

บทที่ 4

ผลการวิจัยและอภิปรายผล

ผู้วิจัยได้ทำการทดลองเพื่อเปรียบเทียบ ขั้นตอนวิธีการย้ายค่าเฉลี่ยแบบเดิม กับขั้นตอนวิธีการที่นำเสนอขึ้นใหม่คือ ขั้นตอนวิธีลำเลียงย้ายค่าเฉลี่ย (TMS) และ ขั้นตอนวิธีลำเลียงย้ายค่าเฉลี่ยแบบครอบคลุม (GTMS) โดยทั่วไปแล้วผลการทำงานของขั้นตอนวิธีการย้ายค่าเฉลี่ยแบบเดิมจะขึ้นอยู่กับค่า σ และการกำหนดค่า Threshold เนื่องจากงานวิจัยนี้มุ่งเน้นไปที่การลดความซ้ำซ้อนเพื่อเพิ่มความเร็วของขั้นตอนวิธีการย้ายตามค่าเฉลี่ย ผู้วิจัยจึงได้ทำการทดลองเปรียบเทียบการเปลี่ยน σ บางค่าเพื่อให้เห็นความแตกต่าง โดยแต่ละการทดลองใช้ตัวแปรเดียวกัน ทดลองบนคอมพิวเตอร์และสิ่งแวดล้อมเดียวกัน ในการทดลองนี้แต่ละขั้นตอนวิธีมีดังนี้คือ MS1 หมายถึง ขั้นตอนวิธีการย้ายค่าเฉลี่ย, MS2 หมายถึง ขั้นตอนวิธีการย้ายค่าเฉลี่ยที่มีการตัดตัวหยุดออก, TMS หมายถึง ขั้นตอนวิธีการลำเลียงย้ายค่าเฉลี่ย, GTMS1 หมายถึง ขั้นตอนวิธีการลำเลียงย้ายค่าเฉลี่ยแบบครอบคลุม ที่กำหนด $\alpha > 1$ ทำให้ได้ผลเดียวกับ TMS และ GTMS2 หมายถึง ขั้นตอนวิธีการลำเลียงย้ายค่าเฉลี่ยแบบครอบคลุมที่กำหนดค่า $\alpha_0 = 0$, $\alpha_p = 0.5$ และ $r_p = 10$ โดย α ในแต่ละรอบคำนวณได้จากสมการ 18 ค่าความถูกต้องวัดได้จากการนับจำนวนข้อมูลที่จัดกลุ่มได้ถูกในแต่ละขั้นตอนวิธีเมื่อเทียบกับการจัดกลุ่มของ MS หาดด้วยจำนวนข้อมูลทั้งหมด ซึ่ง MS1 และ MS2 ได้ผลลัพธ์เดียวกัน ส่วนจำนวนเท่าของการเพิ่มความเร็วเป็นการวัดจำนวนเท่าของเวลาเมื่อเทียบกับ MS เป็นอัตราส่วนระหว่างเวลาของ MS กับแต่ละขั้นตอนวิธี

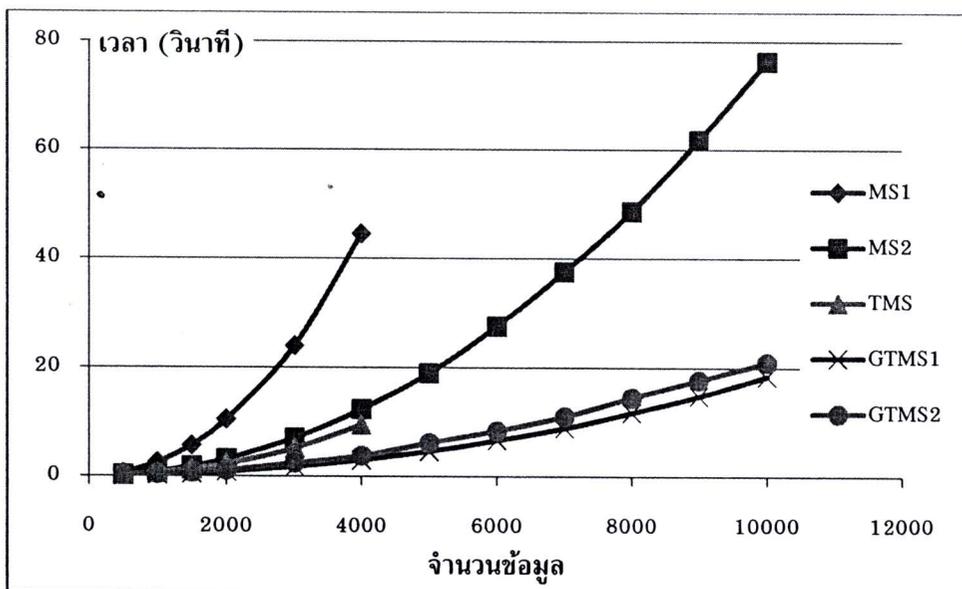
1. การเปรียบเทียบเวลาการทำงานของแต่ละขั้นตอนวิธีเมื่อมีจำนวนข้อมูลเพิ่มขึ้น

เมื่อทำการทดลองในการเปรียบเทียบความเร็วโดยใช้ข้อมูลที่จำลองขึ้นตามหัวข้อ 4.1 ในบทที่ 3 โดยทำการจับเวลาการทำงานของแต่ละขั้นตอนวิธีจำนวน 10 รอบ โดยผู้วิจัยได้ทำการทดลองใช้แบนวิดแตกต่างกัน 2 ค่าคือ 10 และ 2 และได้กำหนดค่า Threshold เท่ากันคือ 10^{-7} โดยข้อมูลที่มียังมีจำนวนตั้งแต่ 5,000 ขึ้นไป TMS ไม่สามารถประมวลผลได้เนื่องจากต้องใช้หน่วยความจำมาก ขณะที่ MS1 ใช้เวลาในการประมวลผลสูงผู้วิจัยจึงไม่ทำการทดลองกับข้อมูลเหล่านั้น

จากตารางที่ 7 ทำการเปรียบเทียบเวลาของขั้นตอนวิธีในการดำเนินการกับข้อมูลทดสอบที่ได้จำลองขึ้นจำนวน 12 ข้อมูลที่มีความแตกต่างกัน และใช้เบตวิดเป็น 10 จะเห็นได้ว่าที่จำนวนข้อมูล 4000 ข้อมูล เป็นจำนวนข้อมูลมากที่สุดที่สามารถเปรียบเทียบเวลาของแต่ละขั้นตอนวิธีได้ทั้งหมด ขั้นตอนวิธี GTMS1 ใช้เวลาน้อยที่สุดในการดำเนินการกับข้อมูล คือ 3.0025 วินาที และ GTMS2 ใช้เวลาในการประมวลผลคือ 3.7918 วินาที ซึ่งใกล้เคียงกับ GTMS1 ส่วน MS1 จะใช้เวลาในการประมวลผลมากที่สุดเมื่อเทียบกับขั้นตอนวิธีอื่น ๆ ซึ่งจะเห็นการเปรียบเทียบเวลาได้อย่างชัดเจนจากภาพที่ 32 โดยแกน x คือ จำนวนข้อมูล ส่วน แกน y คือ เวลา (วินาที) จะเห็นได้ว่า MS1 และ MS2 มีแนวโน้มที่จะใช้เวลาเพิ่มมากขึ้นอย่างรวดเร็วเมื่อข้อมูลมีจำนวนข้อมูลเพิ่มมากขึ้น ขณะที่ TMS นั้นสามารถประมวลผลได้เพียงข้อมูลที่มีจำนวนสมาชิก 4000 ข้อมูล ดังเหตุผลที่กล่าวไปแล้วข้างต้น

ตารางที่ 7 เวลาเฉลี่ยใน 10 รอบ ของแต่ละขั้นตอนวิธีในการดำเนินการกับข้อมูลทดสอบที่มีจำนวนข้อมูลแตกต่างกัน โดยใช้แบนด์วิดท์เท่ากับ 10

จำนวนข้อมูล	เวลาเฉลี่ยใน 10 รอบการทดลอง (วินาที)				
	MS1	MS2	TMS	GTMS1	GTMS2
500	0.6039	0.2536	0.1646	0.0881	0.1605
1000	2.5648	0.8566	0.5622	0.2361	0.3948
1500	5.7622	1.8347	1.2477	0.4816	0.7383
2000	10.4995	3.1768	2.2278	0.8474	1.1246
3000	24.0031	7.0856	5.1297	1.7491	2.4306
4000	44.5456	12.3715	9.5068	3.0025	3.7918
5000	-	18.9691	-	4.5976	6.1617
6000	-	27.5649	-	6.6530	8.4020
7000	-	37.5823	-	8.9072	11.0420
8000	-	48.6477	-	11.7781	14.5361
9000	-	61.7534	-	14.8031	17.7072
10000	-	76.3246	-	18.3958	20.9692



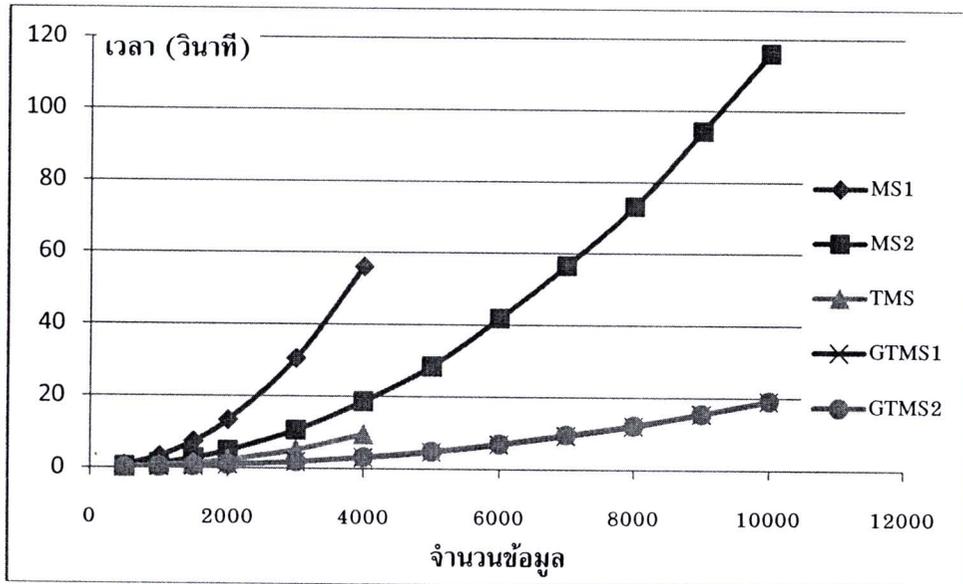
ภาพที่ 32 เปรียบเทียบเวลาเฉลี่ยใน 10 รอบ ของแต่ละขั้นตอนวิธีในการดำเนินการกับข้อมูลทดสอบที่มีจำนวนข้อมูลแตกต่างกันโดยใช้แบนด์วิดท์เท่ากับ 10

ตารางที่ 8 เวลาเฉลี่ยใน 10 รอบ ของแต่ละขั้นตอนวิธีในการดำเนินการกับข้อมูลทดสอบที่มีจำนวนข้อมูลแตกต่างกัน โดยใช้แบนด์วิธเท่ากับ 2

