การประมาณค่าความเอียงหน้าเอกสารสิ่งพิมพ์

The Skew Estimation of Printed Documents

ดร. ยุทธพงษ์ รังสรรค์เสรี * กฤษณะ ชินสาร **

อัลกอริทึมสำหรับการประมาณค่ามุมเอียงของหน้าเอกสารที่นำเสนอ จะเป็นการหาค่าตำแหน่งของจุดดำทางซ้ายสุดของแต่ละบรรทัดภาพเอกสารทั้งหน้าเอกสาร แล้วแบ่งบรรทัดภาพเอกสารออกเป็นบรรทัดตัวหนังสือ โดยพิจารณาจากความสัมพันธ์ของตำแหน่งจุดดำทางซ้ายสุดของบรรทัดภาพเอกสารที่อยู่ติดกัน ทำการรวมบรรทัดตัวหนังสือเป็นพารากราฟของบรรทัดตัวหนังสือ เลือกพารากราฟที่มีจำนวนบรรทัดภาพเอกสารมากที่สุด เป็นตัวแทนในการประมาณค่ามุมเอียงของหน้าเอกสาร ในการประมาณแนวเส้นตรงของขอบซ้ายพารากราฟจะใช้การวิเคราะห์การถดถอย เพื่อหาสมการถดถอยเชิงเส้นเพื่อใช้สำหรับการประมาณค่าความเอียงของภาพเอกสาร การลดสัญญาณรบกวนขนาดเล็กในภาพเอกสาร ใช้วิธีการมอร์โฟโลยีแบบเปิด (Opening Morphology)

In this paper, an algorithm for skew estimation of printed documents is presented. The algorithm will locate the left most pixel of each line in the whole document image. The obtained list of those pixels is then segmented into lines using a continuity criterion, and grouped into paragraphs, respectively. Only the pixels belonging to the longest paragraph are used to estimate the skew angle of the printed page document. In our algorithm, the Regression Analysis based on Least Square Method (LSM) is performed to find Regression Equation for the selected pixels. Before performing skew estimation of printed document, Opening Morphology Method is used for noise reduction on printed pages document.

1. บทนำ

การประมาณค่ามุมเอียงของหน้าเอกสาร เป็นการประมวลผลก่อนการประมวลผลภาพเอกสารจริงที่สำคัญขั้นตอนหนึ่ง ทั้งนี้เนื่องจากว่า ภาพของหน้าเอกสารที่ได้จากการสแกนเพื่อใช้ในการประมวลผลภาพเอกสารอาจมีความเอียง หรือเอกสารยังอยู่ในมุมที่ไม่ควรจะเป็น ซึ่งโดยทั่วไปภาพเอกสารในการทดลองจะมีค่ามุมเอียงเป็นศูนย์องศา ดังนั้นเพื่อให้เกิดความสะดวกและรวดเร็วในการประมวลผลภาพเอกสารจริงจึงควรมีการประมาณค่าความเอียง ของหน้าเอกสารที่จะใช้ในการทดลองก่อน แล้วทำการหมุนภาพเอกสารกลับ ด้วยค่ามุมของความเอียง และทิศทางที่ได้จากการประมาณ

ภาพเอกสารที่สแกนเข้ามาเข้าเก็บในเครื่องคอมพิวเตอร์ในรูปไบนารีภาพเอกสาร บางตำแหน่งจุดภาพหรือกลุ่มของจุดภาพบางกลุ่ม จะเป็นสัญญาณรบกวน (Noise) บนหน้าเอกสาร สัญญาณรบกวนอาจเกิดจากรอยยับของหน้ากระดาษที่นำมาสแกน หรือ รอยเปื่อนหมึกหรือรอยขีดเขียน จุดภาพที่เป็นสัญญาณรบกวนจะเป็นจุดภาพ ที่เราไม่ต้องการใช้ในการประมวลผลภาพเอกสาร เพราะ ถ้านำตำแหน่งจุดภาพของสัญญาณรบกวนเข้ามาร่วมพิจารณาด้วยจะทำให้การประมาณค่ามุมเอียงของหน้าเอกสารผิดพลาดไปจากความเป็นจริงได้ ในงานวิจัยนี้เลือกใช้การขจัดสัญญาณรบกวนโดยใช้วิธีการมอร์โฟโลยีแบบเปิด (Opening Morphology)

