# บทที่ 2
# ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

งานวิจัยนี้มีจุดประสงค์ในการแก้ไขวิธีการประมาณฉากสามมิติโดยการวิเคราะห์ความสัมพันธ์ระหว่างวัตถุ โดยเพิ่มข้อมูลของความสัมพันธ์ที่ครอบคลุมถึงตำแหน่งของวัตถุทั้งหมดภายในรูปภาพ โดยทฤษฎีที่เกี่ยวข้องและงานวิจัยที่เกี่ยวข้องดังนี้

## 2.1 ทฤษฎีที่เกี่ยวข้อง

ในงานวิจัยนี้มีทฤษฎีที่เกี่ยวข้องทั้งหมดสิบทฤษฎี ได้แก่ การประมาณฉากสามมิติโดยการวิเคราะห์ความสัมพันธ์ระหว่างวัตถุ การแบ่งย่อยรูปภาพด้วยการใช้กราฟอย่างมีประสิทธิภาพ ซูเปอร์พิกเซล การแบ่งกลุ่มแบบเบย์ การหาจุดตัดต่ำสุดของกราฟแบบมีต้นทางและปลายทาง กราฟอวัฏจักรระบุทิศทาง การแทนฉากสามมิติด้วยกราฟฉาก การหากราฟเหมือนต้นไม้ใหญ่สุด การฉายแบบทัศนมิติ และการสร้างสามเหลี่ยมแบบเดลาอูไน

### 2.1.1 การประมาณฉากสามมิติโดยการวิเคราะห์ความสัมพันธ์ระหว่างวัตถุ

การประมาณฉากสามมิติโดยการวิเคราะห์ความสัมพันธ์ระหว่างวัตถุ [6,7] เป็นวิธีการสร้างฉากสามมิติโดยการหาความสัมพันธ์ระหว่างวัตถุ หรือ โอพีอาร์ (Object Placement Relation: OPR) โดยในวิธีการนี้ได้แบ่งความสัมพันธ์ระหว่างวัตถุไว้ทั้งหมดหกความสัมพันธ์คือ

1 วัตถุเดียวกัน (Same Object: SO) กล่าวคือวัตถุทั้งสองชิ้นนั้นเป็นวัตถุชนิดเดียวกัน ซึ่งส่งผลให้ความลึกของวัตถุทั้งสองนั้นมีค่าเท่ากัน หรือต่อเนื่องกันไป

2 ระนาบเดียวกัน (Same Plane: SP) กล่าวคือวัตถุทั้งสองชิ้นนั้นวางอยู่ในระนาบเดียวกันแต่เป็นวัตถุคนละชนิดกัน เนื่องจากความแตกต่างของพื้นผิวของวัตถุทั้งสองชนิด โดยความลึกของวัตถุสองชิ้นนั้นจะมีค่าเท่ากันหรือต่อเนื่องกันไป

3 ปิดบัง (Occlude: OC) กล่าวคือวัตถุปิดบังวัตถุอื่นอยู่ โดยความลึกของวัตถุทั้งสองชิ้นจะมีค่าไม่ต่อเนื่องกัน และวัตถุด้านหน้าจะมีความลึกที่น้อยกว่าวัตถุที่ถูกปิดบัง

4 ถูกปิดบัง (be Occluded by: OB) กล่าวคือวัตถุถูกวัตถุอื่นปิดบังอยู่ โดยความลึกของวัตถุทั้งสองชิ้นจะมีค่าไม่ต่อเนื่องกัน และวัตถุด้านหน้าจะมีความลึกที่น้อยกว่าวัตถุที่ถูกปิดบัง

5 วางอยู่ (Place On: PO) กล่าวคือการที่วัตถุทับวัตถุอื่นอยู่ โดยความลึกของวัตถุที่อยู่ด้านบน จะอยู่ระหว่างความลึกสูงสุด และต่ำสุดของความลึกของวัตถุฐาน

6 ถูกวางทับ (be Placed By: PB) กล่าวคือการที่วัตถุถูกทับโดยวัตถุอื่นอยู่ โดยความลึกของวัตถุที่อยู่ด้านบน จะอยู่ระหว่างความลึกสูงสุด และต่ำสุดของความลึกของวัตถุฐาน