จำนวนข้อมูล	เวลาเฉลี่ยใน 10 รอบการทดลอง (วินาที)				
	MS1	MS2	TMS	GTMS1	GTMS2
500	0.8627	0.3844	0.1475	0.0762	0.0997
1000	3.3145	1.2871	0.5767	0.2568	0.2777
1500	7.3611	2.7722	1.2629	0.5100	0.5514
2000	13.3487	4.8984	2.2817	0.9411	0.9801
3000	30.3936	10.6219	5.1574	1.8077	1.8546
4000	56.0479	18.6203	9.5455	3.1118	3.1737
5000	-	28.3838	-	4.7350	4.8546
6000	-	41.8209	-	6.8353	6.9976
7000	-	56.5236	-	9.3324	9.6284
8000	-	73.1108	-	12.0611	12.2612
9000	-	94.2763	-	15.4965	15.5447
10000	-	116.0083	-	19.0510	19.1749

จากตารางที่ 8 ทำการเปรียบเทียบเวลาของขั้นตอนวิธีในการดำเนินการกับข้อมูลทดสอบที่ได้จำลองขึ้นจำนวน 12 ข้อมูลที่มีความแตกต่างกัน และใช้แบนด์วิธเป็น 2 จะเห็นได้ว่าที่จำนวนข้อมูล 4000 ข้อมูลเป็นจำนวนข้อมูลมากที่สุดที่สามารถเปรียบเทียบเวลาของแต่ละขั้นตอนวิธีได้ทั้งหมด ขั้นตอนวิธี GTMS1 ใช้เวลาน้อยที่สุดในการดำเนินการกับข้อมูล คือ 3.1118 และ GTMS2 ใช้เวลาในการประมวลผลคือ 3.1737 ซึ่งใกล้เคียงกับ GTMS1 ส่วน MS1 จะใช้เวลาในการประมวลผลมากที่สุดเมื่อเทียบกับขั้นตอนวิธีอื่น ๆ ซึ่งจะเห็นการเปรียบเทียบเวลาได้อย่างชัดเจนจากภาพที่ 33 จะเห็นได้ว่า MS1 และ MS2 มีแนวโน้มที่จะใช้เวลาเพิ่มมากขึ้นอย่างรวดเร็วเมื่อข้อมูลมีจำนวนข้อมูลเพิ่มมากขึ้น และ GTMS1 และ GTMS2 มีการใช้เวลาในการประมวลผลใกล้เคียงกันมาก



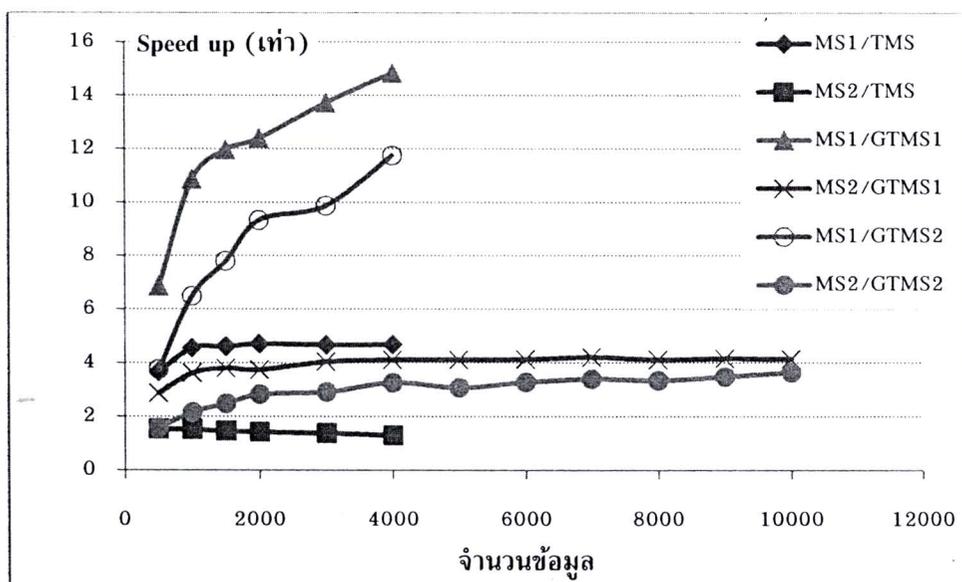


ภาพที่ 33 เปรียบเทียบเวลาเฉลี่ยใน 10 รอบ ของแต่ละขั้นตอนวิธีในการดำเนินการกับข้อมูลทดสอบที่มีจำนวน ข้อมูลแตกต่างกัน โดยใช้แบนด์วิดท์เท่ากับ 2

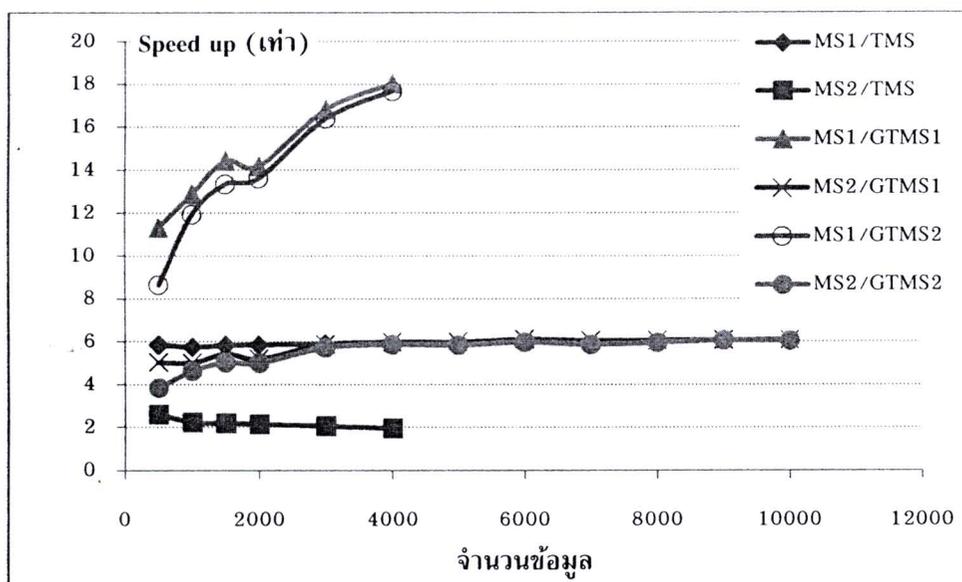
ตารางที่ 9 จำนวนเท่าของการเพิ่มความเร็วของแต่ละขั้นตอนวิธีในการดำเนินการกับข้อมูลทดสอบที่มีจำนวน ข้อมูลแตกต่างกัน โดยใช้แบนด์วิดท์เท่ากับ 10

จำนวน ข้อมูล	จำนวนเท่าของการเพิ่มความเร็ว (เท่า)					
	TMS		GTMS1		GTMS2	
	MS1/TMS	MS2/TMS	MS1/GTMS1	MS2/GTMS1	MS1/GTMS2	MS2/GTMS2
500	3.67	1.54	6.86	2.88	3.76	1.58
1000	4.56	1.52	10.86	3.63	6.50	2.17
1500	4.62	1.47	11.96	3.81	7.81	2.49
2000	4.71	1.43	12.39	3.75	9.34	2.82
3000	4.68	1.38	13.72	4.05	9.88	2.92
4000	4.69	1.30	14.84	4.12	11.75	3.26
5000	-	-	-	4.13	-	3.08
6000	-	-	-	4.14	-	3.28
7000	-	-	-	4.22	-	3.40
8000	-	-	-	4.13	-	3.35
9000	-	-	-	4.17	-	3.49
10000	-	-	-	4.15	-	3.64

จากตารางที่ 9 เป็นการพิจารณาการเพิ่มความเร็วของแต่ละขั้นตอนวิธีเมื่อเทียบกับขั้นตอนวิธีการย้ายตามค่าเฉลี่ยเดิม (MS1) และ ขั้นตอนวิธีการย้ายค่าเฉลี่ยที่มีการตัดตัวหยุดออก (MS2) ซึ่งได้ทำการวิเคราะห์จำนวนเท่าของการเพิ่มความเร็วจากข้อมูลในตารางที่ 7 ที่ใช้แบนด์วิดท์เท่ากับ 10 เพื่อทำการหาว่าขั้นตอนวิธีใดที่จะสามารถเพิ่มความเร็วมากที่สุด จะเห็นได้ว่าที่จำนวนข้อมูล 4000 ข้อมูล เป็นจำนวนข้อมูลมากที่สุดที่สามารถเปรียบเทียบเวลาของแต่ละขั้นตอนวิธีได้ทั้งหมด ปรากฏว่าขั้นตอนวิธี GTMS1 จะให้จำนวนเท่าของการเพิ่มความเร็วมากที่สุดคือ 14.84 เท่า และ 4.12 เท่า เมื่อเทียบกับทั้ง MS1 และ MS2 ตามลำดับ จากภาพที่ 34 โดยแกน x แสดงจำนวนข้อมูล ส่วนแกน y แสดงจำนวนเท่า จะเห็นได้อย่างชัดเจนว่า GTMS1 และ GTMS2 นั้นมีแนวโน้มการเพิ่มความเร็วของแต่ละขั้นตอนวิธีเมื่อเทียบกับ MS1 นั้นเพิ่มขึ้นอย่างรวดเร็วตั้งแต่ 6.86 ถึง 14.84 เท่า เมื่อข้อมูลมีจำนวนเพิ่มขึ้นจาก 500 ถึง 4000 ข้อมูล ส่วน GTMS1 และ GTMS2 ที่เทียบกับ MS2 นั้นเพิ่มขึ้นในช่วงแรกและมีแนวโน้มการเพิ่มความเร็วเกือบจะคงที่ประมาณ 4 เท่า เมื่อข้อมูลมีจำนวนเพิ่มมากขึ้น



ภาพที่ 34 จำนวนเท่าของการเพิ่มความเร็วของแต่ละขั้นตอนวิธีในการดำเนินการกับข้อมูลทดสอบที่มีจำนวนข้อมูลแตกต่างกัน โดยใช้แบนด์วิดท์เท่ากับ 10



ภาพที่ 35 จำนวนเท่าของการเพิ่มความเร็วของแต่ละขั้นตอนวิธีในการดำเนินการกับข้อมูลทดสอบที่มีจำนวนข้อมูลแตกต่างกัน โดยใช้แบนด์วิธเท่ากับ 2

ตารางที่ 10 จำนวนเท่าของการเพิ่มความเร็วของแต่ละขั้นตอนวิธีในการดำเนินการกับข้อมูลทดสอบที่มีจำนวน ข้อมูลแตกต่างกัน โดยใช้แบนด์วิธเท่ากับ 2

จำนวน ข้อมูล	จำนวนเท่าของการเพิ่มความเร็ว (เท่า)					
	TMS		GTMS1		GTMS2	
	MS1/TMS	MS2/TMS	MS1/GTMS1	MS2/GTMS1	MS1/GTMS2	MS2/GTMS2
500	5.85	2.61	11.32	5.04	8.65	3.85
1000	5.75	2.23	12.91	5.01	11.93	4.63
1500	5.83	2.20	14.43	5.44	13.35	5.03
2000	5.85	2.15	14.18	5.21	13.62	5.00
3000	5.89	2.06	16.81	5.88	16.39	5.73
4000	5.87	1.95	18.01	5.98	17.66	5.87
5000	-	-	-	5.99	-	5.85
6000	-	-	-	6.12	-	5.98
7000	-	-	-	6.06	-	5.87
8000	-	-	-	6.06	-	5.96
9000	-	-	-	6.08	-	6.06
10000	-	-	-	6.09	-	6.05

