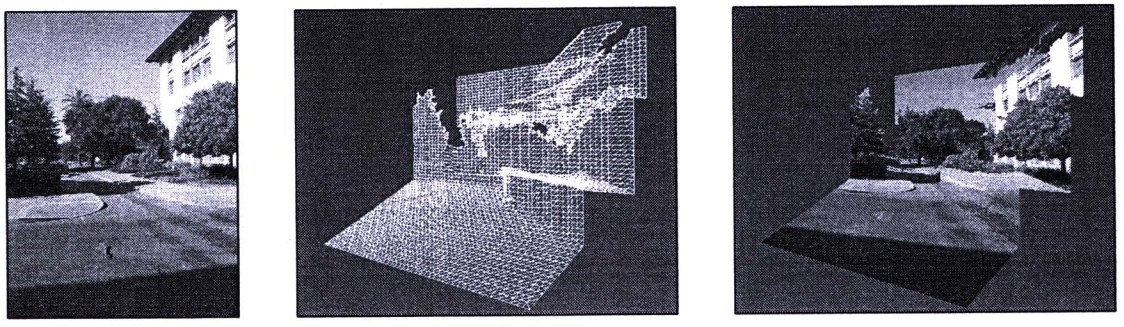
บทที่ 1
บทนำ

1.1 ที่มาและความสำคัญของปัญหา



รูปที่ 1.1 แสดงการสร้างฉากสามมิติจากรูปภาพเดี่ยว

ในปัจจุบัน การแสดงผลเป็นสามมิติได้ถูกนำมาใช้อย่างแพร่หลาย เนื่องมาจากการพัฒนาด้านเครื่องมือการแสดงผล ดังเช่น ภาพยนตร์สามมิติ โทรทัศน์สามมิติ รวมไปถึงเครื่องมือพกพาต่างๆ แต่อย่างไรก็ตามการสร้างฉากสามมิติเพื่อนำไปใช้นั้นกลับเป็นเรื่องที่ยุ่งยากมากกว่า โดยต้องอาศัยผู้เชี่ยวชาญและเวลา ในการสร้างฉากสามมิติเพียงหนึ่งฉากส่งผลให้ผู้ใช้ไม่สามารถเข้าถึงได้มากเท่าที่ควร

การใช้ภาพในการสร้างฉากสามมิตินั้นเป็นวิธีหนึ่งซึ่งช่วยจัดการปัญหาความยุ่งยากเหล่านั้นออกไป วิธีที่กล่าวมานี้ได้ถูกนำไปใช้ในโปรแกรมประยุกต์ต่างๆ อาทิเช่น เช่น มุมมองถนนในกูเกิลแมป (Google map street view) หรือ โฟโต้ซินท์ [1] (photosynth) ของไมโครซอฟท์ ซึ่งสามารถสร้างฉากสามมิติที่มีความเสมือนจริงได้จากภาพที่ได้ในสถาณที่ต่างๆ ถึงอย่างไรโปรแกรมประยุกต์ดังกล่าวต้องการภาพจำนวนมาก และเครื่องมือพิเศษในการสร้างซึ่งเป็นไปได้ยากสำหรับการใช้ทั่วไป

การใช้ภาพเพียงภาพเดียวในการสร้างภาพสามมิติจึงได้ถูกนำมาใช้ในการแก้ไขปัญหาเนื่องจากภาพเดี่ยวนั้นเข้าถึงได้ง่าย และมีอยู่เป็นจำนวนมาก แต่อย่างไรก็ตามปัญหาที่กล่าวมาเป็นปัญหาที่แก้ได้ยาก เนื่องจากข้อมูลภาพนั้นเป็นเพียงข้อมูลสีในสองมิติ การที่จะสร้างฉากที่มีสามมิติได้แก่ ความกว้าง ความยาว และความลึก จากภาพเพียงภาพเดียวนั้นไม่สามารถรู้ถึงข้อมูลความลึกของภาพได้โดยตรง จึงได้มีผู้นำเสนอวิธีการในการแก้ไขปัญหานี้จำนวนมาก อาทิเช่น การใช้ข้อมูลแสงเงาโดยอาศัยการเปลี่ยนแปลงความเข้มสีภายในภาพใน

การหาส่วนกลับสมการหาเวกเตอร์ปกติของพื้นผิว ผ่านสมการการสะท้อนแสงของวัตถุ หรือการสร้างฐานข้อมูลขนาดใหญ่เพื่อทำการเก็บข้อมูลสามมิติแล้วนำมาเทียบกับภาพนำเข้า เพื่อหาโมเดลที่ตรงกัน แต่อย่างไรก็ตามวิธีการดังกล่าวนั้นยังมีข้อจำกัดอยู่หลายประการ ไม่ว่าจะเป็นความสามารถที่สามารถใช้งานได้กับวัตถุบางประเภท และแหล่งกำเนิดแสงบางชนิด หรือไม่สามารถทำงานได้หากไม่มีข้อมูลของวัตถุที่อยู่ในภาพ

