

การคัดเลือกแบนด์ของภาพไฮเปอร์สเปกตรัลเพื่อการจำแนก

นารีนารถ รักสุนทร

คณะเทคโนโลยีอุตสาหกรรม มหาวิทยาลัยราชภัฏสวนสุนันทา

บทคัดย่อ

งานวิจัยนี้นำเสนอวิธีการคัดเลือกแบนด์เพื่อการจำแนกภาพไฮเปอร์สเปกตรัล การคัดเลือกแบนด์นี้ ใช้สารสนเทศร่วมระหว่างแบนด์ และแผนที่อ้างอิงความถูกต้องภาคพื้นดิน และสหสัมพันธ์ระหว่างแบนด์ การคัดเลือกแบนด์ในงานวิจัยนี้ ประกอบด้วย 2 ขั้นตอนที่สำคัญ ขั้นตอนที่ 1 เป็นการกำหนดเขตของแบนด์ที่สำคัญ เรียกเขตนี้ว่า เขตของแบนด์ที่มีโอกาสจะถูกคัดเลือก เขตนี้สามารถหาได้จากสารสนเทศร่วมระหว่างแบนด์แต่ละอัน กับแผนที่อ้างอิงความถูกต้องภาคพื้นดิน ขั้นตอนที่ 2 เป็นการลดความซ้ำซ้อนระหว่างแบนด์ โดยใช้สัมประสิทธิ์สหสัมพันธ์ระหว่างแบนด์ แบนด์ที่ถูกคัดเลือกจะถูกเก็บไว้ในเขตของแบนด์ที่คัดเลือก เขตนี้ประกอบด้วยแบนด์แรก ที่อยู่ในเขตของแบนด์ที่มีโอกาสจะถูกคัดเลือก แบนด์ลำดับที่สองจะถูกเลือกจากแบนด์ ที่อยู่ในเขตของแบนด์ที่มีโอกาสจะถูกคัดเลือก ลำดับถัดไป โดยที่แบนด์นั้น ให้ค่าสัมประสิทธิ์สหสัมพันธ์กับแบนด์ ที่อยู่ในเขตของแบนด์ที่คัดเลือกต่ำกว่าค่าขีดเริ่ม การคัดเลือกแบนด์ลำดับถัดไป ก็ทำในทำนองเดียวกัน จนกระทั่งได้จำนวนแบนด์ครบตามจำนวนที่ต้องการ ผลการทดลองแสดงให้เห็นว่า วิธีที่นำเสนอนี้เป็นวิธีที่มีประโยชน์ สำหรับการคัดเลือกแบนด์ เมื่อใช้จำนวนแบนด์เพียง 50% ของแบนด์ทั้งหมดสำหรับการจำแนกพบว่า ความถูกต้องโดยรวม และค่าเฉลี่ยความถูกต้อง มีค่าสูงกว่าการใช้จำนวนแบนด์ทั้งหมด และสัมประสิทธิ์แคปปามีค่าเท่ากัน โดยมีค่า 81.08%, 82.93% และ 0.78 ตามลำดับ นอกจากนี้แล้ว ถ้าจำนวนแบนด์ที่ใช้ในการจำแนกน้อยกว่า 50% ของแบนด์ทั้งหมดแล้ว ความถูกต้องของผลการจำแนกลดลงไม่มากนัก เมื่อเปรียบเทียบกับ การลดลงของจำนวนแบนด์ เช่น เมื่อใช้จำนวนแบนด์เพียง 10% ของแบนด์ข้อมูลทั้งหมด ค่าความถูกต้องโดยรวม ค่าเฉลี่ยความถูกต้อง และสัมประสิทธิ์แคปปามีค่าลดลงเป็น 69.39%, 66.32% และ 0.65 ตามลำดับ

คำสำคัญ : การลดขนาดมิติข้อมูล, การคัดเลือกแบนด์, การจำแนก, ภาพไฮเปอร์สเปกตรัล

Band Selection for Hyperspectral Image Classification

Nareenart Raksuntorn

The Faculty of Industrial Technology, Suan Sunandha Rajabhat University

ABSTRACT

This research presents a band selection technique for hyperspectral image classification. This technique utilizes mutual information between bands and the ground truth reference map, and correlation coefficients between bands. Band selection in this research comprises two major steps. The first step is to define a set of informative bands, called a candidate set. This set can be identified via mutual information between each band and the ground truth reference map. The second step is to reduce the redundancy between bands using correlation coefficients between bands. Selected bands will be kept in a band selection set. This set comprises the first band presenting in the candidate set. The second band is chosen from the next band presenting in the candidate set that offers a correlation coefficient below the preset threshold. The next band selection can be done in the same criteria until the number of bands is met the requirement. The results show that the proposed technique is useful for band selection. When only 50% of all the bands are utilized for classification, the overall accuracy and average accuracy are greater than those from using all the bands, and kappa coefficients are equal, which are 81.08%, 82.93%, and 0.78, respectively. Moreover, if the number of bands used for classification is less than 50% of all the bands, the classification accuracy is not lower too much when compared with the decrease of the number of bands, such as when only 10% of all the bands are used, the overall accuracy, the average accuracy, and kappa coefficient are dropped to 69.39%, 66.32%, and 0.65, respectively.

