หอสมุดกลาง จุฬาลงกรณมหาวิทยาลัย

บทที่ 2

วิธีดำเนินการวิจัย

ในการศึกษาการเปลี่ยนแปลงพื้นที่ป่าไม้โดยใช้เครื่องคอมพิวเตอร์วิเคราะห์เทปข้อมูลดาวเทียมสำรวจทรัพยากรนั้น ได้มีการศึกษาถึงลักษณะการจัดข้อมูลในเทปเพื่อความถูกต้องในการเขียนโปรแกรมสั่งเครื่องคอมพิวเตอร์ให้วิเคราะห์ข้อมูลตามต้องการ

ลักษณะและการจัดข้อมูลในเทป

ลักษณะและการจัดข้อมูลในเทปนั้น สามารถแบ่งออกตามประเภทของข้อมูลได้ 3 ประเภท ดังนี้

1. ID RECORD

หรือที่เรียกว่า HEADER RECORD เป็นข้อมูลส่วนแรกของเทปแต่ละ strip[1] มีข้อมูลดังกล่าวอยู่ 40 bytes ซึ่งเป็นข้อมูลแบบ binary และ EBCDIC รายละเอียดเกี่ยวกับข้อมูล ID RECORD มีดังนี้คือ

byte ที่ 1-12 Scene/Frame ID จะอยู่ในรูปของจำนวนวันนับตั้งแต่ดาวเทียมเริ่มเข้าสู่วงโคจร จนถึงวันที่ถ่ายภาพ

" 13-16 Tape Sequencing Number เช่น เทป 1of 4 (stirp ที่ 1) เทป 2 of 4 (strip ที่ 2)

" 17-18 Data Record Length

" 19-26 Binary Frame ID ก็คือ Scene/Frame ID ที่เป็น binary

---

[1] พื้นที่ภาพถ่ายแต่ละภาพเท่ากับ 185 x 185 ตร.กม. จะแบ่งความยาวออกเป็น 4 ส่วนเท่า ๆ กันคือ 46.5 x 185 ตร.กม. แต่ละส่วนนี้เรียกว่า Strip และการบันทึกข้อมูลลงเทปจะบันทึกทีละ Strip เรียงกันไป

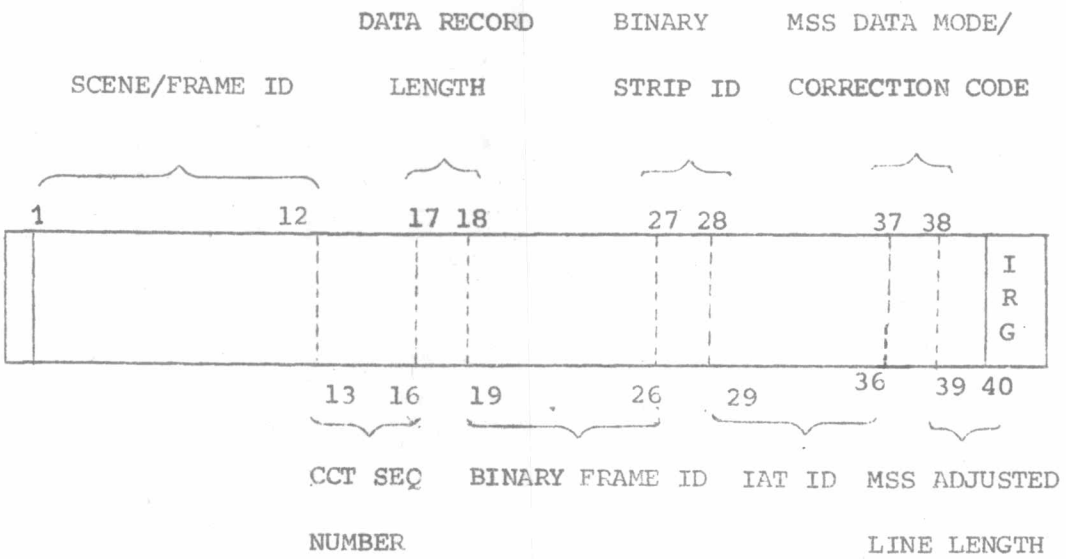
byte ที่ 27-28 Binary Strip ID

" 29-36 Image Anotation Tape (IAT) ID

" 37-38 MSS Data Mode/Correction Code มีระหัสทั้งหมด 16 bits มีค่าเป็น 0 หรือ 1 เพื่อบอกลักษณะของข้อมูล เช่น เป็นข้อมูลแบบ decompress หรือไม่ มี line length adjustment หรือไม่