วิธีการหนึ่งที่มีความสามารถในการสร้างฉากสามมิติจากภาพเดี่ยวอย่างมีประสิทธิภาพคือการใช้โมเดลการเรียนรู้ของคอมพิวเตอร์ (Machine Learning) โดยในวิธีนี้สามารถแบ่งการทำงานได้สามขั้นตอนหลัก ได้แก่ ขั้นตอนแรกจะทำการใช้วิธีการแบ่งภาพเพื่อแบ่งภาพออกเป็นส่วน ย่อยๆที่เรียกว่าซูเปอร์พิกเซล ขั้นตอนที่สองคือการแบ่งกลุ่ม โดยจะทำการสกัดค่าจุดเด่นของแต่ละซูเปอร์พิกเซล ซึ่งจากข้อมูลที่ได้จะถูกนำมาใช้เป็นข้อมูลสำหรับการแบ่งกลุ่มเพื่อหาข้อมูลในการคืนสภาพข้อมูลสามมิติจากแต่ละซูเปอร์พิกเซล ขั้นตอนสุดท้ายจะทำการคืนสภาพออกมาเป็นฉากสามมิติ ซึ่งตัวอย่างของงานวิจัยในด้านนี้เช่น วิธีการที่ทำการแบ่งกลุ่มของซูเปอร์พิกเซลออกเป็นสามกลุ่มได้แก่ พื้น, ท้องฟ้า และส่วนตั้งฉาก [2] แล้วนำมาสร้างฉากสามมิติ โดยใช้วิธีการแบบตัด และพับ และวิธีการที่ทำการหาตำแหน่งและทิศทางของแต่ละซูเปอร์พิกเซล [3-5] โดยหาการวางตัวของส่วนต่างๆในภาพ แต่วิธีการที่กล่าวมาก็ยังไม่สามารถใช้ได้ในฉากที่มีความซับซ้อน หรือมีวัตถุที่มีการทับซ้อนอยู่กันเป็นจำนวนมาก

วิธีการที่ใช้ความสัมพันธ์ระหว่างวัตถุ [6,7] (Object placement relation: OPR) ได้ถูกนำเสนอขึ้น โดยในกำหนดให้ซูเปอร์พิกเซลหนึ่ง แทนส่วนของวัตถุ หรือวัตถุ แล้วนำมาหาความสัมพันธ์ระหว่างซูเปอร์พิกเซลข้างเคียง เพื่อนำไปสร้างฉากสามมิติ โดยวิธีการดังกล่าวสามารถสร้างฉากสามมิติ ที่สามารถใช้งานได้กับฉากที่มีความซับซ้อนมากกว่าได้ แต่ผลลัพธ์ที่ได้จากวิธีการนี้ยังมีความผิดพลาดบางส่วน ดังเช่นมีความขัดแย้งกันเองในความสัมพันธ์ในกลุ่มของซูเปอร์พิกเซล และผลลัพธ์ที่ได้ยังมีความไม่สมบูรณ์ เนื่องจากมีการนิยามเพียงแค่ความสัมพันธ์แบบเฉพาะส่วน (Local) ซึ่งไม่ได้มีการหาความสัมพันธ์แบบวงกว้าง (Global) จึงทำให้ไม่สามารถระบุความสัมพันธ์ระหว่างซูเปอร์พิกเซลที่ไม่ได้อยู่ติดกัน

งานวิจัยนี้จึงนำเสนอการปรับปรุงการนิยามความสัมพันธ์ระหว่างวัตถุ โดยการสร้างความสัมพันธ์แบบวงกว้างโดยใช้โครงสร้างข้อมูลแบบกราฟ ซึ่งสร้างจากการแบ่งส่วนของฉากจากภาพนำเข้า เพื่อนำไปใช้กำหนดความสัมพันธ์ระหว่างวัตถุทั้งฉาก และใช้ความสัมพันธ์ที่ได้มาช่วยในการการประมาณค่าความลึก และสร้างฉากสามมิติ เพื่อปรับปรุงคุณลักษณะของฉากสามมิติที่ถูกสร้าง

1.2 วัตถุประสงค์ของการวิจัย

เพื่อเสนออัลกอริทึมการประมาณฉากสามมิติโดยการวิเคราะห์ความสัมพันธ์ระหว่างวัตถุ โดยการเพิ่มความสัมพันธ์แบบวงกว้างที่มีความแม่นยำในการแบ่งกลุ่ม โดยไม่เกิดความขัดแย้งระหว่างความสัมพันธ์ และปรับปรุงคุณภาพของฉากสามมิติที่สร้าง ซึ่งทำการวัดผลโดยการใช้ค่าผิดพลาดของความลึก และความสมบูรณ์ของรูปภาพที่ได้

1.3 ขอบเขตของการวิจัย

1 ข้อมูลนำเข้าเป็นรูปภาพขนาด 640×480 พิกเซล ที่ไม่มีความโปร่งใส โดยวิธีการสามารถทำงานได้ในรูปทิวทัศน์ ที่ไม่ประกอบด้วยส่วนเพดาน และไม่มีวัตถุที่ลอยอยู่ในฉาก รวมไปถึงมุมกล้องจะต้องเป็นแนวตรงขนานกับเส้นขอบฟ้า