Keywords : Dimensionality reduction, band selection, classification, hyperspectral imagery

บทนำ

ภาพไฮเปอร์สเปกตรัล (Hyperspectral Image) เป็นภาพที่ได้รับความนิยม และถูกนำมาประยุกต์ใช้งานอย่างแพร่หลาย เนื่องจากภาพไฮเปอร์สเปกตรัล เป็นภาพที่มีความละเอียดเชิงคลื่นสูง มีขนาดตั้งแต่ 20 แบนด์จนถึงหลายร้อยแบนด์ ข้อมูลที่อยู่ในภาพถูกเก็บไว้อย่างต่อเนื่องที่มีความยาวคลื่นทุก ๆ 10 นาโนเมตร พิสัยความยาวคลื่นอยู่ระหว่าง 400 – 2,500 นาโนเมตร ภาพไฮเปอร์สเปกตรัลชนิด Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) ประกอบด้วย 224 แบนด์ ความละเอียดเชิงพื้นที่ประมาณ 20 เมตร ขณะที่ภาพไฮเปอร์สเปกตรัลชนิด Hyperspectral Digital Imagery Collection Experiment (HYDICE) ประกอบด้วย 210 แบนด์ มีความละเอียดเชิงพื้นที่อยู่ระหว่าง 0.75 – 3.75 เมตร และภาพไฮเปอร์สเปกตรัล Spaceborne Hyperion ประกอบด้วย 172 แบนด์ มีความละเอียดเชิงพื้นที่ 30 เมตร ด้วยคุณสมบัตินี้จึงตอบสนองต่อความต้องการในการวิเคราะห์ข้อมูลภาพได้อย่างหลากหลาย เช่น การค้นหาเป้าหมายโดยเฉพาะวัตถุที่มีขนาดเล็ก การค้นหาแร่ธาตุที่ค้นพบได้ยาก หรือมีอยู่น้อย การจำแนกวัตถุโดยเฉพาะวัตถุที่เกิดจากการสร้างของมนุษย์ เป็นต้น เนื่องจากภาพไฮเปอร์สเปกตรัลประกอบด้วยจำนวนแบนด์มาก จึงทำให้ขนาดของข้อมูลภาพมีขนาดใหญ่ ทำให้การวิเคราะห์ภาพไฮเปอร์สเปกตรัล ต้องการใช้ทรัพยากรสูง (ทั้งด้านเวลา ความซับซ้อนของการคำนวณ และหน่วยความจำ) ดังนั้น หากทำการลดขนาดมิติ (Dimensionality Reduction) ของภาพไฮเปอร์สเปกตรัล ก่อนนำไปประมวลผลแล้ว ความต้องการใช้ทรัพยากรดังกล่าวจะลดลง จุดมุ่งหมายของการลดขนาดมิติของภาพไฮเปอร์สเปกตรัล คือ การคัดเลือกแบนด์ที่เหลือจำนวนน้อยที่สุด โดยที่ข้อมูลที่สำคัญยังคงเหลืออยู่มากที่สุด เพื่อ

ลดการใช้ทรัพยากรมากเท่าที่สามารถจะทำได้ ดังนั้น แบนด์ที่ถูกคัดเลือกนี้ จึงควรเป็นแบนด์ที่เก็บข้อมูลที่สำคัญ ซึ่งในงานวิจัยนี้จะเรียกแบนด์นี้ว่า แบนด์ที่มีความสำคัญ

โดยทั่วไปการลดขนาดมิติภาพไฮเปอร์สเปกตรัล แบ่งออกเป็นสองกลุ่มใหญ่ ๆ คือ กลุ่มที่ลดขนาดมิติโดยการลดจำนวนแบนด์ จากแบนด์ของภาพทั้งหมด และกลุ่มที่ลดขนาดมิติโดยการลดจำนวนแบนด์ จากแบนด์ของภาพที่ถูกแปลง (Chang, 2013) การลดขนาดมิติของกลุ่มแรกเป็นวิธีการคัดเลือกแบนด์ที่มีความสำคัญ จากแบนด์ของภาพทั้งหมด จากนั้นนำแบนด์ที่ถูกคัดเลือกไว้เหล่านั้น ไปใช้แทนข้อมูลภาพนั้น โดยทั่วไปแล้วในการจำแนกข้อมูลภาพ ถ้ามีการใช้จำนวนแบนด์มากกว่า จะส่งผลให้ความถูกต้องของผลการจำแนกสูงขึ้น แต่เนื่องจากแบนด์ที่อยู่ติดกันของภาพไฮเปอร์สเปกตรัล ส่วนใหญ่จะมีสหสัมพันธ์ (Correlation) กันสูง ดังนั้น การใช้จำนวนแบนด์มากกว่า ไม่อาจรับประกันได้ว่าความถูกต้องของผลการจำแนก จะสูงกว่าการเลือกใช้แบนด์เพียงบางส่วนเสมอไป (Yuan, et.al., 2015) สำหรับการลดขนาดมิติกลุ่มที่สองเป็นการแปลงข้อมูลภาพ โดยใช้วิธีทางสถิติ เพื่อลดสหสัมพันธ์ระหว่างแบนด์ จากนั้นทำการคัดเลือกแบนด์ จากข้อมูลที่ถูกแปลงมาเพียงบางส่วน เช่น การแปลงข้อมูลโดยใช้วิธีการวิเคราะห์องค์ประกอบหลัก (Principle Component Analysis, PCA) เมื่อได้ข้อมูลที่ถูกแปลงแล้ว จะเลือกใช้แบนด์ที่เป็นองค์ประกอบหลัก ในลำดับต้น ๆ เช่น องค์ประกอบหลักที่ 1 – 10 เพื่อใช้ในการจำแนกข้อมูล

งานวิจัยนี้ได้เลือกใช้วิธีการลดขนาดมิติกลุ่มแรก โดยการคัดเลือกแบนด์ที่มีความสำคัญ จากแบนด์ของภาพทั้งหมด เนื่องจากการใช้ข้อมูลที่นำมาจากแบนด์ของภาพดั้งเดิม จะยังคงความหมายทางกายภาพของภาพเดิมไว้ ขั้นตอนวิธี

