การจัดทำดัชนีค้นคืนภาพถ่ายดาวเทียมด้วยการจัดกลุ่มและค่าดัชนีพืชพรรณ

Indexing remote sensing image retrieval using clustering and vegetation indices

ภาวิดา ยะโสวงษ์ และ ชุลีรัตน์ จรัสกุลชัย

ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยเกษตรศาสตร์

Email: pavida.dadarz@gmail.com , fscichj@ku.ac.th

ABSTRACT – Currently, Satellite images has been increasingly and available to download. Additionally, retrieving by latitude and longitude is not sufficiently, a new retrieval method is challenged in this area. Retrieving by specified type of object in remote sensing image such as vegetation, river etc is required. To support this process, an automatic indexing to classify object is need to study. This paper presents techniques of indexing SMMS satellite images retrieval which is a type of multi- spectral images. K-mean clustering technique and vegetation indices are explored with five types of objects. Then indexing techniques are compared between using vegetation index technique and clustering combine with vegetation index technique.

KEY WORDS -- satellite images retrieval; indexing; K-mean; vegetation index

บทคัดย่อ – ปัจจุบันภาพถ่ายดาวเทียมมีปริมาณเพิ่มมากและสามารถดาวโหลดได้ การค้นคืนด้วยการระบุค่าละติดจูด และลองติจูดไม่เพียงพอต่อความต้องการ การค้นคืนในรูปแบบอื่นเป็นปัญหางานวิจัยในสาขานี้ เช่นต้องการทราบวัตถุที่อยู่บนภาพ ไม่ว่าจะเป็น แหล่งน้ำ หรือ พื้นที่เพาะปลูก เป็นต้น งานวิจัยนี้ได้นำเสนอเทคนิคการทำดัชนีคำค้นของภาพถ่ายดาวเทียมSMMS ซึ่งเป็นภาพถ่ายประเภทมัลติสเปกตรัม โดยนำเสนอเทคนิคการจัดกลุ่มแบบ K-mean และใช้ค่าดัชนีพืชพรรณในการเทียบหาค่าสเปกตรัมจากภาพถ่ายเพื่อใช้ในการจำแนกประเภทวัตถุออกเป็น 5 ประเภท จากนั้นจัดทำดัชนีคำค้นด้วยค่าดัชนีพืชพรรณเทียบกับการจัดกลุ่มผนวกกับการใช้ค่าดัชนีพืชพรรณร่วมด้วย

คำ สำ คัญ --การค้นคืนภาพถ่ายดาวเทียม; ดัชนีคำค้น; K-mean; ค่าดัชนีพืชพรรณ

1. บทนำ

ในปัจจุบันข้อมูลภาพถ่ายดาวเทียมมีมากขึ้นทุกวัน การค้นคืนภาพถ่ายดาวเทียมโดยทั่วไปใช้เงื่อนไขทางภูมิศาสตร์ เช่น ละติจูด ลองติจูด หรือวันที่ถ่ายภาพ อย่างไรก็ตามข้อมูลในภาพถ่ายดาวเทียม มีได้หลากหลาย เช่น เป็นพื้นที่เพาะปลูก พื้นที่ที่อยู่อาศัย พื้นที่หนองน้ำ หรือ ภูเขา เป็นต้น ความต้องการในการค้นคืนดังกล่าวไม่สามารถระบุคำค้นด้วยเงื่อนไขทางภูมิศาสตร์ ถ้าค้นค้นในรูปแบบเดิมโดยการสุ่มเลือก ผู้ใช้จะต้องเสียเวลาและไม่สามารถค้นคืนในสิ่งที่ต้องการได้อย่างรวดเร็วและตรงความต้องการ

ในการจัดทำระบบค้นคืนสำหรับภาพถ่ายดาวเทียมนี้จำเป็นต้องมีการจัดทำดัชนีคำค้นเพื่อใช้ในการค้นคืนภาพ ซึ่งในการจัดทำดัชนีคำค้นเราจำเป็นต้องทราบข้อมูลสเปกตรัมไลบรารีของพื้นที่หรือวัตถุในภาพเพื่อที่จะนำมาจำแนกประเภทวัตถุได้ แต่ด้วยข้อจำกัดของภาพถ่ายดาวเทียมที่มีอยู่เป็นภาพถ่ายที่มีข้อมูลสเปกตรัมไลบรารีของแต่ละวัตถุน้อย

ในงานวิจัยนี้จึงได้เสนอเทคนิคการจัดทำดัชนีคำค้นแบบอัตโนมัติ โดยใช้ขั้นตอนวิธีแบบกฎเกณฑ์ โดยใช้ค่าดัชนีพืชพรรณในการจำแนกวัตถุและการจัดกลุ่มแบบอัตโนมัติผนวกกับค่าดัชนีพืชพรรณ เพื่ออำนวยความสะดวกให้กับ

ผู้ใช้งานที่ไม่ทราบค่า ละติจูด ลองติจูด แต่ต้องการทราบประเภทพื้นที่ที่ต้องการค้น เช่น ต้องการค้นหาพื้นที่แหล่งน้ำ ป่าไม้ เป็นต้น ในระบบการค้นคืนภาพถ่ายดาวเทียมจะทำการค้นคืนโดยการป้อนข้อมูลที่ผู้ใช้ต้องการ จากนั้นระบบจะทำการดึงข้อมูลที่ผู้ใช้ต้องการออกมาจากฐานข้อมูล ซึ่งภาพดังกล่าวเป็นภาพถ่ายดาวเทียมประเภทมัลติสเปกตรัมของดาวเทียม SMMS [1] ทั้งนี้ในฐานข้อมูลจะมีการจัดทำข้อมูลสเปกตรัมไลบรารีของบริเวณพื้นที่บางประเภทไว้แล้ว เช่น พื้นที่ป่าไม้ แหล่งน้ำ แหล่งเพาะปลูก เป็นต้น

