บทที่ 3

การออกแบบและพัฒนาโปรแกรม

งานวิจัยนี้เป็นการพัฒนา ระบบออนไลน์สำหรับการรู้จำตัวพิมพ์อักษรไทยและตัวพิมพ์อักษรอังกฤษ ขึ้นเพื่อให้ระบบคอมพิวเตอร์สามารถรับข้อมูลเข้าได้อย่างรวดเร็วและถูกต้องยิ่งขึ้น โดยระบบนี้จะทำการรับข้อมูลตัวพิมพ์อักษรไทย และตัวพิมพ์อักษรอังกฤษด้วยเครื่องกวาดตรวจด้วยแสง (Optical Scanner) แทนการรับข้อมูลเข้าทางแป้นพิมพ์ และจะนำข้อมูลที่รับเข้านี้ซึ่งอยู่ในรูปของบิตเมตริกซ์ซึ่งจะอยู่ในหน่วยความจำ (Memory) มาทำการเรียนรู้ (Training) และทำการรู้จำ (Recognize) ตัวอักษรที่ได้นั้นในทันที โดยผลลัพธ์ที่ได้คือแฟ้มข้อมูลที่สามารถนำไปใช้งานได้ตามต้องการ

ขั้นตอนการวิจัยจะแบ่งเป็น 2 ขั้นตอน ดังนี้

ขั้นตอนที่ 1 ขั้นตอนการเรียนรู้ (Training)

ขั้นตอนที่ 2 ขั้นตอนการรู้จำ (Recognize)

3.1 <u>ขั้นตอนการเรียนรู้</u> (Training)

ขั้นตอนการเรียนรู้เป็นขั้นตอนในการเรียนรู้ตัวอักษรต้นแบบ เพื่อสร้างแฟ้มข้อมูลของตัวอักษรต้นแบบขึ้น เพื่อนำมาใช้เปรียบเทียบกับตัวอักษรที่จะทำการรู้จำต่อไป

โดยมีขั้นตอนการทำงานดังนี้

3.1.1 <u>รับข้อมูลตัวพิมพ์อักษรไทยและตัวพิมพ์อักษรอังกฤษ</u>

การรับข้อมูลตัวพิมพ์อักษรไทยและตัวพิมพ์อักษรอังกฤษนี้ จะใช้เครื่องกวาดตรวจด้วยแสง (Optical Scanner) โดยข้อมูลที่ได้จะอยู่ในหน่วยความจำ โดยเครื่องกวาดตรวจด้วยแสง จะรับข้อมูลเข้ามาแสดงผลที่หน้าจอ ซึ่งอยู่ในรูปของหน่วยความจำชนิดหนึ่งนั้น คือ หน่วยความจำของจอภาพ (VIDEO RAM) หลังจากนั้น จะนำข้อมูลจากหน่วยความจำของจอภาพมาเก็บไว้ใน หน่วยความจำของข้อมูล (DATA MEMORY) ซึ่งเป็นค่าตัวแปร เพื่อนำไปใช้ต่อไป

3.1.2 กำจัดสัญญาณรบกวนและแก้ไขเพิ่มเติมเส้นตรง

การกำจัดสัญญาณรบกวนและการแก้ไขเพิ่มเติมเส้นตรงจะเป็นเทคนิคที่ทำการกำจัดส่วนเกินของจุดต่างของข้อมูลนำเข้า และเพิ่มเติมส่วนที่เป็นเส้นตรงให้ครบถ้วนสมบูรณ์ ซึ่งจะมีผลต่อการทำงานในขั้นตอนต่อไป เพื่อให้การทำงานในขั้นตอนต่อไปดีขึ้น ในขั้นตอนนี้แบ่งออกได้ 2 ขั้นตอน คือ

ขั้นตอนที่ 1 การกำจัดสัญญาณรบกวน
ขั้นตอนที่ 2 การแก้ไขเพิ่มเติมเส้นตรง

ขั้นตอนที่ 1 การกำจัดสัญญาณรบกวน

การกำจัดสัญญาณรบกวนจะใช้หลักการของเมตริกซ์ โดยจะนำจุดที่ต้องการตรวจสอบว่าเป็นสัญญาณรบกวนหรือไม่นั้นมาเปรียบเทียบกับเมตริกซ์ ซึ่งในที่นี้สัญญาณรบกวนมี 6 รูปแบบ โดยการเปรียบเทียบจะเริ่มเปรียบเทียบกับเมตริกซ์รูปแบบที่ 1 ไปเรื่อยๆจนถึงรูปแบบที่ 6 ถ้าจุดที่ต้องการเปรียบเทียบไม่ตรงกับเมตริกซ์รูปแบบใด จะถือว่าจุดนั้นไม่เป็นสัญญาณรบกวน แต่ถ้าหากพบว่าสัญญาณรบกวนนั้นตรงกับรูปแบบใดรูปแบบหนึ่ง จะถือว่าจุดนั้นเป็นสัญญาณรบกวน ก็จะทำการกำจัดจุดนั้นออก และจะต้องทำการตรวจสอบจุดที่อยู่รอบๆจุดที่พบว่าเป็นสัญญาณรบกวนทั้ง 8 จุดต่อไป โดยจุดทั้ง 8 จุดนั้นคือจุดดังที่ปรากฏในเมตริกซ์ต่อไปนี้

1	2	3
4	1	5
6	7	8

1 คือ จุดที่ต้องการตรวจสอบว่าเป็นสัญญาณรบกวนหรือไม่จะต้องไม่เป็นช่องว่าง

1 - 8 คือ จุดที่อยู่รอบๆทั้ง 8 จุด

โดยจะเริ่มตรวจสอบจุดที่ 1 ไปเรื่อยๆจนถึงจุดที่ 8 เมื่อพบว่าจุดใดเป็นสัญญาณรบกวนก็จะทำการกำจัดจุดนั้นและทำการตรวจสอบจุดที่อยู่รอบๆจุดนั้นเช่นกัน เมื่อทำการตรวจสอบ จุดที่อยู่รอบๆทั้ง 8 จุดแล้วไม่พบว่าจุดใดเป็นสัญญาณรบกวนแล้ว ก็จะทำการตรวจสอบจุดต่อไปคือ จุดที่อยู่ทางขวามือของจุดที่ได้ตรวจสอบแล้ว ในที่นี้คือจุดที่ 5 ซึ่งอยู่ทางขวาของจุด 1 (Hiranvanichakorn p., 1982)

แบบที่ 1

0	1	0
	1	

```
      11111111111
     1111111111111
    1111      1111111
   111          11111
  111            11111
 111              1111
111111    1       1111
111111111111      1111
 11111111111      1111
  11111           1111
  1111            1111
  1111            1111
  1111            1111
  1111            1111
  1111            1111
  1111            1111
  1111            1111
  1111            1111
  1111            1111
```

1 คือ จุดที่ไม่เป็นช่องว่าง และตรวจสอบว่า เป็นสัญญาณรบกวน หรือไม่

0 คือ จุดที่เป็นช่องว่าง

1 คือ จุดที่ไม่เป็นช่องว่าง

ช่องว่างคือไม่สนใจว่าเป็นจุดประเภทใด

รูปที่ 3.1 แสดงการกำจัดสัญญาณรบกวนรูปแบบที่ 1

แบบที่ 2

	1	
0	<u>1</u>	0

<u>1</u> คือ จุดที่ไม่เป็นช่องว่าง
และตรวจสอบว่า
เป็นสัญญาณรบกวน
หรือไม่

0 คือ จุดที่เป็นช่องว่าง

1 คือ จุดที่ไม่เป็นช่องว่าง

ช่องว่างคือไม่สนใจว่าเป็นจุดประเภทใด

```
      11111111111
     1111111111111
    1111      1111111
   111          11111
  111             11111
 111               1111
111111             1111
111111111111       1111
 11111111111       1111
  11111   1        1111
  1111             1111
  1111             1111
  1111             1111
  1111             1111
  1111             1111
  1111             1111
  1111             1111
  1111             1111
  1111             1111
```

รูปที่ 3.2 แสดงการกำจัดสัญญาณรบกวนรูปแบบที่ 2

แบบที่ 3

	0	
0	1	
0	0	

```
        11111111111
       11111111111111
      1111      1111111
     111          11111
    111            11111
   111              1111
  1111111           1111
   11111111111      1111
   1111111111       1111
    11111           1111
    1111            1111
    1111            1111
   11111           11111
    1111            1111
    1111            1111
    1111            1111
    1111            1111
    1111            1111
    1111            1111
    1111            1111
```

1 คือ จุดที่ไม่เป็นช่องว่าง และตรวจสอบว่าเป็นสัญญาณรบกวนหรือไม่

0 คือ จุดที่เป็นช่องว่าง

1 คือ จุดที่ไม่เป็นช่องว่าง

ช่องว่างคือไม่สนใจว่าเป็นจุดประเภทใด

รูปที่ 3.3 แสดงการกำจัดสัญญาณรบกวนรูปแบบที่ 3

แบบที่ 4

	0	
	<u>1</u>	0
	0	0

<u>1</u> คือ จุดที่ไม่เป็นช่องว่าง และตรวจสอบว่าเป็นสัญญาณรบกวนหรือไม่

0 คือ จุดที่เป็นช่องว่าง

1 คือ จุดที่ไม่เป็นช่องว่าง

ช่องว่างคือไม่สนใจว่าเป็นจุดประเภทใด

```
                        11111
                       11111
     111111           11111
   1111111111        11111
  111111111111      1111
 11111111111111  11111
11111  11111111111111
1111    111 11111111
 111    111 11111111
 1111   111 1111111111
  1111 111  1111   111
   1111111 1111    1111
    1111   1111    1111
           1111     1111
           1111     111
            1111    111
            111     111
            111     111
            111     111
            111     111
            111     1111
            1111    1111
            1111    1111
         111111111111111
         11111111111111
         1111111111111
```

รูปที่ 3.4 แสดงการกำจัดสัญญาณรบกวนรูปแบบที่ 4

แบบที่ 5

1	0	
0	<u>1</u>	
0		

<u>1</u> คือ จุดที่ไม่เป็นช่องว่าง
และตรวจสอบว่า
เป็นสัญญาณรบกวน
หรือไม่

0 คือ จุดที่เป็นช่องว่าง

1 คือ จุดที่ไม่เป็นช่องว่าง

ช่องว่างคือไม่สนใจว่าเป็นจุดประเภทใด

```
                          11111
                         11111
     111111             11111
   1111111111          11111
  111111111111        1111
 11111111111111     11111
11111  11111111111111
1111    111 11111111
1111    111 11111111
11 11   111 1111111111
  1111 1111  1111   111
    1111111 1111    1111
     1111   1111    1111
            1111     111
            1111     111
             111     111
             111     111
             111     111
             111     111
             111     111
             111    1111
            1111    1111
            1111    1111
        111111111111111
        11111111111111
        1111111111111
```

รูปที่ 3.5 แสดงการกำจัดสัญญาณรบกวนรูปแบบที่ 5

แบบที่ 6

	0	
1	1	1
	0	

```
                       11111
                      11111
     111111          11111
   1111111111       11111
  111111111111     1111
 11111111111111  11111
11111  11111111111111
1111    111 11111111
1111    111 11111111
 1111   111 1111111111
  1111 1111  1111   111
    111111111111    1111
     1111   1111    1111
            1111     111
            1111     111
             111     111
             111     111
             111     111
             111     111
             111     111
             111    1111
            1111    1111
            1111    1111
        111111111111111
        11111111111111
        1111111111111
```

1 คือ จุดที่ไม่เป็นช่องว่าง และตรวจสอบว่าเป็นสัญญาณรบกวนหรือไม่

0 คือ จุดที่เป็นช่องว่าง

1 คือ จุดที่ไม่เป็นช่องว่าง

ช่องว่างคือไม่สนใจว่าเป็นจุดประเภทใด

รูปที่ 3.6 แสดงการกำจัดสัญญาณรบกวนรูปแบบที่ 6

จุฬาลงกรณ์มหาวิทยาลัย

ขั้นตอนที่ 2 การแก้ไขเพิ่มเติมเส้นตรง

การแก้ไขและเพิ่มเติมเส้นตรงจะใช้หลักการของเมตริกซ์เหมือนการกำจัดสัญญาณรบกวน เมื่อพบว่าจุดนั้นเป็นต้องทำการแก้ไขเพิ่มเติมเส้นตรงแล้วจะทำการแก้ไขหรือเพิ่มเติมจุดนั้น และตรวจสอบจุดที่อยู่รอบๆจุดนั้นทั้ง 8 จุด ว่าต้องแก้ไขหรือเพิ่มเติมเส้นตรงหรือไม่ ถ้าต้องทำก็ให้แก้ไขเพิ่มเติมจุดที่พบใหม่ และตรวจสอบจุดที่อยู่รอบๆจุดใหม่นี้ทั้ง 8 จุดว่าต้องแก้ไขหรือเพิ่มเติมเส้นตรงหรือไม่ทำเช่นนี้ไปเรื่อยๆ จนไม่พบจุดที่ต้องแก้ไขเพิ่มเติมเส้นตรงทั้ง 8 จุด จึงกลับมาที่จุดแรกสุด เพื่อตรวจสอบจุดต่อไปเรื่อยๆ ซึ่งการแก้ไขเพิ่มเติมเส้นตรง มี 18 รูปแบบ ในการแก้ไขเพิ่มเติมเส้นตรง จะเริ่มตรวจสอบตั้งแต่รูปแบบที่ 1 ไปตามลำดับ จนถึงรูปแบบที่ 18 ถ้าตรงกับรูปแบบใดก็จะถือว่าจุดนั้นคือจุดที่ต้องเพิ่มเติมเส้นตรง ตัวอย่างเช่น เมื่อเริ่มตรวจสอบที่รูปแบบที่ 1 ปรากฏว่าจุดนั้นไม่ตรงกับรูปแบบที่ 1 ดังนั้นจะต้องนำจุดนั้นไปตรวจสอบกับกับรูปแบบที่ 2 ต่อไป ถ้าจุดนั้นไม่ ตรงกับรูปแบบที่ 2 จะทำการตรวจสอบรูปแบบที่ 3 และ 4 และ 5 ไปจนถึง รูปแบบที่ 18 ตามลำดับ จนกระทั่งพบว่าจุดนั้นตรงกับรูปแบบใด รูปแบบหนึ่ง เช่น ตรงกับรูปแบบที่ 5 ก็จะถือว่าจุดนั้นเป็น จุดที่ต้องเพิ่มเติมเส้นตรง หรือจนกระทั่งตรวจสอบรูปแบบที่ 18 แล้วจุดนั้นไม่ตรงกับรูปแบบใด ก็จะถือว่าจุดนั้นไม่เป็นจุดที่ต้องเพิ่มเติมเส้นตรง (Hiranvanichakorn p., 1982)

0 คือ จุดที่เป็นช่องว่าง และ
จุดที่ต้องแก้ไขเพิ่มเติมเส้นตรง

0 คือ จุดที่เป็นช่องว่าง

1 คือ จุดที่ไม่เป็นช่องว่าง

ช่องว่างคือไม่สนใจว่าเป็นจุดประเภทใด

รูปแบบที่ 1

	0	1	
	0	<u>0</u>	1
	0	0	1
	0	1	

```
                       11111
                      11111
     111111          11111
   1111111111       11111
  111111111111     1111
 11111 11111111  11111
11111   1111111111111
1111    111 11111111
 111    111 11111111
 1111   111 1111111111
   1111 111  1111   111
    1111111 1111    1111
     1111   1111    1111
            1111     111
            1111     011
             111      11
             111     111
             011     111
              11     111
             111     111
             111    1111
            1111    1111
            1111    1111
         111111111111111
         11111111111111
        1111111111111
```

รูปที่ 3.7 แสดงเมตริกซ์ของการแก้ไขเพิ่มเติมเส้นตรงรูปแบบที่ 1

รูปแบบที่ 2

	0	1	
	0	<u>0</u>	1
	0	1	

```
                       11111
                      11111
     111111          11111
   1111111111       11111
  111111111111     1111
 11111 11111111  11111
11111   1111111111111
1111    111 11111111
 111    111 1111111
 1111   111 1111111111
   1111 111  1111   111
    1111111 1111    1111
     1111   1111    1111
            1111     111
            1111     111
             111     011
             111     111
             111     111
             011     111
             111     111
             111    1111
            1111    1111
            1111    1111
         111111111111111
         11111111111111
        1111111111111
```

รูปที่ 3.8 แสดงเมตริกซ์ของการแก้ไขเพิ่มเติมเส้นตรงรูปแบบที่ 2

รูปแบบที่ 3

		1	0
	1	0	0
	1	0	0
		1	0

```
      11111111111
     11111111111111
    1111      1111111
   111          11111
  111            11111
 111              1111
111111            1111
111111111111      1111
 11111111111      1111
  11111           1111
  1111            1111
  1110            1111
  111             1111
  1111            1111
  1111            1111
  1111            1110
  1111            111
  1111            1111
  1111            1111
  1111            1111
```

รูปที่ 3.9 แสดงเมตริกซ์ของการแก้ไขเพิ่มเติมเส้นตรงรูปแบบที่ 3

รูปแบบที่ 4

	1	0	
1	0	0	
	1	0	

```
                       11111
     111111          11111
   1111111111       11111
  111111111111     1111
 11111 11111111  11111
11111   1111111111111
1111    111 11111111
 111    111 11111111
 1111   111 1111111111
   1111 111  1111   111
    1111111 1111    1111
     1111   1111    1111
            1111     111
            1111     111
             111     110
             111     111
             111     111
             110     111
             111     111
             111    1111
            1111    1111
            1111    1111
         111111111111111
         11111111111111
        1111111111111
```

รูปที่ 3.10 แสดงเมตริกซ์ของการแก้ไขเพิ่มเติมเส้นตรงรูปแบบที่ 4

รูปแบบที่ 5

0	0	0	
1	0	1	
	1		

```
      11111011111
     11111111111111
    1111          1111111
   111              11111
  111                11111
 111                  1111
111111                1111
11111111011           1111
 11111111111          1111
  11111               1111
  1111                1111
  1111                1111
  1111                1111
  1111                1111
  1111                1111
  1111                1111
  1111                1111
  1111                1111
  1111                1111
  1111                1111
```

รูปที่ 3.11 แสดงเมตริกซ์ของการแก้ไขเพิ่มเติมเส้นตรงรูปแบบที่ 5

จุฬาลงกรณ์มหาวิทยาลัย

รูปแบบที่ 6

	1		
1	0	1	
0	0	0	

```
        11111111111
       111111011111111
      1111        1111111
     111             11111
    111               11111
   111                 1111
  111111               1111
  111111111111         1111
   11111110111         1111
    11111              1111
    1111               1111
    1111               1111
    1111               1111
    1111               1111
    1111               1111
    1111               1111
    1111               1111
    1111               1111
    1111               1111
    1111               1111
```