* อาจารย์ คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

** นักศึกษาปริญญาโท หลักสูตรวิทยาการคอมพิวเตอร์และเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

Baird [1] ได้เสนอ “Alignment Measure Approach” เป็นอัลกอริทึมในการประมาณค่าความเอียงของหน้าเอกสารที่น่าสนใจ เขาใช้การโปรเจคชั่นค่ามุมเอียงจากจุดกึ่งกลางของกรอบตัวหนังสือที่ทำการแบ่งไว้แล้วก่อนหน้าแล้ว ไปตั้งฉากกับแนวแกนที่กำหนดไว้ แล้วรวมความถี่สะสมของมุมที่เกิดขึ้นบนแนวแกนที่รองรับการโปรเจคชั่น เลือกเอาค่ามุมความเอียงที่มีความถี่สะสมจากการโปรเจคชั่นที่สูงที่สุดเป็นตัวบอกค่าประมาณมุมเอียงของทั้งหน้าเอกสาร วิธีนี้สามารถกำหนดค่ามุมเอียงได้ใกล้เคียงกับความเป็นจริง และในรายงานวิจัยบอกว่าโดยเฉลี่ยแล้ว เวลาในการประมาณค่ามุมเอียงจะใช้ประมาณ 8.5 วินาทีของหน่วยประมวลผลกลาง (CPU Time) ต่อหนึ่งหน้าเอกสาร โดยพัฒนาโปรแกรมบนเครื่อง DEC VAX 11/750 แต่เวลาที่เสนอมานี้ ไม่ได้รวมเวลาที่เขาใช้ในการแบ่งตัวหนังสือทั้งหน้าเอกสารออกจากกัน ซึ่งถ้ารวมเข้าด้วยกันแล้ว จะต้องใช้เวลาของหน่วยประมวลผลกลางมาก

อัลกอริทึมในการประมาณค่ามุมเอียงที่ได้นำเสนอนี้ พัฒนามาจากแนวคิดพื้นฐานของวิธี Left Margin Search ของ Dengel [2] ที่ใช้การหาค่าตำแหน่งจุดดำทางซ้ายสุดของแต่ละบรรทัดของภาพเอกสาร มาประยุกต์ใช้ในการประมาณค่ามุมเอียงของหน้าเอกสาร โดย Dengel จะหาแนวเส้นตรงที่ยาวที่สุดของตำแหน่งจุดดำทางซ้ายสุดของแต่ละบรรทัดภาพของเอกสาร มาใช้ในการประมาณค่ามุมเอียงของหน้าเอกสารทั้งหน้า ซึ่งในการทำเช่นนี้ถ้าตำแหน่งจุดดำที่หามาได้เป็นสัญญาณรบกวนไม่ใช่แนวขอบซ้ายของตัวหนังสือ ก็จะทำให้การประมาณค่าความเอียงเกิดความผิดพลาดได้

จากปัญหาที่เกิดขึ้นกับวิธีการของ Dengel ผู้วิจัยจึงเลือกวิธีการลดสัญญาณรบกวนขนาดเล็ก ๆ ในภาพหน้าเอกสารโดยใช้มอร์โฟโลยีแบบเปิด และในประมาณแนวเส้นตรงของตำแหน่งจุดดำทางซ้ายของแต่ละบรรทัดภาพเอกสารเพื่อใช้ในการประมาณค่ามุมเอียงของหน้าเอกสาร ผู้วิจัยได้เลือกวิธีการปรับแนวเส้นตรงของตำแหน่งจุดดำ โดยใช้การวิเคราะห์การถดถอย (Regression Analysis) ซึ่งเป็นวิธีที่ไม่ซับซ้อน แต่สามารถให้คำตอบในการประมาณแนวเส้นตรงได้สูง ดังนั้นอัลกอริทึมที่นำเสนอจึงสามารถทำงานได้อย่างรวดเร็ว

2. ขั้นตอนการทำงาน

2.1 การแปลงแบบมอร์โฟโลยี

Sonka[3] อธิบายการแปลงแบบมอร์โฟโลยี (Morphology Transformation) ว่าเป็นกระบวนการสำหรับการใช้ศึกษาพฤติกรรมเซตจุดดำในภาพเอกสาร Gonzalez[4] บอกว่ามอร์โฟโลยีทางคณิตศาสตร์ (Mathematical Morphology) คือการประยุกต์ใช้ทฤษฎีเซตสำหรับแยกส่วนประกอบต่าง ๆ ในภาพเอกสาร เช่น ใช้ในการหาขอบของวัตถุในข้อมูลภาพ กำหนดให้เซตของจุดดำในภาพเอกสารแบบไบนารี (Binary Image) อยู่ในสเปซสองมิติ : E^2 และ