จากตารางที่ 10 เป็นการพิจารณาการเพิ่มความเร็ของแต่ละขั้นตอนวิธีเมื่อเทียบกับขั้นตอนวิธีการย้ายตามค่าเฉลี่ยเดิม (MS1) และ ขั้นตอนวิธีการย้ายค่าเฉลี่ยที่มีการตัดตัวหยุดออก (MS2) ซึ่งได้ทำการวิเคราะห์จำนวนเท่าของการเพิ่มความเร็จากข้อมูลในตารางที่ 8 ที่ใช้แบนด์วิดท์เท่ากับ 2 เพื่อทำการหาว่าขั้นตอนวิธีใดที่จะสามารถเพิ่มความเร็มากที่สุด จะเห็นได้ว่าที่จำนวนข้อมูล 4000 ข้อมูล เป็นจำนวนข้อมูลมากที่สุดที่สามารถเปรียบเทียบเวลาของแต่ละขั้นตอนวิธีได้ทั้งหมด ปรากฏว่าขั้นตอนวิธี GTMS1 จะให้จำนวนเท่าของการเพิ่มความเร็มากที่สุดคือ 18.01 เท่า และ 5.98 เท่าเมื่อเทียบกับทั้ง MS1 และ MS2 ตามลำดับ จากภาพที่ 35 จะเห็นได้อย่างชัดเจนว่า GTMS1 และ GTMS2 นั้นมีแนวโน้มการเพิ่มความเร็ของแต่ละขั้นตอนวิธีเมื่อเทียบกับ MS1 นั้นเพิ่มขึ้นอย่างรวดเร็วเมื่อข้อมูลมีจำนวนเพิ่มมากขึ้น ส่วน GTMS1 และ GTMS2 ที่เทียบกับ MS2 นั้นมีแนวโน้มการเพิ่มความเร็เกือบคงที่เมื่อข้อมูลมีจำนวนเพิ่มมากขึ้น รวมทั้ง TMS เมื่อเทียบกับ MS1 นั้นมีแนวโน้มการเพิ่มความเร็เกือบคงที่เมื่อข้อมูลมีจำนวนเพิ่มมากขึ้นด้วยเช่นกัน

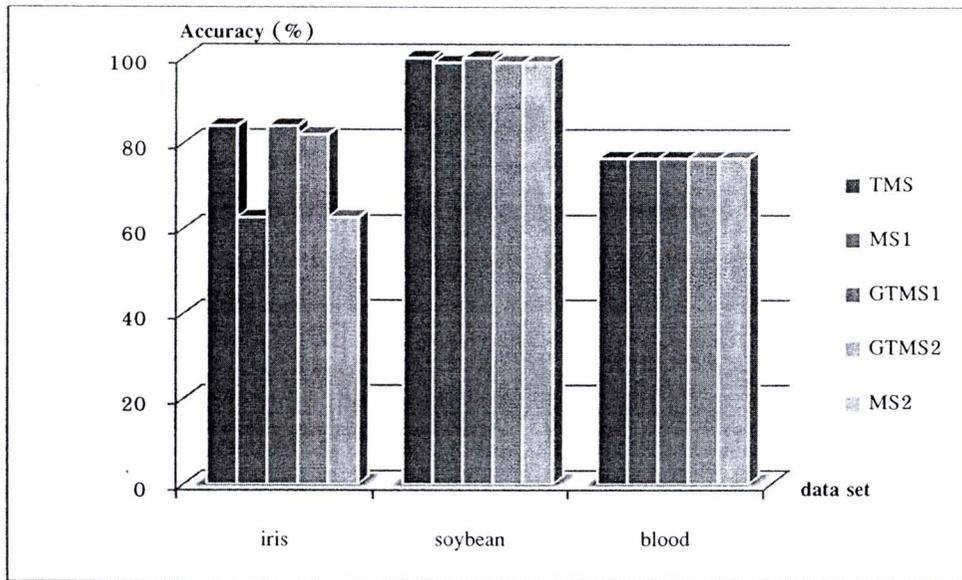
จะเห็นได้จากการเปรียบเทียบเวลาการทำงานของแต่ละขั้นตอนวิธีเมื่อมีจำนวนข้อมูลเพิ่มขึ้นนั้น ขั้นตอนวิธี GTMS1 ซึ่งเป็นขั้นตอนวิธีการลำเลียงย้ายค่าเฉลี่ยแบบครอบคลุม ที่กำหนด $\alpha > 1$ ทำให้ได้ผลลัพธ์เดียวกับ TMS นั้นเป็นขั้นตอนวิธีที่ดำเนินการเร็วที่สุดเมื่อเทียบกับขั้นตอนวิธีการย้ายตามค่าเฉลี่ยเดิม แสดงว่ามีจำนวนข้อมูลจำนวนมากที่ต้องย้ายไปในเส้นทางเดียวกันและขั้นตอนวิธีที่นำเสนอขึ้นได้ลดการคำนวณจุดเหล่านี้ทำให้เวลาในการดำเนินการน้อยกว่ามาก ซึ่งผลลัพธ์ของขั้นตอนวิธีทั้งหมดที่ได้ทดลองกับข้อมูลชุดนี้สามารถจัดกลุ่มข้อมูลได้ถูกต้องทุกขั้นตอนวิธี โดยจากผลการทดลองสามารถเพิ่มความเร็ได้สูงสุด 18.01 เท่า เมื่อเทียบกับ MS1 และ 5.98 เท่าเมื่อเทียบกับ MS2 ของการทดลองที่ใช้แบนด์วิดท์เท่ากับ 2

2. การประยุกต์ใช้กับการจัดกลุ่มข้อมูล

เมื่อทำการทดลองในการเปรียบเทียบความเร็โดยใช้ข้อมูล Iris, Soybean และ Blood รายละเอียดของข้อมูลตามหัวข้อ 4.2 ในบทที่ 3 โดยได้ทำการจับเวลาการทำงานของแต่ละขั้นตอนวิธีจำนวน 10 รอบ โดยผู้วิจัยได้กำหนดค่า Threshold เท่ากันคือ 10^{-7} โดยแบนด์วิดท์ที่ใช้ในการทดลองนี้เนื่องจากงานวิจัยนี้ไม่ได้ศึกษาคำนวณโดยตรง ผู้ทดลองจึงได้ทำการทดลองหลายครั้งเพื่อเลือกแบนด์วิดท์ที่ให้ผลลัพธ์ที่ดีที่สุดสำหรับแต่ละข้อมูล ผลการทดลองแสดงดังต่อไปนี้

ตารางที่ 11 ความถูกต้องของการจัดกลุ่มของข้อมูล iris soybean และ blood เมื่อใช้ขั้นตอนวิธีแบบต่าง ๆ

ข้อมูล	ความถูกต้องของการจัดกลุ่ม (%)				
	TMS	MS1	GTMS1	GTMS2	MS2
iris	84	62.67	84	82	62.67
soybean	100	98.87	100	98.87	98.87
blood	76.47	76.47	76.47	76.47	76.47



ภาพที่ 36 ความถูกต้องของการจัดกลุ่มของข้อมูล iris, soybean และ blood เมื่อใช้ขั้นตอนวิธีแบบต่าง ๆ

ค่าความถูกต้องที่วัดจากการทดลองนี้หาได้จากจำนวนข้อมูลที่ถูกจัดอยู่ในกลุ่มที่ถูกต้องเทียบกับข้อมูลทั้งหมดซึ่งการจัดกลุ่มข้อมูลเหล่านี้เกิดจากการดำเนินการตามขั้นตอนวิธีการย้ายค่าเฉลี่ยแบบต่าง ๆ จากนั้นจัดกลุ่มข้อมูลโดยใช้การเชื่อมต่อกันของกราฟ โดยสร้างกราฟจากข้อมูลที่อยู่ใกล้กันน้อยกว่าครึ่งหนึ่งของแบนด์วิดท์ให้มันเส้นเชื่อมถึงกัน ความถูกต้องของการจัดกลุ่มแสดงในภาพที่ 36 จะเห็นได้ว่า ข้อมูล iris ในขั้นตอนวิธี TMS และ GTMS1 จัดกลุ่มข้อมูลได้ถูกต้องที่สุดเมื่อเทียบกับขั้นตอนวิธีอื่น โดยมีความถูกต้อง 84 % ขณะที่ MS1 และ MS2 ให้ค่าความถูกต้องของการจัดกลุ่มน้อยกว่าโดยมีค่าเพียง 62.67 % ดังแสดงในตาราง 11 พิจารณาข้อมูล soybean ขั้นตอนวิธี TMS และ GTMS1 สามารถจัดกลุ่มได้ถูกต้องทั้งหมด 100 % ขณะที่ MS1, MS2 และ GTMS2 นั้นจะให้ค่าความถูกต้องของการจัดกลุ่มน้อยกว่าคือ 98.87 % และข้อมูล blood นั้นทุกขั้นตอนวิธีจัดกลุ่มได้ถูกต้องเท่ากันทั้งหมดคือ 76.47 %

เมื่อพิจารณาเวลาในการดำเนินการจัดกลุ่มข้อมูลของแต่ละขั้นตอนวิธีโดยเวลาที่วัดในการทดลองนี้วัดรวมไปถึงการจัดกลุ่มข้อมูล ซึ่งแต่ละขั้นตอนวิธีใช้ขั้นตอนวิธีการจัดกลุ่มแบบเดียวกัน ผลการทดลองแสดงดังต่อไปนี้

ตารางที่ 12 เวลาและจำนวนเท่าของการเพิ่มความเร็ของแต่ละขั้นตอนวิธีเมื่อเทียบกับ MS1 และ MS2 ในการดำเนินการกับข้อมูล iris โดยใช้แบนด์วิดท์เท่ากับ 12

MS1 Time (second)	MS2 Time (second)	TMS			GTMS1			GTMS2		
		Time (second)	speedup		Time (second)	speedup		Time (second)	speedup	
			MS1	MS2		MS1	MS2		MS1	MS2
0.7597	0.3446	0.0451	16.85	7.65	0.0337	22.54	10.22	0.1488	5.10	2.32

ข้อมูลของดอก iris นี้ใช้ข้อมูลดั้งเดิมที่ได้จากฐานข้อมูลของ UCI ได้ผลการทดลองดังแสดงในตารางที่ 12 จะเห็นได้ว่า GTMS1 เป็นขั้นตอนวิธีที่เร็วที่สุดและยังมีความถูกต้องของการจัดกลุ่มมากที่สุด โดยเมื่อเทียบกับทั้ง MS1 และ MS2 สามารถเพิ่มความเร็วของการจัดกลุ่มได้ถึง 22.54 เท่า และ 10.22 เท่าตามลำดับ

ตารางที่ 13 เวลาและจำนวนเท่าของการเพิ่มความเร็วของแต่ละขั้นตอนวิธีเมื่อเทียบกับ MS1 และ MS2 ในการดำเนินการกับข้อมูล Soybean โดยใช้แบนด์วิดท์เท่ากับ 0.5

MS1 Time (second)	MS2 Time (second)	TMS			GTMS1			GTMS2		
		Time (second)	speedup		Time (second)	speedup		Time (second)	speedup	
			MS1	MS2		MS1	MS2		MS1	MS2
0.1985	0.0636	0.0635	3.13	1.00	0.0244	8.13	2.61	0.0571	3.48	1.11

จากตารางที่ 13 ข้อมูลโรคของถั่วเหลือง ในแต่ละมิติผู้ทดลองได้ทำการปรับค่าข้อมูลให้อยู่ในช่วง 0 ถึง 1 จากผลการทดลองนี้ GTMS1 เป็นขั้นตอนวิธีที่เร็วที่สุดและยังมีความถูกต้องของการจัดกลุ่มมากที่สุด เช่นเดียวกับการทดลองกับข้อมูล iris โดยเมื่อเทียบกับทั้ง MS1 และ MS2 สามารถเพิ่มความเร็วของการจัดกลุ่มได้ถึง 8.13 เท่า และ 2.61 เท่าตามลำดับ

ตารางที่ 14 เวลา และจำนวนเท่าของการเพิ่มความเร็วของแต่ละขั้นตอนวิธีเมื่อเทียบกับ MS1 และ MS2 ในการดำเนินการกับข้อมูล Blood โดยใช้แบนด์วิดท์เท่ากับ 0.04