รูปที่ 3.12 แสดงเมตริกซ์ของการแก้ไขเพิ่มเติมเส้นตรงรูปแบบที่ 6

รูปแบบที่ 7

	1	1	
1	<u>0</u>	0	1
0	0	0	0

```
      11111111111
     11111111111111
    1111     1111111
   111         11111
  111            11111
 111              1111
111111            1111
111111111111      1111
 11111110 11      1111
  11111           1111
  1111            1111
  1111            1111
  1111            1111
  1111            1111
  1111            1111
  1111            1111
  1111            1111
  1111            1111
  1111            1111
  1111            1111
```

รูปที่ 3.13 แสดงเมตริกซ์ของการแก้ไขเพิ่มเติมเส้นตรงรูปแบบที่ 7

รูปแบบที่ 8

0	0	0	0
1	<u>0</u>	0	1
	1	1	

```
      11111111111
     11111111111111
    1111      1111111
   111          11111
  111             11111
 111               1111
111111             1111
111111110 11       1111
 11111111111       1111
  11111            1111
  1111             1111
  1111             1111
  1111             1111
  1111             1111
  1111             1111
  1111             1111
  1111             1111
  1111             1111
  1111             1111
  1111             1111
```

รูปที่ 3.14 แสดงเมตริกซ์ของการแก้ไขเพิ่มเติมเส้นตรงรูปแบบที่ 8

รูปแบบที่ 9

1	1	1	0
1	0	0	0
1	1	1	0

```
      11111111111
     11111111111111
    1111       1111111
   111           11111
  111             11111
 111               1111
111111             1111
111111111111       1111
 11111111111       1111
  11111            1111
  1111             1111
  1111             1111
  110              1111
  1111             1111
  1111             1111
  1111             1111
  1111             1111
  1111             1111
  1111             1111
  1111             1111
```

รูปที่ 3.15 แสดงเมตริกซ์ของการแก้ไขเพิ่มเติมเส้นตรงรูปแบบที่ 9

รูปแบบที่ 10

0	1	1	1
0	0	0	1
0	1	1	1

```
      11111111111
     11111111111111
    1111          1111111
   111              11111
  111                11111
 111                  1111
111111                1111
111111111111          1111
 11111111111          1111
  11111               1111
  1111                1111
  1111                1111
  011                 1111
  1111                1111
  1111                1111
  1111                1111
  1111                1111
  1111                1111
  1111                1111
  1111                1111
```

รูปที่ 3.16 แสดงเมตริกซ์ของการแก้ไขเพิ่มเติมเส้นตรงรูปแบบที่ 10

รูปแบบที่ 11

0	1	1	
0	1	0	
0	1		

```
      11111111111
     11111111111111
    1111      1111111
   111          11111
  111            11111
 111              1111
111111            1111
111111111111      1111
 11111111111      1111
  11111           1111
  1111            1111
  1111            10
  1111            1
  1111            1
  1111            1
  1111            1111
  1111            1111
  1111            1111
  1111            1111
  1111            1111
```

รูปที่ 3.17 แสดงเมตริกซ์ของการแก้ไขเพิ่มเติมเส้นตรงรูปแบบที่ 11

รูปแบบที่ 12

	1	1	0
	0	1	0
		1	0

```
      11111111111
     11111111111111
    1111      1111111
   111          11111
  111            11111
 111              1111
111111            1111
111111111111      1111
 11111111111      1111
  11111           1111
  1111            1111
  1111             01
  1111              1
  1111              1
  1111              1
  1111            1111
  1111            1111
  1111            1111
  1111            1111
  1111            1111
```

รูปที่ 3.18 แสดงเมตริกซ์ของการแก้ไขเพิ่มเติมเส้นตรงรูปแบบที่ 12

รูปแบบที่ 13

0	0	0	
1	1	1	
1	<u>0</u>		

```
      11111111111
     11111111111111
    1111      1111111
   111          11111
  111             11111
 111               1111
111111             1111
111111111111       1111
 1111110           1111
  11111            1111
  1111             1111
  1111             1111
  1111             1111
  1111             1111
  1111             1111
  1111             1111
  1111             1111
  1111             1111
  1111             1111
  1111             1111
```

รูปที่ 3.19 แสดงเมตริกซ์ของการแก้ไขเพิ่มเติมเส้นตรงรูปแบบที่ 13

รูปแบบที่ 14

1	0		
1	1	1	
0	0	0	

```
      11111111111
     11111111111111
    1111       1111111
   111           11111
  111             11111
 111               1111
111111             1111
111111110          1111
 11111111111       1111
  11111            1111
  1111             1111
  1111             1111
  1111             1111
  1111             1111
  1111             1111
  1111             1111
  1111             1111
  1111             1111
  1111             1111
  1111             1111
```

รูปที่ 3.20 แสดงเมตริกซ์ของการแก้ไขเพิ่มเติมเส้นตรงรูปแบบที่ 14

รูปแบบที่ 15

1	1		
0	1	0	
	1	1	

```
      11111111111
     1111111111111
    1111     1111111
   111         11111
  111           11111
 111             1111
111111           1111
111111111111     1111
 11111111111     1111
  11111          1111
  1111           01
  1111            111
  1111           1111
  1111           1111
  1111           1111
  1111           1111
  1111           1111
  1111           1111
  1111           1111
  1111           1111
```

รูปที่ 3.21 แสดงเมตริกซ์ของการแก้ไขเพิ่มเติมเส้นตรงรูปแบบที่ 15

รูปแบบที่ 16

	1	1	
0	1	<u>0</u>	
1	1		

```
           11111111111
          11111111111111
         1111        1111111
        111            11111
       111              11111
      111                1111
     111111              1111
     111111111111        1111
      11111111111        1111
       11111             1111
       1111               10
       1111              11
       1111              11
       1111              1111
       1111              1111
       1111              1111
       1111              1111
       1111              1111
       1111              1111
       1111              1111
```

รูปที่ 3.22 แสดงเมตริกซ์ของการแก้ไขเพิ่มเติมเส้นตรงรูปแบบที่ 16

รูปแบบที่ 17

	0	1	
1	1	1	
1	0		

```
          111111111
        1111111111111
       1111     1111111
      1111         11111
     1111           11111
     111             1111
   111111            1111
   110  1111111      1111
        1111111      1111
     1111            1111
    1111             1111
    1111             1111
    1111             1111
    1111             1111
    1111             1111
    1111             1111
    1111             1111
    1111             1111
    1111             1111
    1111             1111
```

รูปที่ 3.23 แสดงเมตริกซ์ของการแก้ไขเพิ่มเติมเส้นตรงรูปแบบที่ 17

I17083035

รูปแบบที่ 18

1	0		
1	1	1	
	0	1	

```
       111111111
     1111111111111
    1111     1111111
   1111          11111
  1111            11111
 1111              1111
1111111110         1111
111111111111       1111
 11111111 11       1111
 11111             1111
 1111              1111
 1111              1111
 1111              1111
 1111              1111
 1111              1111
 1111              1111
 1111              1111
 1111              1111
 1111              1111
 1111              1111
```

รูปที่ 3.24 แสดงเมตริกซ์ของการแก้ไขเพิ่มเติมเส้นตรงรูปแบบที่ 18

จะเห็นว่าการกำจัดสัญญาณรบกวนและแก้ไขเพิ่มเติมเส้นตรง จะช่วยในการปรับเส้นของตัวอักษรให้ตรงขึ้นด้วยเช่น

ตัวอักษรก่อน
กำจัดสัญญาณรบกวนและแก้ไขเพิ่มเติมเส้นตรง

```
     1
 111111
 111111
  111111
 111111
1111111
 111111
    1
```

ตัวอักษรหลัง
กำจัดสัญญาณรบกวนและแก้ไขเพิ่มเติมเส้นตรง

```
111111
111111
111111
111111
111111
111111
```

เส้นแสดงขอบของตัวอักษรก่อน
กำจัดสัญญาณรบกวนและแก้ไขเพิ่มเติมเส้นตรง

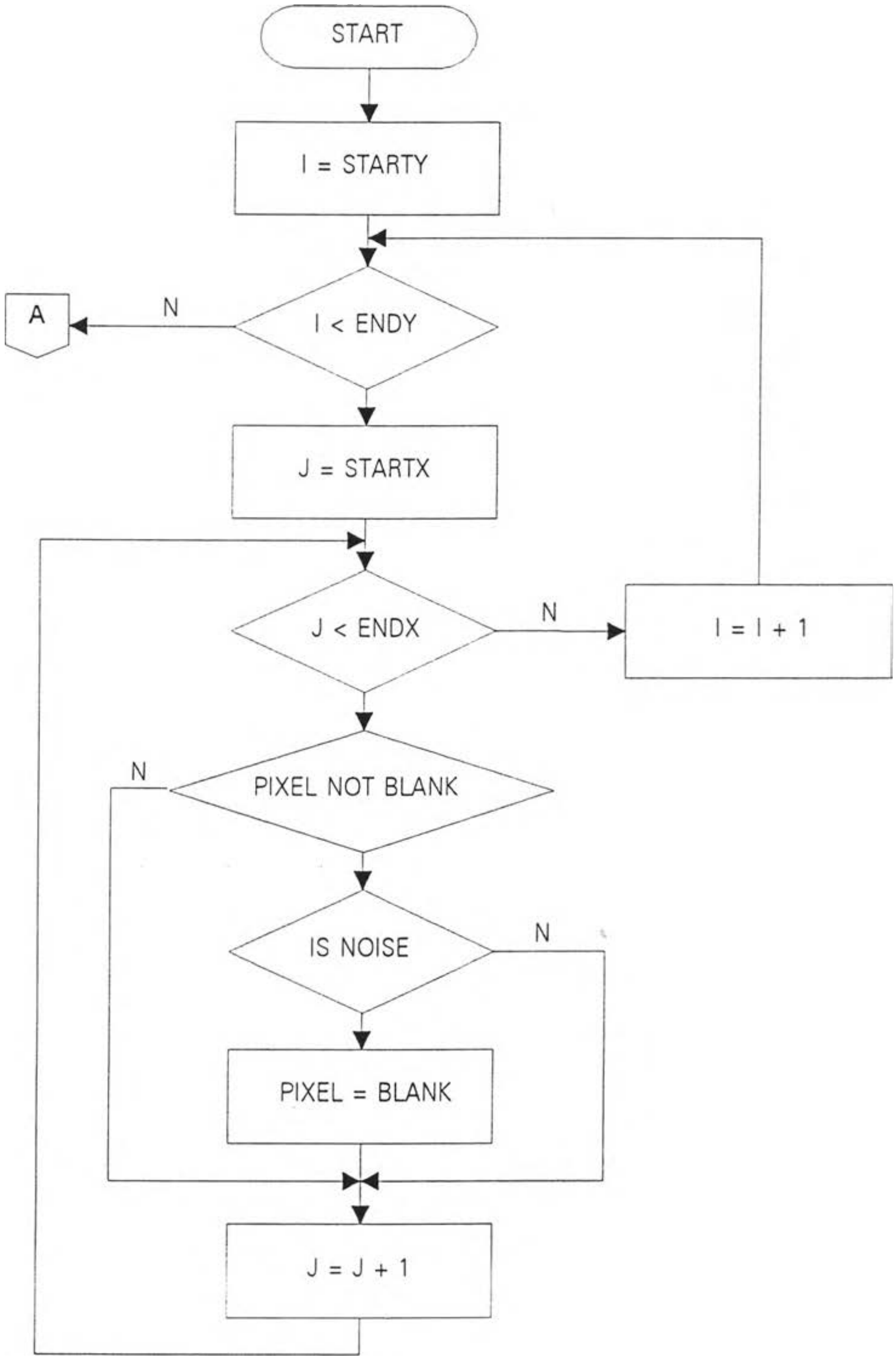
```
    1
 111 11
 1    1
  1    1
 1    1
1     1
 11 111
   1
```

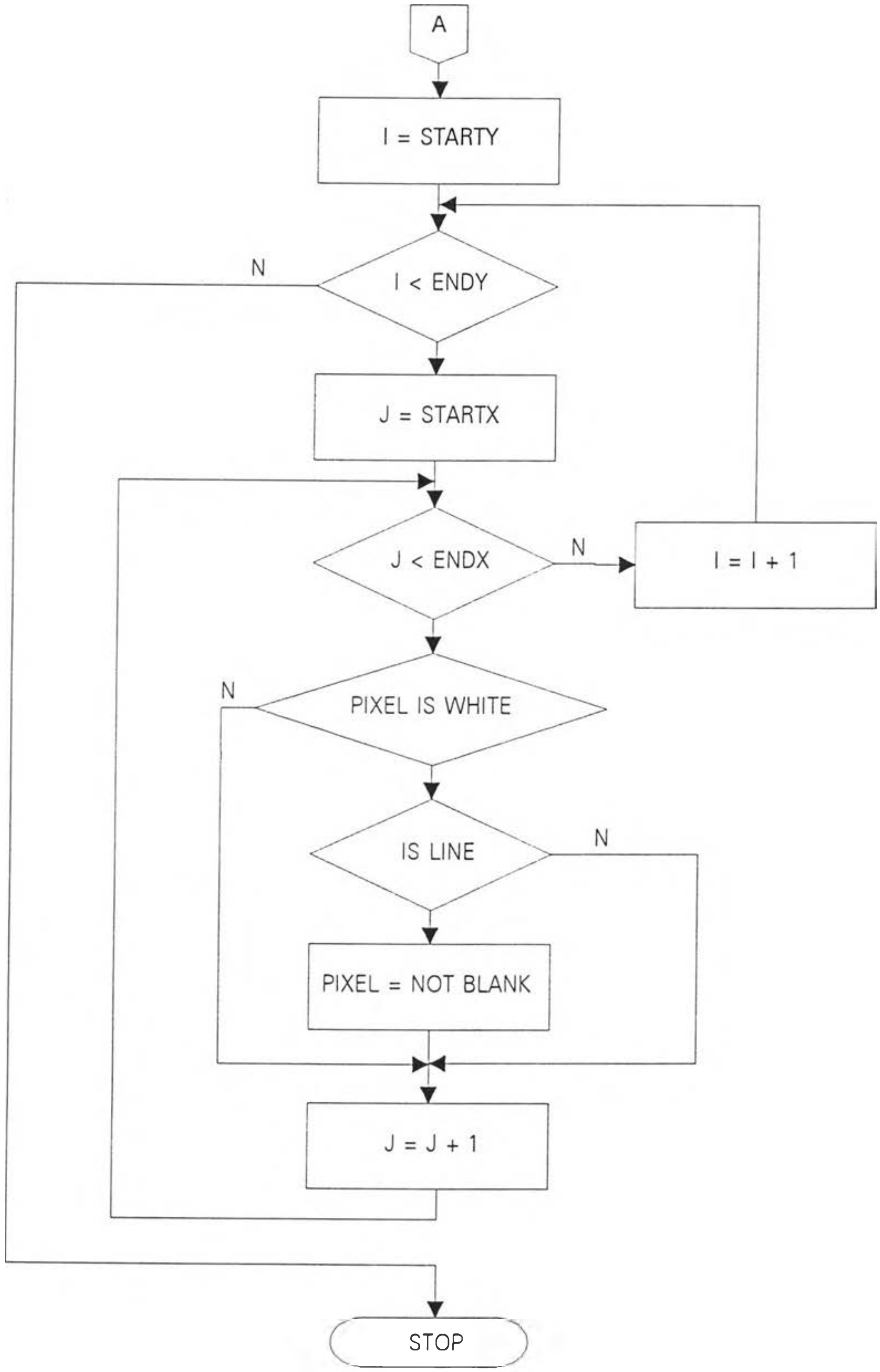
เส้นแสดงขอบของตัวอักษรหลัง
กำจัดสัญญาณรบกวนและแก้ไขเพิ่มเติมเส้นตรง

```
111111
1    1
1    1
1    1
1    1
111111
```

รูปที่ 3.25 ตัวอย่างการเปรียบเทียบการกำจัดสัญญาณรบกวนและเส้นแสดงขอบที่หาได้

รูปที่ 3.26 ผังงานแสดงการจัดสัญญาณรบกวนและแก้ไขเส้นตรง




A
I = STARTY
I < ENDY
N
J = STARTX
J < ENDX
N
I = I + 1
PIXEL IS WHITE
N
IS LINE
N
PIXEL = NOT BLANK
J = J + 1
STOP

3.1.3 หาจุดเริ่มต้นของตัวอักษร

เป็นการตรวจสอบจุดต่างๆที่ประกอบขึ้นเป็นตัวอักษรเพื่อหาว่าจุดใด คือ จุดแรกที่พบ ซึ่งจุดนั้นต้องไม่เป็นช่องว่าง และไม่ใช่จุดที่เป็นเส้นแสดงขอบ หรือเป็นจุดที่อยู่ภายในตัวอักษรตัวอื่น หรือ จุดที่ประกอบเป็นตัวอักษรตัวอื่น โดยการหาจุดเริ่มต้นของตัวอักษรนั้น จะเริ่มที่จุดที่อยู่ในสดมภ์แรกของแถวแรก หรือจุดที่อยู่ใน สดมภ์ที่ 0 แถวที่ 0 ส่วนการหาจุดเริ่มต้นของตัวอักษรต่อไป จะเริ่มที่ จุดเริ่มต้นของตัวอักษรก่อนหน้า และจะหาไปเรื่อยๆจนพบจุดเริ่มต้นของตัวอักษรต่อไป

3.1.4 วิเคราะห์เส้นแสดงขอบของตัวอักษร

เส้นแสดงขอบของอักษร (Contour) หมายถึงเส้นที่ประกอบขึ้นเป็นตัวอักษร เทคนิคนี้จะนำโครงสร้างความโค้งของอักษรมาเป็นลักษณะสำคัญใน การวิเคราะห์ ซึ่งลักษณะสำคัญได้แก่ ส่วนโค้ง จำนวนหัว ตำแหน่งของหัว

การหาส่วนโค้ง จะหาจากเส้นขอบของตัวอักษร โดยใช้การการเปลี่ยนแปลงของมุมองศาระหว่างจุดหนึ่งไปยังจุดหนึ่งในตัวอักษรนั้น ซึ่งจะทำให้ได้ส่วนโค้งของตัวอักษร และจำนวนส่วนโค้งของตัวอักษร รวมทั้งรูปแบบของส่วนโค้งด้วย

