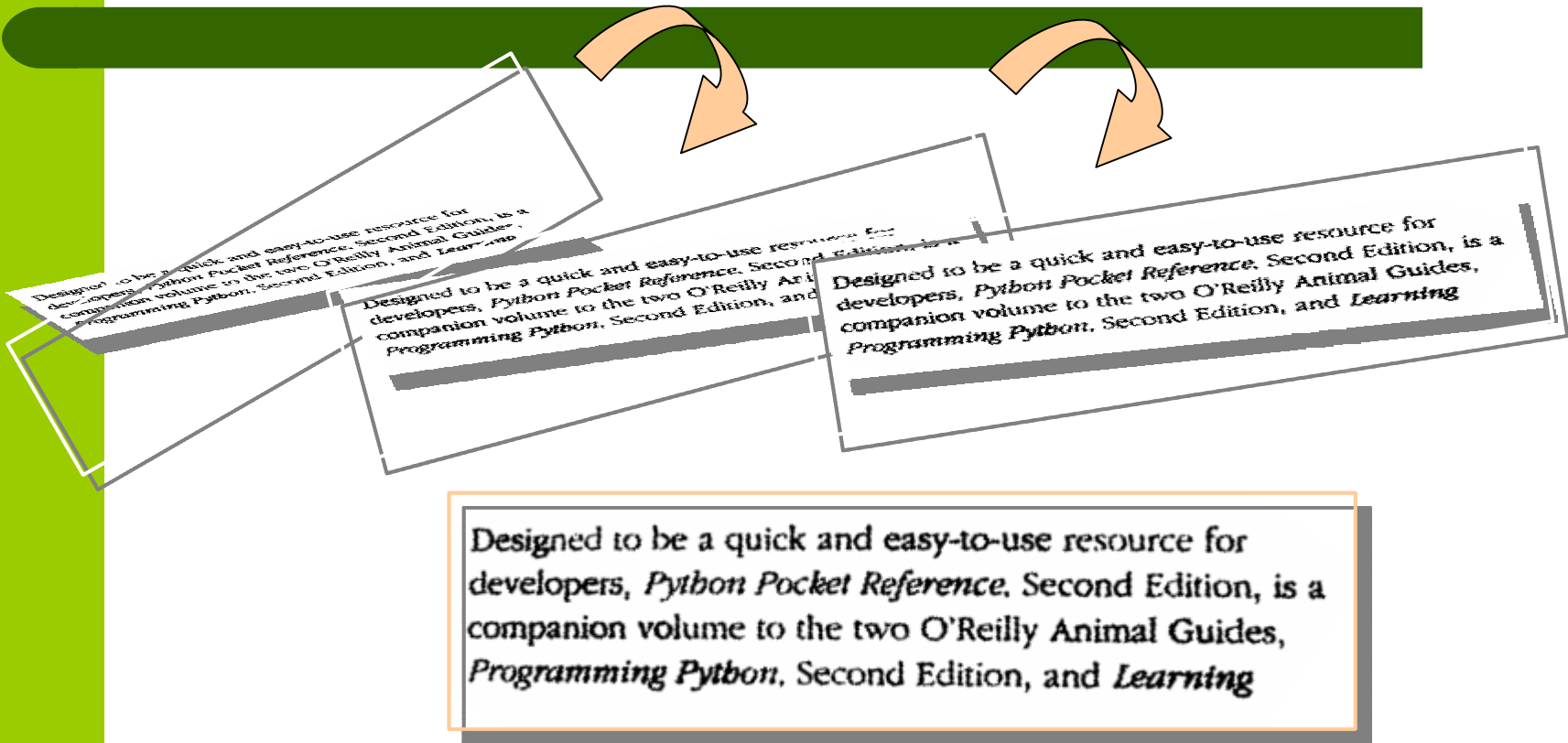
$$A_{\theta} = \sum_{y=5}^{Y-5} \begin{cases} 1 & \text{if } d(y-5) < d(y) \text{ and } d(y+5) < d(y) \\ 0 & \text{otherwise} \end{cases}$$