MS1 Time (second)	MS2 Time (second)	TMS			GTMS1			GTMS2		
		Time (second)	speedup		Time (second)	speedup		Time (second)	speedup	
			MS1	MS2		MS1	MS2		MS1	MS2
7.3956	1.8565	0.4401	16.80	4.22	0.1815	40.75	10.23	0.3856	19.18	4.81

ในการทดลองนี้ ข้อมูลของการบริจาคเลือดที่เมือง Hsin Chu ประเทศไต้หวัน ในแต่ละมิติผู้ทดลองได้ทำการปรับค่าข้อมูลให้อยู่ในช่วง 0 ถึง 1 เช่นเดียวกับข้อมูลที่สอง ได้ผลการทดลองดังแสดงในตารางที่ 14 จะเห็นได้ว่า GTMS1 เป็นขั้นตอนวิธีที่เร็วที่สุดและยังมีความถูกต้องของการจัดกลุ่มมากที่สุดเหมือนกับสองการทดลองที่แล้ว โดยเมื่อเทียบกับทั้ง MS1 และ MS2 สามารถเพิ่มความเร็วของการจัดกลุ่มได้ถึง 40.75 เท่า และ 10.23 เท่าตามลำดับ



3. การประยุกต์ใช้กับการแยกส่วนประกอบของภาพทั่วไป

เมื่อทำการทดลองแยกส่วนประกอบของภาพโดยใช้แบนด์วิดท์ที่แตกต่างกัน และกำหนดค่าพารามิเตอร์ต่าง ๆ ตามหัวข้อที่ 4.3 ในบทที่ 3 ทำการวัดเวลาเฉลี่ยของการทดลองทั้ง 10 รอบ ได้ผลดังตารางที่ 15 พร้อมทั้งจำนวนเท่าของการเพิ่มความเร็วแสดงได้ผลดังตารางที่ 16 รวมถึง ค่าความถูกต้องของแต่ละขั้นตอนวิธีเมื่อเทียบกับขั้นตอนวิธีการย้ายตามค่าเฉลี่ยแบบมาตรฐานแสดงในตารางที่ 17

ตารางที่ 15 เวลาที่ใช้ในการประมวลผลของใช้ในการแยกส่วนประกอบภาพ

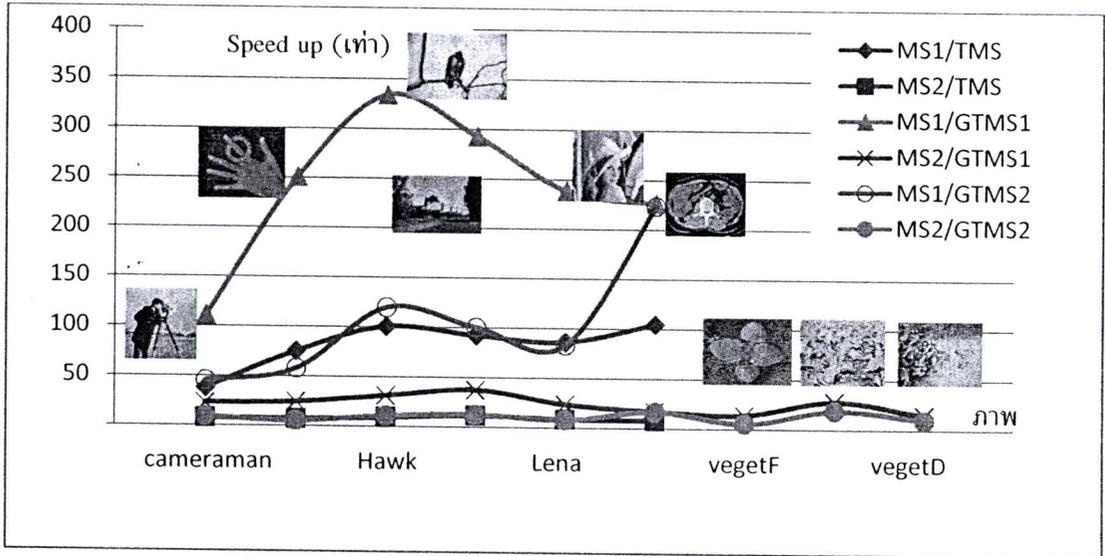
ข้อมูล	s	เฉลี่ยใน 10 รอบ(วินาที)				
		MS	MS2	TMS	GTMS1	GTMS2
Hand	500	747.73	73.19	9.92	2.98	12.92
	200	643.89	64.35	10.17	3.15	9.2
Hawk	500	1442.8	134.22	14.35	4.32	12.05
	200	1477.17	252.16	14.7	4.72	10.62
Houses	500	500	75.61	12.36	3.82	13.2
	200	1215.16	154.81	13.19	4.15	12.11
Lena	500	304.6	49.2	5.57	1.84	6.3
	200	577.83	57.29	6.72	2.42	6.97
cameraman	400	124.07	25.51	4.457	1.53	3.78
	200	174.04	36.10	4.579	1.57	3.84
small bowel loops	100	456.48	36.49	4.876	2.16	2.46
	90	501.23	38.06	4.783	2.21	2.24
Vegetable(F)	100	-	5327.93	-	604.82	1822.47
	50	-	8499.35	-	606.04	1525.82
Vegetable(D)	100	-	8044.08	-	476.49	1085.72
	50	-	13274.20	-	475.38	709.43
Vegetable(B)	100	-	3964.05	-	458.62	796.45
	50	-	6729.63	-	455.16	662.88

ตารางที่ 16 จำนวนเท่าของการเพิ่มความเร็วของการแยกส่วนประกอบภาพของขั้นตอนต่าง ๆ เมื่อเทียบกับขั้นตอนวิธีย้ายตามค่าเฉลี่ยเดิม

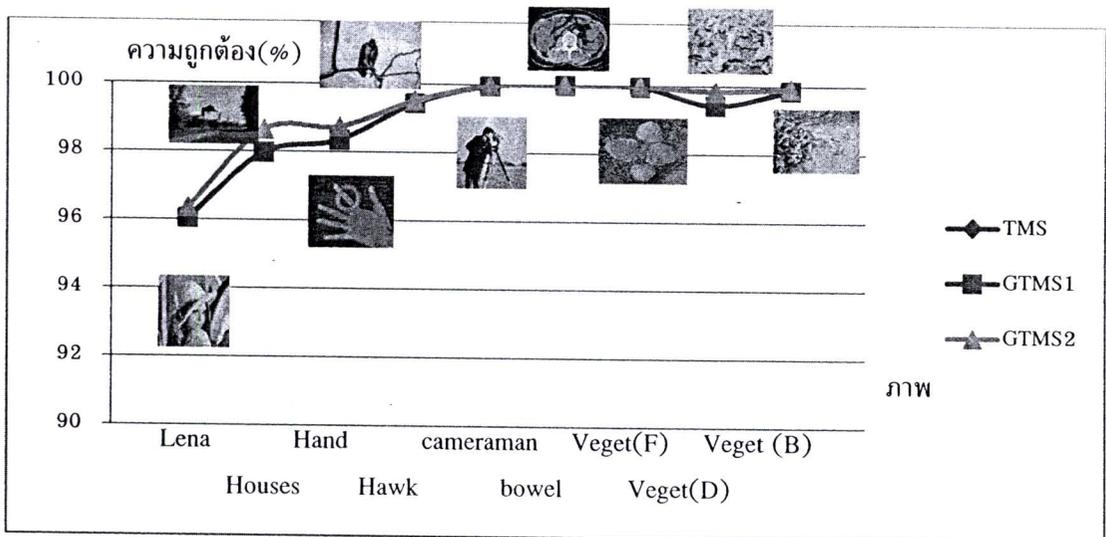
ข้อมูล	s	Speedup(เท่า)					
		TMS เมื่อเทียบกับ		GTMS1 เมื่อเทียบกับ		GTMS2 เมื่อเทียบกับ	
		MS1	MS2	MS1	MS2	MS1	MS2
Hand	500	75.38	7.38	250.92	24.56	57.87	5.66
	200	63.31	6.33	204.41	20.43	69.99	6.99
Hawk	500	100.54	9.35	333.98	31.07	119.73	11.14
	200	100.49	17.15	312.96	53.42	139.09	23.74
Houses	500	40.45	6.12	130.89	19.79	37.88	5.73
	200	92.13	11.74	292.81	37.30	100.34	12.78
Lena	500	54.69	8.83	165.54	26.74	48.35	7.81
	200	85.99	8.53	238.77	23.67	82.90	8.22
cameraman	400	27.84	5.72	81.21	16.70	32.85	6.75
	200	38.00	7.88	110.91	23.00	45.35	9.41
small bowel	100	93.62	7.48	211.51	16.91	185.64	14.84
loops	90	104.79	7.96	226.72	17.22	223.80	17.00
Vegetable(F)	100	-	-	-	8.81	-	2.92
	50	-	-	-	14.02	-	5.57
Vegetable(D)	100	-	-	-	16.88	-	7.41
	50	-	-	-	27.92	-	18.71
Vegetable(B)	100	-	-	-	8.64	-	4.98
	50	-	-	-	14.79	-	10.15

จำนวนเท่าของการเพิ่มความเร็วเป็นการวัดจำนวนเท่าของเวลาเมื่อเทียบกับขั้นตอนวิธีการย้ายตามค่าเฉลี่ย เป็นอัตราส่วนระหว่างเวลาของขั้นตอนวิธีการย้ายตามค่าเฉลี่ยกับแต่ละขั้นตอนวิธีที่นำเสนอขึ้น จะเห็นว่าวิธีที่นำเสนอขึ้นใหม่ คือ GTMS1 นี้สามารถเพิ่มความเร็วในการแยกส่วนประกอบภาพได้สูงถึง 81 – 333 เท่าเมื่อเทียบกับ ขั้นตอนวิธีการย้ายค่าเฉลี่ยแบบไม่ตัดตัวที่หยุดออก (MS1) และ 8 – 53 เท่าเมื่อเทียบกับ ขั้นตอนวิธีการย้ายค่าเฉลี่ยแบบที่ตัดตัวที่หยุดออก (MS2) ซึ่ง GTMS1 จะไม่พิจารณาทิศทางของ ผู้ขนส่ง และผู้พ่วง และมีความคลาดเคลื่อนเล็กน้อยดังแสดงในภาพที่ 33 ซึ่งหากต้องการให้ความคลาดเคลื่อนเมื่อเทียบกับขั้นตอนวิธีการย้ายค่าเฉลี่ยลดลงก็สามารถเพิ่มการพิจารณาทิศทางของ ผู้ขนส่ง และผู้พ่วงเข้าไปด้วย ดังใน

ตารางที่ 17 ของ GTMS2 จะสามารถลดความคลาดเคลื่อนอยู่ในช่วง 0.01 ถึง 4.4 เปอร์เซ็นต์ ซึ่งลดลงจาก TMS และ GTMS1 ที่มีความคลาดเคลื่อนอยู่ในช่วง 0.03 ถึง 5.88 เปอร์เซ็นต์ ซึ่งค่าความถูกต้องนี้วัดได้โดยนับจำนวนข้อมูลที่จัดกลุ่มถูกแต่ละขั้นตอนวิธีเมื่อเทียบกับการจัดกลุ่มของ MS หากด้วยจำนวนข้อมูลทั้งหมด ซึ่ง MS1 และ MS2 ได้ผลลัพธ์เดียวกัน และแสดงค่าความถูกต้องเปรียบเทียบกันทั้งสามขั้นตอนวิธีดังแสดงในภาพที่ 37