งานวิจัยเสนอการจัดทำดัชนีค้นคืนสองแบบคือใช้ข้อมูลดัชนีพรรณพืชมาช่วยในการวิเคราะห์พื้นที่ภาพถ่ายดาวเทียม ซึ่งค่าดัชนีพืชพรรณเป็นค่าที่ใช้วัดมวลชีวภาพที่ปกคลุมพื้นดิน โดยวัดจากการสะท้อนของสองช่วงคลื่น คือ คลื่นสีแดง และคลื่นอินฟราเรดใกล้ ซึ่งคลอโรฟิลล์ในพืชสีเขียวจะดูดซับในคลื่นสีแดง[2] และสะท้อนออกมามากในคลื่นอินฟราเรดใกล้เพราะฉะนั้นพื้นที่ที่มีพืชพรรณมากจะมีคลื่นสีแดงน้อยและคลื่นอินฟราเรดใกล้จะมีค่าสูง ค่าดัชนีพืชพรรณจะอยู่ในช่วง -1 ถึง 1 โดยพื้นที่ที่มีพืชพรรณสีเขียวอยู่หนาแน่นจะมีค่าเข้าใกล้ 1 เช่น พื้นที่ป่าไม้ ส่วนค่าที่ติดลบหรือเข้าใกล้ -1 มากที่สุด คือพื้นที่ที่มีพืชพรรณอยู่น้อยมากหรือไม่มีเลยเช่น แม่น้ำหรือทะเล เป็นต้น

อีกวิธีที่นำมาใช้เพื่อรองรับข้อมูลวัตถุในภาพที่ไม่มีข้อมูลดัชนีพรรณพืช คือ การจัดกลุ่มแบบ K-mean เนื่องจากการจัดกลุ่มแบบ K-mean มีข้อดีคือ ง่าย ใช้เวลาน้อย และจะใช้กับข้อมูลที่ไม่มีผู้สอนระบบ(unsupervised learning)[3] จากนั้นเปรียบเทียบค่าความเหมือน(similarity)ของแต่ละภาพในฐานข้อมูลกับค่าสเปกตรัมไลบรารี หากมีค่าใกล้เคียงกับสเปกตรัมไลบรารีใด ก็จะทำการเก็บข้อมูลพื้นที่เหล่านั้นลงในฐานข้อมูล เมื่อผู้ใช้ต้องการค้นคืนก็จะทำการดึงภาพที่ผู้ใช้ต้องการมาแสดงบนเว็บไซต์ วิธีการค้นคืนเช่นนี้ จะอำนวยความสะดวกให้กับผู้ใช้ที่ไม่รู้ค่า ละติจูด ลองติจูดได้ และสามารถแสดงสีเฉพาะพื้นที่ในภาพที่ผู้ใช้เลือกได้ จากนั้นทำการเปรียบเทียบผลลัพธ์การแสดงผลพื้นที่ในภาพของวิธีการหาค่าดัชนีพืชพรรณ NDVI และวิธีการจัดกลุ่มแบบ K-mean

2. งานวิจัยที่เกี่ยวข้อง

การค้นคืนภาพถ่าย สามารถค้นคืนได้ 2 วิธี คือ การค้นคืนโดยใช้ชื่อภาพ (filename) แต่ข้อมูลภาพในปัจจุบันมีขนาดใหญ่ขึ้นและการค้นคืนด้วยชื่อเป็นข้อมูลที่ไม่ใช่ลักษณะจริงๆของภาพทำให้วิธีการนี้ไม่เหมาะสม อีกทั้งคำจำกัดความในการตั้งชื่อภาพของแต่ละคนมีความแตกต่างกัน อาจทำให้ได้ภาพที่ไม่ตรงกับต้องตามความต้องการ,การค้นคืนโดยใช้ภาพถ่าย (query image) หรือใช้เนื้อหาของภาพถ่ายแบบ CBIR (Content-Based Image Retrieval)[4] ซึ่งเป็นการค้นคืนโดยดึงลักษณะเฉพาะของภาพออกมา เช่น สี รูปร่างและพื้นผิวของวัตถุภายในภาพ เป็นต้น และใช้ข้อมูลเหล่านี้เป็นดัชนีภาพจัดเก็บลงฐานข้อมูล เมื่อต้องการค้นคืนภาพ ผู้ใช้จะทำการเลือกภาพคำขอ จากนั้นระบบจะทำการดึงเอาลักษณะเฉพาะของภาพคำขอมาทำการเปรียบเทียบกับลักษณะเฉพาะของภาพที่อยู่ในฐานข้อมูล เพื่อค้นหาภาพที่มีลักษณะเฉพาะคล้ายกับภาพสอบถามมากที่สุดออกมา เป็นต้น

Li Wei และ Wang Weihong และ Lu Feng [5] ได้เสนอโครงสร้างโมเดลหลักๆในการค้นคืนภาพถ่ายดาวเทียมเป็นลำดับชั้น เรียงลำดับจากต่ำไปสูง ทั้งหมด 5 ลำดับชั้น ดังนี้

1. Metadata layer - เป็นคุณสมบัติพื้นฐานของภาพถ่ายดาวเทียม เช่น ค่าละติจูด ลองติจูด ช่วงคลื่น,ความละเอียดของพื้นที่,เวลาในการถ่ายภาพ เป็นต้น ข้อมูลด้านนี้ใช้สำหรับการค้นคืนแบบคำค้น(keyword search) เช่น ค้นคืนโดยผู้ใช้ระบุวันที่ในการถ่ายภาพ หรือ ระบุละติจูด ลองติจูดของพื้นที่ [1] เป็นต้น