โดยวิธีการนี้จะประกอบไปด้วยขั้นตอนวิธีห้าขั้นตอนคือ ขั้นตอนการแบ่งภาพ ขั้นตอนการสกัดค่าจุดเด่น ขั้นตอนการแบ่งกลุ่ม ขั้นตอนการประมาณค่าความลึก และขั้นตอนการสร้างฉากสามมิติ

ขั้นตอนการแบ่งภาพ (Segmentation) ในวิธีการนี้จะแบ่งภาพออกเป็นส่วนย่อยเรียกว่าซูเปอร์พิกเซลโดยใช้วิธีการแบ่งส่วนย่อยรูปภาพด้วยการใช้กราฟอย่างมีประสิทธิภาพ ซึ่งเมื่อได้ซูเปอร์พิกเซลแล้ว จะทำการปรับปรุงคุณภาพของซูเปอร์พิกเซลโดยการนำซูเปอร์พิกเซลที่มีลักษณะเรียวยาว และขรุขระไปรวมกับซูเปอร์พิกเซลข้างเคียง โดยมีหลักเกณฑ์ในการวัดคือ ความเรียวยาวจะทำการวัดโดยใช้ค่าพื้นที่ต่อเส้นรอบรูป และความขรุขระวัดโดยใช้จำนวนจุดวิกฤติตามเส้นขอบของซูเปอร์พิกเซล

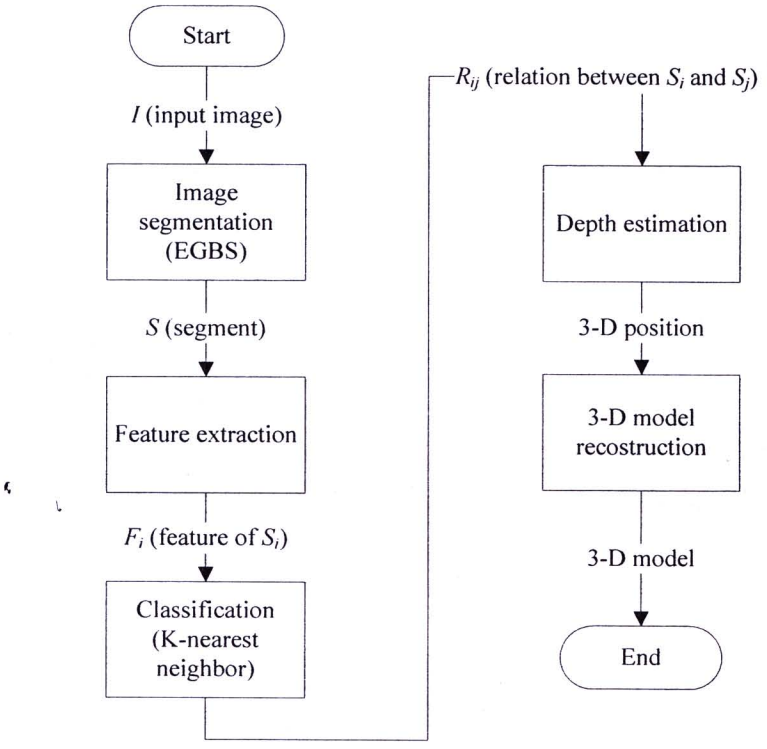
การสกัดค่าจุดเด่น (Feature Extraction) คือการสกัดค่าจุดเด่นเพื่อใช้ในการแบ่งกลุ่ม ซึ่งสกัดออกมาจากซูเปอร์พิกเซล และความแตกต่างระหว่างจุดเด่นของคู่ซูเปอร์พิกเซลข้างเคียง โดยแบ่งค่าจุดเด่นออกเป็นสี่รูปแบบเพื่อใช้ในการแยกความสัมพันธ์ในรูปแบบ

ต่างๆ ได้แก่ จุดเด่นทางลายพื้นผิว (texture) ใช้ในการแบ่งความสัมพันธ์แบบวัตถุเดียวกัน จุดเด่นทางเส้นขอบ (edge) ใช้ในการแบ่งความสัมพันธ์แบบปิดบัง และถูกปิดบัง จุดเด่นทางเส้นตรง (straight line) ใช้ในการแบ่งความสัมพันธ์แบบระนาบเดียวกัน และจุดเด่นทางตำแหน่ง (position) ใช้ในการแบ่งความสัมพันธ์แบบวางบน และถูกวางทับ