- X คือ เซตของตำแหน่งจุดภาพที่มีค่าเป็น 1 ซึ่งในที่นี้กำหนดให้เป็นจุดดำ
- X^c คือ เซตของตำแหน่งจุดภาพทีมีค่าเป็น 0 ซึ่งในที่นี้กำหนดให้เป็นจุดขาว และ
- B คือ เซตของการแปลงข้อมูลแบบมอร์โฟโลยี (Structuring Element)

การแปลงแบบมอร์โฟโลยี เป็นการนำเซตของตำแหน่งจุดภาพที่มีค่าของจุดภาพในภาพเอกสารที่มีค่าเป็น 0 มาทำการโอเปอร์เรตกับเซตของการแปลงข้อมูลแบบมอร์โฟโลยี มีกระบวนการในการแปลงที่น่าสนใจ 2 อย่าง คือ

2.1.1 Erosion

Erosion : เป็นการลดจำนวนสมาชิกของเซตของตำแหน่งจุดดำในภาพเอกสาร : X เมื่อพิจารณ

ประกอบกับเซต B สัญลักษณ์ที่ใช้เป็นเครื่องหมายลบ $\ominus$ ดังสมการ

$$X \ominus B = \{ d \in E^2 : d+b \in X ; b \in B \} \quad (1)$$

การ Erosion จะทำให้จำนวนของจุดภาพที่เป็นจุดดำในภาพเอกสาร : X มีจำนวนลดลง หรือกล่าวอีกอย่างคือว่าเป็นการขยายขนาดของช่องว่าง หรือจำนวนจุดภาพที่เป็นจุดขาว ดังนั้นจะเห็นได้ว่า ถ้ากลุ่มของจุดภาพดำมีขนาดเล็กมาก เช่น สัญญาณรบกวน กลุ่มจุดภาพที่เป็นจุดดำนั้นก็ถูกขจัดออกจากภาพเอกสารไปโดยอัตโนมัติ แต่ถ้ากลุ่มของจุดดำมีขนาดใหญ่จะทำให้ช่องไฟระหว่างกลุ่มของจุดดำมีค่ามากขึ้น

2.1.2 Dilation

Dilation : การเพิ่มจำนวนสมาชิกของเซตของตำแหน่งจุดดำลงไปในภาพเอกสาร เมื่อพิจารณาประกอบกับเซต B สัญลักษณ์ที่ใช้เป็นเครื่องหมายบวก $\oplus$ ดังสมการ

$$X \oplus B = \{ d \in E^2 : d = x+b ; x \in X \text{ and } b \in B \} \quad (2)$$

การ Dilation จะทำให้กลุ่มของจุดภาพที่เป็นจุดดำในภาพเอกสารมีขนาดใหญ่ขึ้น ทำให้ช่องไฟระหว่างกลุ่มจุดดำมีขนาดเล็กลง

2.1.3 มอร์โฟโลยีแบบเปิดและแบบปิด

มอร์โฟโลยีแบบปิด คือ การแปลงมอร์โฟโลยี ที่เริ่มจากการ Dilation แล้วตามด้วยการ Erosion ดังสมการ

$$X \bullet B = (X \oplus B) \ominus B \quad .....(3)$$

การทำมอร์โฟโลยีแบบปิดจะทำให้ขนาดของสัญญาณรบกวนขนาดเล็กออกได้เมื่อผ่านกระบวนการ Erosion และเมื่อทำการ Dilation เซตของจุดดำบางจุดที่เคยถูกกัดเซาะออกไปในตอนแรกก็จะถูกเติมกลับเข้ามาทำให้คุณภาพของภาพเอกสารดีขึ้น

มอร์โฟโลยีแบบเปิด คือ การแปลงมอร์โฟโลยี ที่เริ่มจากการ Erosion แล้วตามด้วยการ Dilation ดังสมการ

$$X \circ B = (X \ominus B) \oplus B \quad .....(4)$$

การทำมอร์โฟโลยีแบบเปิดจะทำให้สามารถขจัดสัญญาณรบกวนขนาดเล็กออกได้เมื่อผ่านกระบวนการ Erosion และเมื่อทำการ Dilation เซตของจุดดำบางจุดที่เคยถูกกัดเซาะออกไปในตอนแรกก็จะถูกเติมกลับเข้ามาทำให้คุณภาพของภาพเอกสารดีขึ้น