" 39-40 MSS Adjusted Line Length

รายละเอียดการเรียงลำดับข้อมูลเหล่านี้แสดงได้ในรูปที่ 3



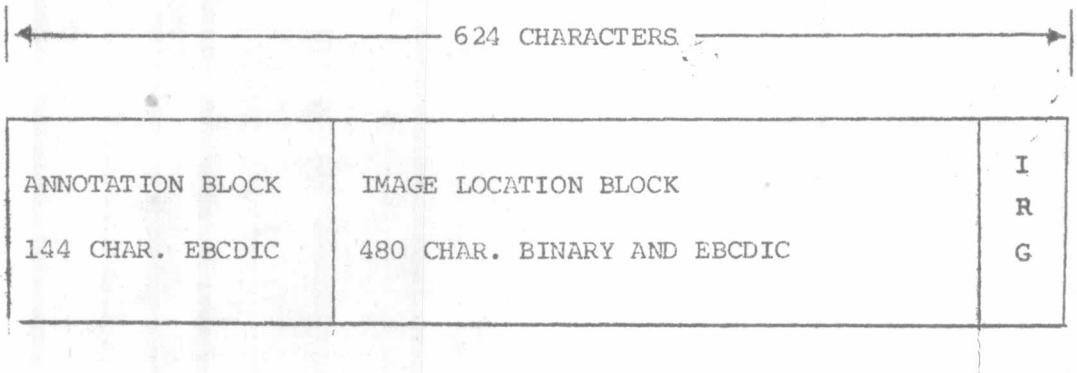
รูปที่ 3 ลำดับข้อมูลต่าง ๆ ใน ID RECORD

2. ANNOTATION RECORD

เป็นข้อมูลส่วนที่สองของเทปแต่ละ strip มีข้อมูลดังกล่าวอยู่ 624 ตัวอักษร (ดังรูปที่ 4) 144 ตัวอักษรแรกเรียกว่า Annotation Block เป็นข้อมูลแบบ EBCDIC ทั้งหมด ข้อมูลนี้จะบอกให้ทราบเกี่ยวกับภาพถ่ายแต่ละภาพ เช่น วันที่ที่ถ่ายภาพ Format

Center (จะบอกตำแหน่งของเส้นรุ้ง เส้นแวง ที่เป็นศูนย์กลางของภาพ) Nadir (ระยะตั้งฉากจากจุดสังเกตถึงดาวเทียม) Sun Elevation (มุมเงยของดวงอาทิตย์) Sun Azimuth (มุมตามแนวราบของดวงอาทิตย์)

ส่วนที่เหลืออีก 480 ตัวอักษร เป็น Image Location Block ข้อมูลส่วนนี้มีทั้งแบบ binary และ EBCDIC ซึ่งให้รายละเอียดเพิ่มเติมเกี่ยวกับภาพ เช่น Tick Mark Location



รูปที่ 4 ข้อมูลใน ANNOTATION BLOCK

3. VIDEO DATA RECORD

เป็นข้อมูลเกี่ยวกับภาพที่ถ่ายทั้งหมด ซึ่งเป็นข้อมูลที่ปรับให้เป็นตัวเลขเรียบร้อยแล้ว โดยที่ค่าแต่ละจุดจะแทนค่าความเข้มแสง ซึ่งมีค่าตั้งแต่ 0 - 63 (ถ้าเป็นข้อมูลแบบ linear) หรือตั้งแต่ 0 - 127 (ถ้าข้อมูลเป็นแบบ decompressed)

การบันทึกข้อมูลส่วนนี้ลงใน CCT จะบันทึกทีเดียวทั้ง 4 แบนด์ โดยเรียงข้อมูลแต่ละแบนด์ทีละ 2 จุด จุดละ 1 byte ดังนี้

Pt. 1	Pt. 2	Pt. 1	Pt. 2	Pt. 1	Pt.2	Pt. 1	Pt. 2 ฯลฯ
Chan A	Chan A	Chan B	Chan B	Chan C	Chan C	Chan D	Chan D
1 byte	1 byte	1 byte	1 byte	1 byte	1 byte	1 byte	1 byte

ข้อมูลของภาพหนึ่ง ๆ จะมีความกว้าง 3000 - 3450 bytes และมีความยาว 2340 scan lines และจะแบ่งข้อมูลแต่ละภาพของแต่ละแบนด์ออกเป็น 4 ส่วน (strip) แล้วบันทึกข้อมูลเรียงกันไปทีละ Strip (ดูรูปที่ 5)

ถ้าเป็นเทป 7 track จะบันทึกข้อมูลได้ม้วนละ strip

" 9 track 800 bpi บันทึกข้อมูลได้ม้วนละ 2 strips

" 9 track 1600 bpi บันทึกข้อมูลทั้ง 4 strips ในม้วนเดียวกัน

ลักษณะการจัดข้อมูลดังกล่าวข้างต้นนี้ จะไม่สะดวกในการวิเคราะห์ข้อมูลโดยตรง จำต้องจัดเรียงข้อมูลใหม่เสียก่อน (ซึ่งจะกล่าวโดยละเอียดในหัวข้อต่อไป) จึงจะนำไปใช้งานต่อไปได้