การคัดเลือกแบนด์ในงานวิจัยนี้ ใช้พื้นฐานจากทั้งวิธีสารสนเทศร่วม (Mutual Information, MI) (Guo, *et al.*, 2006) และสหสัมพันธ์ระหว่างแบนด์

วัตถุประสงค์ของการวิจัย

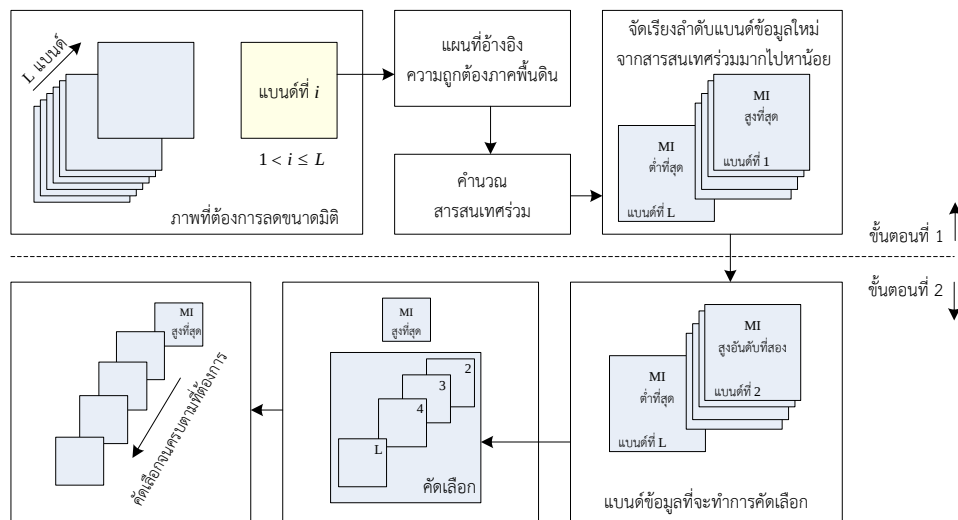
งานวิจัยนี้มีวัตถุประสงค์เพื่อ นำเสนอวิธีการคัดเลือกแบนด์ สำหรับการจำแนกข้อมูลภาพไฮเปอร์สเปกตรัล

ระเบียบวิธีวิจัย

เนื่องจากงานวิจัยนี้ คัดเลือกแบนด์ที่สำคัญจากแบนด์ของภาพทั้งหมด เพื่อใช้แทนข้อมูลภาพในการจำแนก ดังนั้น เซตของแบนด์ที่คัดเลือก จะประกอบด้วยแบนด์ที่มีความสำคัญ และแบนด์ที่ถูกคัดเลือกเหล่านี้ จะมีสหสัมพันธ์

ระหว่างแบนด์ต่ำ หากเซตของแบนด์ที่คัดเลือกประกอบด้วยแบนด์ ที่เก็บข้อมูลที่สำคัญของภาพไว้ได้แล้ว ผลการจำแนกโดยใช้แบนด์ที่คัดเลือกเหล่านี้ จะมีค่าใกล้เคียงกับผลการจำแนก เมื่อใช้แบนด์ของภาพทั้งหมด แผนภาพบล็อกขั้นตอนวิธีในการคัดเลือกแบนด์ ที่นำเสนอในงานวิจัยนี้ แสดงดังภาพที่ 1 และสามารถสรุปเป็นขั้นตอนที่สำคัญสองขั้นตอน คือ ขั้นตอนแรกเป็นการกำหนดเซตของแบนด์ที่มีความสำคัญ ซึ่งต่อไปในบทความนี้จะเรียกว่า เซตของแบนด์ที่มีโอกาสจะถูกคัดเลือก (ส่วนบนของภาพที่ 1) และขั้นตอนที่สองเป็นการลดความซ้ำซ้อนระหว่างแบนด์ (ส่วนล่างของภาพที่ 1)

สำหรับขั้นตอนแรก ในการกำหนดเซตของแบนด์ที่มีความสำคัญ แบนด์ที่มีความสำคัญในงานวิจัยนี้ หมายถึงแบนด์ที่มีสารสนเทศร่วมกับ



ภาพที่ 1 แผนภาพบล็อกขั้นตอนวิธีในการคัดเลือกแบนด์

แผนที่อ้างอิงความถูกต้องภาคพื้นดินสูง ซึ่งใช้หลักการเดียวกับงานวิจัยของ Guo และคณะฯ (Guo, *et al.*, 2006) ซึ่งต่อไปในงานวิจัยนี้ จะเรียกสั้น ๆ ว่า วิธีที่พัฒนาโดย Guo และคณะฯ กล่าวคือ แบนด์ที่มีความสำคัญมากที่สุดจะเป็น

แบนด์ที่มีสารสนเทศร่วมกับแผนที่อ้างอิงความถูกต้องภาคพื้นดินสูงที่สุด และแบนด์ที่มีความสำคัญลำดับถัด ๆ ไป ก็จะเป็นแบนด์ที่มีสารสนเทศร่วมสูงลำดับที่สอง และลำดับถัด ๆ ไปตามลำดับ ดังนั้นขอบเขตของงานวิจัยนี้ จึงจะ