ภาพที่ 37 จำนวนเท่าของการเพิ่มความเร็วของการแยกส่วนประกอบภาพของขั้นตอนต่าง ๆ เมื่อเทียบกับขั้นตอนวิธีย้ายตามค่าเฉลี่ยเดิม



ภาพที่ 38 ความสามารถในการแยกส่วนประกอบของภาพได้ถูกต้องเมื่อเทียบกับ MS

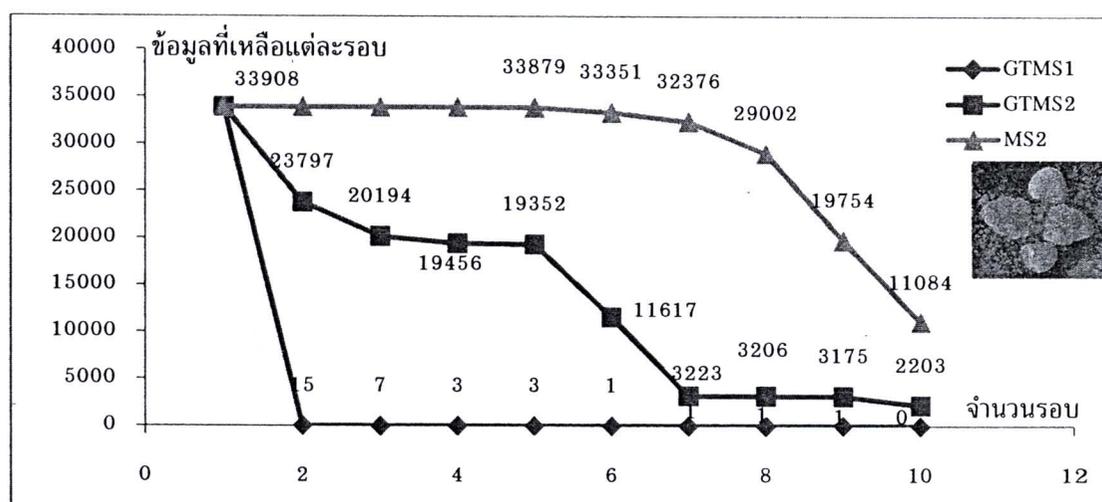
ตารางที่ 17 ค่าความสามารถในการแยกส่วนประกอบของภาพได้ถูกต้องเมื่อเทียบกับขั้นตอนวิธีการย้ายตามค่าเฉลี่ยเดิม

ข้อมูล	s	Accuracy (%)		
		TMS	GTMS1	GTMS2
Hand	500	98.35	98.35	98.74
	200	97.13	97.13	98.77
Hawk	500	99.44	99.44	99.52
	200	99.2	99.2	99.57
Houses	500	97.97	97.97	98.64
	200	99.52	99.52	99.87
Lena	500	96.04	96.04	96.36
	200	94.12	94.12	95.6
cameraman	400	99.96	99.96	99.96
	200	99.94	99.94	99.52
small bowel loops	100	100	100	100
	90	100	100	100
Vegetable(F)	100	-	99.97	99.99
	50	-	99.96	99.96
Vegetable(D)	100	-	99.43	99.85
	50	-	99.6	99.68
Vegetable(B)	100	-	99.87	99.94
	50	-	99.77	99.89

ทั้งนี้หากต้องการให้มีความคลาดเคลื่อนน้อยลงก็สามารถลดหรือเพิ่มค่า α ได้แต่จะทำให้การเพิ่มความเร็วน้อยลงตามไปด้วย เนื่องจาก GTMS2 จะมีการคำนวณใกล้เคียงกับ MS เดิมมากขึ้น ทำให้มีจำนวนข้อมูลที่เหลือในการดำเนินการในแต่ละรอบเหลือจำนวนมาก ดังจะเห็นในตารางที่ 18 จำนวนข้อมูลที่เหลือในแต่ละรอบของ GTMS1 ลดลงอย่างมาก ซึ่งจากภาพที่ 39 เป็นการนำภาพ Vegetable (F) เมื่อใช้ $\sigma = 100$ จะเห็นได้อย่างชัดเจนว่าในการคำนวณในรอบที่ 2 จะเหลือข้อมูลเพียง 15 ข้อมูล จากทั้งหมด 33,908 ข้อมูล ในขณะที่ GMS2 ข้อมูลเหลือ 23,797 ข้อมูล และตัวอย่างผลการแยกส่วนประกอบของภาพโดยใช้ขั้นตอนวิธีที่นำเสนอขึ้นแสดงตามภาพที่ 41 และ 42

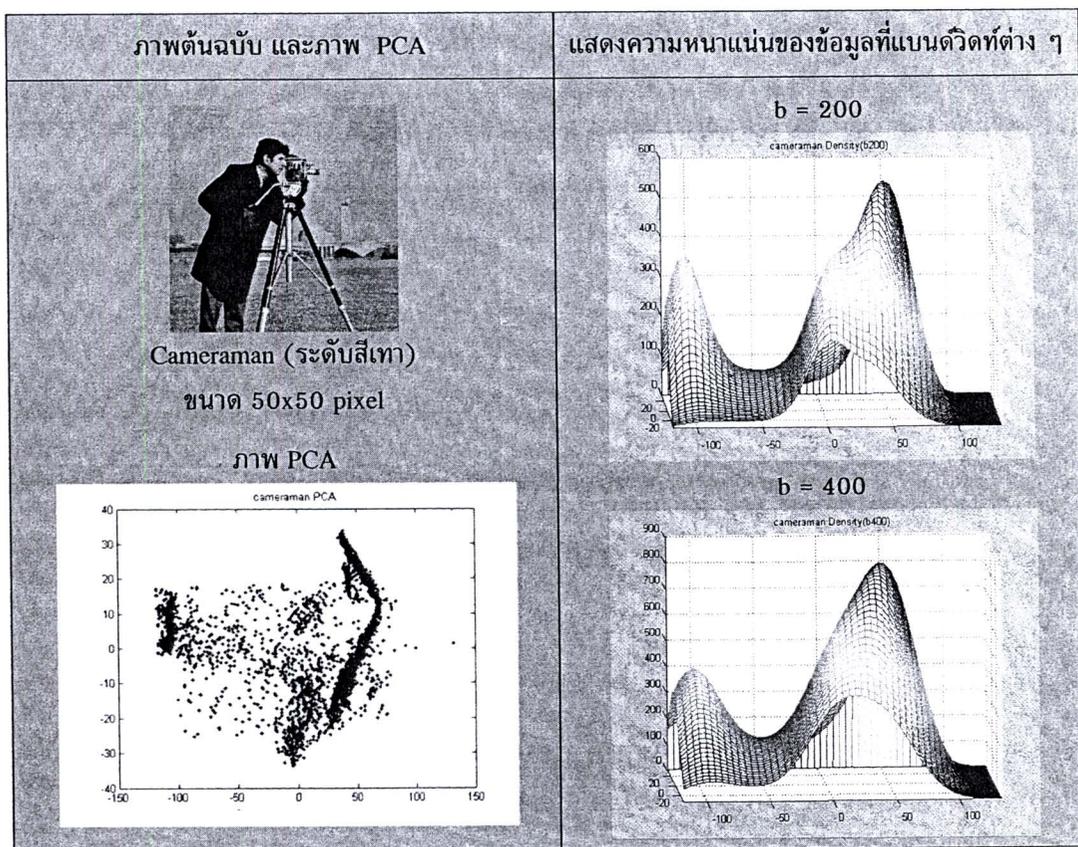
ตารางที่ 18 จำนวนข้อมูลที่เหลืออยู่ใน 5 รอบแรกของแต่ละขั้นตอนวิธี

ข้อมูล	σ	MS2	TMS	GMS1	GMS2
Hand	500	3100,3100,3100,	3100,82,19,	3100,82,19,	3100,1822,1352,
		3100,3100	10,7	10,7	1224,1094
	200	3100,3100,3100,	3100,231,44,	3100,231,44,	3100,1964,1039,
		3100,3100,	20,14	20,14	759,631
Hawk	500	3750,3750,3750,	3750,157,29,	3750,157,29,	3750,1432,991,
		3750,3750	16,9	16,9	829,749
	200	3750,3750,3750,	3750,505,176,	3750,505,176,	3750,2168,1384,
		3750,3750	88,56	88,56	1087,839
Houses	500	3750,3750,3750,	3750,104,20,	3750,104,20,	3750,2163,1254,
		3750,3750	13,10	13,10	1066,955
	200	3750,3750,3750,	3750,361,97,	3750,361,97,	3750,2638,1655,
		3750,3750	46,27	46,27	1228,966
Lena	500	2500,2500,2500,	2500,211,42,	2500,211,42,	2500,1556,985,
		2500,2500	17,6	17,6	720,593
	200	2500,2500,2500,	2500,487,157,	2500,487,157,	2500,1812,1290,
		2500,2500	74,52	75,52	1016,827
Vegeta	100	33908,33908,33908	-	33908,15,7,	33908,23797,20194,
		,33908,33879	-	3,3	19456,19352
ble(F)	50	33908,33908,33908	-	33908,46,14,	33908,19443,12861,
		,33908,33908	-	10,7	11407,10906



ภาพที่ 39 ตัวอย่างแสดงข้อมูลที่เหลือในแต่ละรอบของข้อมูล Vegetable(F)

ผู้ทดลองได้ทำการพิจารณาลักษณะของข้อมูลที่ใช้ในการทดลองโดยได้ใช้ทฤษฎีการวิเคราะห์องค์ประกอบหลัก (PCA) มาใช้ลดมิติข้อมูล เพื่อพิจารณาลักษณะการกระจายตัวความหนาแน่นของข้อมูล (probability density distribution) ว่าขั้นตอนวิธีที่ใช้นั้นเหมาะสมกับข้อมูลลักษณะแบบใดโดยได้ทำการทดลองโดยนำข้อมูล Cameraman, Hand, Lena, Smallbowelloop, Hawk และ Houses ขนาด 5 มิติ ประกอบด้วยตำแหน่งในแนวแกน x ตำแหน่งในแนวแกน y ค่าความเข้มของสีแดง ค่าความเข้มของสีเขียว และค่าความเข้มของสีน้ำเงิน โดยแทนด้วย (x, y, R, G, B) ตามลำดับ และได้นำข้อมูล ภาพทางการเกษตรคือภาพ B, D และ F ในตารางที่ 6 ขนาด 7 มิติ มาทำการแปลงข้อมูลโดยใช้วิธีการวิเคราะห์องค์ประกอบภาพ (PCA) เพื่อลดมิติข้อมูล โดยนำข้อมูลองค์ประกอบหลักที่สำคัญที่สุดสององค์ประกอบแรกของ PCA มาใช้ในการหาลักษณะการกระจายตัวความหนาแน่นของข้อมูลซึ่งใช้ฟังก์ชันประเมิณความหนาแน่น (probability density function) แบบเกาส์เซียน ได้ผลดังภาพที่ 40 โดยภาพที่แสดงความหนาแน่นของข้อมูลที่แบนตัววัดที่ต่าง ๆ นั้นแกนแรกเป็นองค์ประกอบหลักที่สำคัญที่สุดของ PCA อันดับแรก แกนที่สองเป็นองค์ประกอบหลักที่สำคัญที่สุดของ PCA อันดับสอง และแกนที่สาม คือ ค่าความหนาแน่นของข้อมูล ณ ตำแหน่งต่าง ๆ



ภาพที่ 40 ภาพต้นฉบับ ภาพที่ได้จากการทำ PCA และความหนาแน่นของข้อมูลในแบนตัววัดที่ต่างกัน