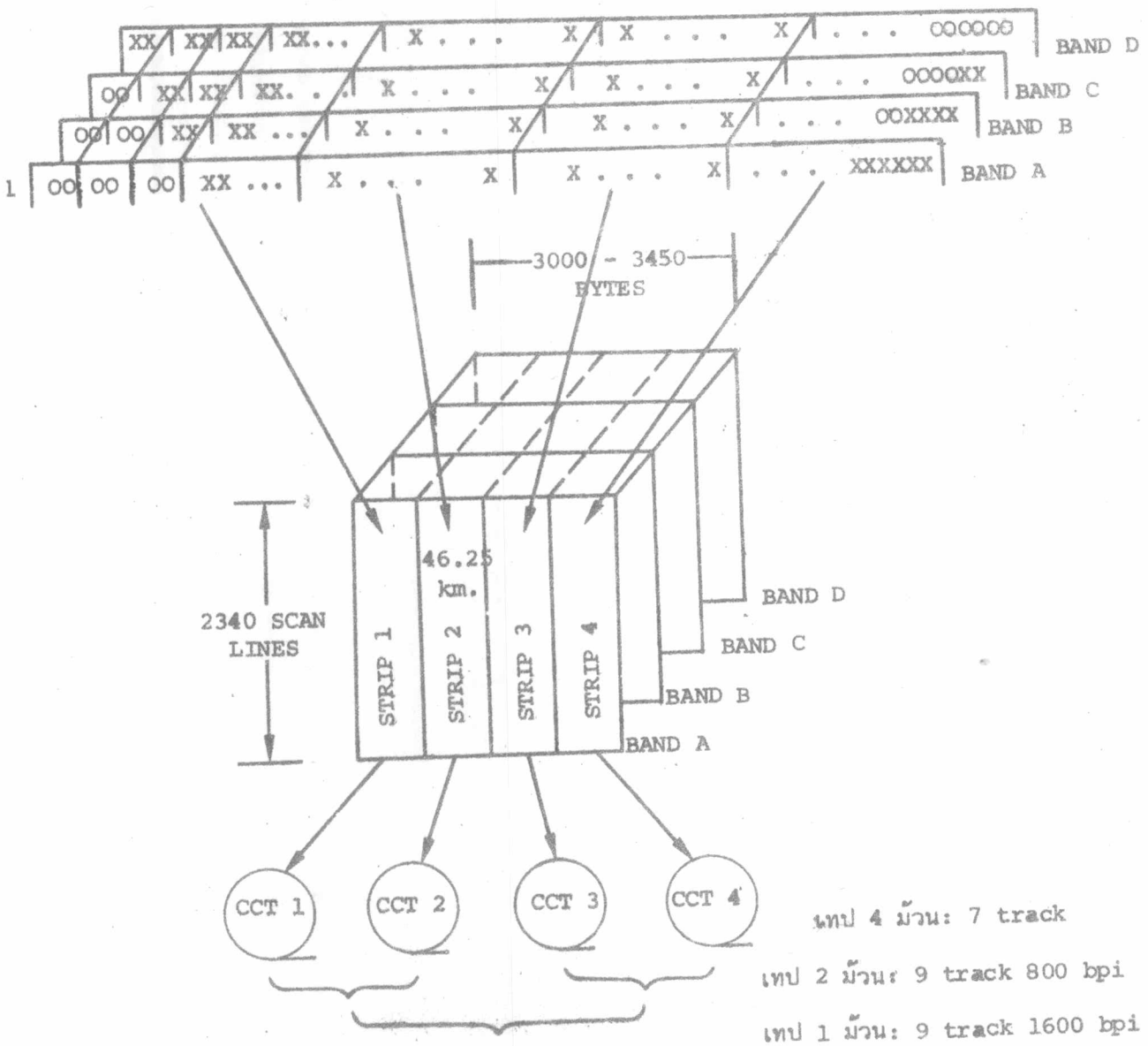
ขั้นตอนการวิเคราะห์และดำเนินการวิจัย

เนื่องจากในประเทศไทยการศึกษาและวิเคราะห์เทปข้อมูลดาวเทียมยังเป็นของใหม่ ดังนั้นในการวิเคราะห์จึงได้นำเอาหลักและแนวทางจากต่างประเทศมาดัดแปลงแก้ไขเพื่อให้เหมาะสมกับการใช้งานในเมืองไทย โดยในช่วงปีที่ผ่านมาได้มีการศึกษาและพัฒนา LIGMALS package ซึ่งเป็น package ที่ใช้งานได้ง่ายและสะดวก และเหมาะกับเครื่องคอมพิวเตอร์ขนาดเล็ก และในวิทยานิพนธ์ฉบับนี้ก็ได้อาศัยแนวทางของ LIGMALS package เป็นรากฐานในการพัฒนา และดัดแปลงแก้ไข ให้เหมาะสมกับงานในวิทยานิพนธ์นี้