ศึกษาเฉพาะกรณีที่ทราบแผนที่อ้างอิงความ
ถูกต้องภาคพื้นดินเท่านั้น

สำหรับขั้นตอนที่สอง เป็นการลดความ
ซ้ำซ้อนระหว่างแบนด์ ขั้นตอนนี้จะทำการลด
จำนวนแบนด์ที่สัมพันธ์กัน หรือลดจำนวนแบนด์
ที่เก็บข้อมูลที่คล้ายกัน กล่าวคือ แบนด์ที่มี
สัมประสิทธิ์สหสัมพันธ์กันสูง จะหมายถึงแบนด์ที่
เก็บข้อมูลความคล้ายกัน หรือมีความสัมพันธ์กัน
มาก ดังนั้น ในกระบวนการคัดเลือกแบนด์จะทำ
การเลือกแบนด์ที่สัมพันธ์กันมากมาเพียงแบนด์
เดียว หรือเพียงบางส่วน เพื่อลดความซ้ำซ้อน
ของแบนด์ข้อมูล กระบวนการคัดเลือกแบนด์
สามารถสรุปเป็นข้อย่อได้ดังนี้

1. เซตของแบนด์ที่คัดเลือก มีค่าเริ่มต้น
เป็นเซตว่าง ดังนี้

$$BSS = \{\emptyset\}$$

เมื่อ BSS คือ เซตของแบนด์ที่คัดเลือก และ $\emptyset$
หมายถึง ไม่มีสมาชิก

2. สำหรับทุกแบนด์ของภาพ จะคำนวณ
ค่าสารสนเทศร่วมระหว่างแบนด์ กับแผนที่อ้างอิง
ความถูกต้องภาคพื้นดิน

3. จัดเรียงลำดับแบนด์ใหม่ จากค่า
สารสนเทศร่วมที่คำนวณได้จากข้อ 2 โดยเรียง
ลำดับจากแบนด์ ที่มีสารสนเทศร่วมมากไปหา
น้อย ซึ่งในงานวิจัยนี้ จะเรียกแบนด์ที่ถูกจัดเรียง
ลำดับใหม่นี้ว่า เซตของแบนด์ที่มีโอกาสจะถูก
คัดเลือก ซึ่งเซตนี้สามารถเขียนได้ดังสมการ

$$CBS = \{B_1 \ B_2 \ \dots \ B_L\}$$

เมื่อ CBS คือ เซตของแบนด์ที่มีโอกาสจะถูกคัด
เลือก เมื่อสารสนเทศร่วมของ $B_1 > B_2 > \dots > B_L$
และ L เป็นจำนวนแบนด์ข้อมูลทั้งหมด

4. แบนด์ B_1 จะถูกคัดเลือกเป็นแบนด์แรก
เนื่องจากมีสารสนเทศร่วมกับแผนที่อ้างอิงความ-

ถูกต้องภาคพื้นดินสูงสุด ดังนั้น

$$BSS = \{B_1\}$$

และเมื่อ B_1 ถูกคัดเลือกแล้ว B_1 จะถูกนำออกจาก
เซต CBS ดังนั้น

$$CBS = \{B_2 \ B_3 \ \dots \ B_L\}$$

5. พิจารณาแบนด์ลำดับแรกไปในเซต
CBS ซึ่งคือ B_2 หาก B_2 มีสัมประสิทธิ์สหสัมพันธ์
กับแบนด์ ที่อยู่ในเซต BSS ต่ำกว่าค่าขีดเริ่มแล้ว
 B_2 จะถูกคัดเลือก และจะถูกนำไปเก็บไว้ในเซต
BSS หากไม่ตรงตามเงื่อนไข ก็ให้เลือกพิจารณา
แบนด์ลำดับถัดไป นั่นคือ B_3 และพิจารณา
จนกระทั่งได้แบนด์ ที่มีสัมประสิทธิ์สหสัมพันธ์กับ
แบนด์ที่อยู่ในเซต BSS ตามกำหนด จากนั้น นำ
แบนด์ลำดับแรกจนถึงแบนด์ที่ถูกคัดเลือก ออก
จากเซต CBS และนำแบนด์ที่ถูกคัดเลือกไปเก็บ
ในเซต BSS ซึ่งขั้นตอนนี้ เซต BSS จะประกอบ
ด้วยสองแบนด์ที่ถูกคัดเลือก

6. ทำข้อ 5 ซ้ำ จนกระทั่งได้จำนวนแบนด์ที่
อยู่ในเซต BSS ครบตามกำหนด

7. จำแนกข้อมูลโดยใช้แบนด์ที่อยู่ในเซต
BSS ด้วยวิธีเมทริกซ์ค่าสหสัมพันธ์

ผลการวิจัย

ผลการวิจัยแบ่งออกเป็น 2 หัวข้อย่อย
ได้แก่ ข้อมูลภาพไฮเปอร์สเปกตรัลที่ใช้ในการ
ทดลอง และผลการทดลอง รายละเอียดแสดง
ดังต่อไปนี้

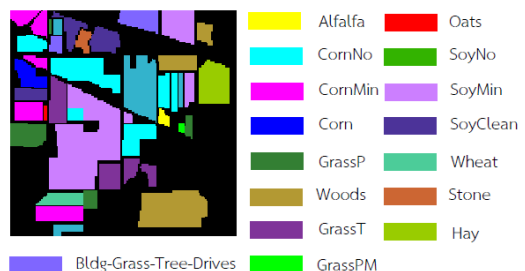
ข้อมูลภาพที่ใช้ในการทดลอง

ภาพไฮเปอร์สเปกตรัล ที่ใช้ในการทดลองนี้
เป็นภาพที่ได้จาก เครื่องตรวจวัดจากระยะไกล
AVIRIS ขนาด 145×145 จุดภาพ จำนวน 220
แบนด์ สามารถดาวน์โหลดได้จากเว็บไซต์ของ
มหาวิทยาลัย Purdue ที่ <https://engineering.purdue.edu/AVIRIS/>