การหาส่วนหัวของตัวอักษร จะทำลักษณะเดียวกับการหาเส้นแสดงขอบของตัวอักษร แต่จะแตกต่างกันตรงที่ การหาส่วนหัวของตัวอักษรนั้น จะหาเส้นขอบภายในตัวอักษร โดยที่เส้นขอบนั้นจะต้องไม่เป็นเส้นเดียวกับเส้นแสดงขอบของตัวอักษร ซึ่งในการหาส่วนหัวของตัวอักษรนี้ จะต้องเริ่มหาจุดเริ่มต้นของส่วนหัวก่อนเช่นเดียวกันกับ การหาเส้นแสดงขอบของตัวอักษร เมื่อได้ส่วนหัวของตัวอักษร ก็จะทำให้ได้ จำนวนหัว ตำแหน่งของส่วนหัว และ ขนาดของตัวอักษร

จากการศึกษาพบว่าเทคนิคเส้นแสดงขอบของอักษร (Contour) มีความเหมาะสมที่จะนำมาใช้ในการวิเคราะห์อักษรไทยเนื่องจากอักษรไทยมีโครงสร้างที่ซับซ้อนประกอบด้วยส่วนโค้งเป็นส่วนใหญ่ ดังนั้นการวิจัยนี้จึงเลือกใช้เทคนิคเส้นแสดงขอบของอักษร (Contour) มาใช้ในการวิเคราะห์ โดยก่อนที่จะทำการวิเคราะห์นี้จะต้องกำหนดจุดเริ่มต้นของตัวอักษรก่อน

การหาเส้นแสดงขอบของตัวอักษรจะเริ่มจากจุดเริ่มต้นของตัวอักษรและจะทำการตรวจสอบจุดต่างๆต่อไปว่าจุดใดเป็นเส้นแสดงขอบของตัวอักษร โดยการตรวจสอบนี้จะตรวจสอบจากจุดเริ่มต้นไปทางซ้ายเรื่อยๆหรือในทิศทวนเข็มนาฬิกา และต้องไม่เป็นจุดของเส้นแสดงขอบของตัวอักษรเก่า หรือไม่เป็นจุดของเส้นแสดงขอบของส่วนหัวของตัวอักษร จนกระทั่งกลับมาที่จุดเริ่มต้นที่กำหนดไว้อีกครั้ง

ดังตัวอย่าง

8	7	6
1	X	5
2	3	4

หอสมุดกลาง สถาบันวิทยบริการ จุฬาลงกรณ์มหาวิทยาลัย

x คือ จุดที่เป็นจุดเส้นแสดงขอบของตัวอักษรเดิม

รูปที่ 3.27 แสดงเมตริกซ์การตรวจสอบจุดต่อไปของเส้นแสดงขอบตัวอักษร

เริ่มหาโดย เริ่มจากช่องที่ 1 กำหนดให้ n = 1 วนรอบจุด x ไปเรื่อยๆ จากช่องที่ 1 (n = 1) ไปยังช่องที่ 2 (n = 2) เมื่อพบจุดใหม่ จะถือว่า จุดนั้นคือจุดของเส้นแสดงขอบจุดใหม่ ดังนั้น x คือ จุดที่เป็นจุดเส้นแสดงขอบของตัวอักษรจุดใหม่ นั่นคือจุดในช่องที่ 2 เดิม หลังจากนั้น จะทำการหาจุดของเส้นแสดงขอบจุดต่อไป จะต้องหาค่า n ใหม่ก่อน โดยดูจากค่า n เดิม ดังนี้

ถ้า ค่า n เดิม น้อยกว่า หรือ เท่ากับ 3 ค่า n ใหม่ คือ
ค่า n ใหม่ เท่ากับ ค่า n เดิม + 5

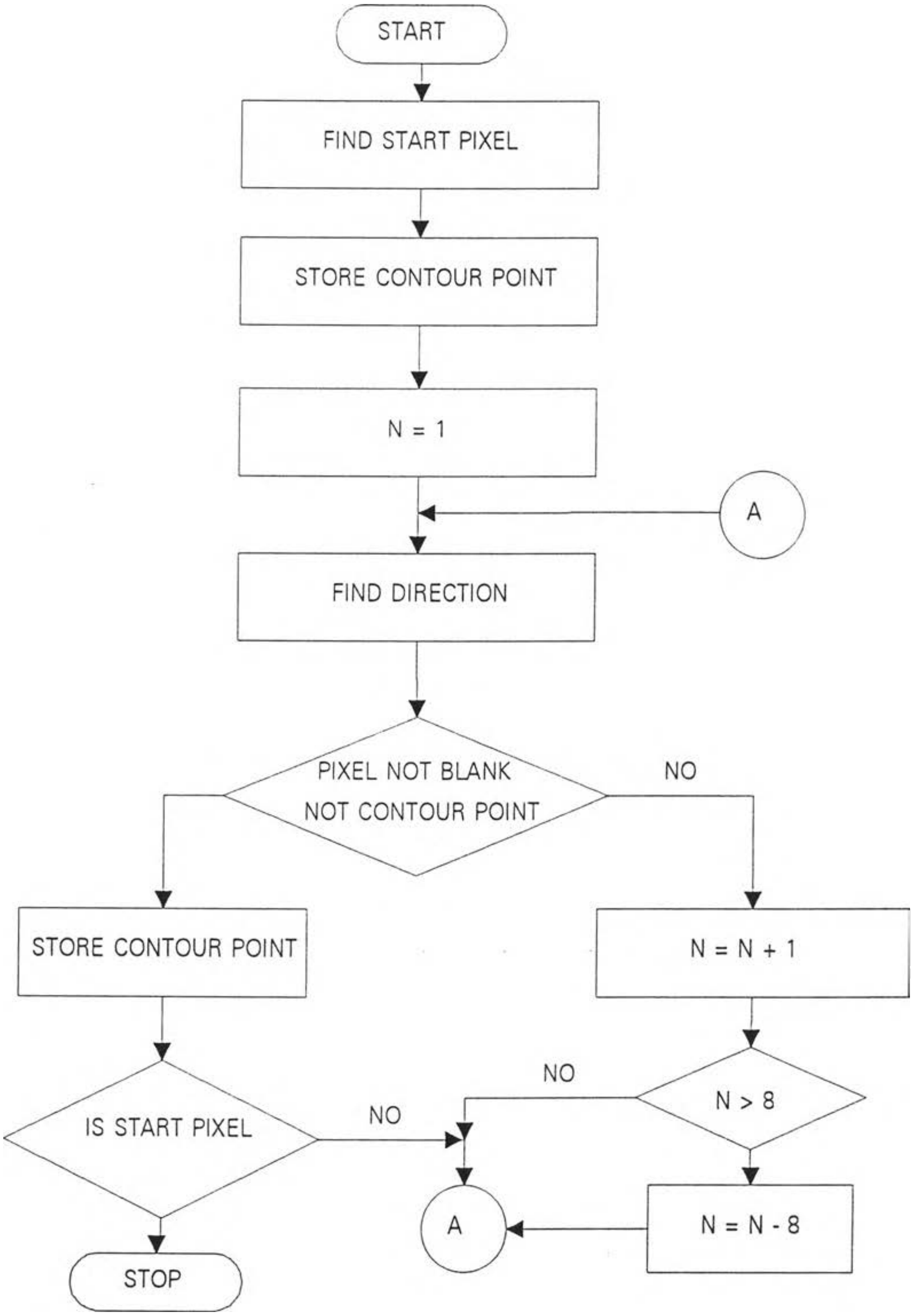
ถ้า ค่า n เดิม มากกว่า 3 ค่า n ใหม่ คือ
ค่า n ใหม่ เท่ากับ ค่า n เดิม - 5

จากตัวอย่างนี้ค่า n เดิมคือ 2 ซึ่งน้อยกว่า 3 ดังนั้น ค่า n ใหม่คือ 7 และการหาจุดของเส้นแสดงขอบของตัวอักษรจุดต่อไป จะต้องเริ่มที่ ช่องที่ 7 (n = 7) และหาไปเรื่อยๆ โดยวนรอบจุด x ไปยัง ช่องที่ 8 (n = 8) และไปยัง ช่องที่ 1 (n = 1) ไปเรื่อยๆ จนพบจุดใหม่ และจะหยุดหา จุดของเส้นแสดงขอบของตัวอักษร เมื่อกลับมายังจุดเริ่มต้น

```
จุดเริ่มต้น  ————————————————————11111111111
                                     1  111111   11
                                    1 11       11   11
                                   1 1          1   1
                                  1 1            1   1
                                 1 1              1  1
                                1   11            1  1
                                1     111111      1  1
                                 1     11111      1  1
                                  1   1           1  1
                                  1  1            1  1
                                  1  1            1  1
                                  1  1            1  1
                                  1  1            1  1
                                  1  1            1  1
                                  1  1            1  1
                                  1  1            1  1
                                  1  1            1  1
                                  1  1            1  1
                                  1111            1111
```

รูปที่ 3.28 แสดงการวิเคราะห์เส้นแสดงขอบของตัวอักษร

รูปที่ 3.29 ผังงานแสดงการวิเคราะห์เส้นขอบของตัวอักษร



3.1.5 กำหนดค่าลักษณะสำคัญของตัวอักษร

การกำหนดค่าลักษณะสำคัญของตัวอักษรนี้จะกำหนดเป็นค่าลักษณะสำคัญต้นแบบสำหรับตัวอักษรต้นแบบนั้น เพื่อใช้เปรียบเทียบต่อไป ซึ่งค่าลักษณะสำคัญเหล่านี้มีความสำคัญมาก

โดยค่าลักษณะสำคัญของตัวอักษรกำหนดไว้ดังนี้

3.1.5.1 รหัสตัวอักษร

รหัสตัวอักษรนี้จะต้องกำหนดเป็นรหัสแอสกี้ (ASCII CODE) เช่น ตัวอักษร ก รหัสตัวอักษรคือ 161

3.1.5.2 จำนวนหัวของตัวอักษร

กำหนดว่าตัวอักษรนั้นมีหัวหรือไม่ หรือมีจำนวนเท่าใด

เช่น ตัวอักษร ก มีจำนวนหัวเท่ากับ 0

ตัวอักษร ข มีจำนวนหัวเท่ากับ 1

ตัวอักษร น มีจำนวนหัวเท่ากับ 2

3.1.5.3 จำนวนส่วนโค้งของตัวอักษร

กำหนดว่าตัวอักษรนั้นมีส่วนโค้งจำนวนเท่าใดในแต่ละรูปแบบโดยมีส่วนโค้งทั้งหมด 4 รูปแบบ ซึ่งจะกล่าวถึงเทคนิคการหาส่วนโค้งภายหลัง

3.1.5.4 ความกว้างของตัวอักษร

ความกว้างของตัวอักษรคือขนาดของตัวอักษรตามแนวนอน

3.1.5.5 ความสูงของตัวอักษร

ความสูงของตัวอักษรคือขนาดของตัวอักษรตามแนวตั้ง

3.1.5.6 <u>ตำแหน่งของหัวของตัวอักษร</u>

ในการวิจัยนี้ส่วนหัวของตัวอักษรคือ ส่วนที่เป็นช่องว่างที่ถูกล้อมรอบด้วยจุดที่เป็นส่วนประกอบของตัวอักษร

ในการหาตำแหน่งของส่วนหัวของตัวอักษรจะกำหนดตำแหน่งของส่วนหัวไว้ 9 ตำแหน่ง โดยจะนำ จุดที่อยู่ทางขวาบนสุดของส่วนหัวนั้นมาหาว่าจุดนั้นอยู่ในตำแหน่งใดของตัวอักษร โดยดูจากตารางที่กำหนด

ตารางจะแบ่งตาม ความกว้างและความสูงของตัวอักษร โดยจะแบ่ง ความกว้าง ออกเป็น 3 ส่วน โดย 2 ส่วนแรกจะมีขนาดเท่ากัน ส่วนที่ 3 จะเป็นส่วนที่เหลือจาก 2 ส่วนแรก และแบ่ง ความสูง ออกเป็น 3 ส่วน โดย 2 ส่วนแรกจะมีขนาดเท่ากัน ส่วนที่ 3 จะเป็นส่วนที่ เหลือจาก 2 ส่วนแรก เช่นกัน ดังนั้นจะได้ตารางที่มี 9 ส่วน และจะดูว่า ส่วนหัวของตัวอักษรแต่ละหัวนั้นอยู่ที่ตำแหน่งใด ของตัวอักษร

เช่น ตัวอักษรตัว น จะพบว่ามีส่วนหัวจำนวน 2 หัว โดยหัวที่ 1 จะอยู่ข้างบน ส่วนหัวที่ 2 จะอยู่ข้างล่าง

กำหนดให้ค่า X คือ ตำแหน่งของส่วนหัวที่ 1 ซึ่งอยู่ในตำแหน่ง (4 , 7)

กำหนดให้ค่า Y คือ ตำแหน่งของส่วนหัวที่ 2 ซึ่งอยู่ในตำแหน่ง (17,24)

จากรูปจะพบว่า ตัวอักษร น มีความสูง = 22 ความกว้าง = 26 แบ่งเป็น 9 ส่วน จะพบว่า ส่วนกว้างแบ่งเป็น 3 ส่วน ที่ตำแหน่ง 1 , 9 , 17 , 26 ส่วนสูงแบ่งเป็น 3 ส่วน ที่ตำแหน่ง 1 , 7 , 14 , 22 ดังรูป

ดังนั้น ส่วนหัวที่ 1 ที่มีค่า X อยู่ที่ตำแหน่ง (4 , 7) จะมีตำแหน่ง = 1

ดังนั้น ส่วนหัวที่ 2 ที่มีค่า Y อยู่ที่ตำแหน่ง (17,24) จะมีตำแหน่ง = 9

```
        12345678901234567890123456
                +       +
1        111111            111
2       111  111          1111
3       11   1111         1111
4       11    X11         1111
5       11    111         1111
6       111   111         1111
7+       11111111         1111
8          111111         1111
9           11111         1111
0            1111         1111
1            1111         1111
2            1111        11111
3            1111      1111111
4+           1111    11111111111
5            1111  11111111111111
6            111111111   111  1111
7            11111111    11    Y11
8            111111      11    111
9            11111       11    111
0            11111        1111111
1            1111         1111111
2             11            1111
```

12345678901234567890123456

+ +

1	2	3
4	5	6
7	8	9

รูปที่ 3.30 แสดงการหาตำแหน่งของหัวของตัวอักษร

3.1.5.7 <u>ระยะห่างระหว่างหัวของตัวอักษร</u>

```
 111111          111
111  111        1111
11   1111       1111
11    X11       1111
11    111       1111
111   111       1111
 11111111       1111
  111111        1111
   11111        1111
    1111        1111
    1111        1111
    1111       11111
    1111     1111111
    1111   11111111111
    1111 11111111111111
    111111111   111  1111
    11111111    11    Y11
    111111      11    111
    11111       11    111
    11111        1111111
    1111         1111111
     11            1111
```

ในกรณีที่ตัวอักษรมีจำนวนหัวเท่ากับ 2 หัว จะหาระยะห่างระหว่างจุดที่อยู่ทางขวาบนสุดของส่วนหัวทั้ง 2 หัว

จากรูป ค่า X คือ จุดทางขวาสุดของส่วนหัวที่ 1

ค่า Y คือ จุดทางขวาสุดของส่วนหัวที่ 2

ดังนั้นระยะห่างระหว่างหัวของตัวอักษร คือ

ระยะห่างระหว่างจุด X กับ จุด Y

โดยเทคนิคนี้เป็นเทคนิคที่ผู้ทำวิจัยได้

นำระยะทางตามระหว่างจุดมาใช้

เพื่อกำหนดเป็นค่าตัวแปรเปรียบเทียบ

รูปที่ 3.31 แสดงการหาระยะห่างระหว่างหัวของตัวอักษร

ในการวิจัยนี้ ได้กำหนดรูปแบบของส่วนโค้งไว้ 4 รูปแบบ คือ

ส่วนโค้งรูปแบบที่ 1 คือ ส่วนโค้งที่มีค่ามุมองศาของจุดเริ่มต้น และ จุดสุดท้ายของส่วนโค้ง มีค่าเป็น บวก ทั้งคู่

ส่วนโค้งรูปแบบที่ 2 คือ ส่วนโค้งที่มีค่ามุมองศาของจุดเริ่มต้น และ จุดสุดท้ายของส่วนโค้ง มีค่าเป็น ลบ ทั้งคู่

ส่วนโค้งรูปแบบที่ 3 คือ ส่วนโค้งที่มีค่ามุมองศาของจุดเริ่มต้น มีค่า เป็น บวก และ จุดสุดท้ายของส่วนโค้ง มีค่าเป็น ลบ

ส่วนโค้งรูปแบบที่ 4 คือ ส่วนโค้งที่มีค่ามุมองศาของจุดเริ่มต้น มีค่า เป็น ลบ และ จุดสุดท้ายของส่วนโค้ง มีค่าเป็น บวก