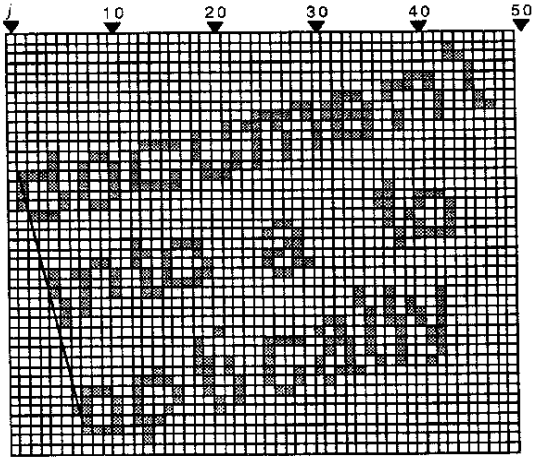
2.2 การกำหนดแนวเส้นตรงเพื่อการประมาณค่ามุมเอียงของหน้าเอกสาร

2.2.1 การหาตำแหน่งจุดดำทางซ้าย

กำหนดโครงสร้างข้อมูลแบบแถวลำดับ V ให้มีขนาดเท่ากับจำนวนบรรทัดของภาพเอกสารทั้งหน้า : n โดย V[i] : (i=1,...,n) จะใช้เก็บค่าตำแหน่งจุดดำทางซ้ายจุดแรกของแต่ละบรรทัดภาพเอกสารของทั้งหน้า ดังแสดงอัลกอริทึมในการหาตำแหน่งจุดดำทางซ้ายดังนี้

```
Algorithm_1 SearchLeftMostPoint
    int V[ImageFile->depth], i,j;
    for(i=0; i<ImageFile->deoth; I++) {
        ReadImageLine( i) to *p;
        for(j=0; j<ImageFile->width; I++) {
            if (p[i] = 0){
                V[i] = j;
                Break;
            }
        }
    }
```

ภาพที่ 1 แสดงขั้นตอนในการหาตำแหน่งจุดดำทางซ้ายของแต่ละบรรทัดภาพเอกสาร



ภาพที่ 2 ตัวอย่างภาพเอกสารที่ได้จากเครื่องสแกนเนอร์

i	1	2	3	...	10	11	...
V[i]	0	43	43	...	21	18	...

ภาพที่ 3 ค่าตำแหน่งจุดดำทางซ้ายของบรรทัดภาพเอกสารในหนึ่งหน้าเอกสารที่เก็บบนแถวลำดับ V[i]

2.2.2 <u>การหาแนวขอบซ้ายของแต่ละบรรทัด</u>

จากค่าตำแหน่งจุดดำทางซ้ายของแต่ละบรรทัดภาพเอกสารที่เก็บในแถวลำดับ V[i] i=1,...,n เราจะพิจารณาความสัมพันธ์กันของตำแหน่งจุดดำทางซ้ายของแต่ละบรรทัดภาพเอกสารที่อยู่ติดกัน ถ้าตำแหน่งจุดดำทางซ้ายของบรรทัดที่อยู่ติดกันมีทิศทางไปในทางเดียวกันเราก็จะรวมเข้าเป็นบรรทัดเดียวกัน โดยพิจารณาจากค่าผลต่างของตำแหน่งจุดดำทางซ้ายของบรรทัดที่อยู่ติดกัน ในขั้นตอนนี้จะสร้างโครงสร้างข้อมูลเพิ่มอีกตัวหนึ่ง : Line สำหรับใช้เป็นตัวแบ่งตำแหน่งจุดดำทางซ้ายใน V[i] ออกเป็นกลุ่มบรรทัดของตัวหนังสือ โครงสร้าง Line มีส่วนประกอบ 2 ส่วน คือ

Line[j].X[k] สำหรับเก็บค่าตำแหน่งของจุดดำทางซ้ายของบรรทัดภาพเอกสารที่สามารถรวมกลุ่มกันเป็นบรรทัดตัวหนังสือที่ j และ