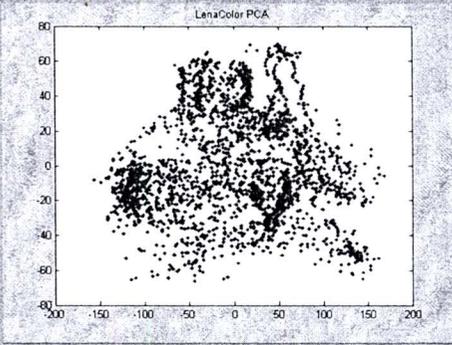
ภาพต้นฉบับ และภาพ PCA



Lena (ภาพสี)

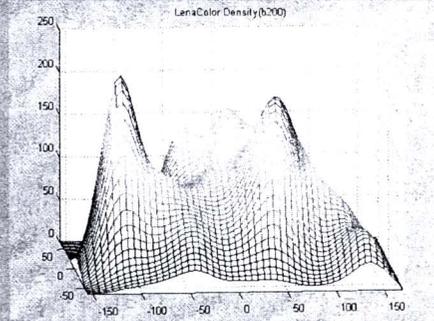
ขนาด 50x50 pixel

ภาพ PCA

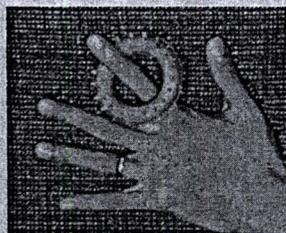
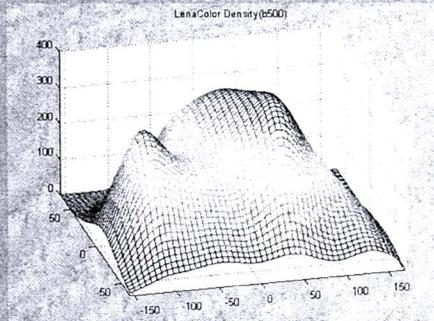


แสดงความหนาแน่นของข้อมูลที่แบนด์วิดท์ต่าง ๆ

b = 200



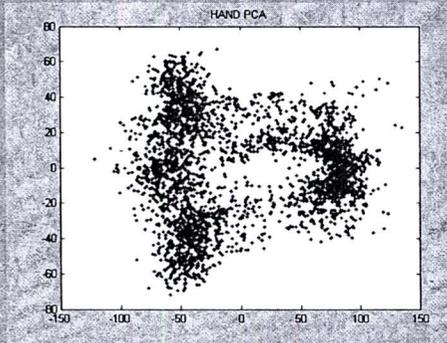
b = 500



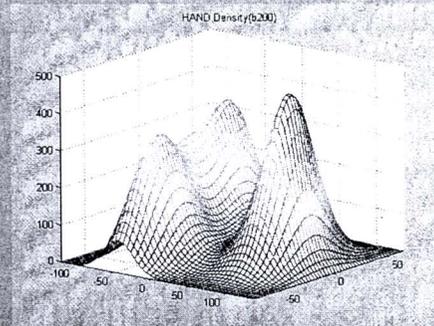
Hand (ภาพสี)