ขั้นตอนการจัดกลุ่ม (Classification) โดยในขั้นตอนนี้จะใช้อัลกอริทึมเพื่อนบ้านใกล้ที่สุด (K-nearest Neighbor) ในการหาความสัมพันธ์ระหว่างคู่ซูเปอร์พิกเซล

ขั้นตอนการหาประมาณความลึก (Depth Estimation) โดยการหาจุดรวมสายตา โดยหาส่วนที่เป็นพื้น และท้องฟ้า ในรูปเพื่อประมาณค่าความลึกของตำแหน่งที่ซูเปอร์พิกเซลสัมผัสกับพื้น โดยจะประมาณค่า

ขั้นตอนการสร้างฉากสามมิติ (Depth Estimation) ในขั้นตอนนี้จะนำข้อมูลที่ได้ทั้งหมดมาทำการประมาณค่าฉากสามมิติ โดยจะการเลือกจุดยอด (Vertex) จากจุดวิกฤตในขอบของซูเปอร์พิกเซล แล้วทำการขึ้นรูปโดยใช้วิธีการสร้างสามเหลี่ยมแบบเดลาอูไน

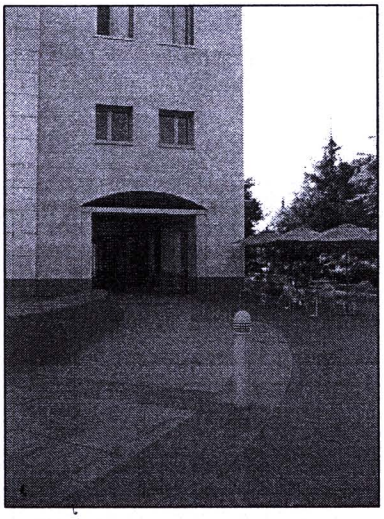



รูปที่ 2.1 ผังการทำงานของวิธีการประมาณฉากสามมิติโดยการวิเคราะห์ความสัมพันธ์ระหว่างวัตถุ

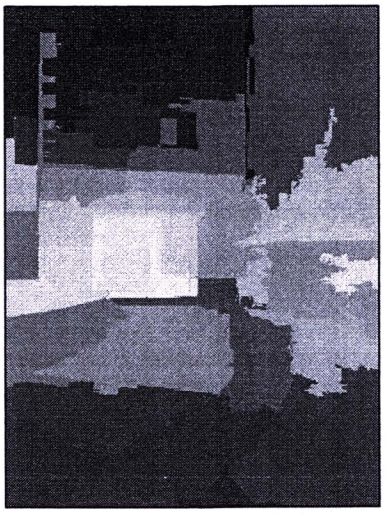
### 2.1.2 การแบ่งย่อยรูปภาพด้วยการใช้กราฟอย่างมีประสิทธิภาพ หรืออีจีบีเอส

การแบ่งย่อยรูปภาพจากรูปคือ คือ กระบวนการแบ่งรูปภาพดิจิทัลออกเป็นเซตของจุดภาพ (Sets of pixels) โดยที่จุดภาพที่อยู่ในเซตมีคุณสมบัติที่ใกล้เคียงกัน

ในการแบ่งกลุ่มด้วยการแบ่งกลุ่มโดยการใช้กราฟอย่างมีประสิทธิภาพ [8] จะทำการมองพิกเซลภาพเป็นแต่ละจุดภายในเส้นกราฟ แล้วทำการจับกลุ่มพิกเซลที่มีความคล้ายคลึงกันออกเป็นกลุ่ม ซึ่งวิธีการนี้จะเป็นวิธีแบบแบ่งส่วนเกิน (over-segment) กล่าวคือจะพยายามไม่ให้มีซูเปอร์พิกเซลที่แบ่งได้ประกอบไปด้วยวัตถุภายในภาพมากกว่าหนึ่งชิ้น ซึ่งตรงกับสมมติฐานของงานวิจัย