2. Original pixel layer - เป็นข้อมูลพิกเซลของภาพที่สะท้อนคุณสมบัติของวัตถุบนพื้นดิน เราสามารถค้นคืนภาพโดยการเปรียบเทียบแต่ละพิกเซลในภาพที่ต้องการค้นหากับภาพเป้าหมาย แต่การเปรียบเทียบเช่นนี้ อาจมีข้อจำกัดด้านความละเอียดเชิงพื้นที่ของภาพถ่าย ซึ่งถ่ายภาพควรจะมีความละเอียดเท่ากันหรือเป็นภาพถ่ายจากกล้องดวงเดียวกัน เพื่อป้องกันความผิดพลาด เป็นต้น

3. Image feature layer - เป็นรายละเอียดของภาพ เช่น สี พื้นผิว รูปร่าง ความระเอียดเชิงพื้นที่ และคุณสมบัติในการมองเห็นอื่นๆ เป็นต้น เราสามารถค้นคืนภาพโดยการคำนวณความเหมือนระหว่างภาพที่ต้องการค้นคืนกับภาพในฐานข้อมูล เช่น Hong Liu และ Xiaohong Yu [6] เสนอการจัดทำแอพลิเคชันสำหรับการค้นคืนภาพถ่าย โดยทำการจัดกลุ่มรูปภาพในฐานข้อมูลแบบ K-mean และวิเคราะห์ค่าความเหมือนของแต่ละภาพในฐานข้อมูลก่อน เมื่อผู้ใช้ใส่ภาพที่ต้องการค้นเข้ามาระบบจะนำภาพมาสกัดและไปเปรียบเทียบกับค่าความเหมือนกับภาพในฐานข้อมูล และแสดงผลการค้นคืนของภาพ

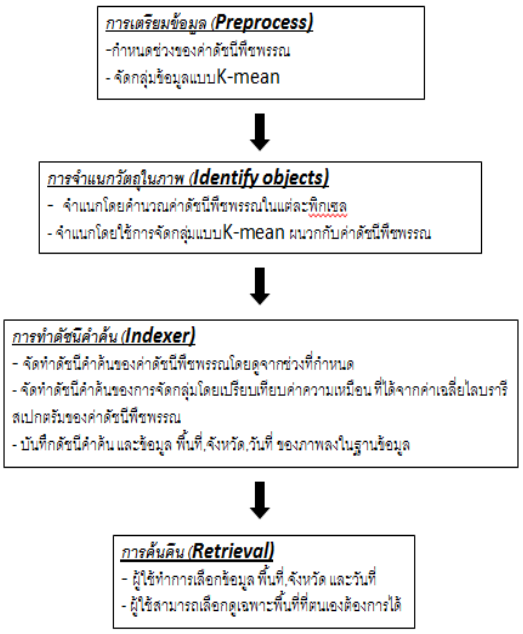
ให้ผู้ใช้ , Priti maheshwary [7] เสนอการค้นคืนภาพถ่ายดาวเทียมแบบมัลติสเปกตรัมความละเอียด 23.5 เมตร โดยมีการจัดทำฐานข้อมูลพื้นที่ของภาพถ่ายดาวเทียมไว้ 3 กลุ่ม คือ ทุ่งหญ้า,แหล่งน้ำ,พืชพรรณ และจะวิเคราะห์ภาพที่นำมาค้นด้วยค่าสี พื้นผิว และค่าดัชนีพืชพรรณ จากนั้นวิเคราะห์ค่าความเหมือนของภาพที่นำมาค้นกับภาพในฐานข้อมูล และดึงภาพที่มีความเหมือนกับภาพในฐานข้อมูลมากที่สุดออกมาตามความต้องการของให้ผู้ใช้ และ Jayant Mishra Anubhav Sharma และ Kapil Chaturvedi [8] เสนอการเพิ่มประสิทธิภาพการค้นคืนภาพถ่ายแบบ Content-based image retrieval(CBIR) เนื่องจากวิธี CBIR ซึ่งมีการเปรียบเทียบภาพโดยใช้คุณลักษณะของสีใช้เวลานานและยาก จึงมีการเพิ่มประสิทธิภาพโดยอาศัยการทำฮีสโทแกรมของสีร่วมกับการจัดกลุ่มแบบ K-mean มาช่วยในการจัดกลุ่มภาพที่มีสีคล้ายกัน ทำให้การเปรียบเทียบทำได้รวดเร็วและง่ายขึ้น เป็นต้น

4. Geographical feature layer - เป็นข้อมูลเชิงพื้นที่ ได้จากการรวมกันของข้อมูลรูปภาพและข้อมูลเวกเตอร์เชิงพื้นที่ เราสามารถดึงข้อมูลเชิงพื้นที่ เช่น รูปร่าง ขนาด ระยะห่างระว่างวัตถุ ออกมาได้โดยใช้การค้นคืนในข้อมูลเวกเตอร์เชิงพื้นที่ เช่น จากงานวิจัยที่ Li Wei และ Wang Weihong และ Lu Feng ได้ทำคือการค้นคืนภาพถ่ายดาวเทียมโดยอาศัยคุณสมบัติทางภูมิศาสตร์(Geographical feature) คุณสมบัติทางความหมาย (Semantic feature) และการจับคู่คุณสมบัติของรูปภาพ(Image feature matching) โดยใช้ตรรกศาสตร์ เช่น AND, OR, NOT ในการวิเคราะห์คำค้นที่ผู้ใช้พิมพ์เข้ามา เพื่อให้ได้ผลลัพธ์ในการค้นคืน ซึ่งผลการค้นคืนสามารถค้นคืนวัตถุในภาพได้ เช่น ต้องการค้นคืนบริเวณที่อยู่อาศัยที่ใกล้แหล่งน้ำไม่เกิน 1 กิโลเมตร เป็นต้น