ขั้นตอนในการวิเคราะห์และดำเนินงานวิจัยมีดังนี้คือ

1. จัดเรียงข้อมูลใหม่ (Reformatting)

เป็นขั้นตอนการเตรียมข้อมูลใน CCT ให้พร้อมที่จะนำมาใช้งาน



รูปที่ 5 ระบบการเก็บข้อมูลของแต่ละภาพใน CCT

---

[1] X คือ video data

O คือ registration fell characters ซึ่งจะบอกให้ทราบว่าเป็นแบนด์ใดด้วย เช่น แบนด์ที่สี่จะเริ่มด้วย RFC 3 bytes แต่ละ byte มีค่าเป็น FF (ฐาน 16) ซึ่งมีค่าเป็น 256 (ฐาน 10)

เนื่องจากลักษณะการจัดข้อมูล (video data) ที่อยู่ใน CCT ดังกล่าวมาแล้วข้างต้น ถึงแม้ว่าจะสามารถนำมาวิเคราะห์ได้โดยตรงก็ตามแต่จะไม่ให้ความสะดวกเต็มที่ ดังนั้นจึงมีการจัดเรียงข้อมูลเหล่านี้เสียใหม่ ซึ่งทำให้ใช้งานได้ง่ายขึ้น โดยจัดเรียงข้อมูลที่ละจุดของทุก ๆ channel หรือแบนด์ให้อยู่ติดกัน ข้อมูลที่อยู่ใน CCT แต่ละ byte นั้น FORTRAN จะไม่สามารถอ่านโดยตรงได้ จึงได้จัดเรียงข้อมูลเหล่านี้เสียใหม่ให้เป็น 3 byte integer เช่น ค่า 256 เมื่อเขียนเป็น integer จะอยู่ในลักษณะ F2F5F5 ซึ่งนอกจากจะสะดวกต่อการใช้งานต่อไปแล้ว ยังสามารถใช้กับเครื่องคอมพิวเตอร์เครื่องอื่น ๆ ได้ด้วย (transportable) ดังนั้นข้อมูล 1 byte ใน CCT เมื่อจัดเรียงข้อมูลใหม่แล้ว จะขยายเป็น 3 bytes (ดูรูปที่ 6) แล้วเขียนลงในเทปม้วนใหม่

EROS DATA TAPE FORMAT

HEADER RECORD	ANNOTATION RECORD	Pt. 1 ChanA 1byte	Pt. 2 ChanA 1byte	Pt. 1 ChanB 1byte	Pt.2 ChanB 1byte	Pt. 1 ChanC 1byte	Pt. 2 ChanC 1byte	Pt. 1 ChanD 1byte	Pt. 2 ChanD 1byte

OU-LIGMALS TAPE FORMAT

HEADER RECORD	ANNOTATION RECORD	Pt. 1 ChanA 3byte	Pt. 2 ChanB 3byte	Pt. 1 ChanC 3byte	Pt. 2 ChanD 3byte	Pt. 1 ChanA 3byte	Pt. 2 ChanB 3byte	Pt. 1 ChanC 3byte	Pt. 2 ChanD 3byte

รูปที่ 6 เปรียบเทียบระหว่างข้อมูลก่อนและหลังจัดเรียงใหม่

2. การเลือก Bench Mark (หา Coordinating)

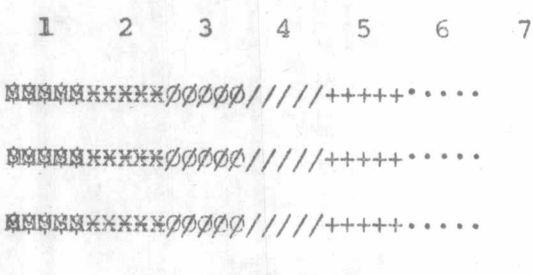
ในวิทยานิพนธ์นี้ เนื่องจากต้องศึกษาข้อมูลเปรียบเทียบเนื้อที่ป่าไม้ในบริเวณเดียวกันแต่ต่างปีกัน ข้อมูลใน CCT ที่ต่างปีกันนี้จะคลุมพื้นที่ต่างกันเล็กน้อย ดังนั้นจึงจำเป็นต้อง

หาจุดพิกัด (Coordinate) ที่จะให้ภาพของทั้งสองปี ทับกันพอดี โดยเลือกพื้นที่ที่มีลักษณะเด่น เช่น แนวแม่น้ำ เกาะ ส่วนเว้าโค้งของชายหาด มาเป็น bench mark