2 การแสดงผลแบบจำลองออกมาเป็นรูปแบบสามมิติ โดยเป็นรูปแบบวีอาเอ็มแอล 2.0 (VRML2.0) โดยที่แบบจำลองที่ได้จะไม่มีการคำนึงถึงทิศทางของเวกเตอร์ปกติ

3 ข้อมูลที่ใช้ในขั้นตอนการเรียนรู้ ประกอบด้วยรูปภาพจริงจากฐานข้อมูลของงานวิจัยการประมาณค่าความลึกจากรูปเดี่ยว [3-5]

4 งานวิจัยนี้จะทำการวัดผลทั้งหมดทั้งหมดสองด้านได้แก่ ความแม่นยำในการแบ่งกลุ่ม และความผิดพลาดในการประมาณฉากสามมิติ โดยนำค่าที่ได้มาเปรียบเทียบกับวิธีการประมาณฉากสามมิติโดยใช้ความสัมพันธ์ระหว่างวัตถุ ประกอบด้วยรูปภาพจริงจากฐานข้อมูลของงานวิจัยการประมาณค่าความลึกจากรูปเดี่ยว [3-5]

1.4 ประโยชน์ที่ได้รับ

วิธีการที่นำเสนอได้เพิ่มความสัมพันธ์ระหว่างวัตถุแบบวงกว้าง ซึ่งสามารถสามารถเพิ่มความแม่นยำในการแบ่งกลุ่ม และปรับปรุงคุณภาพของผลลัพธ์ที่เป็นฉากสามมิติเมื่อเปรียบเทียบกับวิธีการเดิม โดยผลลัพธ์ที่ได้สามารถนำไปประยุกต์ใช้ในการควบคุมหุ่นยนต์และในอัลกอริทึมการตรวจจับวัตถุ

1.5 วิธีดำเนินการวิจัย

การดำเนินงานวิจัยนี้ใช้ระยะเวลาทั้งหมด 22 เดือนโดยเริ่มต้นตั้งแต่เดือนมิถุนายน 2552 ถึงเดือนมีนาคม 2554 โดยประกอบไปด้วยทั้งหมดเจ็ดขั้นตอนได้แก่ศึกษางานวิจัย และทฤษฎีที่เกี่ยวข้อง วิเคราะห์และออกแบบอัลกอริทึม ออกแบบโปรแกรมประยุกต์ที่ใช้ทดสอบ พัฒนาและทดสอบโปรแกรมประยุกต์ ทดลองและประเมินผล สรุปผลการดำเนินงาน จัดทำวิทยานิพนธ์

	เดือนเริ่มต้น	ระยะเวลา (เดือน)	มิ.ย.-52	ก.ค.-52	ส.ค.-52	ก.ย.-52	ต.ค.-52	พ.ย.-52	ธ.ค.-52	ม.ค.-53	ก.พ.-53	มี.ค.-53	เม.ย.-53	พ.ค.-53	มิ.ย.-53	ก.ค.-53	ส.ค.-53	ก.ย.-53	ต.ค.-53	พ.ย.-53	ธ.ค.-53	ม.ค.-54	ก.พ.-54	มี.ค.-54
ศึกษางานวิจัย และทฤษฎีที่เกี่ยวข้อง	ก.ค.-52	5																						
วิเคราะห์ และออกแบบอัลกอริทึม	ก.ย.-52	7																						
ออกแบบโปรแกรมประยุกต์ที่ใช้ทดสอบ	ม.ค.-52	5																						
พัฒนา และทดสอบโปรแกรมประยุกต์	พ.ค.-53	8																						
ทดลองและประเมินผล	ต.ค.-53	5																						
สรุปผลการดำเนินงาน	ม.ค.-54	2																						
จัดทำวิทยานิพนธ์	ก.พ.-54	2																						

รูปที่ 1.2 แสดงแผนภาพแกนท์แสดงระยะเวลาการดำเนินการวิจัย

1.6 ผลงานตีพิมพ์จากงานวิจัย

งานวิจัยนี้ได้รับการตีพิมพ์เป็นบทความทางวิชาการดังนี้

- "Adaptive 3-D Scene Construction from Single Image Using Extended Object Placement Relation" โดย อาชว์ สรรพอาษา, ณัชพล ฟูตระกูล และ พิษณุ คนองชัยยศ ในงานประชุมวิชาการ "8th ACM SIGGRAPH international Conference on Virtual-Reality Continuum and Its Applications in industry": ซึ่งจัดขึ้นที่ โตเกียว ประเทศญี่ปุ่น ในระหว่างวันที่ 14-15 กุมภาพันธ์ 2553

- "Global Object Placement Relation for Improving 3-D Scene Construction from Single Image" โดย อาชว์ สรรพอาษา และ พิษณุ คนองชัยยศ ในงานประชุมวิชาการ "The eighth annual international conference organized by Electrical Engineering/ Electronics, Computer, Telecommunications and Information Technology ": ซึ่งจัดขึ้นที่ ขอนแก่น ประเทศไทย ในระหว่างวันที่ 17-19 พฤษภาคม 2554