5. Semantic layer – มีการเก็บข้อมูลที่เป็นลักษณะเฉพาะของแต่ละพื้นที่ในภาพเอาไว้ ซึ่งเราสามารถค้นคืนภาพได้โดยอาศัยการจับคู่หรือใช้กลไกทางความรู้ดึงข้อมูลเหล่านั้นออกมา ยกตัวอย่างเช่น สามารถค้นคืนวัตถุในภาพที่ประเภทการใช้ที่ดินบริเวณนั้นเป็นที่สำหรับการใช้ถนนได้ เป็นต้น

3. วิธีการนำเสนอ

ขั้นตอนในการจัดทำดัชนีคำค้นในงานวิจัยนี้แบ่งคำค้นตามพื้นที่หรือวัตถุที่ได้จากค่าดัชนีพืชพรรณคือ แบ่งเป็น 5 กลุ่มได้แก่ พื้นที่ป่าไม้ แหล่งน้ำ แหล่งที่อยู่อาศัย พื้นดินแห้งแล้ง และพื้นที่เพาะปลูก โดยใช้ค่าดัชนีพืชพรรณเพื่อใช้ในการระบุพื้นที่ในภาพ และใช้อัลกอริธึม K-mean ในการเปรียบเทียบค่าความเหมือนของแต่ละกลุ่มในภาพเพื่อระบุหาพื้นที่ ดังรูปที่ 1. มีขั้นตอนการทำงานดังนี้



รูปที่ 1. แสดง prototype ของการค้นคืนภาพถ่ายดาวเทียมโดยการจำแนก

3.1 การจัดเตรียมข้อมูลเพื่อการจำแนกวัตถุ

ด้วยข้อจำกัดในการจัดหาไลบรารีสเปกตรัมสำหรับแต่ละวัตถุยังไม่มีการนำเสนอและเผยแพร่ให้ใช้งานสาธารณะ งานวิจัยนี้จึงใช้ค่าดัชนีพืชพรรณในการการจำแนกวัตถุเบื้องต้น 5 วัตถุด้วยกัน ซึ่งค่าดัชนีพืชพรรณ คือ ค่าที่ใช้วัดมวลชีวภาพที่ปกคลุมพื้นดินของภาพถ่ายดาวเทียม โดยวัดจากการสะท้อนของสองช่วงคลื่น คือ คลื่นสีแดง และคลื่นอินฟราเรดใกล้[9] มีค่าระหว่าง –1 ถึง 1 ซึ่งบริเวณพื้นที่ที่มีพืชพรรณสีเขียวมาก ค่าดัชนีพืชพรรณจะสูงและเข้าใกล้ 1 เช่น ป่าไม้ ส่วนบริเวณพื้นที่ที่ไม่มีพืชพรรณสีเขียวอยู่เลยหรือมีอยู่น้อยมาก ค่าดัชนีพืชพรรณจะเข้าใกล้ -1 เช่น แหล่งน้ำ ทะเล เป็นต้น

ขั้นตอนในการจัดเตรียมไลบรารีสำหรับจำแนกวัตถุ 5 ประเภทด้วยค่าดัชนีพืชพรรณจะทำการนำข้อมูลในภาพถ่ายดาวเทียมมาหาค่าดัชนีพืชพรรณในแต่ละพิกเซลเพื่อระบุหากลุ่มพื้นที่ที่ต้องการในภาพถ่ายดาวเทียม โดยระบบจะนำช่วง

คลื่นสีแดงและช่วงคลื่นอินฟราเรดใกล้ไปคำนวณหาค่าดัชนีพืชพรรณ ดังสมการที่1.

$$NDVI = (NIR-RED)/(NIR+RED) \quad (1)$$

เมื่อ NIR = ช่วงคลื่นอินฟราเรดใกล้ และ RED = ช่วงคลื่นสีแดง

จากนั้นเปรียบเทียบค่าที่คำนวณได้กับช่วงในฐานข้อมูล หากค่าที่คำนวณได้อยู่ในช่วงที่กำหนด ก็จะถูกบันทึกพื้นที่นั้นลงในฐานข้อมูล จะทำให้ทราบว่า ในภาพนั้นมีพื้นที่ใดบ้าง โดยกำหนดช่วงดัชนีพืชพรรณของแต่ละกลุ่มในฐานข้อมูล ดังตารางที่1.

ตารางที่ 1. แสดงช่วงของค่าดัชนีพืชพรรณ NDVI

ช่วงดัชนีพืชพรรณ	ความหมาย
-1 < NDVI <= -0.1	ช่วงที่มีพืชพรรณสีเขียวน้อยมากหรือไม่มีเลย ได้แก่ พื้นที่แหล่งน้ำ
-0.06 < NDVI <= -0.03	ช่วงของพื้นที่แหล่งที่อยู่อาศัย
0.1 < NDVI <= 0.2	ช่วงของพื้นที่ดินแห้งแล้ง
0.3 < NDVI <= 0.59	ช่วงของพื้นที่เพาะปลูก
0.6 < NDVI <= 1	ช่วงที่มีพื้นที่ที่มีพืชพรรณสีเขียวหนาแน่น ได้แก่ พื้นที่ป่าไม้

แต่ละช่วงของค่าดัชนีพืชพรรณในตารางที่1. ได้ทำการกำหนดขึ้นมาโดยทดสอบกับภาพถ่ายดาวเทียม SMMS เพื่อดูว่าแต่ละช่วงที่เราได้กำหนดมานั้นตรงกับบริเวณพื้นที่ที่เราต้องการหรือไม่โดยอ้างอิงและปรับลดจากงานวิจัยอื่น[9] หากค่าแต่ละช่วงยังไม่ครอบคลุมบริเวณที่ต้องการดีพอก็ทำการปรับค่าช่วงบริเวณนั้นเรื่อยๆจนกว่าค่าดัชนีพืชพรรณที่กำหนดจะเหมาะสมที่สุดกับภาพถ่ายดาวเทียม SMMS และจึงได้ตั้งช่วงดัชนีพืชพรรณเหล่านี้เป็นเกณฑ์ในการทำงานวิจัย