ซึ่งจะกล่าวถึงรายละเอียดภายหลัง

3.1.5.8 <u>จำนวนส่วนโค้งทั้งหมด</u>

กำหนดว่าตัวอักษรมีจำนวนส่วนโค้งทั้งหมด เท่าใด

3.1.5.9 <u>จำนวนส่วนโค้งรูปแบบที่ 1</u>

กำหนดว่าตัวอักษรมีจำนวนส่วนโค้งรูปแบบที่ 1 เท่าใด

3.1.5.10 <u>จำนวนส่วนโค้งรูปแบบที่ 2</u>

กำหนดว่าตัวอักษรมีจำนวนส่วนโค้งรูปแบบที่ 2 เท่าใด

3.1.5.11 <u>จำนวนส่วนโค้งรูปแบบที่ 3</u>

กำหนดว่าตัวอักษรมีจำนวนส่วนโค้งรูปแบบที่ 3 เท่าใด

3.1.5.12 <u>จำนวนส่วนโค้งรูปแบบที่ 4</u>

กำหนดว่าตัวอักษรมีจำนวนส่วนโค้งรูปแบบที่ 4 เท่าใด

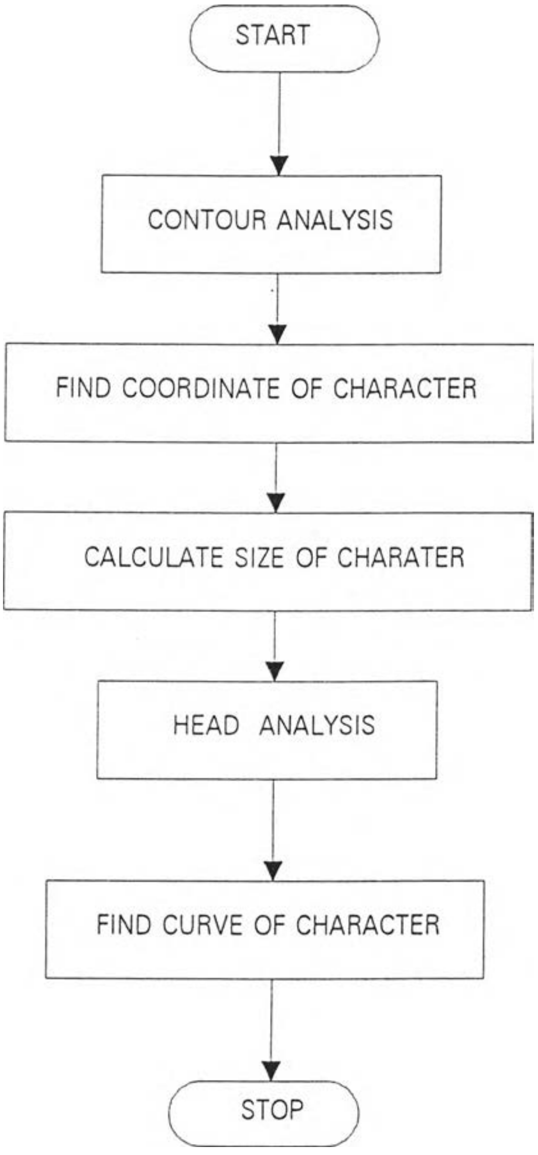
3.1.5.13 <u>จำนวนเส้นตรงตามแนวนอน</u>

กำหนดว่าตัวอักษรมีจำนวนเส้นตรงตามแนวนอนเท่าใด

3.1.5.14 <u>จำนวนเส้นตรงตามแนวตั้ง</u>

กำหนดว่าตัวอักษรมีจำนวนเส้นตรงตามแนวตั้งเท่าใด

รูปที่ 3.32 ผังงานแสดงการกำหนดค่าลักษณะสำคัญ



3.1.6 กำหนดกลุ่มของตัวอักษรต้นแบบ

การกำหนดกลุ่มของตัวอักษร เป็นการกำหนดว่าตัวอักษรนั้นๆควรจะอยู่ในกลุ่มตัวอักษรต้นแบบใด

โดยการวิจัยได้กำหนดกลุ่มของตัวอักษรต้นแบบ ดังนี้

กลุ่มที่ 1

กลุ่มตัวอักษรที่ไม่มีหัว
อัตราส่วนระหว่างความสูงต่อความกว้างของตัวอักษร
น้อยกว่า 1 หรือ เท่ากับ 1

กลุ่มที่ 2

กลุ่มตัวอักษรที่ไม่มีหัว
อัตราส่วนระหว่างความสูงต่อความกว้างของตัวอักษร
น้อยกว่า 1.5 หรือ เท่ากับ 1.5 แต่ มากกว่า 1

กลุ่มที่ 3

กลุ่มตัวอักษรที่ไม่มีหัว
อัตราส่วนระหว่างความสูงต่อความกว้างของตัวอักษร
น้อยกว่า 2 หรือ เท่ากับ 2 แต่ มากกว่า 1.5

กลุ่มที่ 4

กลุ่มตัวอักษรที่ไม่มีหัว
อัตราส่วนระหว่างความสูงต่อความกว้างของตัวอักษร
มากกว่า 2

กลุ่มที่ 5

กลุ่มตัวอักษรที่มีหัว เพียงหัวเดียว
อัตราส่วนระหว่างความสูงต่อความกว้างของตัวอักษร
น้อยกว่า 1 หรือ เท่ากับ 1

[library stamp]

กลุ่มที่ 6

กลุ่มตัวอักษรที่มีหัว เพียงหัวเดียว
อัตราส่วนระหว่างความสูงต่อความกว้างของตัวอักษร
น้อยกว่า 1.5 หรือ เท่ากับ 1.5 แต่ มากกว่า 1

กลุ่มที่ 7

กลุ่มตัวอักษรที่มีหัว เพียงหัวเดียว
อัตราส่วนระหว่างความสูงต่อความกว้างของตัวอักษร
น้อยกว่า 2 หรือ เท่ากับ 2 แต่ มากกว่า 1.5

กลุ่มที่ 8

กลุ่มตัวอักษรที่มีหัว เพียงหัวเดียว
อัตราส่วนระหว่างความสูงต่อความกว้างของตัวอักษร
มากกว่า 2

กลุ่มที่ 9

กลุ่มตัวอักษรที่มีหัว 2 หัว
อัตราส่วนระหว่างความสูงต่อความกว้างของตัวอักษร
น้อยกว่า 1 หรือ เท่ากับ 1

กลุ่มที่ 10

กลุ่มตัวอักษรที่มีหัว 2 หัว
อัตราส่วนระหว่างความสูงต่อความกว้างของตัวอักษร
น้อยกว่า 1.5 หรือ เท่ากับ 1.5 แต่ มากกว่า 1

กลุ่มที่ 11

กลุ่มตัวอักษรที่มีหัว 2 หัว
อัตราส่วนระหว่างความสูงต่อความกว้างของตัวอักษร
น้อยกว่า 2 หรือ เท่ากับ 2 แต่ มากกว่า 1.5

<u>กลุ่มที่ 12</u>

กลุ่มตัวอักษรที่มีหัว 2 หัว
อัตราส่วนระหว่างความสูงต่อความกว้างของตัวอักษร
มากกว่า 2

<u>กลุ่มที่ 13</u>

กลุ่มตัวอักษรที่มีหัว 3 หัว
อัตราส่วนระหว่างความสูงต่อความกว้างของตัวอักษร
น้อยกว่า 1 หรือ เท่ากับ 1

<u>กลุ่มที่ 14</u>

กลุ่มตัวอักษรที่มีหัว 3 หัว
อัตราส่วนระหว่างความสูงต่อความกว้างของตัวอักษร
น้อยกว่า 1.5 หรือ เท่ากับ 1.5 แต่ มากกว่า 1

<u>กลุ่มที่ 15</u>

กลุ่มตัวอักษรที่มีหัว 3 หัว
อัตราส่วนระหว่างความสูงต่อความกว้างของตัวอักษร
น้อยกว่า 2 หรือ เท่ากับ 2 แต่ มากกว่า 1.5

<u>กลุ่มที่ 16</u>

กลุ่มตัวอักษรที่มีหัว 3 หัว
อัตราส่วนระหว่างความสูงต่อความกว้างของตัวอักษร
มากกว่า 2

<u>กลุ่มที่ 17</u>

กลุ่มตัวอักษรที่มีหัว 4 หัว
อัตราส่วนระหว่างความสูงต่อความกว้างของตัวอักษร
น้อยกว่า 1 หรือ เท่ากับ 1

<u>กลุ่มที่ 18</u>

กลุ่มตัวอักษรที่มีหัว 4 หัว
อัตราส่วนระหว่างความสูงต่อความกว้างของตัวอักษร
น้อยกว่า 1.5 หรือ เท่ากับ 1.5 แต่ มากกว่า 1

<u>กลุ่มที่ 19</u>

กลุ่มตัวอักษรที่มีหัว 4 หัว
อัตราส่วนระหว่างความสูงต่อความกว้างของตัวอักษร
น้อยกว่า 2 หรือ เท่ากับ 2 แต่ มากกว่า 1.5

<u>กลุ่มที่ 20</u>

กลุ่มตัวอักษรที่มีหัว 4 หัว
อัตราส่วนระหว่างความสูงต่อความกว้างของตัวอักษร
มากกว่า 2

3.1.7 <u>เก็บข้อมูลไว้ในแฟ้มข้อมูลตัวอักษรต้นแบบ</u>

เมื่อได้ข้อมูลต่างๆเรียบร้อยแล้วจะนำข้อมูลนั้นไปเก็บไว้ในแฟ้มข้อมูลตัวอักษรต้นแบบ เพื่อใช้เปรียบเทียบต่อไป

แฟ้มตัวอักษรต้นแบบประกอบด้วย

ASCII	รหัสตัวอักษร
HIGH	ความสูงของตัวอักษร
WIDE	ความกว้างของตัวอักษร
HIGH/WIDE	อัตราส่วนของความสูงต่อความกว้าง
LINE X	จำนวนเส้นตรงตามแนวนอน
LINE Y	จำนวนเส้นตรงตามแนวตั้ง
CURVE	จำนวนส่วนโค้ง
CURVE 1	จำนวนส่วนโค้งรูปแบบที่ 1
CURVE 2	จำนวนส่วนโค้งรูปแบบที่ 2
CURVE 3	จำนวนส่วนโค้งรูปแบบที่ 3
CURVE 4	จำนวนส่วนโค้งรูปแบบที่ 4
HEAD	จำนวนหัวทั้งหมด
LOC 1	ตำแหน่งของหัวที่ 1
DIS 1	ระยะห่างระหว่างหัวที่ 1 กับหัวที่ 1
LOC 2	ตำแหน่งของหัวที่ 2
DIS 2	ระยะห่างระหว่างหัวที่ 2 กับหัวที่ 1
LOC 3	ตำแหน่งของหัวที่ 3
DIS 3	ระยะห่างระหว่างหัวที่ 3 กับหัวที่ 1
LOC 4	ตำแหน่งของหัวที่ 4
DIS 4	ระยะห่างระหว่างหัวที่ 4 กับหัวที่ 1

แฟ้มข้อมูลของตัวอักษรต้นแบบที่ไม่มีหัว

ASCII	HIGH	WIDE	LINE X	LINE Y	CURVE	CURVE 1	CURVE 2	....

......	CURVE 3	CURVE 4	HEAD	HIGH/WIDE

แฟ้มข้อมูลของตัวอักษรต้นแบบที่มีหัว 1 หัว

ASCII	HIGH	WIDE	LINE X	LINE Y	CURVE	CURVE 1	CURVE 2	....

......	CURVE 3	CURVE 4	HEAD	LOC 1	DIS 1	HIGH/WIDE

แฟ้มข้อมูลของตัวอักษรต้นแบบที่มีหัว 2 หัว

ASCII	HIGH	WIDE	LINE X	LINE Y	CURVE	CURVE 1	CURVE 2	....

......	CURVE 3	CURVE 4	HEAD	LOC 1	DIS 1	LOC 2	DIS 2	HIGH/WIDE

แฟ้มข้อมูลของตัวอักษรต้นแบบที่มีหัว 3 หัว

ASCII	HIGH	WIDE	LINE X	LINE Y	CURVE	CURVE 1	CURVE 2	CURVE 3	....

......	CURVE 4	HEAD	LOC 1	DIS 1	LOC 2	DIS 2	LOC 3	DIS 3	HIGH/WIDE

แฟ้มข้อมูลของตัวอักษรต้นแบบที่มีหัว 4 หัว

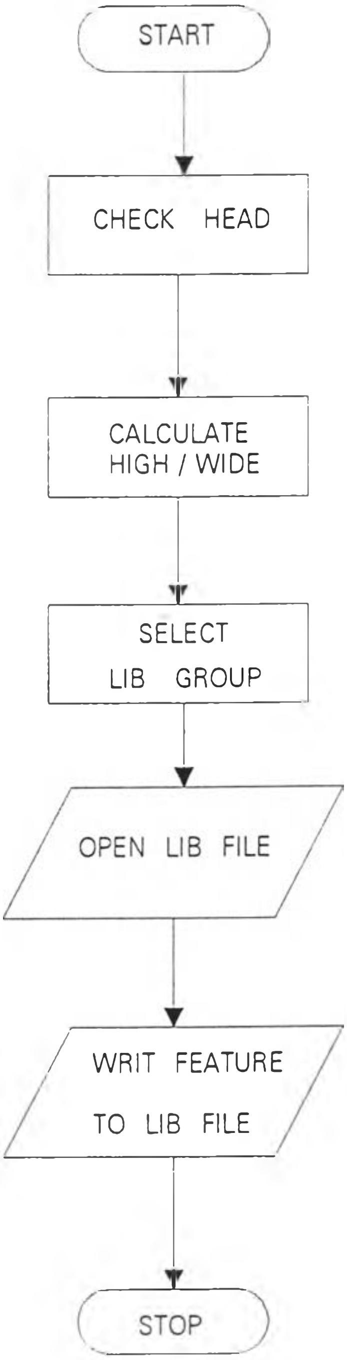
ASCII	HIGH	WIDE	LINE X	LINE Y	CURVE	CURVE 1	CURVE 2	....

......	CURVE 3	CURVE 4	HEAD	LOC 1	DIS 1	LOC 2	DIS 2	....

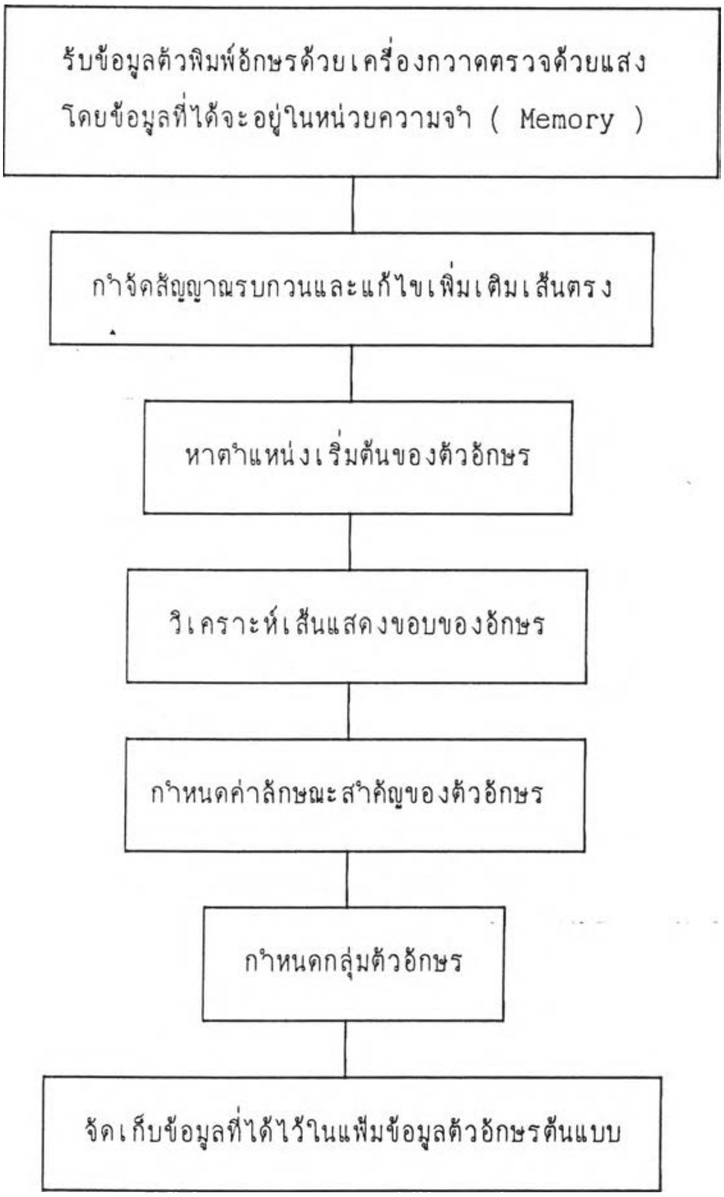
.....	LOC 3	DIS 3	LOC 4	DIS 4	HIGH/WIDE

รูปที่ 3.33 แสดงการเก็บแฟ้มข้อมูลทั้ง 5 รูปแบบ

รูปที่ 3.34 ผังงานแสดงการเก็บข้อมูลไว้ในแฟ้มข้อมูลตัวอักษรต้นแบบ



รูปที่ 3.35 แสดงขั้นตอนการทำงานการเรียนรู้ตัวอักษร



3.2 <u>ขั้นตอนการรู้จำ</u> (Recognize)

ขั้นตอนการรู้จำ เป็นขั้นตอนในการรับข้อมูลตัวอักษรที่จะทำการรู้จำ แล้วนำมาหาค่าตัวแปรต่างๆ มาเปรียบเทียบกับตัวอักษรต้นแบบ เพื่อหาว่าตัวอักษรนั้น คือตัวใด เพื่อนำไปสร้างเป็นแฟ้มข้อมูล เพื่อนำไปใช้งานต่อไป

โดยมีขั้นตอนการทำงานดังนี้

3.2.1 <u>รับข้อมูลตัวพิมพ์อักษรไทยและตัวพิมพ์อักษรอังกฤษ</u>

การรับข้อมูลตัวพิมพ์อักษรไทยและตัวพิมพ์อักษรอังกฤษนี้ จะใช้เครื่องกวาดตรวจด้วยแสง (Optical Scanner) โดยข้อมูลที่ได้จะอยู่ในหน่วยความจำ โดยเครื่องกวาดตรวจด้วยแสง จะรับข้อมูลเข้ามาแสดงผลที่หน้าจอ ซึ่งอยู่ในรูปของหน่วยความจำชนิดหนึ่งนั้น คือหน่วยความจำของจอภาพ (VIDEO RAM) หลังจากนั้นจะนำข้อมูลจากหน่วยความจำของจอภาพมาเก็บไว้ใน หน่วยความจำของข้อมูล (DATA MEMORY) ซึ่งเป็นค่าตัวแปร เพื่อนำไปใช้ต่อไป เช่นเดียวกับขั้นตอนการเรียนรู้

3.2.2 <u>กำจัดสัญญาณรบกวนและแก้ไขเพิ่มเติมเส้นตรง</u>

การกำจัดสัญญาณรบกวนและการแก้ไขเพิ่มเติมเส้นตรงจะเป็นเทคนิคที่ทำการกำจัดส่วนเกินของจุดต่างของข้อมูลนำเข้า และเพิ่มเติมส่วนที่เป็นเส้นตรงให้ครบถ้วนสมบูรณ์ ซึ่งจะมีผลต่อการทำงานในขั้นตอนต่อไป เพื่อให้การทำงานในขั้นตอนต่อไปดีขึ้น เช่นเดียวกับขั้นตอนการเรียนรู้

3.2.3 <u>หาจุดเริ่มต้นของตัวอักษร</u>

เป็นการตรวจสอบจุดต่างๆที่ประกอบขึ้นเป็นตัวอักษรเพื่อหาว่าจุดใด คือ จุดแรกที่พบ ซึ่งจุดนั้นต้องไม่เป็นช่องว่าง และไม่ใช่จุดที่เป็นเส้นแสดงขอบ หรือเป็นจุดที่อยู่ภายในตัวอักษรตัวอื่น หรือ จุดที่ประกอบเป็นตัวอักษรตัวอื่น โดยการหาจุดเริ่มต้นของตัวอักษรนั้น จะเริ่มที่จุดที่อยู่ในสดมภ์แรกของแถวแรก หรือจุดที่อยู่ใน สดมภ์ที่ 0 แถวที่ 0 ส่วนการหาจุดเริ่มต้นของตัวอักษรต่อไป จะเริ่มที่ จุดเริ่มต้นของตัวอักษรก่อนหน้า และจะหาไปเรื่อยๆจนพบจุดเริ่มต้นของตัวอักษรต่อไป เช่นเดียวกับขั้นตอนการเรียนรู้