Line[j].Y[k] สำหรับเก็บค่าบรรทัดภาพเอกสารของหมายเลขบรรทัดตัวหนังสือที่ j

จำนวนของบรรทัดของตัวหนังสือที่ได้จากอัลกอริทึมที่นำเสนอจะขึ้นอยู่กับ แนวของตำแหน่งจุดดำทางซ้าย และความต่อเนื่องกันของตำแหน่งจุดดำทางซ้ายของแต่ละบรรทัดภาพเอกสารที่อยู่ติดกัน กล่าวคือ ถ้านำอัลกอริทึมนี้ไปทดสอบกับภาพเอกสารที่มีแนวขอบซ้ายของหน้ากระดาษเป็นเส้นตรง เช่น ภาพเอกสารที่เป็นตาราง ก็จะได้จำนวนบรรทัดตัวหนังสือเพียงบรรทัดเดียว แต่ถ้าแนวขอบซ้ายของหน้าเอกสารมีการเว้นวรรคเป็นช่วงตามบรรทัดของหนังสือ จำนวนบรรทัดของตัวหนังสือก็จะมากขึ้นตาม ดังแสดงผลการทดลองในภาพที่ 4

k	1	2	3	4	5
Line[1].X[k]	2	2	2	2	3
Line[1].Y[k]	15	16	17	18	19

k	1	2	3	4	5	6
Line[2].X[k]	5	5	5	6	6	6
Line[2].Y[k]	25	26	27	28	29	30

ภาพที่ 4 แสดงความสัมพันธ์ระหว่างค่าตำแหน่งจุดดำทางซ้าย กับค่าบรรทัดภาพเอกสาร ของบรรทัดตัวหนังสือที่ 1 และ 2

2.2.3 <u>การหาแนวขอบซ้ายของพารากราฟ</u>

เมื่อแบ่งบรรทัดตัวหนังสือออกจนครบทั้งหน้าเอกสารแล้ว พิจารณารวมแนวของซ้ายของบรรทัดตัวหนังสือที่มีแนวโน้มของตำแหน่งจุดดำไปในทางเดียวกันเข้าเป็นแนวขอบซ้ายของพารากราฟเดียวกัน โดยพิจารณาจากค่าระยะห่างเฉลี่ยของตำแหน่งจุดดำทางซ้ายของแต่ละบรรทัดตัวหนังสือที่ห่างจากขอบ

ซ้ายของหน้าเอกสาร ถ้าค่าระยะห่างเฉลี่ยของ 2 บรรทัดตัวหนังสือที่อยู่ติดกันมีแนวโน้มไปในทางเดียวกัน ให้เรียงต่อสองบรรทัดตัวหนังสือนั้นเป็นพารากราฟเดียวกัน

ถ้าค่าระยะห่างเฉลี่ยของแต่ละบรรทัดตัวหนังสือไม่สอดคล้องหรือไม่มีแนวโน้มไปในทางเดียวกัน ให้สร้างพารากราฟสำหรับเก็บแนวขอบซ้ายของตัวหนังสือขึ้นมาใหม่ วนรอบการทำงานไปจนกว่าจะครบทุกบรรทัดของตัวหนังสือ

จากขั้นตอนที่กล่าวมาจะเห็นได้ว่า ในหนึ่งหน้าเอกสารที่นำมาวิเคราะห์อาจแบ่งออกได้เพียงหนึ่งพารากราฟ หรือหลายพารากราฟก็ได้ขึ้นอยู่กับ ระยะห่างเฉลี่ยของตำแหน่งจุดดำทางซ้ายของแต่ละบรรทัดตัวหนังสือ

l	1	2	3	...	10	11	...
P[1].X[l]	2	2	2	...	6	6	...
P[1].Y[l]	15	16	17	...	29	30	...