ก)



ข)

รูปที่ 2.2 แสดงผลลัพธ์การแบ่งย่อยรูปภาพด้วยการใช้กราฟอย่างมีประสิทธิภาพ ก) ภาพนำเข้า ข) ผลลัพธ์จากการแบ่งกลุ่มโดยแต่ละสีแสดงถึงส่วนย่อยที่ได้

### 2.1.3 ซูเปอร์พิกเซล

ซูเปอร์พิกเซล $S_i$ คือกลุ่มของจุดภาพที่มีคุณสมบัติใกล้เคียงกันที่อยู่ติดกัน โดยที่จุดภาพแต่ละจุดจะสามารถอยู่ได้ในซูเปอร์พิกเซลเพียงกลุ่มเดียว งานวิจัยนี้ทำการแบ่งซูเปอร์พิกเซลโดยการใช้วิธีการแบ่งรูปด้วยกราฟอย่างมีประสิทธิภาพ โดยที่เราสามารถนิยามความสัมพันธ์ระหว่างซูเปอร์พิกเซล $S$ และรูปภาพนำเข้า $I$ ได้ดังสมการ

$$S_1, S_2, \dots, S_n = EGBS(I)$$
$$S_i = \{(x,y) \subset I | (x,y) have\ a\ similar\ property\ defined\ by\ EGBS\}$$
$$\text{โดยที่ } \forall_i \forall_j \left(S_i \cap S_j\right) = \emptyset, \bigcup_{i=1}^{n} S_i = I$$

โดยทำการกำหนดของซูเปอร์พิกเซลเพื่อนบ้าน (Neighbor Superpixel) คือซูเปอร์พิกเซลที่เชื่อมต่อกันโดยมีความเชื่อมต่อแบบสี่ (4-connected) ดังสมการ

$$S_i \in Nb(S_j) \leftrightarrow \begin{cases} \exists(x_j \pm 1, y_j) \in S_i \\ \exists(x_j, y_j \pm 1) \in S_i \end{cases}$$

### 2.1.4 การแบ่งกลุ่มแบบเบย์

การแบ่งกลุ่มข้อมูลแบบเบย์ [9] เป็นวิธีการพื้นฐานสำหรับการแบ่งกลุ่มแบบความน่าจะเป็น โดยการประยุกต์ทฤษฎีของเบย์ (Bayes Theorem) คือ $p(A|B) = \frac{p(B|A)p(A)}{p(B)}$ ซึ่งสามารถให้คำตอบออกมาในรูปแบบของความน่าจะเป็นที่ข้อมูลจะตกอยู่ในคลาส ซึ่งการแบ่งกลุ่มแบบเบย์สามารถอธิบายได้ด้วยสมการดังต่อไปนี้ โดยให้ $p(C|F_1 \cdots F_n)$ คือความน่าจะเป็นที่ข้อมูลที่มีจุดเด่น 1 ถึง n จะอยู่ในคลาส *C* และ *Z* คือตัวแปรขยาย

$$p(C|F_1 \cdots F_n) = \frac{1}{Z} p(C) \prod_{i=1}^{n} p(F_i|C)$$

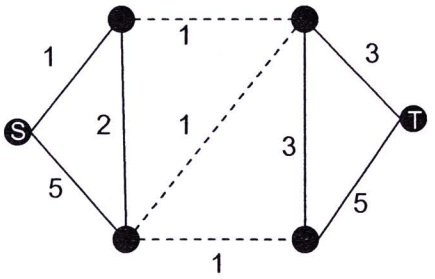
โดยที่การแบ่งกลุ่มแบบเบย์จะทำการเลือกคลาสที่มีความน่าจะเป็นสูงสุดเป็นคลาสของข้อมูล ซึ่งแทนด้วยสมการ

$$classify(f_1 \cdots f_n) = argmax_c p(C = c) \prod_{i=1}^{n} p(F_i = f_i|C = c)$$