3.2.4 วิเคราะห์เส้นแสดงขอบของตัวอักษร

การหาเส้นแสดงขอบของตัวอักษรจะเริ่มจากจุดเริ่มต้นของตัวอักษรและจะทำการตรวจสอบจุดต่างๆต่อไปว่าจุดใดเป็นเส้นแสดงขอบของตัวอักษร โดยการตรวจสอบนี้จะตรวจสอบจากจุดเริ่มต้นไปทางซ้ายเรื่อยๆหรือในทิศทวนเข็มนาฬิกา และต้องไม่เป็นจุดของเส้นแสดงขอบของตัวอักษรเก่า หรือไม่เป็นจุดของเส้นแสดงขอบของส่วนหัวของตัวอักษร จนกระทั่งกลับมาที่จุดเริ่มต้นที่กำหนดไว้อีกครั้ง เช่นเดียวกับขั้นตอนการเรียนรู้

3.2.5 กำหนดค่าลักษณะสำคัญของตัวอักษร

การกำหนดค่าลักษณะสำคัญนี้จะกำหนดขึ้นเพื่อใช้เปรียบเทียบ กับค่าลักษณะสำคัญต้นแบบเพื่อใช้ในการรู้จำ

3.2.6 กำหนดกลุ่มตัวอักษรต้นแบบ

เมื่อกำหนดตัวแปรต่างๆเรียบร้อยแล้ว จากนั้นจะทำการกำหนดกลุ่มของตัวอักษรว่าตัวอักษรนั้นอยู่กลุ่มใด เช่นกลุ่มตัวอักษรที่ไม่มีหัว กลุ่มตัวอักษรที่มีหัวเดียว กลุ่มตัวอักษรที่มีสองหัว และ อัตราส่วนระหว่างความสูงของตัวอักษรต่อความกว้างของตัวอักษร เป็นต้น ซึ่งการกำหนดกลุ่มของตัวอักษรนี้ จะทำให้ตัวอักษรต้นแบบที่จะนำมาเปรียบเทียบการรู้จำมีจำนวนน้อยลง โดยตัวอักษรต้นแบบนี้ต้องได้ทำการเรียนรู้ไว้ก่อนแล้วซึ่งจะมีผลทำให้การเปรียบเทียบเร็วขึ้น

3.2.7 เปรียบเทียบกับตัวอักษรต้นแบบในกลุ่มตัวอักษร

เมื่อกำหนดกลุ่มตัวอักษรต้นแบบ และค่าลักษณะสำคัญที่ใช้เปรียบเทียบเรียบร้อยแล้ว จะนำมาทำการเปรียบเทียบ ในการวิจัยนี้ได้ศึกษาวิธีการเปรียบเทียบดังนี้

3.2.7.1 ไดนามิคโปรแกรมมิ่ง (Dynamic Programming) (Sakoe H. and Chiba S., 1978)

ไดนามิคโปรแกรมมิ่ง คือว่าเป็นการเปรียบเทียบในรูปแบบที่มีการปรับตัวยืดหยุ่น ที่สามารถนำมาใช้ในการหาค่าความแตกต่างที่น้อยที่สุดระหว่างข้อมูลได้ ใน

การเปรียบเทียบแต่ละครั้งจะต้องทำงานวนซ้ำหลายรอบ เพื่อให้ได้ค่าที่ใกล้เคียงมากที่สุด ในรูปแบบที่มีการปรับตัวยืดหยุ่นเหมาะกับชุดข้อมูลที่มีค่าที่ใช้ในการเปรียบเทียบไม่เท่ากัน และมีการเปลี่ยนแปลงจำนวนค่าที่ใช้เปรียบเทียบในข้อมูลแต่ละชุดเสมอ

3.2.7.2 ฟังก์ชันการหาค่าผลรวมความแตกต่างสัมบูรณ์
(Absolute Difference Summation Function)
(Frederick S. Hillier and Gerald J. Lieberman, 1991)

ฟังก์ชันใน การหาค่าผลรวมความแตกต่างสัมบูรณ์ระหว่างค่าสองค่า เป็นการเปรียบเทียบเพื่อหาค่าความแตกต่างที่น้อยที่สุด โดยจะหาค่าสัมบูรณ์ของผลต่างระหว่างค่าเปรียบเทียบของข้อมูล 2 ชุด และนำค่าสัมบูรณ์ที่ได้มาคูณกับค่าน้ำหนักของค่าเปรียบเทียบแต่ละค่าและนำมาหาผลรวม เพื่อนำไปเปรียบเทียบหาค่าที่ใกล้เคียงที่สุด ในการเปรียบเทียบนี้เหมาะกับชุดข้อมูลที่มีค่าที่ใช้ในการเปรียบเทียบเท่ากัน และจะทำงานได้อย่างรวดเร็วเนื่องจากไม่มีการทำงานแบบวนซ้ำ

ในการวิจัยนี้ต้องการความเร็วในการทำงานอย่างมาก รวมถึงจำนวนค่าที่ใช้ในการเปรียบเทียบของข้อมูลในแต่ละชุดมีจำนวนคงที่เสมอ ดังนั้น ไดนามิคโปรแกรมมิ่งจึงไม่เหมาะกับงานวิจัยนี้ เนื่องจากไดนามิคโปรแกรมมิ่งมีการทำงานวนซ้ำหลายรอบ ซึ่งจะทำให้เวลาที่ใช้ในการทำงานแต่ละครั้งเสียเวลามากคือประมาณ 40.25 วินาที ด้วยเครื่องคอมพิวเตอร์ 80486 DX และจำนวนค่าเปรียบเทียบมีการเปลี่ยนแปลงอยู่ตลอดเวลา ส่วนฟังก์ชันการหาค่าผลรวมความแตกต่างสัมบูรณ์ จะใช้เวลาในการทำงานน้อยคือประมาณ 10.26 วินาที ด้วยเครื่องคอมพิวเตอร์ 80486 DX และ มีจำนวนค่าเปรียบเทียบคงที่เสมอ และผลที่ได้จากการวิจัยจะให้ผลความถูกต้องที่ใกล้เคียงกันมาก ดังนั้นในการวิจัยนี้จึงใช้วิธีการเปรียบเทียบแบบฟังก์ชันการหาค่าผลรวมความแตกต่างสัมบูรณ์ มาใช้ในการเปรียบเทียบหาค่าที่ใกล้เคียงที่สุด

เทคนิคการหาค่าความแตกต่างของค่าลักษณะสำคัญระหว่างตัวอักษร ที่จะทำการรู้จำและตัวอักษรต้นแบบที่มีค่าน้อยที่สุดเพื่อให้ได้ข้อมูลผลลัพธ์ที่ต้องการ

Frederick S. Hillier และ Gerald J. Lieberman ได้กล่าวถึงฟังก์ชันในการหาค่าผลรวมความแตกต่างสัมบูรณ์ระหว่างค่าสองค่า ดังนี้

$$Z = \sum_{i=1}^{n} W_i X_i$$

Z คือ ค่าผลรวมความแตกต่างระหว่างค่า 2 ค่า

W_i คือ ค่าน้ำหนัก

X_i คือ ค่าความแตกต่างของส่วนประกอบของข้อมูลทั้ง 2

กำหนดให้

$X_i = | X1_i - X2_i |$

$X1_i$ คือ ค่าของส่วนประกอบของข้อมูลชุดที่ 1

$X2_i$ คือ ค่าของส่วนประกอบของข้อมูลชุดที่ 2

$X1_i = X1_1 , X1_2 , X1_3 , ... , X1_n$

$X2_i = X2_1 , X2_2 , X2_3 , ... , X2_n$

n คือ จำนวนส่วนประกอบของข้อมูล

ค่าน้ำหนักนั้นได้จากการกำหนดค่าตามลำดับความสำคัญของส่วนประกอบต่างๆ ส่วนประกอบใดมีความสำคัญต่อการหาค่าความแตกต่างของข้อมูล ถ้าส่วนประกอบที่มีความสำคัญมากที่สุด จะกำหนดให้ค่าน้ำหนักของส่วนประกอบนั้นมีค่ามากที่สุด และส่วนประกอบที่มีความสำคัญรองลงมา จะกำหนดให้มีค่าน้ำหนักลดลงมาตามลำดับ ในการหาค่าน้ำหนักที่เหมาะสมนั้นจะได้จากการปรับเปลี่ยนค่าน้ำหนัก และนำมาตรวจสอบหาเปอร์เซนต์ความถูกต้องในการเปรียบเทียบ ถ้าพบว่าค่าเปอร์เซนต์ที่หาได้ มีค่ามากกว่าหรือเท่ากับ 90 เปอร์เซนต์ ก็จะถือว่าค่าน้ำหนักที่กำหนดนั้น สามารถนำมาใช้ในการหาค่าผลรวมความแตกต่างได้ (Frederick S., 1990)

การวิจัยนี้ได้นำฟังก์ชันนี้มาใช้ใน การหาค่าความแตกต่างของค่าลักษณะสำคัญระหว่างตัวอักษรที่จะทำการรู้จำและตัวอักษรต้นแบบ เพื่อนำมาเปรียบเทียบหาค่าที่มีความแตกต่างกันน้อยที่สุด

จากฟังก์ชัน

$$Z = \sum_{i=1}^{n} W_i X_i$$

กำหนดให้ $X_i = |X1_i - X2_i|$

ดังนั้น

$$Z = \sum_{i=1}^{n} W_i . |X1_i - X2_i|$$

Z คือ ค่าความแตกต่างระหว่างตัวอักษรต้นแบบกับตัวอักษรที่ทำการรู้จำ

W_i คือ ค่าน้ำหนักของลักษณะสำคัญตัวที่ i ของตัวอักษร

$X1_i$ คือ ค่าลักษณะสำคัญตัวที่ i ของตัวอักษรต้นแบบ

$X2_i$ คือ ค่าลักษณะสำคัญตัวที่ i ของตัวอักษรที่ทำการรู้จำ

n คือ จำนวนลักษณะสำคัญของตัวอักษร ขึ้นอยู่กลุ่มของตัวอักษร

กรณีกลุ่มตัวอักษรที่ไม่มีจำนวนหัว	n = 11
กรณีกลุ่มตัวอักษรที่มีจำนวนหัวเท่ากับ 1 หัว	n = 13
กรณีกลุ่มตัวอักษรที่มีจำนวนหัวเท่ากับ 2 หัว	n = 15
กรณีกลุ่มตัวอักษรที่มีจำนวนหัวเท่ากับ 3 หัว	n = 17
กรณีกลุ่มตัวอักษรที่มีจำนวนหัวเท่ากับ 4 หัว	n = 19

$X1_i = X1_1 , X1_2 , X1_3 , ... , X1_n$

$X2_i = X2_1 , X2_2 , X2_3 , ... , X2_n$

เช่น $X1_1$, $X2_1$ คือ ความกว้างของตัวอักษร

$X6_6$, $X2_6$ คือ จำนวนส่วนโค้งรูปแบบที่ 1 ของตัวอักษร

$X1_{10}$, $X2_{10}$ คือ จำนวนหัวของตัวอักษร

$X1_{11}$, $X2_{11}$ คือ ตำแหน่งของหัวที่ 1 ของตัวอักษร

$X1_{14}$, $X2_{14}$ คือ ระยะห่างระหว่างหัวที่ 1 กับหัวที่ 2

การกำหนดค่าน้ำหนักของลักษณะสำคัญ ได้จากการกำหนดค่าแล้วตรวจสอบหาเปอร์เซนต์ความถูกต้องในการเปรียบเทียบ โดยกำหนดว่าเปอร์เซนต์ความถูกต้องต้องมีค่าตั้งแต่ 90 เปอร์เซนต์ขึ้นไป

โดยในขั้นแรกจะกำหนดค่าน้ำหนักของลักษณะสำคัญของตัวอักษรมีค่าเท่ากับ 1 ทุกค่า พบว่า ในการรู้จำตัวอักษร 314 ตัวอักษร สามารถรู้จำได้ถูกต้อง 235 ตัวอักษร หรือสามารถรู้จำได้ถูกต้อง 74.84 เปอร์เซนต์

จะเห็นได้ว่าผลที่ได้มีเปอร์เซนต์ความถูกต้องน้อยมาก จึงต้องทำการปรับปรุงค่าน้ำหนักของลักษณะสำคัญใหม่ โดยจะกำหนดให้ส่วนหัวของตัวอักษรมีความสำคัญมากที่สุด จำนวนส่วนโค้งรูปแบบต่างๆของตัวอักษร และจำนวนเส้นตรงของตัวอักษร มีความสำคัญรองลงมา และลักษณะสำคัญอื่นๆ มีความสำคัญรองลงมา เนื่องจากในการวิจัยนี้ได้จัดกลุ่มตัวอักษรต้นแบบแยกตามจำนวนหัวของตัวอักษรและอัตราส่วนระหว่างความสูงต่อความกว้างของตัวอักษร ดังนั้น ค่าน้ำหนักของจำนวนหัวของตัวอักษร และค่าน้ำหนักของอัตราส่วนระหว่างความสูงต่อความกว้างของตัวอักษร จึงไม่มีความจำเป็นที่จะต้องใช้ เนื่องจากมีค่าที่เท่ากัน จึงไม่มีผลต่อการหาค่าความแตกต่างระหว่างตัวอักษร

ดังนั้นในการปรับปรุงค่าน้ำหนักของลักษณะสำคัญของตัวอักษรใหม่ ดังนี้

ค่าน้ำหนักของจำนวนเส้นตรงตามแนวนอน	=	2
ค่าน้ำหนักของจำนวนเส้นตรงตามแนวตั้ง	=	2
ค่าน้ำหนักของจำนวนส่วนโค้งรูปแบบที่ 1	=	1.5
ค่าน้ำหนักของจำนวนส่วนโค้งรูปแบบที่ 2	=	1.5
ค่าน้ำหนักของจำนวนส่วนโค้งรูปแบบที่ 3	=	1.5
ค่าน้ำหนักของจำนวนส่วนโค้งรูปแบบที่ 4	=	1.5
ค่าน้ำหนักของตำแหน่งของหัว	=	8
ค่าน้ำหนักของลักษณะสำคัญอื่นๆ	=	1

พบว่า ในการรู้จำตัวอักษร 290 ตัวอักษร สามารถรู้จำได้ถูกต้อง 231 ตัวอักษร หรือสามารถรู้จำได้ถูกต้อง 79.65 เปอร์เซนต์

ในการปรับปรุงค่าน้ำหนักที่จะนำมาใช้ จะต้องเพิ่มค่าหรือลดค่าเป็นจำนวนเท่าของเดิม คือ 0.5 เท่า 1 เท่า 1.5 เท่า 2 เท่า

หลังจากนั้นได้ทำการปรับปรุงค่าน้ำหนักของลักษณะสำคัญของตัวอักษรใหม่ดังนี้

ค่าน้ำหนักของจำนวนเส้นตรงตามแนวนอน	=	2
ค่าน้ำหนักของจำนวนเส้นตรงตามแนวตั้ง	=	2
ค่าน้ำหนักของจำนวนส่วนโค้งรูปแบบที่ 1	=	3
ค่าน้ำหนักของจำนวนส่วนโค้งรูปแบบที่ 2	=	3
ค่าน้ำหนักของจำนวนส่วนโค้งรูปแบบที่ 3	=	3
ค่าน้ำหนักของจำนวนส่วนโค้งรูปแบบที่ 4	=	3
ค่าน้ำหนักของตำแหน่งของหัว	=	16
ค่าน้ำหนักของลักษณะสำคัญอื่นๆ	=	1

พบว่า ในการรู้จำตัวอักษร 495 ตัวอักษร สามารถรู้จำได้ถูกต้อง 421 ตัวอักษร หรือสามารถรู้จำได้ถูกต้อง 85.05 เปอร์เซนต์

หลังจากนั้นได้ทำการปรับปรุงค่าน้ำหนักของลักษณะสำคัญของตัวอักษรใหม่ดังนี้

ค่าน้ำหนักของจำนวนเส้นตรงตามแนวนอน	=	4
ค่าน้ำหนักของจำนวนเส้นตรงตามแนวตั้ง	=	4
ค่าน้ำหนักของจำนวนส่วนโค้งรูปแบบที่ 1	=	6
ค่าน้ำหนักของจำนวนส่วนโค้งรูปแบบที่ 2	=	6
ค่าน้ำหนักของจำนวนส่วนโค้งรูปแบบที่ 3	=	3
ค่าน้ำหนักของจำนวนส่วนโค้งรูปแบบที่ 4	=	3
ค่าน้ำหนักของตำแหน่งของหัว	=	16
ค่าน้ำหนักของลักษณะสำคัญอื่นๆ	=	1

พบว่า ในการรู้จำตัวอักษร 491 ตัวอักษร สามารถรู้จำได้ถูกต้อง 443 ตัวอักษร หรือสามารถรู้จำได้ถูกต้อง 90.23 เปอร์เซนต์

หลังจากนั้นได้ทำการปรับปรุงค่าน้ำหนักของลักษณะสำคัญของตัวอักษรใหม่ดังนี้

ค่าน้ำหนักของจำนวนเส้นตรงตามแนวนอน	=	5
ค่าน้ำหนักของจำนวนเส้นตรงตามแนวตั้ง	=	5
ค่าน้ำหนักของจำนวนส่วนโค้งรูปแบบที่ 1	=	9
ค่าน้ำหนักของจำนวนส่วนโค้งรูปแบบที่ 2	=	9
ค่าน้ำหนักของจำนวนส่วนโค้งรูปแบบที่ 3	=	6
ค่าน้ำหนักของจำนวนส่วนโค้งรูปแบบที่ 4	=	6
ค่าน้ำหนักของตำแหน่งของหัว	=	20
ค่าน้ำหนักของลักษณะสำคัญอื่นๆ	=	1