- Method B

$$B_{\theta} = \sum_{y=0}^Y \begin{cases} 1 & \text{if } d(y) > \frac{\max(d)}{2} \\ 0 & \text{otherwise} \end{cases}$$

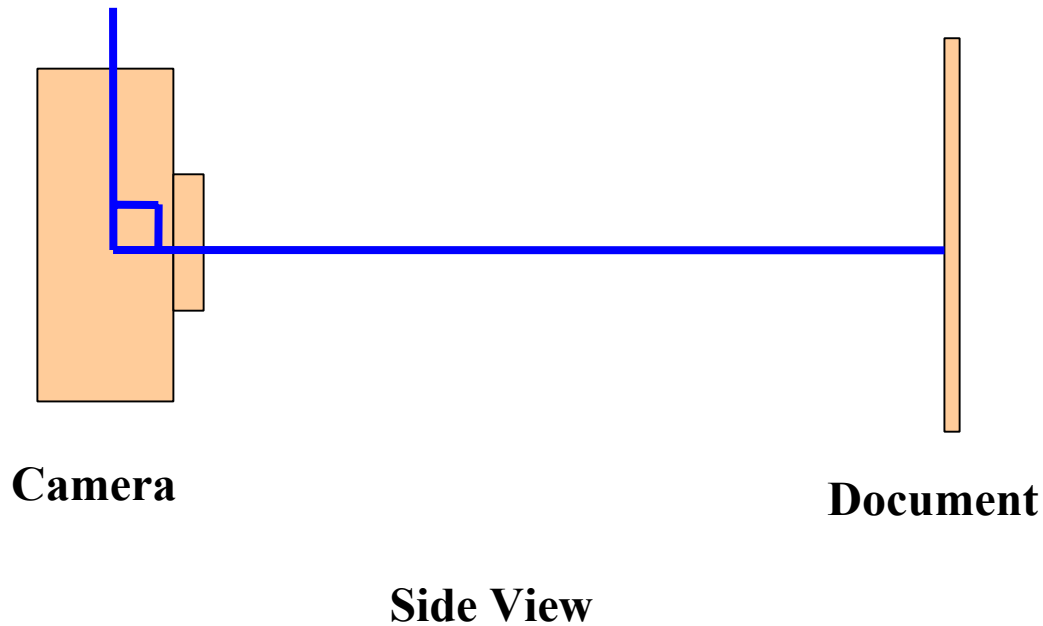
- Method C $C_{\theta} = A_{\theta} + B_{\theta}$

Result



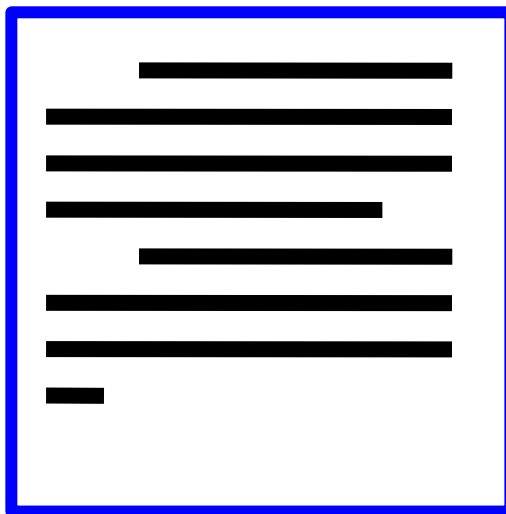
Experiment A

- ทำการบันทึกภาพเอกสาร โดยกล้องทำมุม 90 องศา กับเอกสาร



Experiment A (cont.)