3.2 การจัดกลุ่ม

ในการจัดกลุ่มภาพถ่ายดาวเทียมเพื่อจำแนกหาพื้นที่หรือวัตถุในภาพ จำเป็นต้องมีตัวแทนของไลบรารีสเปกตรัมของแต่ละวัตถุเพื่อใช้ในการจำแนกหาวัตถุในภาพ และเนื่องจากภาพถ่ายดาวเทียม SMMS ที่เราต้องการระบุพื้นที่หรือวัตถุนี้เป็นข้อมูลแบบไม่มีผู้สอนระบบทำให้เราไม่ทราบค่าสเปกตรัมไลบรารีและไม่สามารถจำแนกวัตถุได้ ดังนั้นเราจึงได้จัดทำดัชนีคำค้นด้วยการจัดกลุ่มแบบ K-mean ผนวกกับการใช้ค่าดัชนีพืชพรรณที่ได้จากการกำหนดช่วงในข้อ 3.1 มาหาค่าตัวแทนหรือค่าเฉลี่ยของแต่ละวัตถุ เพื่อใช้ในการจำแนกพื้นที่หรือวัตถุในภาพ ยกตัวอย่างเช่น หากคำนวณค่าดัชนีพืชพรรณของแต่ละพิกเซลออกมาและนำพิกเซลนั้นมาเปรียบเทียบกับค่าดัชนีพืชพรรณที่กำหนดในฐานข้อมูลแล้วพบว่าตรงกับกลุ่มของแหล่งน้ำ ก็จะนำค่าในแต่ละแบนด์ (แบนด์สีเขียว น้ำเงิน แดง และอินฟราเรดใกล้) ของพิกเซลที่เป็นแหล่งน้ำมาหาค่าตัวแทนหรือค่าเฉลี่ยเพื่อกำหนดเป็นสเปกตรัมไลบรารีของข้อมูลแหล่งน้ำ จากนั้นทำการบันทึกสเปกตรัมไลบรารีนั้นลงบนฐานข้อมูล เป็นต้น

3.3 การจัดทำคำค้นให้กับภาพ

เมื่อจัดทำสเปกตรัมไลบรารีของแต่ละพื้นที่หรือวัตถุในภาพเรียบร้อยแล้ว จากนั้นนำตัวแทนของแต่ละกลุ่มมาคำนวณหาค่าความเหมือนกับสเปกตรัมไลบรารีของพื้นที่ที่ได้จัดทำขึ้น ซึ่งสมการที่ใช้ในการเปรียบเทียบหาความเหมือน คือ Euclidean distance ดังสมการที่2.

$$D_{(X,Y)} = \sqrt{(X_{red} - Y_{red})^2 + (X_{green} - Y_{green})^2 + (X_{blue} - Y_{blue})^2 + (X_{nir} - Y_{nir})^2)} \quad (2)$$

จากสมการที่ (2) แสดงค่า

Xred = ค่าสเปกตรัมไลบรารีของช่วงคลื่นสีแดงของพื้นที่ใดๆที่เรากำหนด

Yred = ค่าสเปกตรัมเฉลี่ยของช่วงคลื่นสีแดงในแต่ละกลุ่ม

Xgreen = ค่าสเปกตรัมไลบรารีของช่วงคลื่นสีเขียวของพื้นที่ใดๆที่เรากำหนด

Ygreen = ค่าสเปกตรัมเฉลี่ยของช่วงคลื่นสีเขียวในแต่ละกลุ่ม

Xblue = ค่าสเปกตรัมไลบรารีของช่วงคลื่นสีน้ำเงินของพื้นที่ใดๆที่เรากำหนด

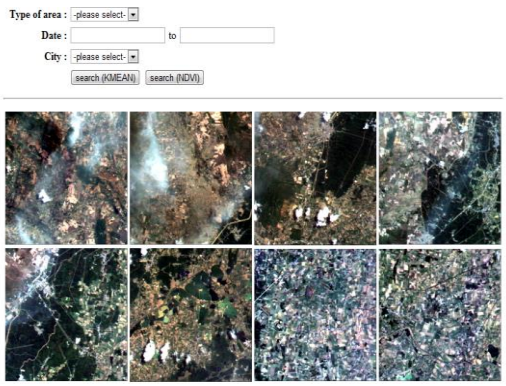
Yblue = ค่าสเปกตรัมเฉลี่ยของช่วงคลื่นสีน้ำเงินในแต่ละกลุ่ม

Xnir = ค่าสเปกตรัมไลบรารีของช่วงคลื่นสีอินฟราเรดใกล้ของพื้นที่ใดๆที่เรากำหนด

Ynir = ค่าสเปกตรัมเฉลี่ยของช่วงคลื่นสีอินฟราเรดใกล้ในแต่ละกลุ่มการจัดทำคำค้นให้กับภาพ