3. การสร้างภาพพิมพ์คอมพิวเตอร์ (Gray Map)

โดยปกติการศึกษาและวิเคราะห์ข้อมูลด้วยเครื่องคอมพิวเตอร์นั้น เราไม่ได้ศึกษาข้อมูลทั้งหมดของภาพแต่ละภาพ แต่จะเลือกศึกษาเฉพาะบริเวณที่ต้องการเท่านั้น ดังนั้นเพื่อให้การวิเคราะห์ข้อมูลในบริเวณที่ต้องการได้สะดวกขึ้นจึงมีการดึงข้อมูลจากเทปที่จัดเรียงข้อมูลใหม่แล้วเฉพาะพื้นที่ที่ต้องการลง disk ก่อน จากนั้นจึงสร้างภาพพิมพ์คอมพิวเตอร์จากข้อมูลใน disk โดยใช้สัญญลักษณ์หรืออักษรต่าง ๆ แทนระดับความเข้มแสง ซึ่งโดยปกติสายตามนุษย์ไม่สามารถแจกแจงความแตกต่างของความเข้มแสงทั้ง 64 ระดับได้ เราจึงแบ่งระดับความเข้มแสงทั้ง 64 ระดับนี้ออกเป็นขั้น ๆ (Level Setting) ส่วนจะแบ่งออกเป็นกี่ขั้นนั้นก็ขึ้นอยู่กับความเหมาะสมและความต้องการของผู้ใช้ด้วย การแบ่งระดับความเข้มแสงออกมากเกินไป ก็จะทำให้การแจกแจงความแตกต่างในแต่ละระดับยากขึ้นด้วย หรือถ้าน้อยเกินไปก็แจกแจงข้อมูลได้ไม่ละเอียด สำหรับในวิทยานิพนธ์นี้จะแบ่งออกเป็น 7 ระดับ แต่ละระดับจะแทนด้วยสัญญลักษณ์ตัวหนึ่ง โดยสัญญลักษณ์ที่ใช้จะมีความเข้มเพิ่มขึ้นหรือลดลงในแต่ละระดับด้วย และถ้าหากจะเน้นสัญญลักษณ์ให้มีความเข้มเพิ่มขึ้นโดยการพิมพ์ซ้ำตัวอักษร (over print) ก็จะช่วยให้สามารถแสดงความแตกต่างในแต่ละระดับได้ดียิ่งขึ้น (ดูรูปที่ 7) จากประสบการณ์ที่ผ่านมาพบว่า ภาพที่พิมพ์โดยใช้เทคนิคการพิมพ์ซ้ำสามารถให้คุณภาพของภาพนี้ได้ดีกว่าแบบไม่พิมพ์ซ้ำ โดยเฉพาะอย่างยิ่งทางด้าน Tone Shadow และ Texture

ปกติการสร้างภาพพิมพ์จะสร้างแบบ negative คือค่าความเข้มแสงที่น้อยจะใช้สัญญลักษณ์ที่เข้มมาก และจะใช้สัญญลักษณ์ที่จางลงมากับค่าความเข้มแสงที่เพิ่มขึ้น เพราะจะทำให้ได้ภาพตรงกับความเป็นจริง



รูปที่ 7 สัญญลักษณ์ที่ใช้พิมพ์ในแต่ละระดับ

การทำ Level Setting นอกจากจะต้องเลือกสัญญลักษณ์ที่ใช้พิมพ์ให้เหมาะสมกับระดับความเข้มแสงแล้ว ความสำคัญอีกอย่างก็คือการกำหนดขอบเขต (boundary) ของแต่ละชั้นว่าจะใช้ค่าจากไหนถึงไหนอีกด้วย ในการกำหนดขอบเขตมีอยู่ 2 วิธีด้วยกันคือ

3.1 การกำหนดเอง (Manual Level Set)

ผู้ใช้จะกำหนดขอบเขตเป็นเท่าไรก็ได้ตามความต้องการหรือความเหมาะสม ซึ่งอาจจะกำหนดให้ทุกระดับมีช่วง (interval) ห่างเท่า ๆ กันก็ได้ แต่โดยปกติพบว่าความเข้มแสงนี้ จะไม่กระจายอยู่ในทุก ๆ ระดับอย่างเท่า ๆ กัน ดังนั้นก็อาจทำ Histogram plot หรือ Densitometry plot ของพื้นที่นั้น ๆ ออกมาดู เพื่อช่วยในการตัดสินใจกำหนดขอบเขตก็ได้

3.2 กำหนดโดยอัตโนมัติ (Automatic Level Set)