purdue.edu/~bieht/MultiSpec/hyperspectral. Htm ภาพไฮเปอร์สเปกตรัลนี้เป็นภาพที่มีแผนที่อ้างอิงความถูกต้องภาคพื้นดิน ดังนั้น จึงเป็นภาพที่สามารถนำมาใช้กับงานวิจัย ที่ต้องการตรวจสอบความถูกต้องของผลการจำแนกได้ ภาพนี้มีความซับซ้อนสูงมาก เพราะเป็นภาพที่ประกอบด้วยวัตถุสูงถึง 16 ชนิด ตารางที่ 1 แสดงรายละเอียดของข้อมูลภาพ ได้แก่ ชนิดของวัตถุ ชื่อและจำนวนตัวอย่างของวัตถุทั้งหมด ที่นำมาจากแผนที่อ้างอิงความถูกต้องภาคพื้นดิน จำนวนจุดภาพสำหรับการฝึกฝน และจำนวนจุดภาพสำหรับการทดสอบ เนื่องจากตัวจำแนกเมทริกซ์มากเลขศูนย์ ใช้เทคนิคการเรียนรู้แบบมีผู้สอน (Supervised Learning) ดังนั้น จึงต้องการตัวอย่างสำหรับการฝึกฝนด้วย ในการทดลองนี้ ได้แบ่งตัวอย่างจุดภาพของวัตถุแต่ละชนิดออกเป็นสองส่วนเท่า ๆ กัน เพื่อใช้ในการฝึกฝนและการทดสอบ พิจารณาตารางที่ 1 พบว่า จำนวนตัวอย่างจุดภาพของวัตถุแต่ละชนิดแตกต่างกัน วัตถุบางชนิดมีจำนวนตัวอย่างมาก และวัตถุบางชนิดมีจำนวนตัวอย่างเพียงเล็กน้อย เช่น ตัวอย่างจุดภาพของ SoyMin มีมากถึง 2,468 จุดภาพ ซึ่ง SoyMin เป็นวัตถุที่มีจำนวนตัวอย่างสูงสุด ในขณะที่ตัวอย่างจุดภาพของ Oats จะมีเพียง 20 จุดภาพ ซึ่ง Oats เป็นวัตถุที่มีจำนวนตัวอย่างต่ำที่สุด ดังนั้น ในงานวิจัยนี้ จึงได้แสดงค่าเฉลี่ยความถูกต้องของผลการจำแนก

เพิ่มเติมด้วย ภาพที่ 2 แสดงแผนที่อ้างอิงความถูกต้องภาคพื้นดิน ของวัตถุทั้ง 16 ชนิด ที่ใช้ในการทดลอง วัตถุแต่ละชนิดถูกแสดงด้วยสีที่แตกต่างกัน ได้แก่ สีเหลือง หมายถึง Alfalfa สีแดง หมายถึง Oats สีฟ้า หมายถึง SoyNo สีเขียว หมายถึง SoyNo สีชมพู หมายถึง CornMin สีม่วง หมายถึง SoyMin สีน้ำเงิน หมายถึง Corn สีดำ หมายถึง SoyClean สีน้ำตาล หมายถึง GrassP สีเทา หมายถึง Wheat สีส้ม หมายถึง Woods สีฟ้าเข้ม หมายถึง Stone สีเขียวเข้ม หมายถึง GrassT สีเขียวอ่อน หมายถึง Hay สีฟ้าเข้ม หมายถึง GrassPM สีฟ้า หมายถึง Bldg-Grass-Tree-Drives

จุดภาพของวัตถุแต่ละชนิดรวมทั้งหมด 10,366 จุดภาพ และพื้นที่บริเวณสีดำเป็นพื้นหลังของภาพ มีจำนวนรวมทั้งสิ้น 10,659 จุดภาพ ซึ่งจุดภาพเหล่านี้จะไม่ถูกนำมาใช้ในการทดลอง



ภาพที่ 2 แผนที่อ้างอิงความถูกต้องภาคพื้นดิน

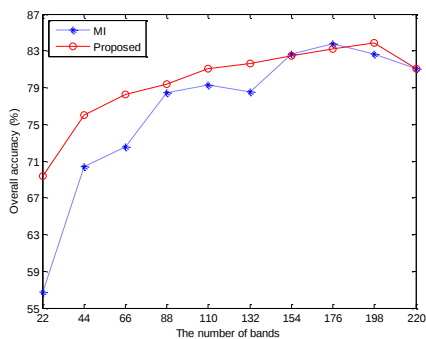
ตารางที่ 1 รายละเอียดของข้อมูลภาพ

ชื่อ	จำนวนจุดภาพทั้งหมด	จำนวนจุดภาพที่ใช้ในการฝึกฝน	จำนวนจุดภาพที่ใช้ในการทดสอบ
Alfalfa	54	27	27
CornNo	1,434	717	717
CornMin	834	417	417
Corn	234	117	117
GrassP	497	249	248
GrassT	747	374	373
GrassPM	26	13	13
Hay	489	245	244
Oats	20	10	10
SoyNo	968	484	484
SoyMin	2,468	1,234	1,234
SoyClean	614	307	307
Wheat	212	106	106
Woods	1,294	647	647
Bldg-Grass-Tree-Drives	380	190	190
Stone	95	48	47
รวม	10,366	5,185	5,181