- บันทึกภาพเอกสารโดยหมูนเอกสาร ตามเงื่อนไขพิก้า 10, 30, 50 องศา , ทวนเงื่อนไขพิก้า 20, 40, 60 องศา และ 90 องศา

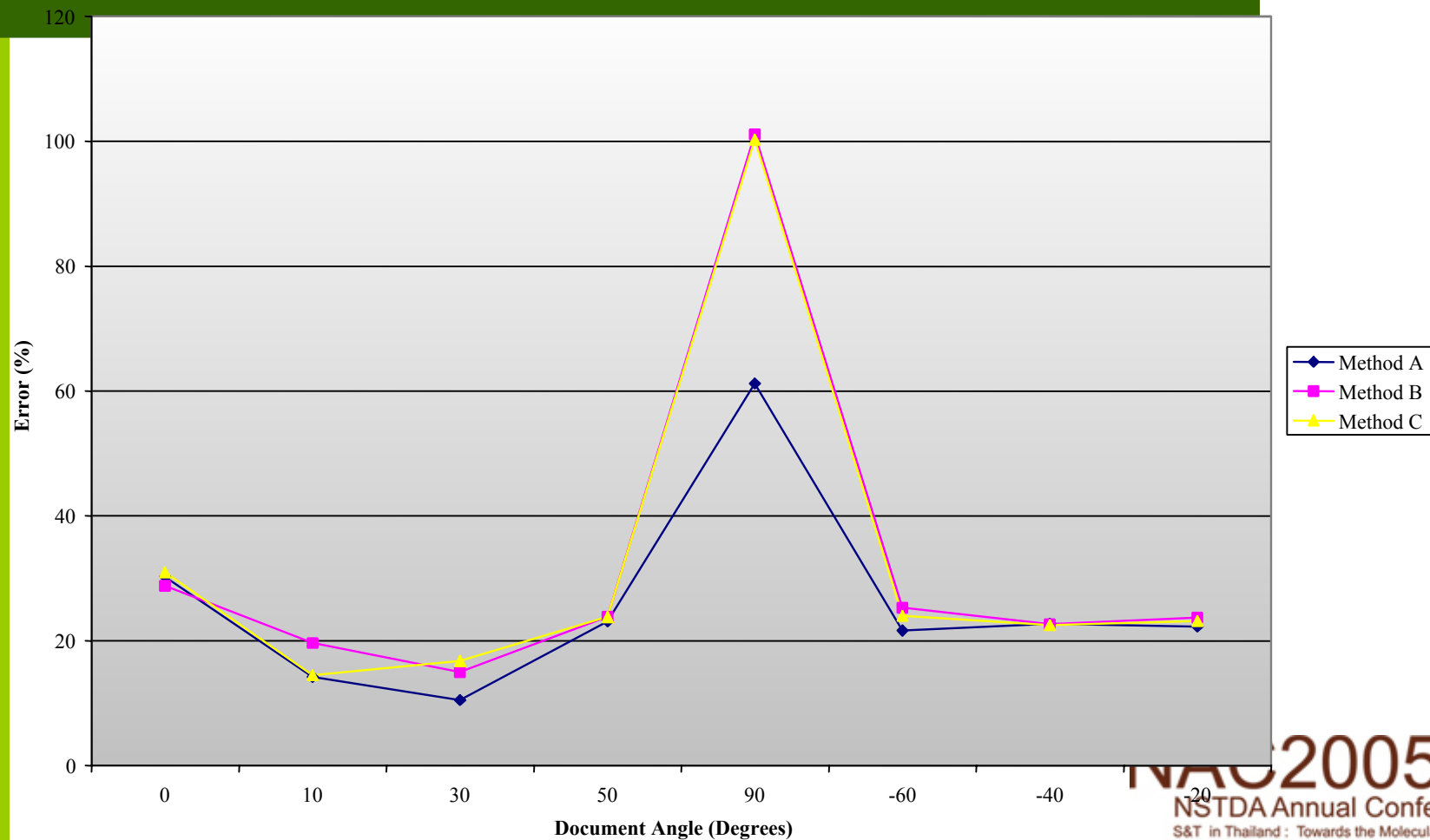


Experiment A (cont.)