ภาพที่ 5 แสดงค่าข้อมูลในพารากราฟที่ 1 ที่มีแนวขอบซ้ายของหลาย ๆ บรรทัดตัวหนังสือ

เมื่อทำการแบ่งพารากราฟตัวหนังสือเสร็จ ให้เลือกพารากราฟของบรรทัดตัวหนังสือที่มีจำนวนบรรทัดภาพเอกสารมากที่สุด เป็นตัวแทนในการประมาณค่าความเอียงของทั้งหน้าเอกสาร ทั้งนี้เพราะต้องการให้ได้ตัวแทนของตำแหน่งจุดดำทางซ้ายให้มากที่สุดเท่าที่จะเป็นไปได้ในการไปประมาณความเอียง เพื่อให้ค่ามุมความเอียงของหน้าเอกสารใกล้เคียงความจริงที่สุด

2.3 การหาแนวความเอียงโดยใช้สมการถดถอย

เนื่องจากตำแหน่งของจุดดำทางซ้ายของแต่ละบรรทัดตัวหนังสือที่มาประกอบกันเป็นพารากราฟของบรรทัดตัวหนังสือเกิดขึ้นอย่างสุ่ม เราไม่สามารถกำหนดค่าตำแหน่งจุดดำทางซ้ายให้เกิดขึ้นในตำแหน่งใดตำแหน่งหนึ่งได้ แต่เราก็ทราบว่าแนวขอบซ้ายของตำแหน่งจุดดำของพารากราฟที่ใช้เป็นตัวแทนในการประมาณค่ามุมเอียงของหน้าเอกสารเกือบจะเป็นแนวเส้นตรงอันเดียวกัน เพื่อการประมาณแนวเส้นตรงให้เกิดความถูกต้อง ผู้วิจัยได้เลือกใช้การวิเคราะห์การถดถอย [4][5] เพื่อประมาณสมการถดถอยของตำแหน่งจุดดำทางซ้ายของพารากราฟที่ถูกเลือกมาเป็นตัวแทนในการประมาณค่ามุมเอียงของทั้งหน้าเอกสาร

การวิเคราะห์การถดถอยเป็นวิธีการทางสถิติที่ใช้ในการตรวจสอบความสัมพันธ์ของข้อมูล 2 ชุด (หรือมากกว่า) ว่าข้อมูล 2 ชุดดังกล่าวนั้นมีความสัมพันธ์กันอย่างไร จะมีการกล่าวถึงตัวแปรตาม : Y จะมีการเปลี่ยนแปลงอย่างไร เมื่อตัวแปรอิสระ : X มีการเปลี่ยนไป การวิเคราะห์หาสมการถดถอยโดยวิธีกำลังสองน้อยที่สุด (Least Square Method : LSM) ถือว่าเป็นวิธีการหนึ่งในการประมาณสมการถดถอยอย่างมาก เพราะเป็นวิธีที่ตรงไปตรงมา ไม่มีความซับซ้อน การประมาณค่า Y-Estimate สามารถหาได้จากสมการ

$$Y_i^{'} = a + b(X_i) \quad ; a = \text{Y-intercept}$$
$$b = \text{Slope}$$

จาก $Y_i = a + b(X_i) + e_i$

$$e_i = Y_i - a - b(X_i)$$

$$\sum e_i^2 = \sum (Y_i - a - b(X_i))^2 \quad \text{(ค่าน้อยที่สุด)}$$

หาอนุพันธ์ของ $\sum e_i^2$ เทียบกับ a และ b กำหนดค่า = 0

$$\frac{\partial \sum e_i^2}{\partial a} = na + b\sum X_i - \sum Y_i = 0$$

$$\therefore \quad \sum Y_i = na + b\sum X_i$$

และ

$$\frac{\partial \sum e_i^2}{\partial b} = a\sum X_i + b\sum X_i^2 - \sum X_i Y_i = 0$$

$$\therefore \sum X_i Y_i = a\sum X_i + b\sum X_i^2$$

ดังนั้นสมการปกติ คือ

$$\sum Y_i = na + b\sum X_i \quad \text{.....(5)}$$

$$\sum X_i Y_i = a\sum X_i + b\sum X_i^2 \quad \text{.....(6)}$$

จะได้ $$b = \frac{\sum X_i Y_i - \frac{\sum X_i \sum Y_i}{n}}{\sum X_i^2 - \frac{(\sum X_i)^2}{n}}$$

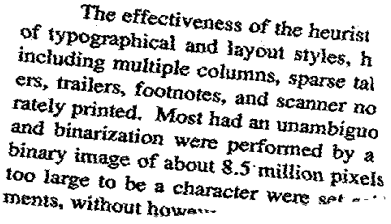
และ $$\theta = \tan^{-1}(b) \quad \text{.....(7)}$$

ทิศทางของการหมุนภาพเอกสารก พิจารณาจากเครื่องหมายของค่าความชันของสมการ (b)