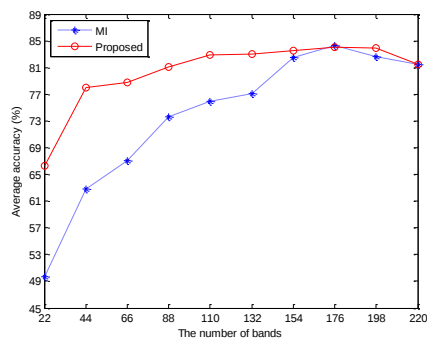
ผลการทดลอง

ในงานวิจัยนี้ ใช้วิธีเมทริกซ์มากเลขศูนย์ที่พัฒนาโดย Chen และคณะฯ (Chen, *et al.*, 2011) เป็นตัวจำแนกข้อมูล เนื่องจากความหลากหลายของจุดภาพตัวอย่าง ตัวจำแนกนี้มีความต้องการใช้จำนวนแบนด์ข้อมูลเพื่อการจำแนกอย่างน้อยที่สุด เท่ากับจำนวนวัตถุที่ต้องการจำแนก นั่นคือ อย่างน้อยที่สุด คือ 16 แบนด์ สำหรับการจำแนกวัตถุจำนวน 16 ชนิด ดังนั้น ในงานวิจัยนี้ จึงได้กำหนดจำนวนแบนด์ที่ใช้ในการจำแนก โดยเริ่มต้นที่ 22 แบนด์ (หรือ 10% ของจำนวนแบนด์ทั้งหมด) จากนั้นเพิ่มจำนวนแบนด์ครั้งละ 10% ไปจนถึง 100% นั่นคือ ผลการทดลองจะประกอบด้วย ผลการจำแนกเมื่อใช้จำนวนแบนด์ 22, 44, 66, 88, 110, 132, 154, 176, 198 และ 220 แบนด์ ในการทดลองได้เปรียบเทียบผลการจำแนก กับวิธีที่พัฒนาโดย Guo และคณะฯ ซึ่งเป็นวิธีที่มีพื้นฐาน จากสารสนเทศเพียงอย่างเดียว

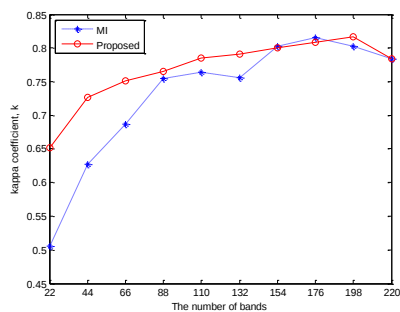
วิธีนี้มีข้อเสีย คือ การกำหนดความกว้างของการคัดเลือกแบนด์ จะขึ้นกับประสบการณ์ของผู้ใช้ในการทดลองนี้ แผนที่อ้างอิงความถูกต้องภาคพื้นดินจะเป็นแผนที่จริง ซึ่งแสดงในภาพที่ 2 ผลการทดลองแสดงดังภาพที่ 3 – 5 โดยคำอธิบายภาพ MI และ Proposed จะหมายถึง วิธีที่มีพื้นฐานจากสารสนเทศร่วมที่พัฒนาโดย Guo และคณะฯ และวิธีที่นำเสนอตามลำดับ ภาพที่ 3 แสดงผลการเปรียบเทียบความถูกต้องโดยรวมของผลการจำแนก ภาพที่ 4 แสดงผลการเปรียบเทียบค่าเฉลี่ยความถูกต้อง ของผลการจำแนก และภาพที่ 5 แสดงผลการเปรียบเทียบสัมประสิทธิ์แคปปา รายละเอียดแสดงดังตารางที่ 2 ผู้สนใจสามารถศึกษา วิธีการคำนวณค่าเหล่านี้ได้จากเอกสารอ้างอิง (Congalton, 1991) ตารางที่ 3 – 4 แสดงผลการจำแนกของวัตถุแต่ละชนิดของวิธีที่พัฒนาโดย Guo และคณะฯ และวิธีที่นำเสนอตามลำดับ



ภาพที่ 3 ความถูกต้องโดยรวม



ภาพที่ 4 ค่าเฉลี่ยความถูกต้อง



ภาพที่ 5 สัมประสิทธิ์แคปปา

ตารางที่ 2 เปรียบเทียบผลการจำแนกวิธีที่พัฒนาโดย Guo และคณะฯ และวิธีที่นำเสนอ

จำนวนแบนด์ที่ใช้ ในการจำแนก	ความถูกต้องโดยรวม (%)		ค่าเฉลี่ยความถูกต้อง (%)		สัมประสิทธิ์แคปปา	
	วิธี Guo และคณะฯ	วิธีที่นำเสนอ	วิธี Guo และคณะฯ	วิธีที่นำเสนอ	วิธี Guo และคณะฯ	วิธีที่นำเสนอ
22	56.75	69.39	49.65	66.32	0.51	0.65
44	70.37	76.03	62.76	77.96	0.63	0.73
66	72.53	78.21	67.06	78.77	0.69	0.75
88	78.44	79.41	73.68	81.12	0.75	0.77
110	79.25	81.08	75.99	82.93	0.76	0.78
132	78.52	81.63	77.15	83.07	0.76	0.79
154	82.63	82.46	82.45	83.53	0.80	0.80
176	83.81	83.23	84.27	84.02	0.82	0.81
198	82.63	83.88	82.63	83.90	0.80	0.82
220	81.03	81.03	81.44	81.44	0.78	0.78

ตารางที่ 3 ผลการจำแนกของวิธีที่พัฒนาโดย Guo และคณะฯ

ชื่อ	จำนวนจุดภาพที่ใช้ในการทดสอบ	จำนวนแบนด์ที่ใช้และจำนวนจุดภาพที่จำแนกได้ถูกต้อง									
		220 แบนด์	198 แบนด์	176 แบนด์	154 แบนด์	132 แบนด์	110 แบนด์	88 แบนด์	66 แบนด์	44 แบนด์	22 แบนด์
Alfalfa	27	18	18	19	17	15	16	20	19	12	9
CornNo	717	503	503	514	521	492	537	537	525	490	414
CornMin	417	309	317	330	317	297	297	299	275	265	207
Corn	117	79	82	83	81	78	78	62	54	50	40
GrassP	248	234	235	236	236	228	216	218	187	171	112
GrassT	373	364	366	371	368	355	328	329	308	313	233
GrassPM	13	11	11	11	10	10	8	9	5	4	2
Hay	244	232	239	240	241	241	238	237	231	227	208
Oats	10	8	8	9	9	7	7	3	3	3	3
SoyNo	484	388	400	407	393	387	398	378	326	297	194
SoyMin	1,234	974	1002	1020	1006	970	1006	988	929	891	760
SoyClean	307	206	217	222	215	182	196	216	186	190	138
Wheat	106	105	105	104	104	102	101	102	103	95	77
Woods	647	609	618	616	607	584	569	560	518	547	475
Bldg-Grass-Tree-Drives	190	114	116	115	111	75	67	65	46	53	35
Stone	47	44	44	45	45	45	44	41	43	38	33