- นำภาพที่ได้ไปประมวลผลด้วย **Pre-Processor** และ **SimpleOCR**
- ทำการเปรียบเทียบผลลัพธ์ที่ได้จากการรู้จำตัวอักษร และ **Ground Truth Data** ด้วยวิธี **Levenshtein Distance**
- คำนวณค่าความผิดพลาด (**Character Accuracy**) และบันทึกผล

$$\% \text{ err} = \frac{LD(\text{ground truth data}, \text{result})}{\text{length of Ground Truth Data (characters)}} \times 100$$

Results : Experiment A



Experiment B

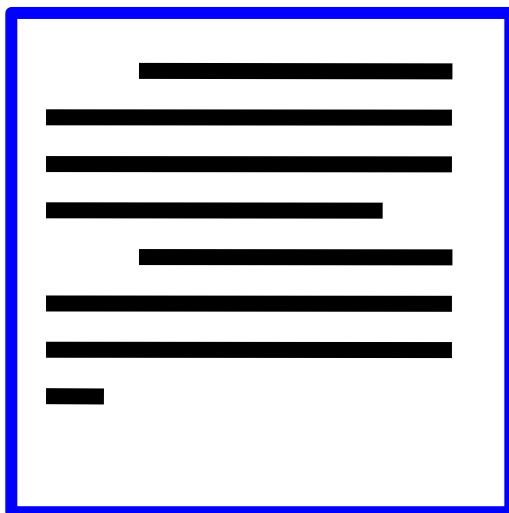
- ทำการบันทึกภาพเอกสาร โดยกล้องทำมุม 90° องศา กับเอกสาร และมุมก้ม $5, 10, 15$ องศา กับเอกสาร