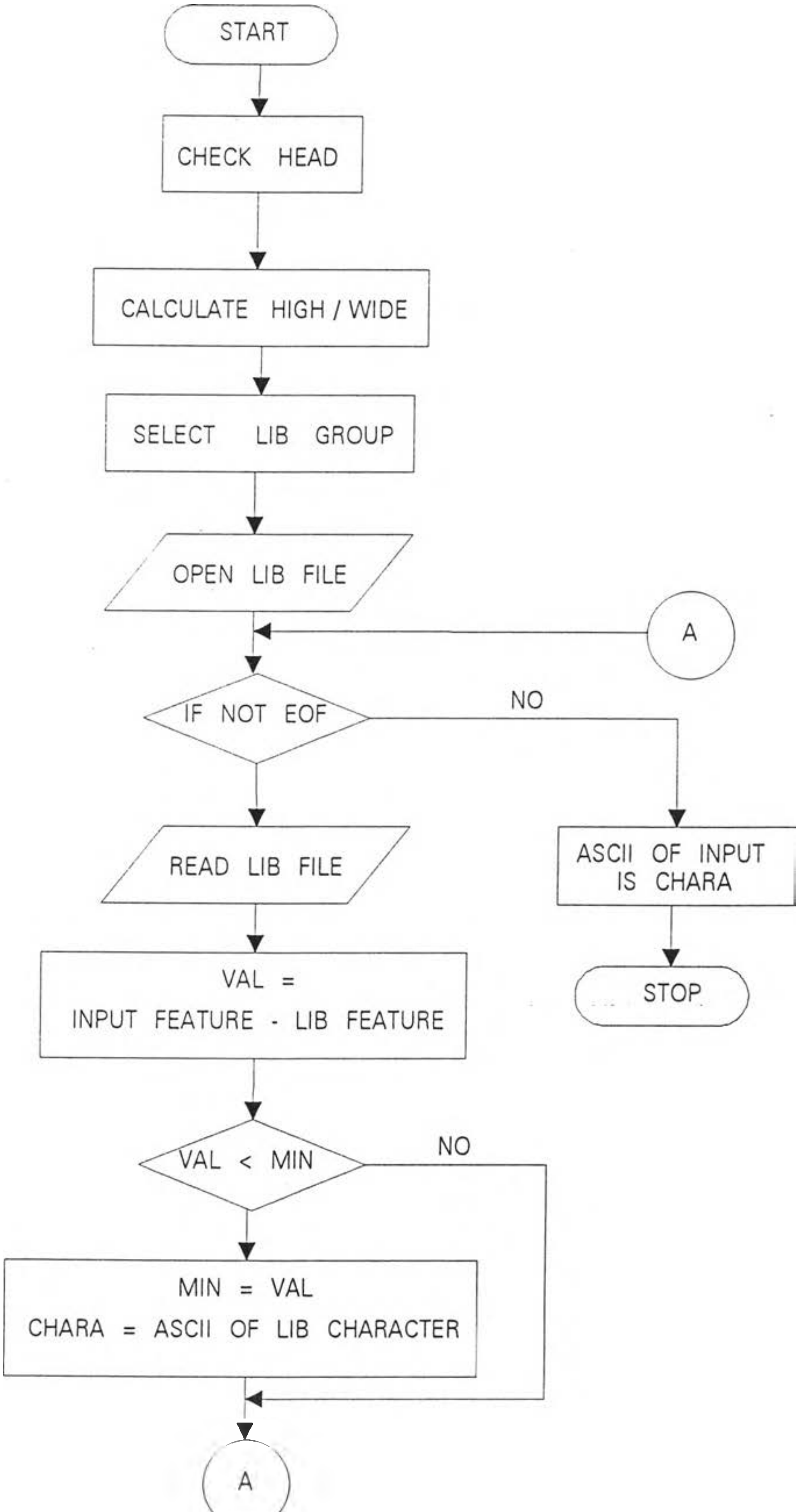
พบว่า ในการรู้จำตัวอักษร 510 ตัวอักษร สามารถรู้จำได้ถูกต้อง 459 ตัวอักษร หรือสามารถรู้จำได้ถูกต้อง 90.00 เปอร์เซนต์

จากการวิจัยพบว่าค่าน้ำหนักของลักษณะสำคัญของตัวอักษรที่เหมาะสมในการหาค่าความแตกต่างระหว่างตัวอักษรต้นแบบกับตัวอักษรที่จะทำการรู้จำ คือ

ค่าน้ำหนักของจำนวนเส้นตรงตามแนวนอน	=	4
ค่าน้ำหนักของจำนวนเส้นตรงตามแนวตั้ง	=	4
ค่าน้ำหนักของจำนวนส่วนโค้งรูปแบบที่ 1	=	5
ค่าน้ำหนักของจำนวนส่วนโค้งรูปแบบที่ 2	=	5
ค่าน้ำหนักของจำนวนส่วนโค้งรูปแบบที่ 3	=	3
ค่าน้ำหนักของจำนวนส่วนโค้งรูปแบบที่ 4	=	3
ค่าน้ำหนักของตำแหน่งของหัว	=	15
ค่าน้ำหนักของลักษณะสำคัญอื่นๆ	=	1

โดยค่าความแตกต่างที่หาได้นั้น ค่าใดที่มีค่าน้อยที่สุดก็จะถือว่าค่านั้น คือค่าที่ใกล้เคียงกันมากที่สุดระหว่างตัวอักษรต้นแบบกับตัวอักษรที่ทำการรู้จำ และจะถือว่าตัวอักษรที่ทำการรู้จำมีค่าใกล้เคียงกับตัวอักษรต้นแบบนั้นมากที่สุด หรือตัวอักษรที่รู้จำได้คือ ตัวอักษรต้นแบบที่มีค่าความแตกต่างที่น้อยที่สุดนั้

รูปที่ 3.36 ผังงานแสดงการเปรียบเทียบกับตัวอักษรต้นแบบ



3.2.8 จัดเก็บข้อมูลที่ได้

จะนำผลที่เปรียบเทียบได้มาเก็บไว้ในแฟ้มข้อมูลตัวอักษร ซึ่งสามารถนำมาใช้งานตามต้องการได้

ซึ่งจะเป็นผลที่จะช่วยลดขั้นตอนในการนำข้อมูลเข้าสู่ระบบคอมพิวเตอร์และทำให้ปัญหาทางด้านความผิดพลาดของการนำข้อมูลเข้าลดน้อยลงด้วย

โดยตัวอักษรจะถูกจัดอยู่ในรูปของตัวอักษร 3 ระดับ ซึ่งจะกล่าวถึงรายละเอียดภายหลัง

โดยมีขั้นตอนการจัดเก็บตัวอักษรที่ได้จากการรู้จำไว้ในแฟ้มข้อมูลตัวอักษรดังนี้ คือนำตัวอักษรที่ได้จากการรู้จำมาจัดเก็บไว้ในแถวลำดับ (ARRAY) ของระดับตัวอักษร โดยแบ่งตัวอักษรออกเป็น 4 ระดับดังนี้

ระดับที่ 1 วรรณยุกต์ และ การันต์

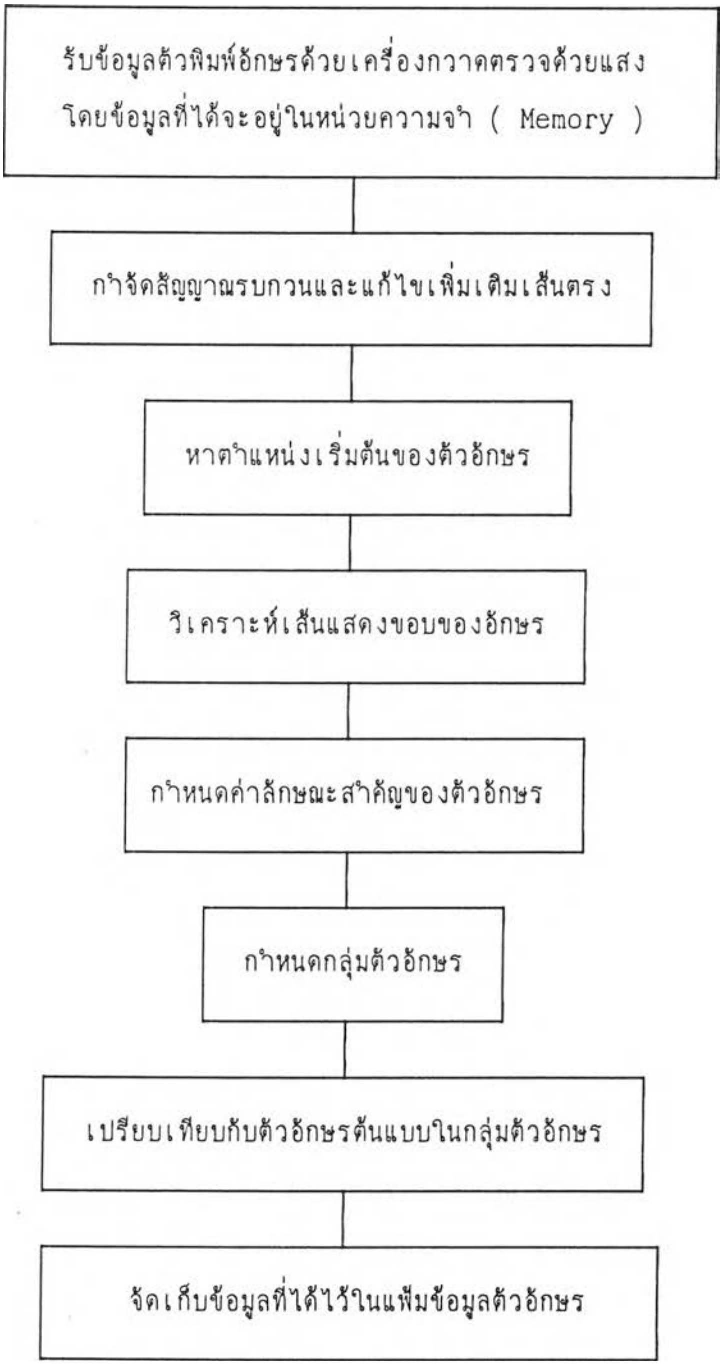
ระดับที่ 2 สระบน

ระดับที่ 3 พยัญชนะ และ สระกลาง และ พยัญชนะพิเศษ

ระดับที่ 4 สระล่าง

เมื่อเก็บตัวอักษรไว้ในแถวลำดับของระดับตัวอักษรหมดทุกตัวอักษรแล้วจะทำการเรียงลำดับข้อมูลของตัวอักษรในแถวลำดับของระดับตัวอักษร ทั้ง 4 ระดับ เมื่อเรียงลำดับเรียบร้อย จะทำการจัดเก็บตัวอักษร โดยจะเริ่มจากการจัดเก็บตัวอักษรตัวที่ 1 ในแถวลำดับตัวอักษรพยัญชนะและทำการเปรียบเทียบตัวอักษรพยัญขนะนั้นกับตัวอักษรตัวที่ 1 ในแถวลำดับตัวอักษรสระบน ถ้าตัวอักษรสระบนนี้อยู่บนตัวอักษรพยัญชนะ ก็จะถือว่าตัวอักษรสระบนตัวนี้ คือสระบนของพยัญชนะ จะทำการจัดเก็บตัวอักษรสระบนนั้นไว้ในแฟ้มข้อมูลตัวอักษร และจะทำการตรวจสอบกับตัวอักษรตัวที่ 1 ในแถวลำดับของสระล่าง ก็จะทำการเปรียบเทียบกับตัวอักษรพยัญชนะว่าตัวอักษรสระล่างนี้อยู่ล่างตัวอักษรพยัญชนะหรือไม่ ถ้าไม่ใช่ก็จะนำตัวอักษรพยัญชนะมาทำการเปรียบเทียบกับตัวอักษรในแถวลำดับของวรรณยุกต์ ถ้าตัวอักษรวรรณยุกต์ในแถวลำดับวรรณยุกต์นี้มีตำแหน่งอยู่บนตัวอักษรพยัญชนะ จะถือว่าตัวอักษรวรรณยุกต์นี้ คือวรรณยุกต์ของพยัญชนะนี้ จะทำการจัดเก็บตัวอักษรวรรณยุกต์นั้นไว้ในแฟ้มข้อมูลตัวอักษร และจะทำการจเก็บตัวอักษรพยัญชนะตัวต่อไปในแถวลำดับของตัวอักษรพยัญชนะ และเปรียบเทียบกับตัวอักษรในแถวลำดับของสระบน และเปรียบเทียบกับตัวอักษรในแถวลำดับของสระล่าง และเปรียบเทียบกับตัวอักษรในแถวลำดับของวรรณยุกต์ ตามลำดับเพื่อตรวจสอบว่าเป็นสระบน หรือสระล่าง หรือวรรณยุกต์ ของพยัญชนะนี้หรือไม่ เพื่อทำการจัดเก็บตัวอักษรไว้ในแฟ้มข้อมูลตัวอักษรต่อไปหรือไม่ และจะทำเช่นนี้ไปจนครบตัวอักษรพยัญชนะทั้งหมด จะถือว่าการจัดเก็บตัวอักษรไว้ในแฟ้มข้อมูลตัวอักษรเรียบร้อย

รูปที่ 3.37 แสดงขั้นตอนการทำงานการรู้จำตัวอักษร



3.3 การหาส่วนโค้ง และรูปแบบของส่วนโค้ง

การหาส่วนโค้ง และรูปแบบของส่วนโค้ง ของตัวอักษรนี้จะทำให้ได้ค่าตัวแปรเกี่ยวกับส่วนโค้ง นั่นคือ จำนวนส่วนโค้ง ซึ่งจะนำไปใช้เปรียบเทียบต่อไป

เทคนิคที่ใช้นี้คือ เทคนิคในการเปลี่ยนจุดที่มีค่าคู่ลำดับ (X , Y) ระหว่างคู่ลำดับของจุดหนึ่ง กับคู่ลำดับของจุดต่อไป ให้เป็นค่ามุมองศา (Degree)

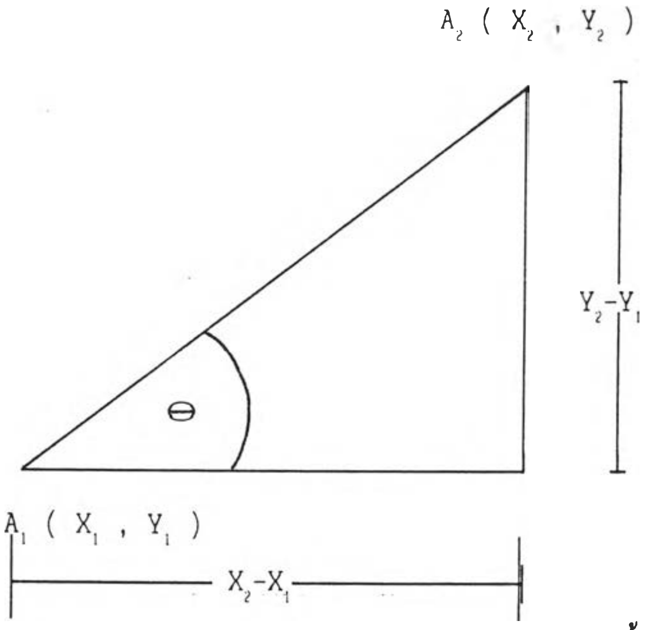
ซึ่งเทคนิคนี้เป็นเทคนิคที่ได้คิดขึ้นมาใหม่สำหรับงานวิจัยนี้ โดยมีขั้นตอนดังนี้

3.3.1 การหาส่วนโค้ง

การหาส่วนโค้งจะทำการเปลี่ยนจุดคู่ลำดับ (X , Y) จากจุดหนึ่งไปยังจุดหนึ่งเป็นค่ามุมองศา (Degree) เนื่องจากจุดแต่ละจุดที่ได้จากการอ่านด้วย เครื่องกวาดตรวจด้วยแสง (Optical Scanner) จะมีตำแหน่งของจุดที่แน่นอน โดยที่ ตำแหน่งของจุดเหล่านั้นจะกำหนด ด้วยคู่ลำดับ (X , Y) ในการหาค่ามุมองศาระหว่างจุดหนึ่ง ไปยังจุดหนึ่งนั้นจะใช้คู่ลำดับ (X , Y) ของจุดทั้งสองมาเปรียบเทียบ เพื่อหามุมองศา โดยใช้สมการทางคณิตศาสตร์

ตัวอย่างเช่น

จุดที่ 1 A_1 มีค่าของคู่ลำดับ (X,Y) คือ (X_1 , Y_1)
จุดที่ 2 A_2 มีค่าของคู่ลำดับ (X,Y) คือ (X_2 , Y_2)



การหาค่ามุมองศา
ใช้สมการทางคณิตศาสตร์

$$\Theta = \text{ARCTAN}\left[\frac{Y_2 - Y_1}{X_2 - X_1}\right]$$

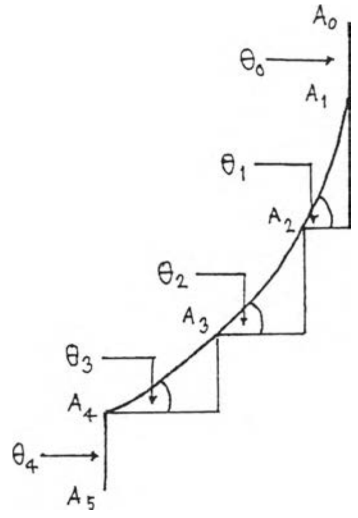
ถ้า $X_2 - X_1 = 0$ กำหนด $\Theta = 90$

รูปที่ 3.38 แสดงตัวอย่างการหามุมของส่วนโค้ง

ในการหาส่วนโค้งจะทำการหามุมองศาระหว่างจุดต่างๆเช่นนี้ไปเรื่อยๆ จนกระทั่งพบว่าค่ามุมองศาที่ได้มีค่าเป็น 0, 90, 180, 270, 360 ก็จะถือว่าจุดนั้นเป็นจุดสุดท้ายของส่วนโค้งนั้น

ตัวอย่างการหาส่วนโค้ง

A_0 - A_5 คือจุดบนส่วนโค้ง



Θ_0 Θ_4 มีค่าเป็น 0 องศา
Θ_1 Θ_2 Θ_3 ไม่มีค่าเป็น
0, 90, 180, 270, 360

เนื่องจาก Θ_1. คือ ค่ามุมองศาเริ่มต้นที่ไม่มีค่าเป็น
0, 90, 180, 270, 360

Θ_4 คือ ค่ามุมองศาที่มีค่าเป็น 0 ที่
พบครั้งแรกหลังจากที่มีค่ามุมองศา
ไม่มีค่าเป็น
0, 90, 180, 270, 360

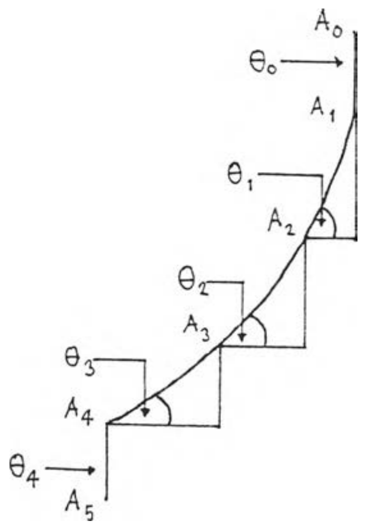
ดังนั้น ส่วนโค้งนี้มีจุดเริ่มต้นที่ A_1
ส่วนโค้งนี้มีจุดสุดท้ายที่ A_4