โดยปกติแล้วข้อมูลจะกระจายอยู่ในระดับต่าง ๆ อย่างไม่เท่ากันและจะมารวม

กันอยู่ที่จุด ๆ หนึ่ง และที่ระดับความเข้มแสงน้อย ๆ กับมาก ๆ นั้นจะมีข้อมูลอยู่น้อย ทำให้ตั้งสมมติฐานได้ว่าข้อมูลในแต่ละพื้นที่นั้น ๆ มีการกระจายแบบ Normal distribution ได้ และการกำหนดขอบเขตก็พยายามให้มีจำนวนข้อมูลในแต่ละระดับเท่า ๆ กัน โดยอาศัยหลักการนี้ในโปรแกรมเพื่อให้เครื่องคอมพิวเตอร์คำนวณขอบเขตแต่ละค่าให้ จากประสบการณ์พบว่าการกำหนด level ตาม Normal Distribution นั้น นอกจากจะสามารถให้ภาพพิมพ์ที่มีคุณภาพที่ดีแล้ว เรายังสามารถนำเอาข้อมูลภาคพื้นดิน (ground truth) มาช่วยปรับแต่ง level ที่กำหนดจากเครื่องคอมพิวเตอร์เหล่านั้น ยังผลให้ได้ภาพพิมพ์ที่ดีขึ้นตรงตามสภาพแวดล้อมที่แท้จริงด้วย

4. หาค่าสถิติต่าง ๆ (Statistics)

การหาค่าสถิตินี้เพื่อใช้เป็นตัวแทนของข้อมูลประเภทต่าง ๆ ที่ต้องการจำแนก โดยการลง training area ก็คือการกำหนดพื้นที่สี่เหลี่ยมเล็ก ๆ หลาย ๆ พื้นที่ให้คลุมข้อมูลแต่ละประเภทที่ต้องการ และจากการสมมติให้ข้อมูลทั้ง 4 แบนด์ มีการกระจายแบบ Multivariate Normal Distribution ทำให้สามารถคำนวณเพื่อประมาณค่าสถิติอันได้แก่ mean vector และ covariance matrix ของข้อมูลแต่ละประเภทได้ และนำค่าสถิติที่ได้นี้ไปใช้ในการจำแนกภาพต่อไป

ในขั้นนี้นอกจากจะคำนวณค่าสถิติต่าง ๆ แล้ว ยังมีการทดสอบค่าสถิติที่ได้ว่ามีคุณสมบัติตามสมมติฐานหรือไม่ โดยการทำ Histogram plot ของข้อมูลแต่ละประเภท แล้วยังมีการสเก็ตภาพการกระจายข้อมูลเพื่อดูว่า ค่าสถิติเหล่านั้นจะเป็นตัวแทนที่ดีหรือไม่อีกด้วย

5. การลองจำแนกภาพ (Preclassification)

เป็นขั้นการวิเคราะห์เพื่อทดสอบค่าสถิติที่ได้อีกครั้งหนึ่งว่าให้ผลการจำแนกภาพตามต้องการหรือไม่ โดยทดลองจำแนกภาพบริเวณใดบริเวณหนึ่ง (หรือหลายแห่ง) ที่เราพอจะทราบอยู่แล้วว่า ควรให้ผลการจำแนกเป็นเช่นไร หากยังไม่ได้ผลตามต้องการ ก็จะต้องมีการแก้ไข เช่น หาค่าสถิติจาก training area ที่เลือกใหม่ เป็นต้น

6. การจำแนกภาพ (Classification)

เมื่อทดสอบค่าสถิติจนเป็นที่พอใจแล้ว จึงนำค่าสถิติเหล่านั้นมาจำแนกภาพกับข้อมูลทั้งหมดอีกครั้งหนึ่ง

ส่วนที่เกี่ยวกับทฤษฎีและวิธีการประมาณค่า parameters ของการจำแนกภาพแบบ Maximum Likelihood ที่เลือกใช้ในวิทยานิพนธ์นี้จะได้กล่าวถึงโดยละเอียดในบทต่อไป และขั้นตอนการทำงานของวิทยานิพนธ์นี้สามารถสรุปด้วยผังภาพในรูปที่ 8

ลักษณะป่าไม้ และประเภทป่าไม้ที่จะใช้จำแนกภาพ

ป่าไม้ที่ต้องการศึกษานี้ อยู่ในพื้นที่ภาคตะวันออกของประเทศไทย ซึ่งพอจะสรุปภูมิประเทศของภาคตะวันออกได้คือ

ภาคตะวันออกประกอบด้วยจังหวัดทั้งหมด 7 จังหวัด คือ จังหวัดนครนายก ฉะเชิงเทรา ปราจีนบุรี ชลบุรี ระยอง จันทบุรี และ จังหวัดตราด มีเนื้อที่ทั้งหมดประมาณ 36,394 ตร.กม. ลักษณะภูมิประเทศโดยทั่วไปเป็นที่ราบลุ่ม ถัดเข้ามาพื้นที่จะมีลักษณะเป็นลูกคลื่น มีเขาเตี้ยเป็นหย่อม ๆ มีป่าโปร่ง และทุ่งหญ้า ตอนเหนือของภาคเป็นภูเขาสูง ส่วนตอนล่างเป็นชายฝั่งทะเล และลักษณะป่าไม้ที่พบในภาคนี้ มีดังนี้คือ[1,2]