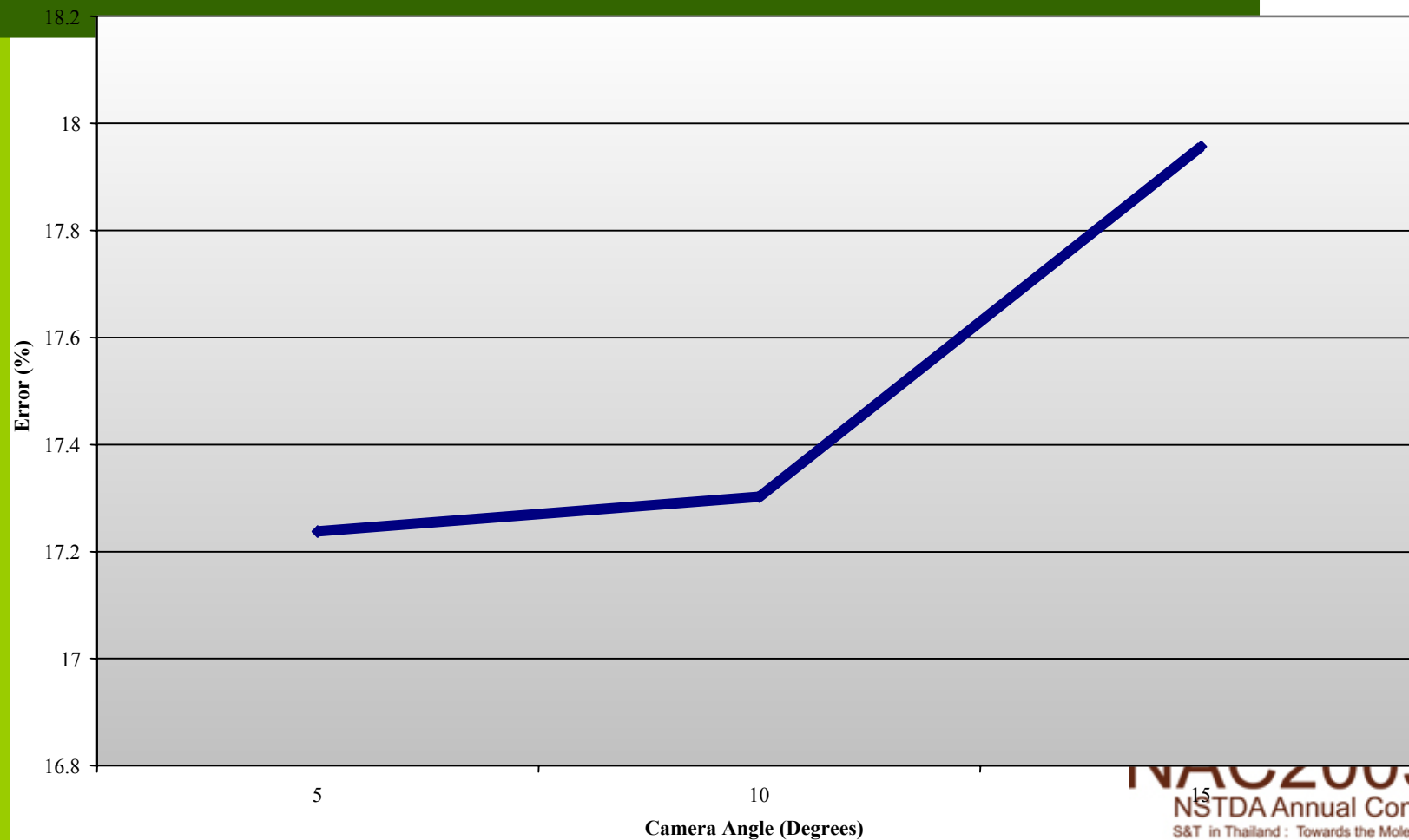
Side View

Experiment B (cont.)

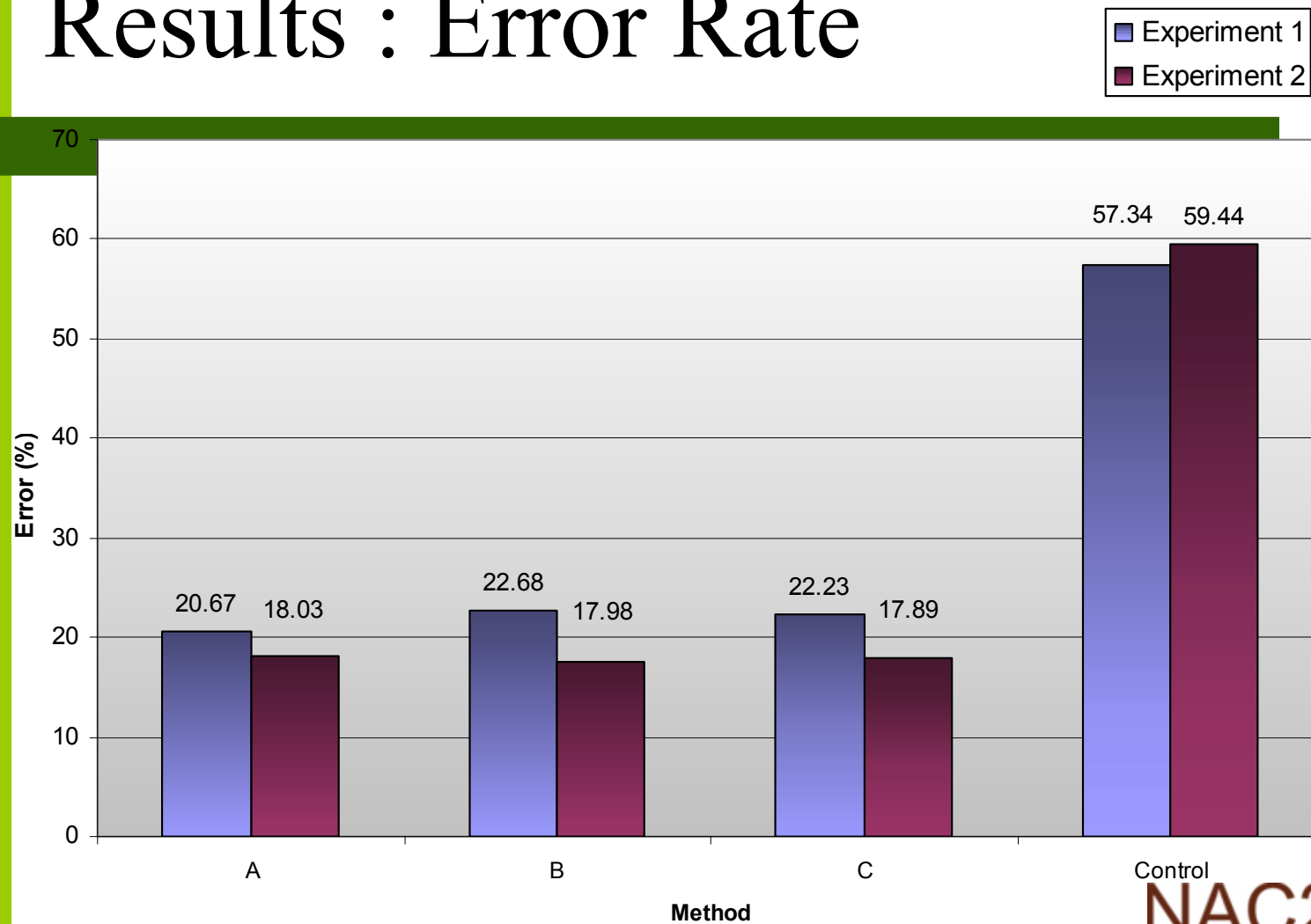
- บันทึกภาพเอกสารโดยหมูนเอกสาร ตามเงื่อนไขพิก้า 10, 20, 30, 40 และ 50 องศา