ซึ่งการแบ่งกลุ่มแบบเบย์นั้นจะแสดงถึงความน่าจะเป็นที่เหตุการณ์ที่เราสนใจจะเกิดขึ้นโดยเทียบกับเหตุการณ์ที่เกิดขึ้นในอดีต ซึ่งข้อดีของการใช้วิธีแบบความน่าจะเป็นในการแบ่งกลุ่มคือสามารถนำไปประกอบกับโมเดลความน่าจะเป็นอื่นได้ง่าย และสามารถเปรียบเทียบค่าความถูกต้องระหว่างคำตอบหลายคำตอบได้ชัดเจน

### 2.1.5 การหาจุดตัดต่ำสุดของกราฟแบบมีต้นทางและปลายทาง

จุดตัดต่ำสุดของกราฟคือกลุ่มของเส้นเชื่อมภายในกราฟที่มีผลรวมของน้ำหนักเส้นเชื่อมน้อยที่สุด ซึ่งเมื่อตัดออกไปแล้วจะทำการแบ่งกราฟออกเป็นสองกราฟย่อย โดยการหาจุดตัดแบบมีต้นทาง (Source) และปลายทาง (Target) จะทำการกำหนดจุดยอดสองจุดยอด ซึ่งกราฟย่อยที่ได้จะต้องทำการแบ่งแยกจุดยอดทั้งสองจุดออกจากกัน การแก้ปัญหาดังกล่าวสามารถกระทำได้โดยใช้ขั้นตอนวิธีของฟอร์ดและเฟอร์กูสัน [10]



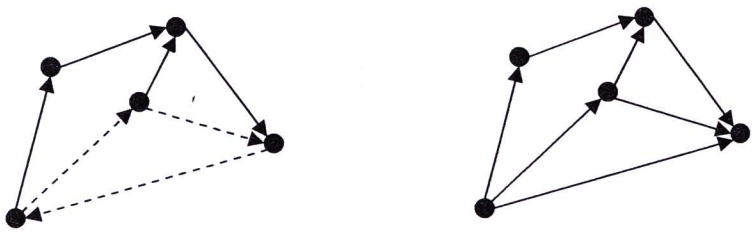

สำนักงานคณะกรรมการวิจัยแห่งชาติ ห้องสมุดงานวิจัย

รูปที่ 2.3 ผลลัพธ์การหาจุดตัดต่ำสุดของระหว่างจุดยอด S และ T โดยเส้นประแสดงเส้นเชื่อมที่ถูกตัดออก

### 2.1.6 กราฟอวัฏจักรระบุทิศทาง

กราฟอวัฏจักรระบุทิศทางคือกราฟระบุทิศทางที่ไม่มีวัฏจักร (Strongly Connected Component) กล่าวคือไม่มีเส้นทางจากจุดยอดใดภายในกราฟระบุทิศทางที่เชื่อมต่อกับตัวเอง โดยกราฟชนิดนี้สามารถสร้างการเรียงตัวระหว่างจุดยอดได้ โดยการตรวจหาวัฏจักรจากกราฟระบุทิศทางสามารถกระทำได้โดยการใช้ขั้นตอนวิธีของทาจัน [11] ในการค้นหากราฟที่มีวัฏจักร

โดยการหากราฟอวัฏจักรระบุทิศทางที่มีผลรวมของน้ำหนักสูงที่สุดจากกราฟอวัฏจักรใดๆจะอาศัยขั้นตอนวิธีการหาเซตของเส้นเชื่อมย้อนกลับ (Feedback Arc Set) โดยปัญหาดังกล่าวเป็นปัญหาเอ็นพีบริบูรณ์ (NP-Complete Problem) ซึ่งใช้เวลาในการคำนวณมาก ในงานวิจัยนี้จึงได้เลือกใช้ขั้นตอนวิธีการหากราฟเหมือนต้นไม้ใหญ่สุดในการหากราฟอวัฏจักรสูงที่สุดแทน



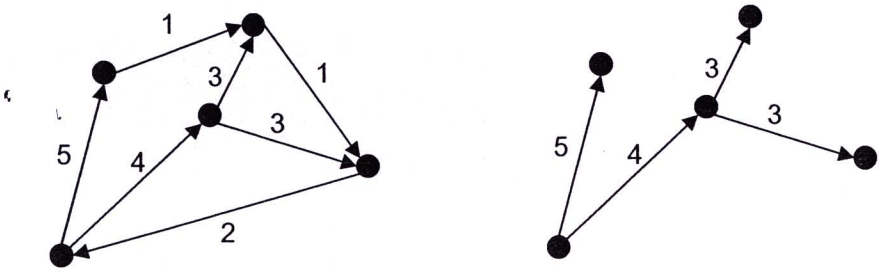
ก) ข)