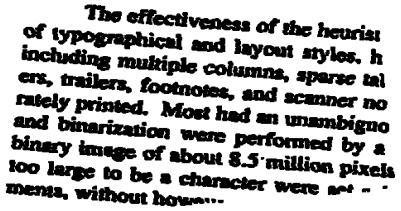
3. การทดลอง

ในงานวิจัยนี้ จะเป็นการทดลองกับภาพเอกสารที่เป็นไบนารี, เอกสารมีหลายพารากราฟ, หลายคอลัมน์, ใช้รูปแบบตัวหนังสือในการพิมพ์ที่แตกต่างกัน, ขนาดตัวหนังสือที่แตกต่างกัน, และหน้าเอกสารมาจากสื่อสิ่งพิมพ์ที่แตกต่างกัน

การสแกนเอกสารใช้เครื่องสแกนเนอร์ HEWLETT PACKARD รุ่น ScanJet 3C ด้วยความละเอียดในการสแกน 200 จุดต่อนิ้ว สแกนเป็นภาพเอกสาร 2 ระดับ มีค่ามุมเอียงของหน้าเอกสารที่แตกต่างกัน ค่ามุมเอียงมีทั้งในแนวทวนเข็ม และตามเข็มนาฬิกา ภาพเอกสารในการทดลองจำนวน 26 ภาพ ภาพเอกสารขนาด 600 X 500 จุดภาพ


The effectiveness of the heurist
of typographical and layout styles, h
including multiple columns, sparse tal
ers, trailers, footnotes, and scanner no
rately printed. Most had an unambiguo
and binarization were performed by a
binary image of about 8.5 million pixels
too large to be a character were set
ments, without how

ภาพที่ 6 ภาพเอกสารรูปแบบ 1 ที่ใช้ในการทดลองมีความเอียงในทิศทวนเข็มนาฬิกา


The effectiveness of the heurist
of typographical and layout styles, h
including multiple columns, sparse tal
ers, trailers, footnotes, and scanner no
rately printed. Most had an unambiguo
and binarization were performed by a
binary image of about 8.5 million pixels
too large to be a character were set
ments, without how

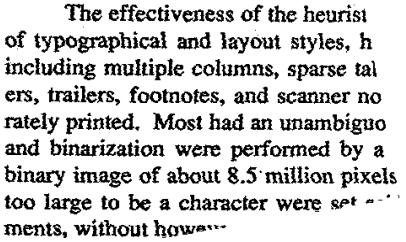
ภาพที่ 7 ภาพเอกสารเมื่อผ่านกระบวนการ Opening Morphology เพื่อลดสัญญาณรบกวน

จากภาพที่ 6 และภาพที่ 7 จะเห็นความแตกต่างของข้อมูลภาพที่ชัดเจน คือในภาพที่ 6 ที่เป็นภาพเริ่มต้นของการทดลองจะมีจุดภาพขนาดเล็กที่อยู่นอกแนวของตัวหนังสือซึ่งเราจะถือว่าเป็นสัญญาณรบกวนขนาดเล็ก ถ้าจุดภาพดังกล่าวเป็นจุดที่อยู่ในแนวบรรทัดของตัวหนังสือจริง และอยู่ก่อนหน้าที่จะถึงแนวขอบซ้ายของบรรทัดตัวหนังสือ ถ้านำจุดนั้นเข้ามาประกอบการพิจารณาก็จะทำให้การคำนวณผิดพลาดไปได้ จึงควรขจัดจุดนั้นออกไปด้วยการทำ Erosion แต่เนื่องจากอัลกอริทึมที่นำเสนอเป็นไปอย่างอัตโนมัติ การขจัดจุดดำจะส่งผลกระทบต่อขอบของตัวหนังสือ ทำให้ขาดหายไปไม่สมบูรณ์ ผู้วิจัยจึงได้เพิ่มขั้นตอนการทำ Dilation สำหรับการเพิ่มจำนวนจุดดำในภาพเอกสารเมื่อเพิ่มจำนวนจุดดำเข้าไปแล้วทำให้คุณภาพของภาพเอกสารดีขึ้น (จำนวนเซตของจุดดำมากขึ้น) ดังแสดงในภาพที่ 7