Results : Experiment B



Results : Error Rate



Conclusions

- ในโครงการนี้ได้นำเสนอวิธีใหม่ในการประมวลผลภาพส่วนหน้า 2 วิธี และนำสองวิธีนี้มาประกอบกันเป็นวิธีการใหม่อีก 1 วิธี เพื่อใช้ในการปรับปรุงภาพถ่ายเอกสารจากกล้อง ซึ่งวิธีใหม่นี้ น่าจะเป็นแนวทางในการนำไปประยุกต์ใช้ต่อไปได้ในอนาคต
- ค่าความผิดพลาดต่ำสุดของการทดลองครั้งที่ 1 ได้จากวิธีการ A เท่ากับ 20.67% และค่าความผิดพลาดต่ำสุดของการทดลองครั้งที่ 2 ได้จากวิธีการ C เท่ากับ 17.89% ซึ่งอาจจะปรับปรุงให้ดีขึ้นได้โดยการเปลี่ยนเงื่อนไขการทดลอง
- ลักษณะภาพที่มีผลต่อความผิดพลาดของการทดลองอย่างชัดเจน คือ ภาพเอกสารที่ทำมุม 90 องศา ซึ่งไม่มีข้อมูลอ้างอิงมากพอที่จะระบุได้ว่าภาพที่หมุนไปแล้วเป็นภาพหัวกลับหรือไม่

Applications

- Mobile OCR for the Blind



Applications (cont.)