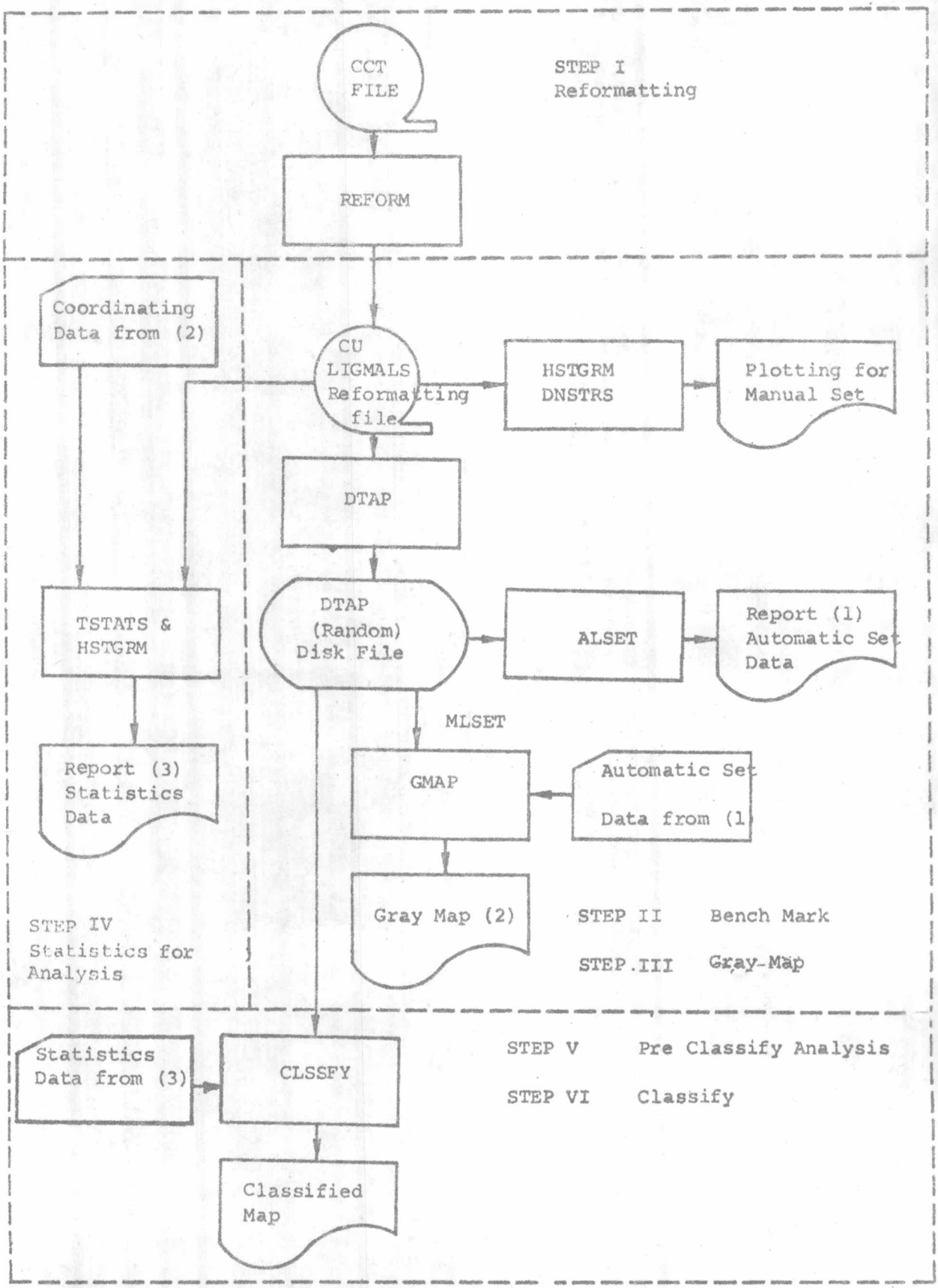
1. ป่าดงดิบ (Tropical Evergreen Forest)

ป่าชนิดนี้มักขึ้นตามบริเวณ หุบเขา เนินเขา พื้นที่ริมแม่น้ำลำธาร หรือบนเขาสูงที่มีความชุ่มชื้นมาก ต้นไม้ที่มีอยู่จะมีลำต้นสูงใหญ่ขึ้นอยู่แน่นทึบ ระเกะระกะ เบียดชิดติดกันจนแสงอาทิตย์ไม่สามารถส่องถึงพื้นได้ในเวลากลางวัน ทำให้ไม้พื้นล่างขึ้นอยู่รก และแน่นทึบ เป็นเหตุให้ป่าชุ่มชื้นอยู่ตลอดเวลา และมีอุณหภูมิต่ำ