ขนาด 62x50 pixel

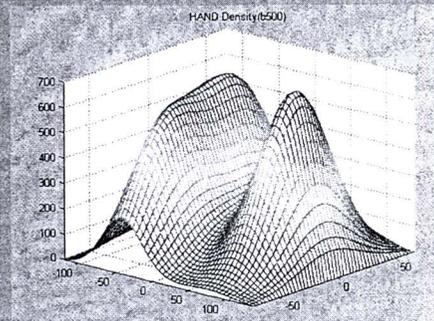
ภาพ PCA



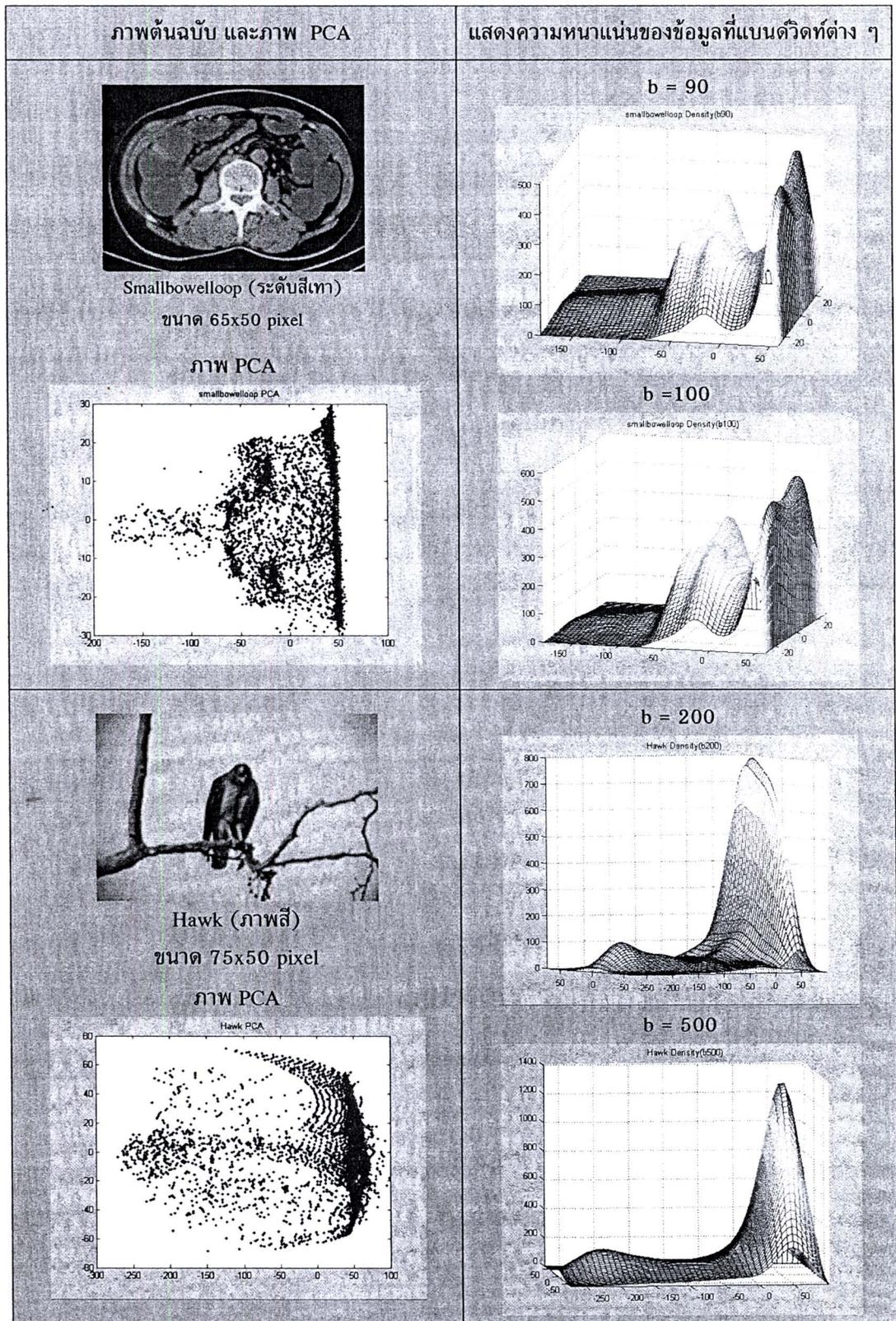
b = 200



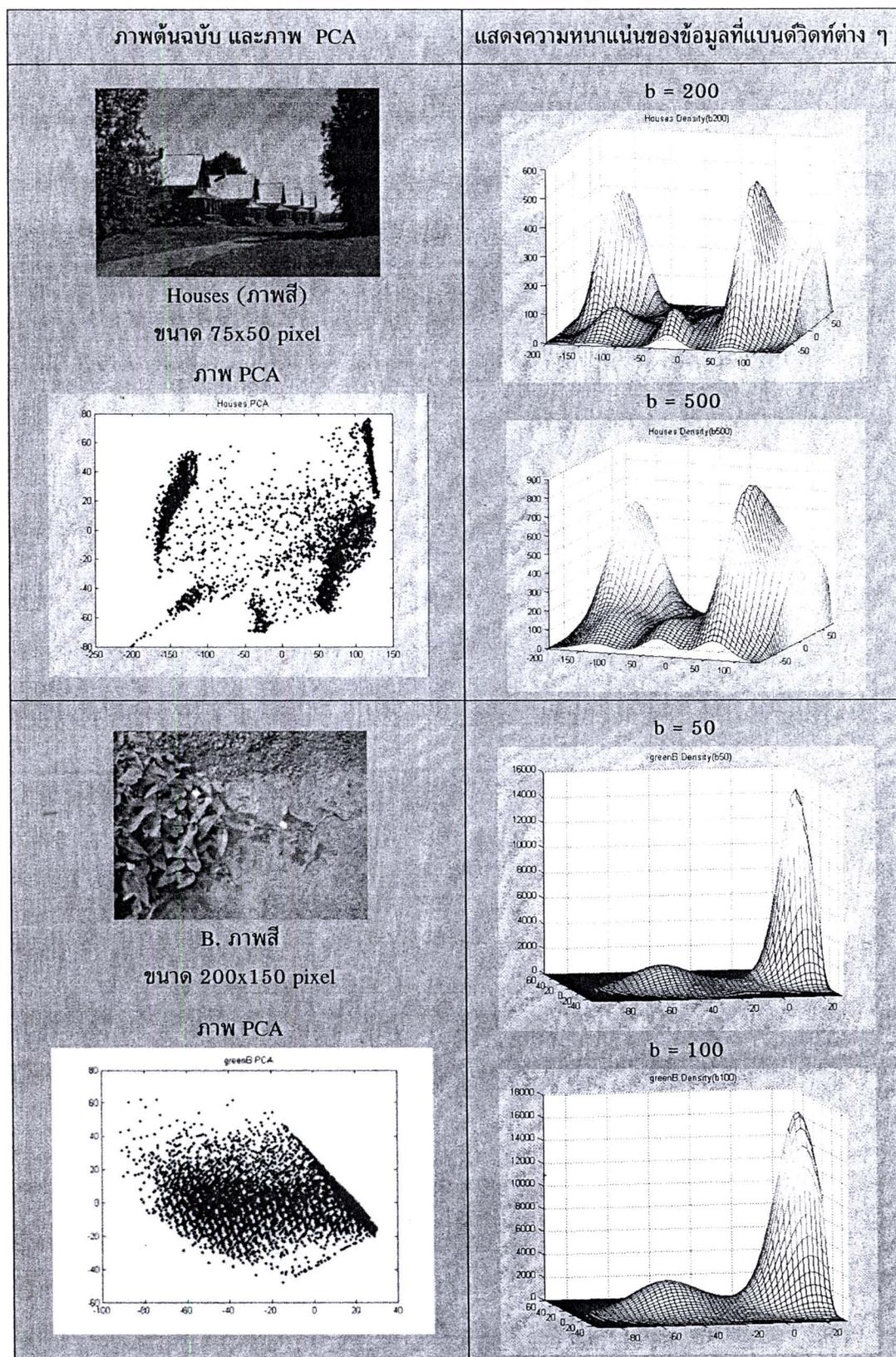
b = 500



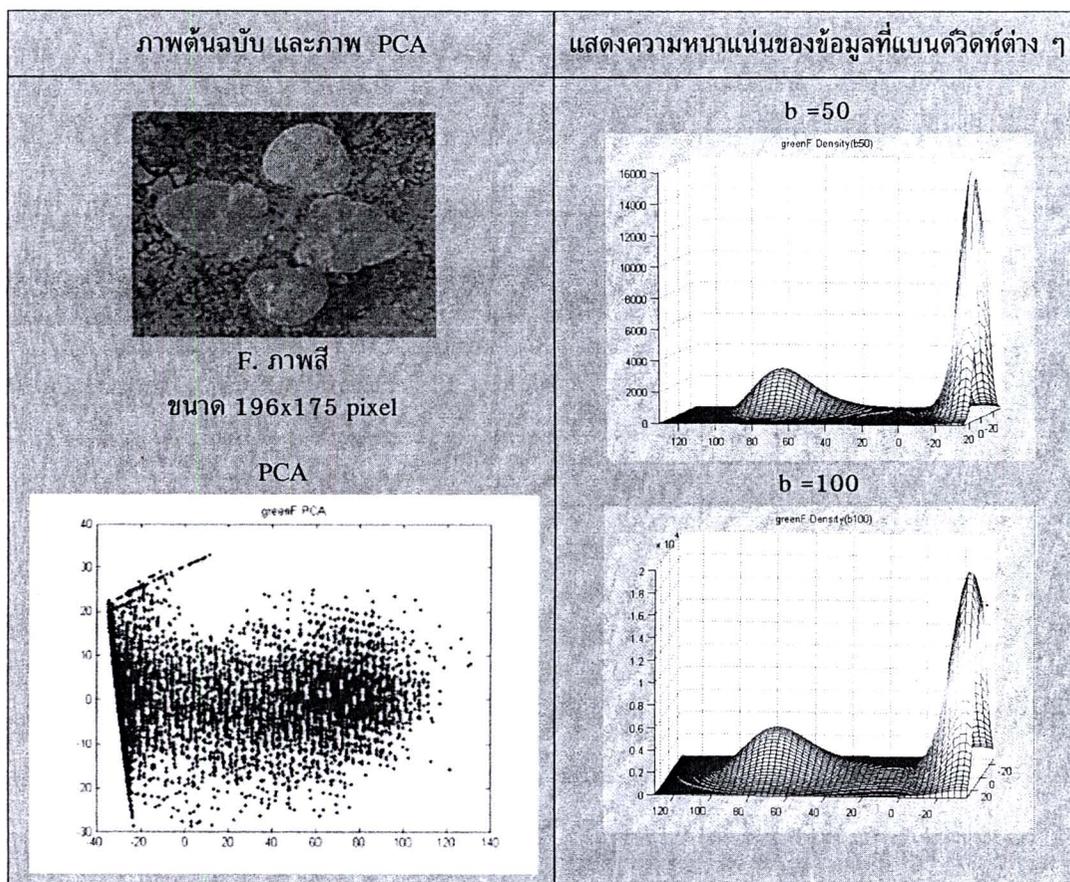
ภาพที่ 40 ภาพต้นฉบับ ภาพที่ได้จากการทำPCA และความหนาแน่นของข้อมูลในแบนด์วิดท์ที่ต่างกัน (ต่อ)



ภาพที่ 40 ภาพต้นฉบับ ภาพที่ได้จากการทำPCA และความหนาแน่นของข้อมูลในแบนตวิดท์ที่ต่างกัน (ต่อ)

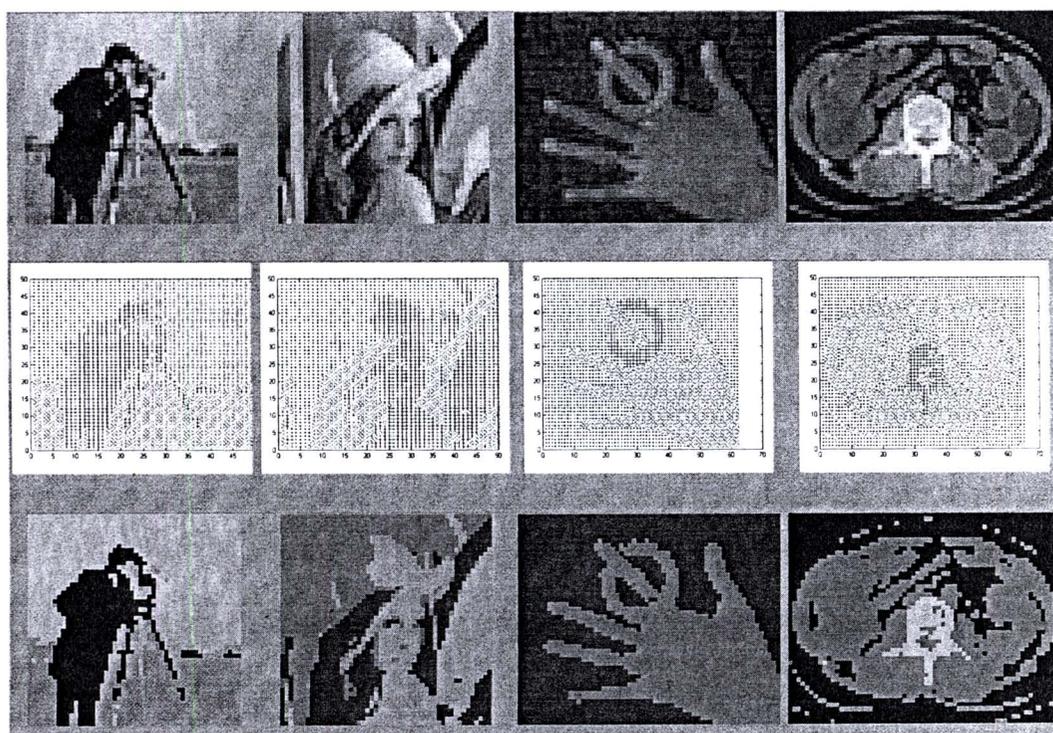


ภาพที่ 40 ภาพต้นฉบับ ภาพที่ได้จากการทำPCA และความหนาแน่นของข้อมูลในแบนด์วิดท์ที่ต่างกัน (ต่อ)

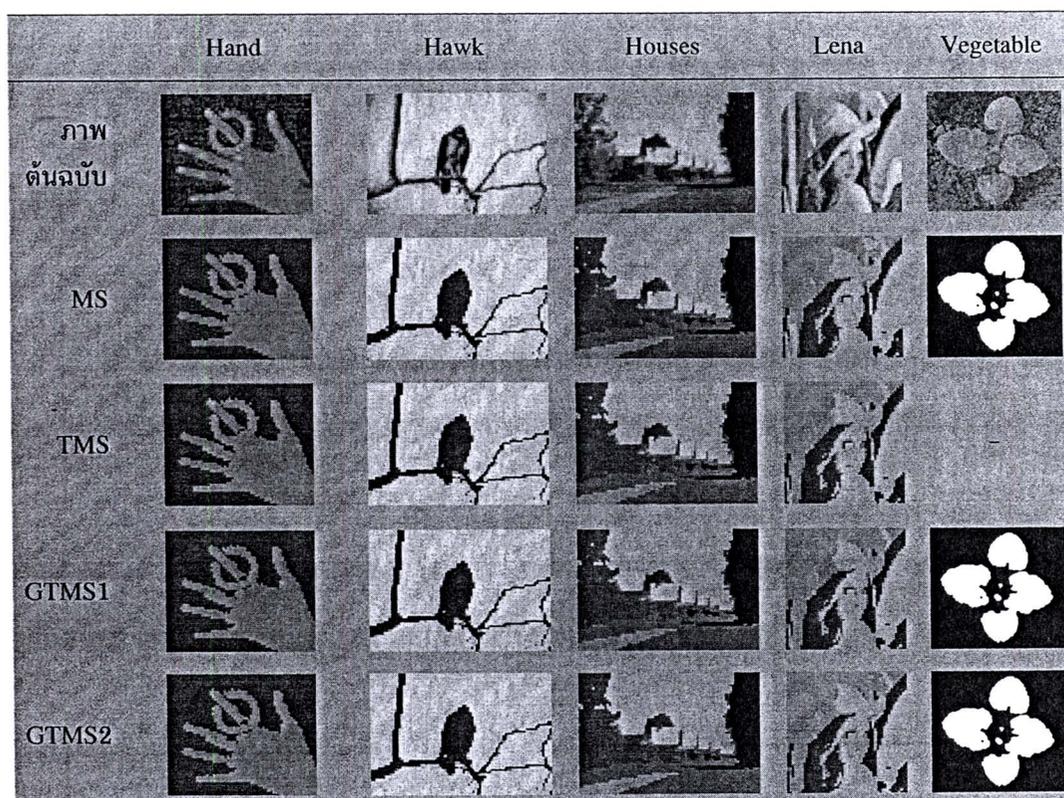


ภาพที่ 40 ภาพต้นฉบับ ภาพที่ได้จากการทำPCA และความหนาแน่นของข้อมูลในแบนตัวที่ต่างกัน (ต่อ)

จากภาพที่ 40 จะเห็นว่าภาพแต่ละภาพจะแสดงความหนาแน่นของข้อมูลแตกต่างกัน โดยภาพ Hawk และภาพทางการเกษตรคือภาพ B, D และ F จะให้ภาพที่แสดงความหนาแน่นของกลุ่มข้อมูลแตกต่างกันสูงอย่างชัดเจนโดยมีกลุ่มข้อมูลเพียงสองกลุ่ม ส่วนภาพอื่น ๆ นั้นจะเห็นได้ถึงความหนาแน่นของกลุ่มข้อมูลมีค่าใกล้เคียงกันไม่แตกต่างกันมากพร้อมกับข้อมูลมีหลายกลุ่มข้อมูล เมื่อนำภาพดังกล่าวมาพิจารณาร่วมกับการทดลองหา speed up จะเห็นได้ว่าภาพ Hawk จะให้ speed up โดยใช้ขั้นตอนวิธี GTMS1 เมื่อเทียบกับ MS1 สูงสุดคือ 333.98 เท่า พร้อมทั้งจะให้ speed up โดยใช้ขั้นตอนวิธี GTMS1 เมื่อเทียบกับ MS2 สูงสุดคือ 53.42 เท่า ดังนั้นลักษณะภาพที่เหมาะสมกับขั้นตอนวิธีนั้นที่น่าเสนอขึ้นนี้ควรเป็นภาพที่มีลักษณะความหนาแน่นของข้อมูลแต่ละกลุ่มที่แตกต่างกันมากจะทำให้การคำนวณนั้นใช้เวลาที่เร็วกว่าวิธีเดิมมาก ส่วนข้อมูลอื่น ๆ ที่นำมาทดลองซึ่งบางรูปก็มีความหนาแน่นของกลุ่มข้อมูลไม่แตกต่างกันมาก เมื่อใช้ขั้นตอนที่นำเสนอขึ้นก็ยังคงเพิ่มความเร็วได้เช่นกัน ซึ่งสอดคล้องกับความหมายของขั้นตอนวิธีการย้ายตามค่าเฉลี่ยที่ย้ายตามเส้นทางไปยังจุดที่มีความหนาแน่นสูงที่สุดนั่นเอง