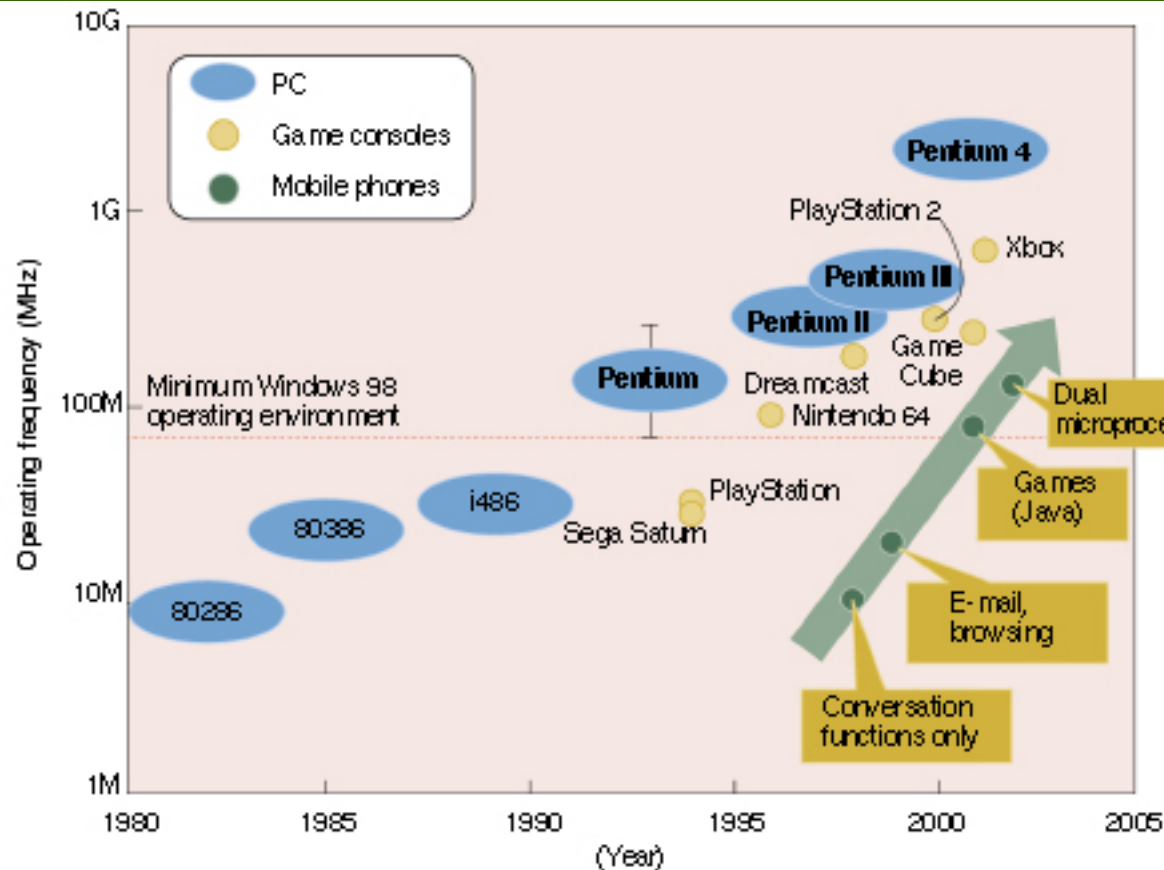


Fig 1 Mobile Phone Microcontroller Operating Frequency over 100MHz
Changes in operating frequency for PC and game systems microprocessors, and mobile phone microcontrollers. Some mobile phones began to mount dual microcontrollers in 2002; one for applications like games and image processing, the other for baseband processing. The SH-Mobile from Hitachi is the faster application microcontroller, at 133MHz. This is comparable with the first Pentium chips and game system microprocessors, as far as speed is concerned. The operating frequency of mobile phone microcontrollers began to rise sharply from 1999, when e-mail and Internet access functions spread, and bit width doubled from 16 to 32 bits. Until then, mobile phone processing was almost entirely baseband processing, at about 10MHz. Note that performance of microprocessors and microcontrollers depends on more than just operating frequency. Microcontrollers in mobile phones do not have the floating point functions that are implemented in PCs or game system chips.

Applications (cont.)

a) Optical character recognition



b) Infrared communication



Fig 2 New OCR and Infrared Communication Functions in Mova N504iS (a) Reading character information from a business card with the OCR function. (b) Can also be used as a TV remote controller.

Suggestions

- ควรทำการศึกษาในกรณีที่มุมกล้องมีมุมก้ม/เงย ที่แตกต่างกันออกไป
- ควรจะทำการทดลองกับกลุ่มตัวอย่างที่มีขนาดใหญ่ขึ้น
- ควรทำการศึกษาในกรณีที่กล้องถ่ายภาพ มีระยะโฟกัสคงที่ หรือกล้องถ่ายภาพที่ไม่มีโหมดแมนูอัลในการถ่ายภาพ