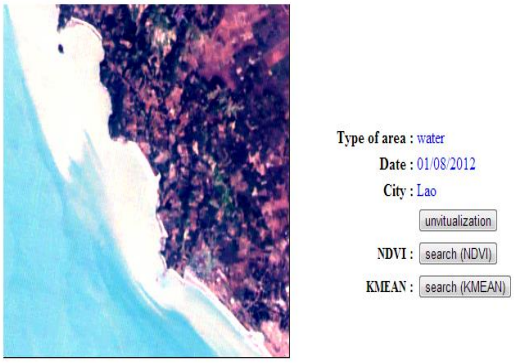
4. ผลการทดลอง

การทดลองนี้ใช้ภาพถ่ายดาวเทียม SMMS ประเภทมัลติสเปกตรัม มีทั้งหมด 4 ช่วงคลื่น คือ ช่วงคลื่นสีแดง,สีเขียว,สีน้ำเงิน, และอินฟราเรดใกล้ ความละเอียดภาพ 30 เมตรต่อจุด โดยภาพที่นำมาใช้ในการทดลองมีขนาด 350x321 พิกเซล มีการจัดเก็บสเปกตรัมไลบรารีของกลุ่มพื้นที่ที่กำหนดไว้จำนวน 5 กลุ่ม คือ พื้นที่ป่าไม้ แหล่งน้ำ แหล่งที่อยู่อาศัย พื้นดินแห้งแล้ง และพื้นที่เพาะปลูก เป็นต้น งานวิจัยนี้ได้ทดลองค้นคืนและจำแนกวัตถุบนภาพถ่ายดาวเทียมโดยใช้ค่าดัชนีพืชพรรณเทียบกับการจัดกลุ่มผนวกกับการใช้ค่าดัชนีพืชพรรณ ดังรูปที่ 2-9.

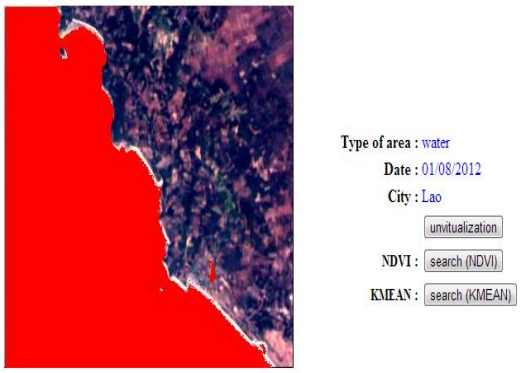


รูปที่ 2. แสดงการค้นคืนภาพถ่ายดาวเทียม โดยผู้ใช้สามารถใส่เงื่อนไขในการค้นคืน 3 ประเภท ได้แก่ ประเภทพื้นที่ ช่วงเวลา และ จังหวัด เป็นต้น

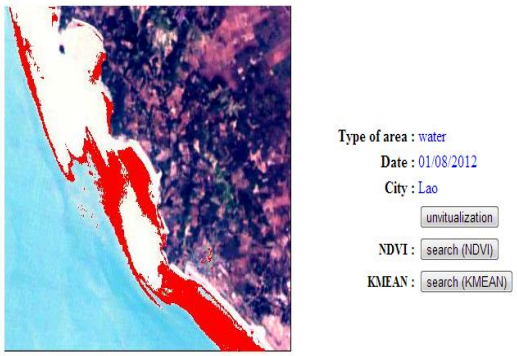
รูปที่ 3. แสดงผลการค้นคืนภาพถ่ายดาวเทียม โดยผู้ใช้เลือกเงื่อนไขในการค้นคืนคือ ต้องการค้นคืนภาพที่มีพื้นที่ประเภทแหล่งน้ำ ระหว่างวันที่ 1ธันวาคม 2011 ถึง วันที่ 31 มกราคม 2012



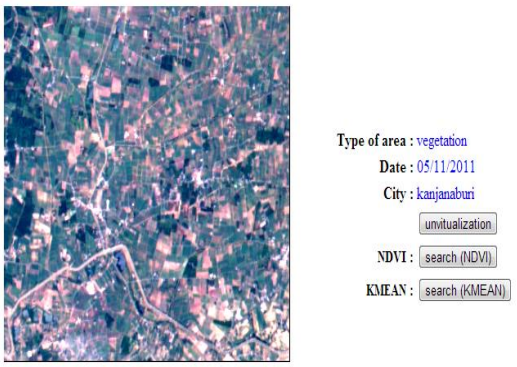
รูปที่ 4. ภาพก่อนแสดงสีเฉพาะพื้นที่แหล่งน้ำ: แสดงภาพที่มีพื้นที่ประเภทแหล่งน้ำตามที่ผู้ใช้ใส่เงื่อนไขในการค้นคืนในรูปที่3. ซึ่งผู้ใช้สามารถเลือกได้ว่าจะแสดงสีของพื้นที่น้ำโดยใช้วิธีการดัชนีพืชพรรณ NDVI หรือวิธีการจัดกลุ่มแบบ K-mean



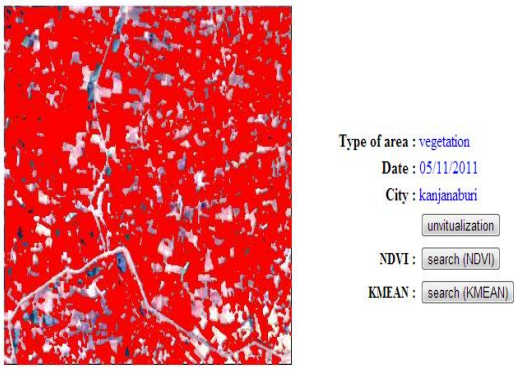
รูปที่ 5. แสดงสีเฉพาะพื้นที่แหล่งน้ำ โดยใช้ค่าดัชนีพืชพรรณ NDVI ในการแสดงผล



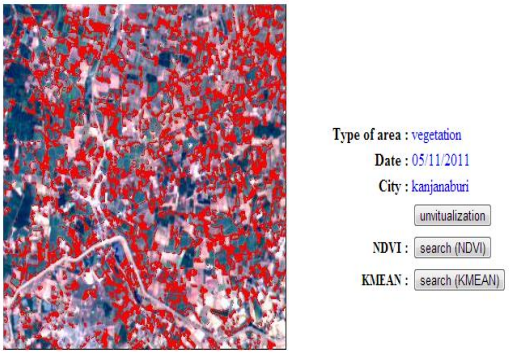
รูปที่ 6. แสดงสีเฉพาะพื้นที่แหล่งน้ำ โดยใช้การจัดกลุ่มแบบ K-mean ในการแสดงผล