รูปที่ 3.39 แสดงตัวอย่างการหาส่วนโค้ง

3.3.2 การหารูปแบบของส่วนโค้ง

เมื่อหาส่วนโค้งได้แล้ว ก็จะทำการหาว่าส่วนโค้งนั้นเป็นส่วนโค้งแบบใด โดยพิจารณาจากค่ามุมองศา (Θ) ที่จุดเริ่มต้น และจุดสุดท้าย

เช่น



จากรูปค่ามุมองศาที่ใช้ในการพิจารณา คือ

Θ_1 และ Θ_3

A_0 - A_5 คือจุดบนส่วนโค้ง

Θ_0 - Θ_4 คือมุมองศา

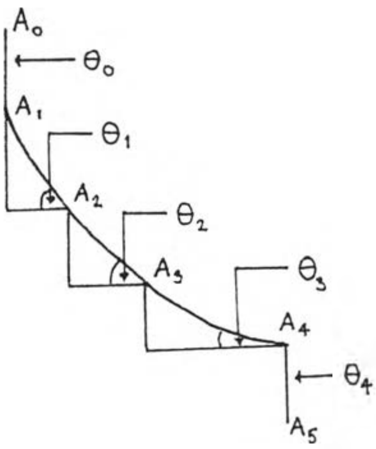
รูปที่ 3.40 แสดงการหารูปแบบของส่วนโค้ง

ในการวิจัยนี้ได้กำหนดส่วนโค้งไว้ 4 รูปแบบ ดังนี้

<u>รูปแบบที่ 1</u>

รูปแบบของส่วนโค้ง ที่มีค่ามุมองศาของจุดเริ่มต้น และ มุมองศาของจุดสุดท้าย มีค่าเป็น บวก เหมือนกันทั้งสองค่า

ดังตัวอย่าง



พบว่า

θ_1 มีค่าเป็น บวก

θ_3 มีค่าเป็น บวก

A_0 - A_5 คือจุดบนส่วนโค้ง

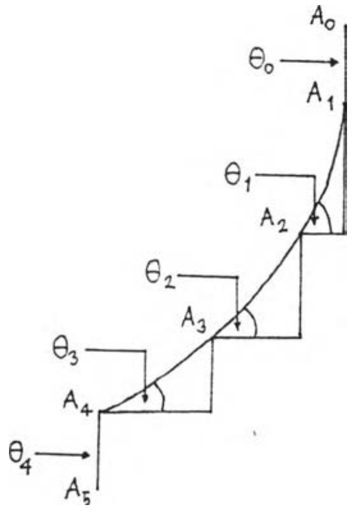
θ_0 - θ_4 คือมุมองศา

รูปที่ 3.41 แสดงส่วนโค้งรูปแบบที่ 1

<u>รูปแบบที่ 2</u>

รูปแบบของส่วนโค้ง ที่มีค่ามุมองศาของจุดเริ่มต้น และมุมองศาของจุดสุดท้าย มีค่าเป็น ลบ เหมือนกันทั้งสองค่า

ดังตัวอย่าง



พบว่า

θ_1 มีค่าเป็น ลบ

θ_3 มีค่าเป็น ลบ

A_0 - A_5 คือจุดบนส่วนโค้ง

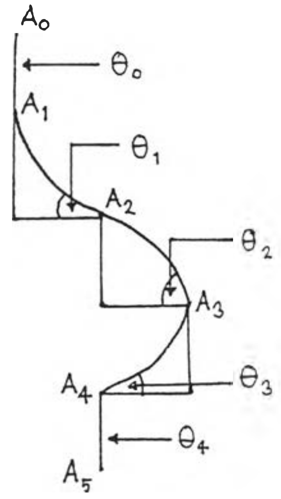
θ_0 - θ_4 คือมุมองศา

รูปที่ 3.42 แสดงส่วนโค้งรูปแบบที่ 2

<u>รูปแบบที่ 3</u>

รูปแบบของส่วนโค้ง ที่มีค่ามุมองศาของจุดเริ่มต้น มีค่าเป็น บวก และ มุมองศาของจุดสุดท้าย มีค่าเป็น ลบ

ดังตัวอย่าง



พบว่า

θ_1 มีค่าเป็น บวก

θ_3 มีค่าเป็น ลบ

A_0 - A_5 คือจุดบนส่วนโค้ง

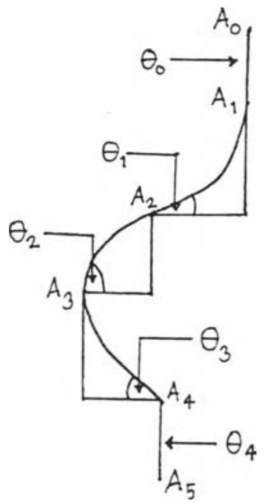
θ_0 - θ_4 คือมุมองศา

รูปที่ 3.43 แสดงส่วนโค้งรูปแบบที่ 3

รูปแบบที่ 4

รูปแบบของส่วนโค้ง ที่มีค่ามุมองศาของจุดเริ่มต้น มีค่าเป็น ลบ และ มุมองศาของจุดสุดท้าย มีค่าเป็น บวก

ดังตัวอย่าง



พบว่า

θ_1 มีค่าเป็น ลบ

θ_3 มีค่าเป็น บวก

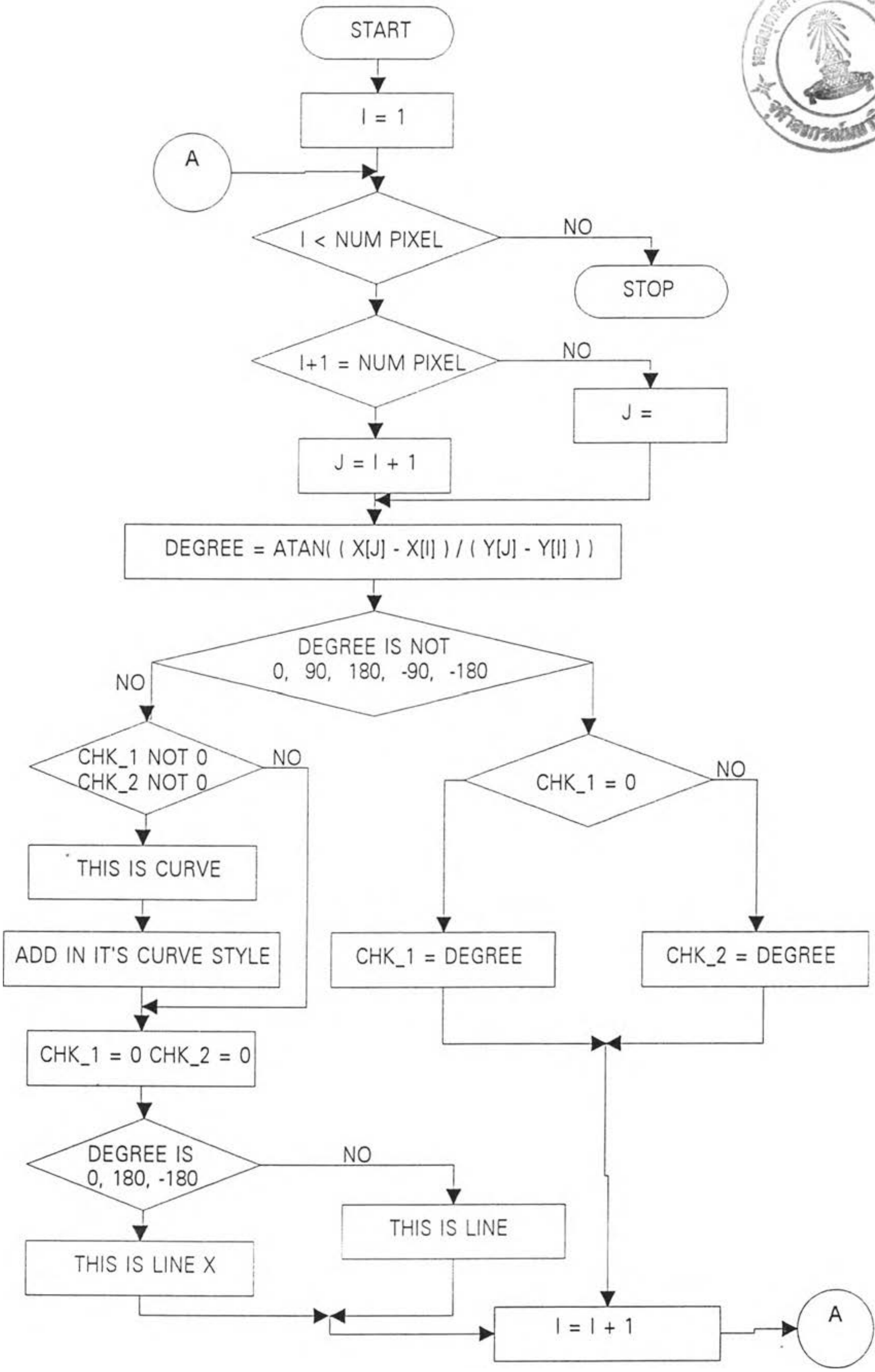
A_0 - A_5 คือจุดบนส่วนโค้ง

θ_0 - θ_4 คือมุมองศา

รูปที่ 3.44 แสดงส่วนโค้งรูปแบบที่ 4

รูปที่ 3.45 ผังงานแสดงการหาส่วนโค้งและรูปแบบของส่วนโค้ง

หอสมุดกลาง สถาบันวิทยบริการ จุฬาลงกรณ์มหาวิทยาลัย



3.4 การหาเส้นตรง

การหาเส้นตรงของตัวอักษรนี้จะทำให้ได้ค่าลักษณะสำคัญคือจำนวนเส้นตรงตามแนวนอน และจำนวนเส้นตรงตามแนวตั้ง ซึ่งจะนำไปใช้เปรียบเทียบต่อไป

เทคนิคที่ใช้นี้ คือ เทคนิคในการเปลี่ยนจุดที่มีค่าคู่ลำดับ (X , Y) ระหว่างคู่ลำดับของจุดหนึ่งกับคู่ลำดับของจุดต่อไป ให้เป็นค่ามุมองศา (Degree) เช่นเดียวกับการหาส่วนโค้ง แต่ต่างกันที่ การหาเส้นตรงนั้นจะใช้ค่ามุมองศาที่เท่ากับ 0, 90, 180, 270, 360 หรือ -90, -180, -270, -360 ส่วนการหาส่วนโค้งจะไม่ใช้ค่ามุมองศาดังกล่าว ส่วนขั้นตอนในการหาเส้นตรงนั้นจะทำการหามุมองศาระหว่างจุดต่างๆเช่นนี้ไปเรื่อยๆ จนกระทั่งพบว่าค่ามุมองศาที่ได้มีค่าไม่เท่ากับมุมองศาเดิม จึงจะถือว่าจุดส่วนนั้นเป็นเส้นตรง

เส้นตรงที่มีมุมองศา 0, -180, 180, -360, 360 คือ เส้นตรงแนวนอน

เส้นตรงที่มีมุมองศา -90, 90, -270, 270 คือ เส้นตรงแนวตั้ง

3.5 การจัดระดับตัวอักษร

เนื่องจากตัวอักษรไทยสามารถ แบ่งได้ 4 ระดับใหญ่ๆ คือ

ระดับที่ 1 วรรณยุกต์ และ การันต์
ระดับที่ 2 สระบน
ระดับที่ 3 พยัญชนะ และ สระกลาง และ พยัญชนะพิเศษ
ระดับที่ 4 สระล่าง

วรรณยุกต์	ได้แก่	่ ้ ๊ ๋ ์
พยัญชนะพิเศษ	ได้แก่	ฤ ฦ ๆ ฯ
สระบน	ได้แก่	ํ ิ ี ึ ื ั ็
สระกลาง	ได้แก่	ะ า เ แ ไ ใ โ
สระล่าง	ได้แก่	ุ ู

การจัดเก็บเมื่อได้ตัวอักษรทั้ง 4 ระดับแล้วจะทำการเรียงลำดับตัวอักษรในแต่ละระดับ โดยการเรียงลำดับนี้จะเรียงลำดับตาม บรรทัด และ สดมภ์ของตัวอักษร เนื่องจากตัวอักษรที่เข้ามาแต่ละตัวจะมี ตำแหน่งที่อยู่ ซึ่งตำแหน่งที่อยู่นี้จะประกอบด้วย แถวบนสุดของตัวอักษร แถวล่างสุดของตัวอักษร สดมภ์ซ้ายสุดของตัวอักษร สดมภ์ขวาสุดของตัวอักษร ตัวอย่างเช่น

```
1234567890123456789012345678901234567890

1                            11111
2                           1   1
3          111111          1   1
4        11      11       1  11
5       1          1     1  1
6      1   11       1  11  1
7     1   1  1   1  111   1
8     1  1    1 1 1      1
9     1  1    1 1 1      1
0     1   1   1 1 1   11111
1      1   1 1  1  1        1
2       11  1  1  1    1111  1
3         1   1   1   1   1  1
4          111    1  1     1 1
5                 11 1     1 1
6                  1 1     1 1
7                  1 1     1 1
8                  1 1     1 1
9                  1 1    1  1
0                 1  1    1  1
1                 1  1    1  1
2              111   111111  1
3              1            1
4              1111111111111
```

จากตัวอย่าง อักษร ช มี

แถวบนสุด คือ 1

แถวล่างสุด คือ 24

สดมภ์ซ้ายสุด คือ 7

สดมภ์ขวาสุด คือ 34

หอสมุดกลาง สถาบันวิทยบริการ จุฬาลงกรณ์มหาวิทยาลัย

```
0123456789012345678901234567890
1            11111111111
2           1  111111   11
3          1 11      11   11
4         1 1          1   1
5       1 1            1   1
6      1 1              1  1
7     1   11            1  1
8     1     111111      1  1
9      1     11111      1  1
0       1   1           1  1
1       1  1            1  1
2       1  1            1  1
3       1  1            1  1
4       1  1            1  1
5       1  1            1  1
6       1  1            1  1
7       1111            1111
```

จากตัวอย่าง อักษร ก มี

แถวบนสุด คือ 1

แถวล่างสุด คือ 17

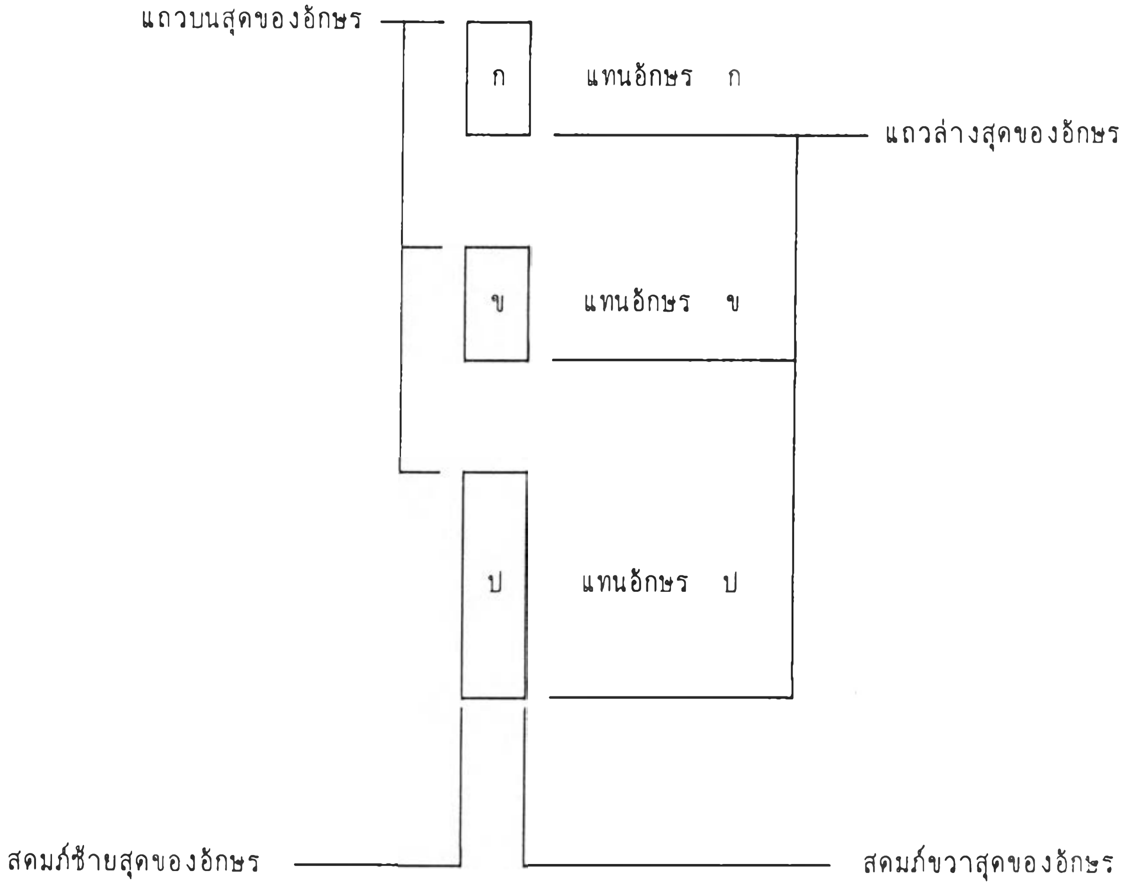
สดมภ์ซ้ายสุด คือ 6

สดมภ์ขวาสุด คือ 27

รูปที่ 3.46 แสดงตัวอักษรที่จะกำหนดแถวและสดมภ์

จากอักษร ก ดามตัวอย่างข้างต้น จะพบว่า มีตำแหน่ง แถวบนสุดของตัวอักษร แถวล่างสุดของตัวอักษร สดมภ์ซ้ายสุดของตัวอักษร สดมภ์ขวาสุดของตัวอักษร เมื่อได้แล้วจะนำมาเปรียบเทียบกับตัวอักษรที่อ่านเข้ามาก่อน โดยขั้นแรกจะเปรียบเทียบว่าเป็นตัวอักษรที่อยู่บรรทัดเดียวกันหรือไม่ โดยใช้หลักการดังนี้

กำหนดให้



รูปที่ 3.47 แสดงการกำหนดแถวและสดมภ์ให้ตัวอักษร

การเปรียบเทียบว่าตัวอักษรนั้นอยู่บรรทัดเดียวกันหรือไม่ จะนำค่าของแถวบนสุดของตัวอักษร และค่าของแถวล่างสุดของตัวอักษรแต่ละตัว มาเปรียบเทียบกัน ตามหลักการนี้