---

[1] บุญชนะ กลั่นคำสอน และธงชัย จารุพพัฒน์, "การสำรวจและศึกษาหาความเปลี่ยนแปลงพื้นที่ป่าไม้ในท้องที่ภาคตะวันออกด้วยภาพถ่ายดาวเทียม," หน้า 9

[2] บุญชนะ กลั่นคำสอน, "พืชพรรณธรรมชาติ," (กรุงเทพมหานคร: กรมป่าไม้, 2518), หน้า 7-9



รูปที่ 8 ผังภาพแสดงขั้นตอนการทำงานในวิทยานิพนธ์

ป่าดงดิบมักพบอยู่ทั่วไปในทุกจังหวัดของภาคนี้ และส่วนใหญ่ยังมีความสมบูรณ์ยกเว้นตามบริเวณที่ราบ ซึ่งมักถูกบุกรุก แผ้วถาง

2. ป่าชายเลน (Mangrove Forest)

ขอบขึ้นอยู่ที่ดินเลนริมทะเล ตามฝั่งทะเล และปากแม่น้ำลำคลองที่มีน้ำทะเลขึ้นถึง ซึ่งจะพบในจังหวัดตราด จันทบุรี ระยอง และชลบุรี

3. ป่าเต็งรัง (Dipterocarp Forest)

ขึ้นอยู่บริเวณที่ต่ำตามเนินเขาหรือภูเขาที่เป็นหินชั้นป่าไม้หนาแน่นและไม่สูง สภาพป่าโดยทั่วไปถูกบุกรุกแผ้วถาง ส่วนใหญ่พบในจังหวัดปราจีนบุรี

4. ป่าชายหาด (Beach Forest)

มีขึ้นอยู่ทั่วไปตามชายหาดทรายริมทะเล มีเนื้อที่ไม่มากนักมีสภาพเป็นป่าโปร่ง

5. ป่าพรุ (Fresh Water Swamp Forest)

ขึ้นอยู่ในบริเวณที่ราบต่ำที่มีน้ำจืดท่วม หรือขังตลอดปี มีลักษณะเป็นป่าโปร่ง ต้นไม้ขึ้นอยู่ห่าง ๆ

ส่วนในพื้นที่ที่จะทำการศึกษานั้น (ดูรูปที่ 13) ลักษณะภูมิประเทศทางตอนบนของภาพจะเป็นที่ราบ ในตอนกลางของภาพพื้นที่จะเป็นลูกคลื่นมีแนวภูเขาเป็นหย่อม ๆ เช่น เขาชะมูน เขาชะเอม เขาชะเมา ในบริเวณตอนบนจนถึงตอนกลางนี้ พื้นที่ส่วนใหญ่จะเป็นป่าดงดิบ (ในรูปจะเห็นเป็นสีดำ) ถัดลงมาจากป่าดงดิบจนถึงแนวชายหาดจะเป็นที่ราบลุ่มพื้นที่ส่วนใหญ่จะใช้เพาะปลูกและทำสวน เช่น ปลูกข้าว มันสำปะหลัง สวนยางพารา สวนผลไม้ เป็นต้น ส่วนบริเวณปากแม่น้ำประแสร์ จะเป็นป่าชายเลน ต่อจากแนวชายหาดลงมาจะเป็นทะเล

ลักษณะป่าไม้ดังกล่าวข้างต้นนี้ หากนำมาใช้แยกประเภทข้อมูล ก็จะประสบกับปัญหาบางประการ เช่น ข้อมูลบางประเภทมีอยู่น้อย หรือมีข้อมูลกระจายมากเกินไป ทำให้การหาค่าสถิติจาก Training Area ที่เป็นตัวแทนของข้อมูลประเภทนั้น ๆ ผิดพลาดไปได้ ดังนั้นเพื่อมิให้สูญเสียความสำคัญในการจำแนกประเภท จึงจะแยกประเภทออกอย่างกว้าง ๆ ดังนี้คือ

1. ป่าทึบ ได้แก่ป่าดงดิบ

2. ป่าขายเลน

3. อื่น ๆ คือส่วนที่ไม่ใช่ป่า ซึ่งรวมถึงพื้นที่เพาะปลูกและทำไร่ทำสวน

4. ทะเล

เนื่องจากในพื้นที่บริเวณที่ใช้ศึกษานั้นไม่มีป่าโปร่งจึงได้ตัดป่าโปร่งออกจากการจำแนกข้อมูลในครั้งนี้