รูปที่ 7. ภาพก่อนแสดงสีเฉพาะพื้นที่เพาะปลูก : แสดงภาพที่มีพื้นที่เพาะปลูกตามที่ผู้ใช้ใส่เงื่อนไขในการค้นคืน



รูปที่ 8. แสดงสีเฉพาะพื้นที่เพาะปลูก โดยใช้ค่าดัชนีพืชพรรณ NDVI ในการแสดงผล

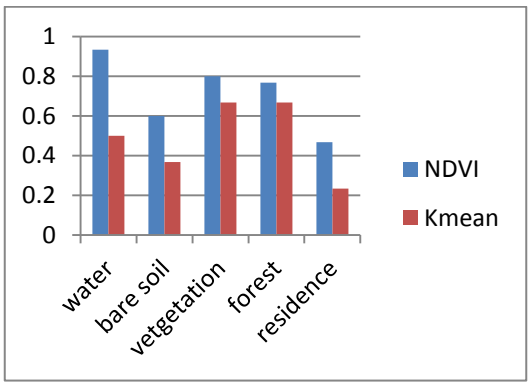


รูปที่ 9. แสดงสีเฉพาะพื้นที่เพาะปลูก โดยใช้การจัดกลุ่มแบบ K-mean ในการแสดงผล

จากการวิเคราะห์ภาพที่ 5-6. แสดงการค้นคืนโดยจำแนกพื้นที่ที่เป็นแหล่งน้ำพบว่า การใช้ค่าดัชนีพืชพรรณสามารถจำแนกพื้นที่ที่เป็นแหล่งน้ำได้ครอบคลุมดีกว่าการจัดกลุ่ม ส่วนการจัดกลุ่มไม่สามารถจำแนกได้ทั้งหมด อาจเป็นเพราะว่าข้อมูลของพื้นที่แหล่งน้ำมีลักษณะแตกต่างกัน เช่น น้ำลึก น้ำตื้น เป็นต้น ดังนั้นการใช้การจัดกลุ่มในการจำแนกพื้นที่อาจมีประโยชน์สำหรับผู้ที่ต้องการศึกษาเชิงลึกทางด้านนิเวศวิทยา

5. การวัดประสิทธิภาพ

ได้ทำการประเมินความถูกต้องของการจำแนกพื้นที่ของ 2 วิธีการ คือ วิธีการหาค่าดัชนีพืชพรรณ และวิธีการจัดกลุ่มแบบ K-Mean โดยทำการสุ่มเลือกพิกเซลในภาพออกมาและทำการเปรียบเทียบกับแต่ละวิธีการว่า สามารถจำแนกพื้นที่นั้นๆได้ถูกต้องหรือไม่ ซึ่งแสดงค่าความถูกต้องของการจำแนกได้ดังรูปที่10.



รูปที่ 10. แสดงค่าความถูกต้องของการจำแนกวัตถุในภาพถ่ายโดยใช้ค่าดัชนีพืชพรรณและการจัดกลุ่ม

จากผลการทดลองพบว่าการใช้ค่าดัชนีพืชพรรณสามารถจำแนกวัตถุได้ดังนี้ สามารถจำแนกพื้นที่แหล่งน้ำได้ 93% พื้นดินแห้งแล้ง 60% พื้นที่เพาะปลูก 80% ป่าไม้ 77% และแหล่งที่อยู่อาศัย 47% การจัดกลุ่มK-mean ผนวกกับค่าดัชนีพืชพรรณสามารถจำแนกวัตถุได้ดังนี้ สามารถจำแนกพื้นที่แหล่งน้ำได้ 50% พื้นดินแห้งแล้ง 37% พื้นที่เพาะปลูก 67% ป่าไม้ 67% และแหล่งที่อยู่อาศัย 23% จากการประเมินความถูกต้องพบว่าค่าดัชนีพืชพรรณสามารถจำแนกกลุ่มพื้นที่แหล่งน้ำได้ดีที่สุด รองลงมาคือ กลุ่มของพื้นที่เพาะปลูกและพื้นที่ป่าไม้ ส่วนกลุ่มพื้นที่ที่มีมีความถูกต้องจากการประเมินผลน้อยที่สุดคือพื้นที่แหล่งที่อยู่อาศัย เนื่องจากค่าดัชนีพืชพรรณที่นำมาใช้จำแนกพื้นที่แหล่งที่อยู่อาศัยมีค่าสเปกตรัมใกล้เคียงกับพื้นที่แหล่งน้ำ เมฆ เป็นต้น จึงทำให้การจำแนกออกมาได้ไม่ดีนัก ส่วนการนำการจัดกลุ่มแบบK-mean มาใช้ในการจำแนกพบว่าพื้นที่เพาะปลูกและพื้นที่ป่าไม้มีความถูกต้องมากที่สุด แต่ยังดีไม่เท่ากับการประเมินความถูกต้องโดยเทียบค่าดัชนีพืชพรรณในแต่ละพิกเซลเนื่องจาก ในการจัดกลุ่มอาจทำให้ข้อมูลบางส่วนหายไปเนื่องจากต้องใช้ค่าเฉลี่ยเป็นไลบรารีสเปกตรัมเพื่อนำไปคำนวณหากลุ่มพื้นที่ ดังนั้นผลลัพธ์ที่ได้จึงมีความถูกต้องลดลง