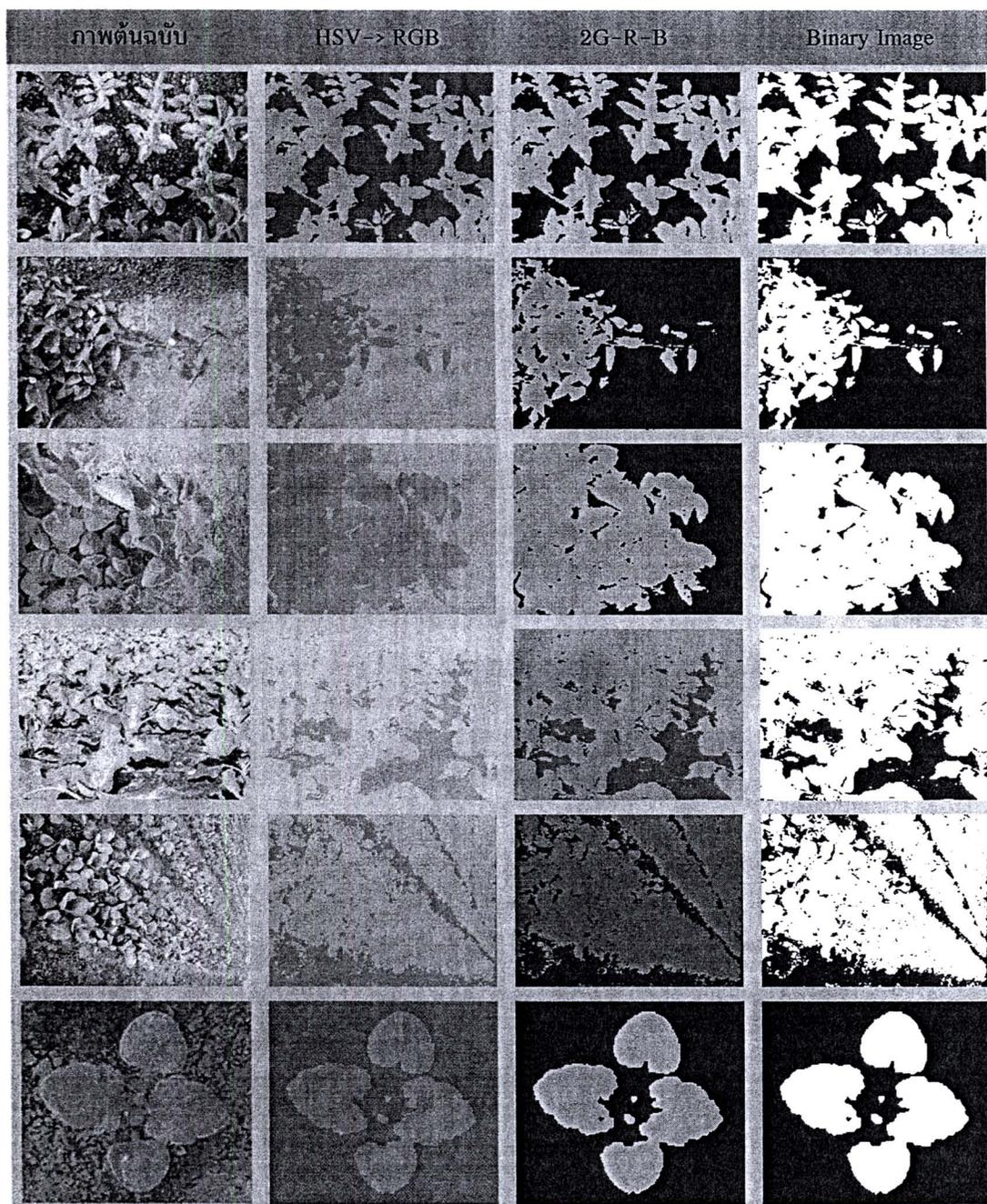
ภาพที่ 41 ตัวอย่างผลการแยกส่วนประกอบของภาพ แฉวแรกเป็นภาพต้นฉบับ แฉวกลางแสดงพิกเซลที่อยู่ในกลุ่มเดียวกัน แฉวล่างแสดงผลการแยกส่วนประกอบของภาพ ของขั้นตอนวิธี GTMS1



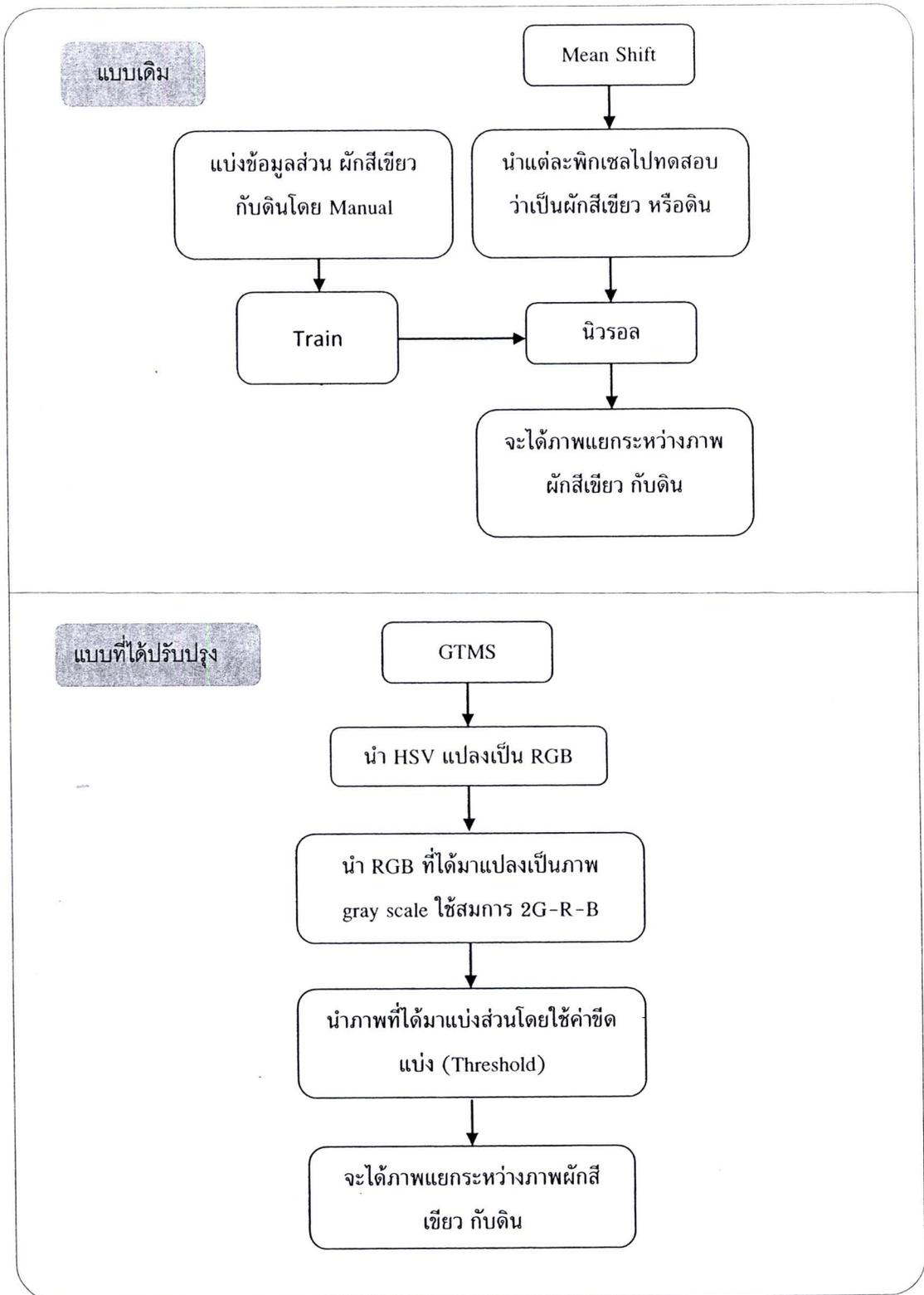
ภาพที่ 42 ตัวอย่างผลของการแยกส่วนประกอบของภาพกับขั้นตอนวิธีที่น่าเสนอขึ้น

4. การประยุกต์ใช้ขั้นตอนวิธีการลำเลียงย้ายตามค่าเฉลี่ยแบบครอบคลุมกับการแยกส่วนประกอบของภาพพืชผลทางการเกษตร

ภาพที่ 40 ด้านล่างเป็นการนำขั้นตอนวิธีการลำเลียงย้ายตามค่าเฉลี่ยแบบครอบคลุมมาประยุกต์ใช้กับการแยกส่วนประกอบของภาพพืชผลทางการเกษตรโดยแสดงผลของการแบ่งส่วนที่เป็นพืชสีเขียวออกจากส่วนอื่น ๆ ซึ่งในงานวิจัยนี้ในขั้นตอนแรกได้ใช้ข้อมูลแทนจุดด้วยเวกเตอร์ในลักษณะเดียวกับในบทความ [12] ซึ่งบทความนี้เป็นส่วนหนึ่งของโครงการใช้คอมพิวเตอร์เพื่อติดตามสถานะก๊าซไนโตรเจนของถั่วเหลืองเพื่อแยกส่วนของพืชออกจากส่วนที่เป็นเงาของพืช และส่วนที่เป็นดิน ส่วนขั้นตอนที่ใช้ในการแยกส่วนประกอบของภาพได้มีการปรับปรุงให้เหมาะสมและมีความรวดเร็ว ดังนี้ ในส่วนที่บทความใช้ ขั้นตอนวิธีการย้ายตามค่าเฉลี่ย ผู้วิจัยได้ใช้ ขั้นตอนวิธีการลำเลียงย้ายตามค่าเฉลี่ยแบบครอบคลุมที่ได้พัฒนาขึ้นในงานวิจัยนี้ และในส่วนที่สองคือการแบ่งบริเวณของภาพที่เป็นสีเขียวออกจากส่วนอื่น ในบทความนี้ได้ใช้เครือข่ายประสาทเทียมในการแบ่ง ซึ่งผู้วิจัยเห็นว่าข้อมูลที่ได้หลังจากที่ดำเนินการตามขั้นตอนวิธีการลำเลียงย้ายตามค่าเฉลี่ยแบบครอบคลุมแล้ว นั้น เป็นข้อมูลที่แยกกันอย่างชัดเจนสามารถที่จะนำข้อมูล สีในระบบ HSV ของตำแหน่งสุดท้ายมาแปลงกลับเป็นข้อมูล สีในระบบ RGB ซึ่งจะได้ภาพที่แยกกันอย่างชัดเจนดังแสดงในภาพที่ 43 คอลัมน์ที่ 2 หลังจากนั้นผู้วิจัยได้แปลงภาพให้เป็นภาพในระดับสีเทา ตามวิธี Excess Green index (ExG) [25, 26] ตามสมการ $I = 2G - R - B$ ดังแสดงในคอลัมน์ที่ 3 จากนั้นผู้วิจัยได้แปลงค่าความสว่างให้อยู่ในช่วง 0 ถึง 1 แล้วใช้ค่าขีดแบ่ง (Threshold) ช่วยในการแบ่งส่วนของภาพโดยกำหนดให้ตำแหน่งที่มีความสว่างมากกว่า 0.1 เป็นตำแหน่งของพืชสีเขียว ซึ่งในการทดลองนี้ผู้วิจัยได้ใช้ค่าขีดแบ่งเดียวกันทุกภาพ จากผลการทดลองพบว่า วิธีที่ผู้วิจัยได้นำเสนอสามารถที่จะแยกส่วนที่เป็นพืชสีเขียวหลังจากดำเนินการตามขั้นตอนวิธีการลำเลียงย้ายตามค่าเฉลี่ยแบบครอบคลุมแล้วได้อย่างชัดเจน โดยผู้วิจัยได้ทำการเปรียบเทียบขั้นตอนวิธีของบทความเดิม กับแนวทางของผู้วิจัยที่ได้นำเสนอขึ้น ในการแยกส่วนประกอบของภาพพืชผลทางการเกษตรดังแสดงในภาพที่ 44



ภาพที่ 43 การนำขั้นตอนวิธีการลำเลียงย้ายตามค่าเฉลี่ยแบบครอบคลุมมาประยุกต์ใช้กับการแยกส่วนประกอบของภาพพืชผลทางการเกษตร



ภาพที่ 44 การเปรียบเทียบขั้นตอนวิธีของบทความเดิม กับแนวทางของผู้วิจัยที่ได้นำเสนอขึ้นในการแยกส่วนประกอบของภาพพืชผลทางการเกษตร