ตารางที่ 4 ผลการจำแนกของวิธีที่นำเสนอ

ชื่อ	จำนวนจุดภาพที่ใช้ในการทดสอบ	จำนวนแบนด์ที่ใช้และจำนวนจุดภาพที่จำแนกได้ถูกต้อง									
		220 แบนด์	198 แบนด์	176 แบนด์	154 แบนด์	132 แบนด์	110 แบนด์	88 แบนด์	66 แบนด์	44 แบนด์	22 แบนด์
Alfalfa	27	18	20	20	20	20	21	20	20	21	13
CornNo	717	503	522	510	499	486	484	463	446	421	360
CornMin	417	309	329	324	314	310	302	290	285	277	221
Corn	117	79	78	78	81	82	83	80	78	71	55
GrassP	248	234	238	238	237	236	233	231	230	228	216
GrassT	373	364	369	369	367	367	365	364	365	362	359
GrassPM	13	11	11	11	11	11	11	11	8	10	6
Hay	244	232	240	239	238	237	236	236	238	238	238
Oats	10	8	8	9	9	9	9	8	8	8	3
SoyNo	484	388	407	404	403	393	396	380	357	347	338
SoyMin	1,234	974	1,013	1,009	999	989	987	970	965	947	832
SoyClean	307	206	228	221	215	212	199	190	186	165	124
Wheat	106	105	105	105	105	105	105	105	105	105	105
Woods	647	609	616	616	616	615	613	607	609	597	587
Bldg-Grass-Tree-Drives	190	114	117	114	114	113	112	113	106	97	97
Stone	47	44	45	45	44	44	45	46	46	45	41

สรุปและอภิปรายผล

เมื่อพิจารณาผลการจำแนกพบว่า ทั้งวิธีที่พัฒนาโดย Guo และคณะฯ และวิธีที่นำเสนอ มีประสิทธิภาพในการจำแนกลดลง เมื่อจำนวนแบนด์ที่ใช้ในการจำแนก ลดลงมากกว่า 30% ของจำนวนแบนด์ทั้งหมด (หรือกล่าวคือ เมื่อใช้แบนด์ที่คัดเลือกต่ำกว่า 154 แบนด์) และเมื่อใช้แบนด์ที่คัดเลือกจำนวน 154, 176 และ 198 แบนด์แล้ว ประสิทธิภาพของผลการจำแนกสูงขึ้น โดยมีค่าความถูกต้องโดยรวม ค่าเฉลี่ยความถูกต้อง และ

สัมประสิทธิ์แคปปา สูงขึ้นกว่าการใช้ข้อมูลภาพทั้งหมด วิธีที่พัฒนาโดย Guo และคณะฯ ให้ผลการจำแนกดีที่สุด เมื่อใช้แบนด์ที่คัดเลือกจำนวน 176 แบนด์ โดยมีค่าความถูกต้องโดยรวม 83.81% ค่าเฉลี่ยความถูกต้อง 84.27% และสัมประสิทธิ์แคปปา 0.82 ในขณะที่วิธีที่นำเสนอให้ผลการจำแนกดีที่สุด เมื่อใช้แบนด์ที่คัดเลือกจำนวน 198 แบนด์ โดยมีค่าความถูกต้องโดยรวม 83.88% และเมื่อใช้แบนด์ที่คัดเลือกจำนวน 176 แบนด์ ค่าเฉลี่ยความถูกต้อง 84.02% และ

สัมประสิทธิ์แคปปา 0.82 ซึ่งจะเห็นได้ว่าทั้งสองวิธีให้ผลใกล้เคียงกัน ทั้งนี้ผลการจำแนกของทั้งสองวิธีสูงขึ้น ในขณะที่จำนวนแบนด์ลดลงเนื่องจากแบนด์ข้อมูลที่เป็นคลื่น ช่วงความยาวที่ตุงกลืนคลื่น และแบนด์ที่มีอัตราสัญญาณต่อสัญญาณรบกวนต่ำ (Signal-to-noise ratio, SNR) ถูกตัดทิ้งจากการกระบวนการคัดเลือก

เมื่อจำนวนแบนด์ถูกลดลงไปมากขึ้น โดยใช้แบนด์ที่คัดเลือกระหว่าง 10% – 60% ของจำนวนแบนด์ทั้งหมด พบว่าวิธีที่นำเสนอให้ผลการจำแนกดีกว่า โดยมีความถูกต้องโดยรวม ค่าเฉลี่ยความถูกต้อง และสัมประสิทธิ์แคปปาสูงกว่า ซึ่งแสดงให้เห็นว่าวิธีที่นำเสนอมีการคัดเลือกแบนด์ ได้อย่างมีประสิทธิภาพมากกว่า