หลักการที่ 1 ในการเปรียบเทียบแถวของตัวอักษร
ถ้า แถวบนสุดของตัวอักษรแรก มีค่าน้อยกว่าหรือเท่ากับ แถวบนสุดของตัวอักษรหลัง
และ แถวบนสุดของตัวอักษรแรก มีค่าน้อยกว่าหรือเท่ากับ แถวล่างสุดของตัวอักษรหลัง
และ แถวล่างสุดของตัวอักษรแรก มีค่าน้อยกว่าหรือเท่ากับ แถวล่างสุดของตัวอักษรหลัง
และ แถวล่างสุดของตัวอักษรแรก มีค่ามากกว่าหรือเท่ากับ แถวบนสุดของตัวอักษรหลัง
จะถือว่า ตัวอักษรตัวแรก และ ตัวอักษรตัวหลัง อยู่บรรทัดเดียวกัน

หลักการที่ 2 ในการเปรียบเทียบแถวของตัวอักษร
ถ้า แถวบนสุดของตัวอักษรแรก มีค่ามากกว่าหรือเท่ากับ แถวบนสุดของตัวอักษรหลัง
และ แถวบนสุดของตัวอักษรแรก มีค่าน้อยกว่าหรือเท่ากับ แถวล่างสุดของตัวอักษรหลัง
และ แถวล่างสุดของตัวอักษรแรก มีค่ามากกว่าหรือเท่ากับ แถวล่างสุดของตัวอักษรหลัง
และ แถวล่างสุดของตัวอักษรแรก มีค่ามากกว่าหรือเท่ากับ แถวบนสุดของตัวอักษรหลัง
จะถือว่า ตัวอักษรตัวแรก และ ตัวอักษรตัวหลัง อยู่บรรทัดเดียวกัน

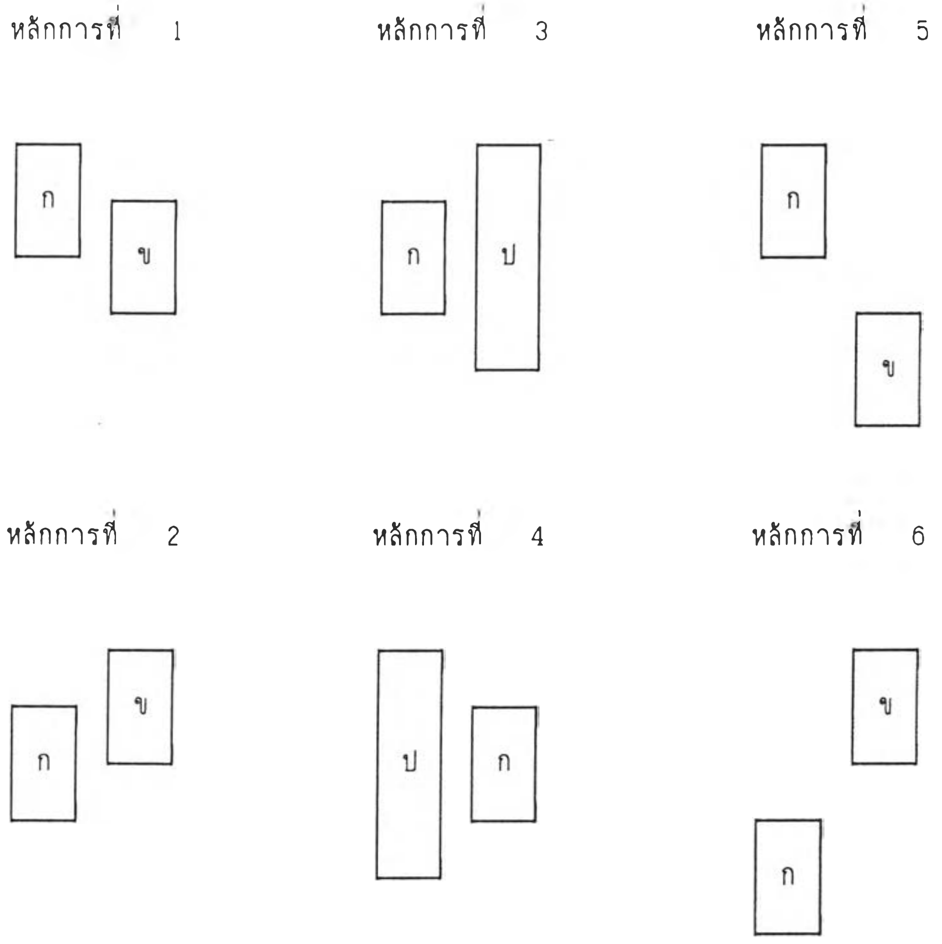
หลักการที่ 3 ในการเปรียบเทียบแถวของตัวอักษร
ถ้า แถวบนสุดของตัวอักษรแรก มีค่ามากกว่าหรือเท่ากับ แถวบนสุดของตัวอักษรหลัง
และ แถวบนสุดของตัวอักษรแรก มีค่าน้อยกว่าหรือเท่ากับ แถวล่างสุดของตัวอักษรหลัง
และ แถวล่างสุดของตัวอักษรแรก มีค่าน้อยกว่าหรือเท่ากับ แถวล่างสุดของตัวอักษรหลัง
และ แถวล่างสุดของตัวอักษรแรก มีค่ามากกว่าหรือเท่ากับ แถวบนสุดของตัวอักษรหลัง
จะถือว่า ตัวอักษรตัวแรก และ ตัวอักษรตัวหลัง อยู่บรรทัดเดียวกัน

หลักการที่ 4 ในการเปรียบเทียบแถวของตัวอักษร
ถ้า แถวบนสุดของตัวอักษรแรก มีค่าน้อยกว่าหรือเท่ากับ แถวบนสุดของตัวอักษรหลัง
และ แถวบนสุดของตัวอักษรแรก มีค่าน้อยกว่าหรือเท่ากับ แถวล่างสุดของตัวอักษรหลัง
และ แถวล่างสุดของตัวอักษรแรก มีค่ามากกว่าหรือเท่ากับ แถวล่างสุดของตัวอักษรหลัง
และ แถวล่างสุดของตัวอักษรแรก มีค่ามากกว่าหรือเท่ากับ แถวบนสุดของตัวอักษรหลัง
จะถือว่า ตัวอักษรตัวแรก และ ตัวอักษรตัวหลัง อยู่บรรทัดเดียวกัน

หลักการที่ 5 ในการเปรียบเทียบแถวของตัวอักษร
ถ้า แถวบนสุดของตัวอักษรแรก มีค่าน้อยกว่า แถวบนสุดของตัวอักษรหลัง
และ แถวบนสุดของตัวอักษรแรก มีค่าน้อยกว่า แถวล่างสุดของตัวอักษรหลัง
และ แถวล่างสุดของตัวอักษรแรก มีค่าน้อยกว่า แถวล่างสุดของตัวอักษรหลัง
และ แถวล่างสุดของตัวอักษรแรก มีค่าน้อยกว่า แถวบนสุดของตัวอักษรหลัง
จะถือว่า ตัวอักษรตัวแรก และ ตัวอักษรตัวหลัง ไม่อยู่บรรทัดเดียวกัน
จะถือว่า ตัวอักษรตัวแรก อยู่บรรทัดบน และ ตัวอักษรตัวหลัง อยู่บรรทัดล่าง

หลักการที่ 6 ในการเปรียบเทียบแถวของตัวอักษร

ถ้า แถวบนสุดของตัวอักษรแรก มีค่ามากกว่าหรือเท่ากับ แถวบนสุดของตัวอักษรหลัง
และ แถวบนสุดของตัวอักษรแรก มีค่ามากกว่าหรือเท่ากับ แถวล่างสุดของตัวอักษรหลัง
และ แถวล่างสุดของตัวอักษรแรก มีค่ามากกว่าหรือเท่ากับ แถวล่างสุดของตัวอักษรหลัง
และ แถวล่างสุดของตัวอักษรแรก มีค่ามากกว่าหรือเท่ากับ แถวบนสุดของตัวอักษรหลัง
จะถือว่า ตัวอักษรตัวแรก และ ตัวอักษรตัวหลัง ไม่อยู่บรรทัดเดียวกัน
จะถือว่า ตัวอักษรตัวแรก อยู่บรรทัดล่าง และ ตัวอักษรตัวหลัง อยู่บรรทัดบน



รูปที่ 3.48 แสดงตัวอย่างของการตรวจสอบว่าตัวอักษรอยู่บรรทัดเดียวกันหรือไม่

เมื่อหาว่าตัวอักษรทั้ง 2 ตัวอยู่บรรทัดเดียวกันหรือไม่แล้ว หลังจากนั้นจะทำการตรวจสอบเพื่อเรียงลำดับก่อนหลัง

ถ้าตัวอักษรทั้ง 2 อยู่บรรทัดเดียวกัน จะทำการตรวจสอบต่อว่า ตัวอักษรใดอยู่หน้า ตัวอักษรใดอยู่หลัง โดยใช้ค่าสดมภ์ซ้ายสุดของตัวอักษรทั้ง 2 ในการเปรียบเทียบ ถ้าตัวอักษรใดมีค่าสดมภ์ซ้ายสุดน้อยกว่า ตัวอักษรนั้นต้องอยู่หน้า

การตรวจสอบว่าสระบน หรือสระล่างเป็นสระของตัวอักษรใด จะใช้หลักการเช่นเดียวกัน แต่จะเปลี่ยนจากการใช้ แถวในการเปรียบเทียบก็จะใช้ สดมภ์ในการเปรียบเทียบ โดยมีหลักการ 4 หลักการ ดังนี้

หลักการที่ 1 ในการเปรียบเทียบสระและตัวอักษร

ถ้า สดมภ์ซ้ายสุดของสระ มีค่ามากกว่าหรือเท่ากับ สดมภ์ซ้ายสุดของตัวอักษร
และ สดมภ์ซ้ายสุดของสระ มีค่าน้อยกว่าหรือเท่ากับ สดมภ์ขวาสุดของตัวอักษร
และ สดมภ์ขวาสุดของสระ มีค่ามากกว่าหรือเท่ากับ สดมภ์ขวาสุดของตัวอักษร
และ สดมภ์ขวาสุดของสระ มีค่ามากกว่าหรือเท่ากับ สดมภ์ซ้ายสุดของตัวอักษร

หลักการที่ 2 ในการเปรียบเทียบสระและตัวอักษร

ถ้า สดมภ์ซ้ายสุดของสระ มีค่าน้อยกว่าหรือเท่ากับ สดมภ์ซ้ายสุดของตัวอักษร
และ สดมภ์ซ้ายสุดของสระ มีค่าน้อยกว่าหรือเท่ากับ สดมภ์ขวาสุดของตัวอักษร
และ สดมภ์ขวาสุดของสระ มีค่าน้อยกว่าหรือเท่ากับ สดมภ์ขวาสุดของตัวอักษร
และ สดมภ์ขวาสุดของสระ มีค่ามากกว่าหรือเท่ากับ สดมภ์ซ้ายสุดของตัวอักษร

หลักการที่ 3 ในการเปรียบเทียบสระและตัวอักษร

ถ้า สดมภ์ซ้ายสุดของสระ มีค่ามากกว่าหรือเท่ากับ สดมภ์ซ้ายสุดของตัวอักษร
และ สดมภ์ซ้ายสุดของสระ มีค่าน้อยกว่าหรือเท่ากับ สดมภ์ขวาสุดของตัวอักษร
และ สดมภ์ขวาสุดของสระ มีค่าน้อยกว่าหรือเท่ากับ สดมภ์ขวาสุดของตัวอักษร
และ สดมภ์ขวาสุดของสระ มีค่ามากกว่าหรือเท่ากับ สดมภ์ซ้ายสุดของตัวอักษร

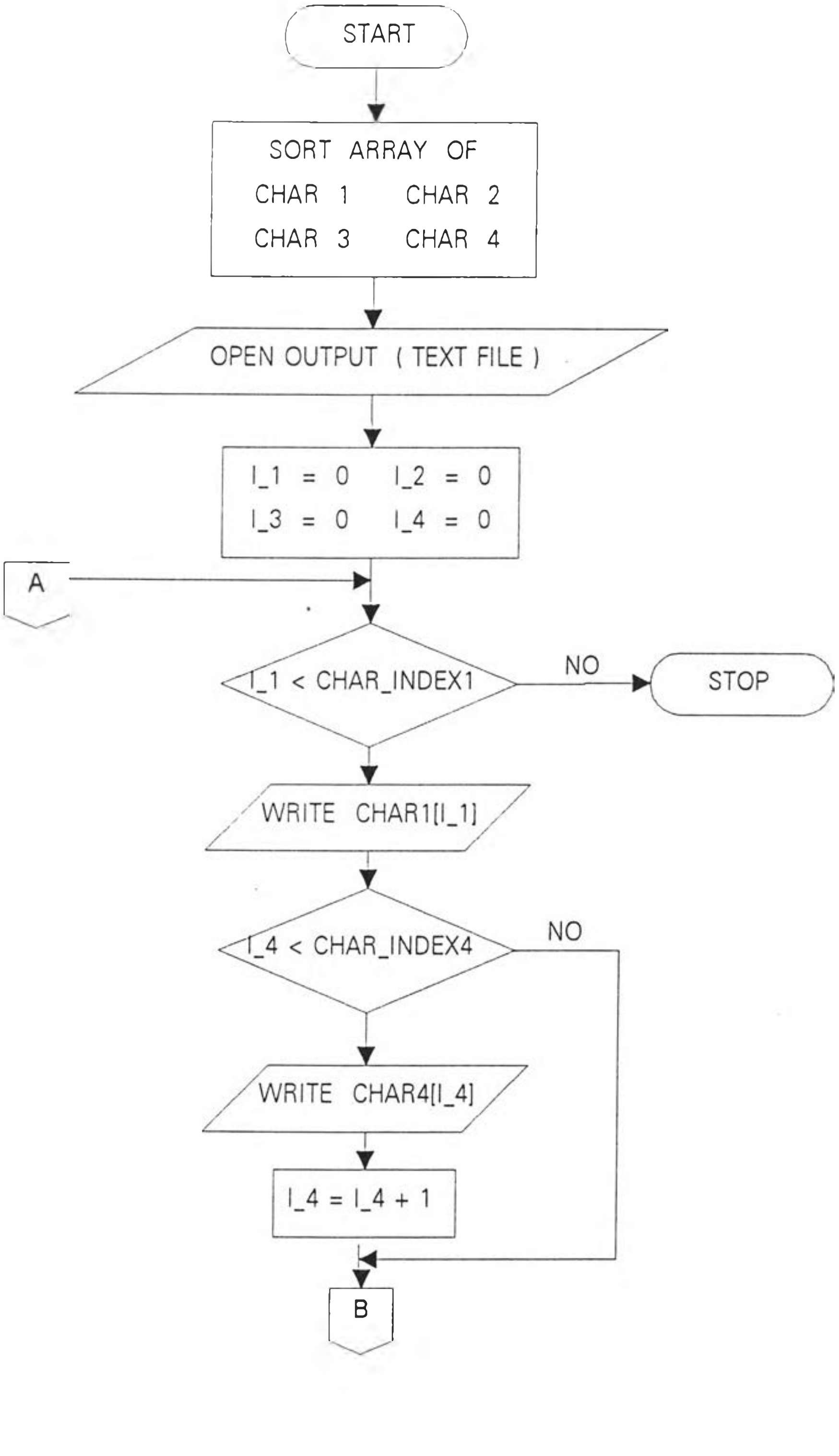
หลักการที่ 4 ในการเปรียบเทียบสระและตัวอักษร

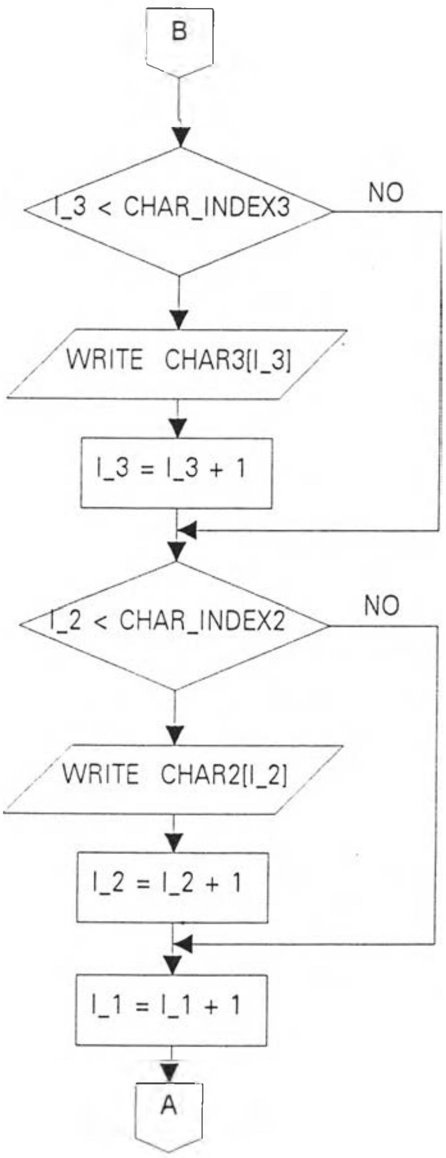
ถ้า สดมภ์ซ้ายสุดของสระ มีค่าน้อยกว่าหรือเท่ากับ สดมภ์ซ้ายสุดของตัวอักษร
และ สดมภ์ซ้ายสุดของสระ มีค่าน้อยกว่าหรือเท่ากับ สดมภ์ขวาสุดของตัวอักษร
และ สดมภ์ขวาสุดของสระ มีค่ามากกว่าหรือเท่ากับ สดมภ์ขวาสุดของตัวอักษร
และ สดมภ์ขวาสุดของสระ มีค่ามากกว่าหรือเท่ากับ สดมภ์ซ้ายสุดของตัวอักษร

จะถือว่า สระนั้นเป็นสระของตัวอักษร เมื่อ

สระนั้นเป็นสระบน และ สระบนนั้นอยู่บรรทัดบน ส่วน ตัวอักษรอยู่บรรทัดล่าง
สระนั้นเป็นสระล่าง และ สระล่างนั้นอยู่บรรทัดล่าง ส่วน ตัวอักษรอยู่บรรทัดบน

รูปที่ 3.49 ผังงานแสดงการจัดระดับตัวอักษร




B
I_3 < CHAR_INDEX3
NO
WRITE CHAR3[I_3]
I_3 = I_3 + 1
I_2 < CHAR_INDEX2
NO
WRITE CHAR2[I_2]
I_2 = I_2 + 1
I_1 = I_1 + 1
A