4. ผลการทดลอง

จากการทดลองกับภาพเอกสาร ทั้ง 26 ภาพ มีมุมความเอียงที่แตกต่างกัน โปรแกรมที่ใช้พัฒนาโดยภาษาซีบนเครื่องไมโครคอมพิวเตอร์ที่ใช้หน่วยประมวลผลกลางรุ่นเพนเทียม ผลปรากฏว่าอัลกอริทึมที่นำเสนอสามารถประมาณค่ามุมเอียงของหน้าเอกสารได้ถูกต้องสูง การประมาณค่ามุมเอียงมีความคาดเคลื่อนไม่เกิน 1 องศา จากรูป 6 แสดงเอกสารที่สแกนเข้ามามีมุมเอียง เมื่อผ่านการประมาณมุมเอียงด้วยวิธี

การที่นำเสนอประมาณค่ามุมเอียงได้ -5 องศา แล้วทำการหมุนกลับด้วยค่ามุม และทิศทางทวนเข็มนาฬิกา ได้ภาพเอกสารใหม่ดังแสดงในภาพที่ 8


The effectiveness of the heurist
of typographical and layout styles, h
including multiple columns, sparse tal
ers, trailers, footnotes, and scanner no
rately printed. Most had an unambiguo
and binarization were performed by a
binary image of about 8.5 million pixels
too large to be a character were set
ments, without how

ภาพที่ 8 แสดงภาพที่ได้จากการหมุนกลับของหน้าเอกสารมุมเอียง - 5 องศา

5. สรุป

อัลกอริทึมที่นำเสนอในงานวิจัยนี้สามารถใช้ประมาณค่ามุมเอียงของหน้าเอกสารได้ทุกประเภท เนื่องจากอัลกอริทึมนี้จะใช้เพียงแนวของจุดดำทางซ้ายของตัวหนังสือตัวแรกในแต่ละบรรทัด นำมาเรียงต่อกันเป็นพารากราฟ เลือกพารากราฟที่มีเซตข้อมูลมากที่สุดเป็นตัวแทนของหน้าเอกสารในการประมาณค่ามุมเอียง จึงใช้เวลาของหน่วยประมวลผลกลางน้อย แต่ให้ความถูกต้องในการประมาณค่ามุมเอียงของหน้าเอกสารได้สูง เพราะใช้การวิเคราะห์การถดถอยสำหรับประมาณสมการถดถอยของแนวขอบของตำแหน่งจุดดำทางซ้ายของตัวหนังสือ

อัลกอริทึมที่นำเสนอนี้ จะสามารถจัดการกับภาพเอกสารไบนารีเท่านั้น, ภาพเอกสารมีกลุ่มของสัญญาณรบกวนเกิดขึ้นได้, และโดยมากเอกสารที่สแกนเข้ามาจะมีค่ามุมเอียงที่เกิดขึ้นเป็นมุมขนาดเล็ก ซึ่งสาเหตุของความเอียงในหน้าเอกสารอาจเกิดเนื่องมาจาก การให้สแกนเนอร์ทำการดึงกระดาษเข้าเครื่องอัตโนมัติ (Auto Feed) หรือเกิดจากการวางหน้ากระดาษไม่ตรง ดังนั้นในงานวิจัยนี้จึงเลือกทดสอบอัลกอริทึมที่นำเสนอกับภาพเอกสารที่มีมุมเอียงอยู่ในช่วงประมาณ $\pm$ 15 องศา

6. เอกสารอ้างอิง

[1] H.S. Baird., “The Skew of printed documents,” Proc. of Society of Photographic Scientists and Engineers, Vol.40, pp. 21-24, 1987.

[2] A. Dengel., “ANASTASIL : A System for Low-level and High-level Geometric Analysis of printed documents,” IN *Structured Document Image Analysis.,* By H.S. Baird., H. Bunke., and K. Yamamoto., Springer-Verlag, pp. 70-98, 1992.

[3] M. Sonka, V. Hlavac, and R. Boyle., *Image Processing, Analysis and Machine Vision,* Chapman & Hall Computing, Boundary Row, London, pp. 422-442, 1993.

[4] S.Nakamura., *Applied Numerical Method IN C.,* Prentice Hall International Inc., pp. 294-303, 1993.

[5] ทัศนีย์ ชังเทศ และ สมภพ ถาวรยิ่ง., *การวิเคราะห์การถดถอยและสหสัมพันธ์,* โรงพิมพ์มหาวิทยาลัยธรรมศาสตร์, กรุงเทพฯ หน้า 18-20, 2530.