พิจารณาตารางที่ 3 พบว่าวิธีที่พัฒนาโดย Guo และคณะฯ สามารถจำแนก Hay ได้ดีที่สุด ถึงแม้ว่าจำนวนแบนด์ถูกลดลงไปถึง 198 แบนด์แล้ว ความถูกต้องของการจำแนกยังสูงถึง 85% แต่วิธีนี้ไม่สามารถจำแนก GrassPM และ Bldg-Grass-Tree-Drives ได้เลย โดยความถูกต้องของการจำแนกมีค่าต่ำกว่า 19% แต่หากพิจารณาตารางที่ 4 พบว่าวิธีที่นำเสนอ สามารถจำแนกข้อมูลหลายชนิดได้ดี ถึงแม้ว่าจำนวนแบนด์ถูกลดลงไปสูงถึง 198 แบนด์ โดยความถูกต้องของการจำแนกยังมีค่าสูงระหว่าง 87% - 99% ซึ่งได้แก่ GrassP, GrassT, Hay, Wheat, Woods และ Stone เมื่อพิจารณาภาพที่ 3 – 5 และตารางที่ 2 พบว่า วิธีที่นำเสนอมีประสิทธิภาพ ในการจำแนกดีกว่า โดยความถูกต้องของการจำแนกโดยรวม ค่าเฉลี่ยความถูกต้องของการจำแนก และสัมประสิทธิ์แคปปามีค่าสูงกว่า ยกเว้นเมื่อใช้จำนวนแบนด์ 154 และ 176 แบนด์ ผลการจำแนกของวิธีที่นำเสนอต่ำกว่าเพียงเล็กน้อย

จากที่ได้กล่าวมาทั้งหมดแล้ว จึงสามารถสรุปโดยรวมได้ว่า วิธีที่นำเสนอสามารถนำมาใช้

ลดขนาดมิติข้อมูลภาพไฮเปอร์สเปกตรัล เพื่อการจำแนกได้ดี ผลการจำแนกลดลงจากการใช้ข้อมูลทุกแบนด์น้อย เมื่อเทียบกับจำนวนแบนด์ที่ลดลง และเมื่อเปรียบเทียบผลการจำแนก กับเทคนิคการลดมิติข้อมูลภาพไฮเปอร์สเปกตรัลอื่น ที่ใช้พื้นฐานสารสนเทศร่วมเหมือนกัน วิธีที่นำเสนอเป็นวิธีที่มีประสิทธิภาพในการจำแนกสูงกว่า โดยมีความถูกต้องโดยรวมของการจำแนกสูงกว่า ค่าเฉลี่ยความถูกต้องของการจำแนกสูงกว่า และสัมประสิทธิ์แคปปาก็มีค่ามากกว่า

ข้อเสนอแนะ

งานวิจัยนี้ ทำการคัดเลือกแบนด์ที่มีความสำคัญ โดยแบนด์ที่มีความสำคัญ เป็นแบนด์ที่มีสารสนเทศร่วม กับแผนที่อ้างอิงความถูกต้องภาคพื้นดินสูง แบนด์ที่ถูกคัดเลือกต่างๆ ต้องมีสหสัมพันธ์กันต่ำ การคัดเลือกแบนด์ในงานวิจัยนี้ กระทำการคัดเลือกแบนด์ครั้งละหนึ่งแบนด์ จนได้จำนวนแบนด์ครบตามต้องการ จะสังเกตได้ว่า เมื่อจำนวนแบนด์ที่ต้องการมีมากขึ้น โอกาสที่ความซ้ำซ้อนระหว่างแบนด์ ก็เพิ่มขึ้นตามไปด้วย ดังนั้น งานวิจัยในอนาคต จึงอาจเพิ่มเงื่อนไขในการคัดเลือกแบนด์ เช่น มีการตรวจสอบผลการจำแนกร่วมด้วย หรือพิจารณาจากค่าความคลาดเคลื่อนของจุดภาพ หลังการเพิ่มแบนด์ในแต่ละครั้งร่วมด้วย อย่างไรก็ตาม หากมีเงื่อนไขมากขึ้น ความซับซ้อนของกระบวนการคำนวณ ก็จะมากขึ้นตามไปด้วย ดังนั้น ผู้ใช้งานควรเลือกให้เหมาะสม กับงานที่นำไปประยุกต์ใช้ในขณะนั้น หากต้องการให้ประสิทธิภาพในการจำแนกสูงขึ้น ก็อาจต้องใช้เวลาในการประมวลผลมากขึ้น

กิตติกรรมประกาศ

งานวิจัยนี้ได้รับทุนอุดหนุนบางส่วนจากมหาวิทยาลัยราชภัฏสวนสุนันทา

References

- Chang, -I C. (2013). **Hyperspectral data processing: algorithm design and analysis**. Singapore: Wiley.
- Chen, Y., Nasrabadi, M. N. and Tranc, D. T. (2011). Hyperspectral image classification using dictionary-based sparse representation. **IEEE Transaction on Geoscience Remote Sensing**, 49(10), 3973 – 3985.
- Congalton, R. T. (1991). A Review of Assessing the Accuracy of classifications of remotely sensed data. **Remote Sensing Environment**, 37, 35 – 46.
- Cover, M. T. and Thomas, A. J. (1991). **Elements of information theory**. New York: Wiley.
- Duda, O. R, Hart, E. P. and Stork, G. D. (2001). **Pattern classification**. (2nd). New York: Wiley.
- Guo, B., Gunn, R. S., Damper, I. R. and Nelson, D. B. J. (2006). Band selection for hyperspectral image classification using mutual information. **IEEE Transaction on Geoscience Remote Sensing, Letter**, 3(4), 522 – 526.
- Yuan, Y., Zhu, G. and Wang, Q. (2015). Hyperspectral band selection by multitask sparsity pursuit. **IEEE Transaction on Geoscience Remote Sensing**, 53(2), 631 – 644.