4. บทสรุป

งานวิจัยนี้นำเสนอการจัดกลุ่มภาพถ่ายดาวเทียมแบบไม่มีผู้สอนระบบ โดยเปรียบเทียบวิธีการค้นคืน 2 วิธี คือ ใช้ค่าดัชนีพืชพรรณในการระบุพื้นที่ในภาพถ่ายดาวเทียมโดยดูจากช่วงที่กำหนดในฐานข้อมูล และใช้การจัดกลุ่มแบบ K-mean โดยมีการจัดทำสเปกตรัมไลบรารีของพื้นที่ที่ต้องการไว้ในฐานข้อมูล จากนั้นนำภาพไปเปรียบเทียบกับสเปกตรัมไลบรารีของพื้นที่นั้นๆ และหาค่าความเหมือนระหว่างสเปกตรัมไลบรารีกับข้อมูลภาพในแต่ละกลุ่ม จากนั้นทำการดึงภาพขึ้นมาแสดงผลตามเงื่อนไขที่ผู้ใช้ต้องการ ซึ่งจากการทดลองพบว่าการค้นคืนภาพถ่ายดาวเทียมโดยใช้ค่าดัชนีพืชพรรณสามารถระบุประเภทพื้นที่ในภาพได้มีประสิทธิภาพกว่าการค้นคืนโดยใช้การจัดกลุ่มแบบ K-mean เนื่องจากการคำนวณค่าดัชนีพืชพรรณจะคำนวณทุกพิกเซล ทำให้ทราบว่าแต่ละพิกเซลตรงกับช่วงใดในแต่ละพื้นที่ เมื่อแสดงผลเฉพาะพื้นที่นั้นทำให้แม่นยำกว่า ส่วนการจัดกลุ่มแบบK-mean อาจทำให้ข้อมูลบางส่วนหายไปเนื่องจากใช้ค่าเฉลี่ยเป็นไลบรารีสเปกตรัมเพื่อไปเปรียบเทียบค่าความเหมือนกับแต่ละกลุ่ม และสิ่งรบกวนภายในภาพอาจทำให้ผลลัพธ์ในการจัดกลุ่มผิดเพี้ยนไป เช่น เมฆ แสง เงา เป็นต้น และความแม่นยำของการประมวลผลภาพถ่ายดาวเทียมขึ้นอยู่กับคุณภาพของภาพถ่ายดาวเทียมด้วย หากภาพถ่ายดาวเทียมมีสิ่งรบกวนมากเกินไปอาจทำให้แต่ละพิกเซลของภาพผิดเพี้ยนจากความเป็นจริง ซึ่งอาจส่งผลให้ความแม่นยำลดน้อยลง นอกจากนี้ยังสามารถนำผลลัพธ์ของการค้นคืนวัตถุในภาพมาช่วยในการตรวจจับการเปลี่ยนแปลงของพื้นที่ในแต่ละช่วงเวลาได้ ยกตัวอย่างเช่น ในพื้นที่บริเวณเดียวกันแต่ช่วงเวลาแตกต่างกัน วันแรกเราค้นคืนพบพื้นที่บริเวณที่มีพืชพรรณสีเขียวอยู่หนาแน่นมาก แต่วันถัดมากลับค้นคืนพบพื้นที่บริเวณสีเขียวอยู่น้อยจนแทบไม่มีเลย เราอาจสันนิษฐานได้ว่าพื้นที่ดังกล่าวอาจโดนบุกรุกหรือเกิดความแห้งแล้ง ทำให้เราสามารถนำข้อมูลเหล่านี้ไปวางแผนการใช้พื้นที่ได้ในอนาคตได้ เป็นต้น และในงานวิจัยในอนาคตอาจมีการนำค่าละติจูด ลองติจูด มาแสดงผลในการค้นคืน

เอกสารอ้างอิง

[1] ศูนย์จัดการและประยุกต์ใช้งานดาวเทียม สถานีรับสัญญาณดาวเทียมจุฬาภรณ์, "การประยุกต์ใช้งานดาวเทียมเอนกประสงค์ขนาดเล็ก(SMMS—Small Multi-Mission Satellite)", Available from: http://smms.eng.ku.ac.th/ [5 January 2012]

[2] Adia S.O and Rabiu A. B (Phd), "Change Detection of Vegetation Cover, using Multi- Temporal Remote Sensing Data and GIS Techniques", *Digital Library for Physics and Astronomy[online],* Available from: http://adsabs.harvard.edu [5 January 2012]

[3] Umesh K Kand Suresha, "Web Image Retrieval using Clustering Approaches*", International Symposium on Devices MEMS Intelligent Systems and Communications ISDMISC2011*, April 2011

[4] Gerald Schaefer, "Content-Based Image Retrieval – Some Basics"*, Advances in Intelligent and Soft Computing,* Vol. 103, 2011, pp. 21-29

[5] Li Wei and Wang Weihong and Lu Feng, “Research on remote sensing image retrieval based on geographical and semantic features”, *International Conference on Image Analysis and Signal Processing - IASP*, 2009, pp. 162–165

[6] Hong Liu and Xiaohong Yu, “ Application Research of k-means Clustering Algorithm in Image Retrieval System”, *Proceedings of the Second Symposium International Computer Science and Computational Technology(ISCSCT '09) Huangshan, P. R. China*, Vol. 26-28, December 2009, pp. 274-277

[7] Priti Maheshwary et. al., “Retrieval of Remote Sensing Images Using Color, Texture and Spectral Features”, *International Journal of Engineering Science and Technology* Vol. 2, No. 9, 2010, pp. 4306-4311

[8] Jayant Mishra Anubhav Sharmaand and Kapil Chaturvedi, “An Unsupervised Cluster-based Image Retrieval Algorithm using Relevance Feedback ”, *International Journal of Managing Information Technology (IJMIT)*, Vol. 3, No. 2, May 2011

[9] U.S. Department of Agriculture Foreign Agricultural Service , “Normalized Difference Vegetation Index (NDVI)”, Available from: http://www.pecad.fas.usda.gov/ [5 January 2012]

Copyright © 2013 by the Journal of Information Science and Technology.