รูปที่ 2.4 กราฟแบบระบุทิศทาง ก) แบบมีวัฏจักร ข) แบบอวัฏจักร

สำนักงานคณะกรรมการวิจัยแห่งชาติ
ห้องสมุดงานวิจัย
วันที่ 18 มิ.ย. 2555
เลขทะเบียน 246931
เลขเรียกหนังสือ

### 2.1.7 การหากราฟเหมือนต้นไม้ใหญ่สุด

กราฟเหมือนต้นไม้ (Arborescence) เป็นกราฟระบุทิศทางที่มีเส้นทางจากจุดยอดถึงจุดยอดอื่นเพียงเส้นทางเดียว โดยที่กราฟเหมือนต้นไม้ทุกกราฟนั้นจะเป็นกราฟอวัฏจักร โดยปัญหาการหากราฟเหมือนต้นไม้ใหญ่สุด คือการหากราฟย่อยเหมือนต้นไม้จากกราฟระบุทิศทางที่มีผลรวมของน้ำหนักในแต่ละเส้นเชื่อมมากที่สุด การหากราฟเหมือนต้นไม้ใหญ่สุดสามารถทำได้โดยการใช้ขั้นตอนวิธีของเอ็ดมอน [12]



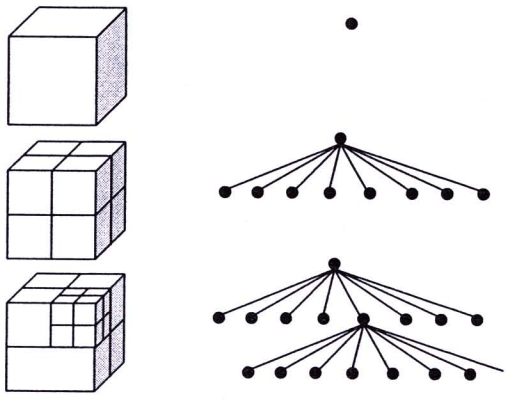

ก) ข)

รูปที่ 2.5 ผลลัพธ์การหากราฟเหมือนต้นไม้ใหญ่สุด ก) ข้อมูลนำเข้า ข) กราฟเหมือนต้นไม้ใหญ่สุดของกราฟนำเข้า

### 2.1.8 การแทนฉากสามมิติด้วยกราฟฉาก

กราฟฉากคือโครงสร้างข้อมูล ซึ่งใช้แสดงความสัมพันธ์ระหว่างวัตถุภายในฉากสามมิติ โดยมักจะมีลักษณะเป็นต้นไม้ที่มีโนดแม่เพียงโนดเดียว ซึ่งการการแปลงค่าในโนดแม่ทั้งหมดจะส่งผลกระทบต่อโนดลูกตามลำดับชั้น ซึ่งรวมไปถึงการกำหนดข้อมูลในลักษณะต่างๆ เช่น โอเพ็นจีแอล (OpenGL) ซึ่งการใช้กราฟฉากสามารถนำไปเพิ่มประสิทธิภาพของขั้นตอนต่างๆในสายท่อกราฟฟิก (graphic pipeline) โดยในงานวิจัยนี้ได้ใช้การแทนกราฟฉากด้วยวิธีต้นไม้อัฐภาค

ต้นไม้แบบอัฐภาคคือโครงสร้างต้นไม้ที่นิยมใช้ในการแทนข้อมูลสามมิติในลักษณะหนึ่ง โดยใช้หลักการแบ่งมิติ (space-partitioning) โดยแบ่งข้อมูลสามมิติโดยการใช้ระนาบตัดในทั้งสามแนวแกน ซึ่งส่งผลให้เกิดการแบ่งข้อมูลออกเป็นแปดส่วนเท่ากัน โดยที่ในแต่ละส่วนสามารถใช้หลักการเดียวกันเพื่อแยกย่อยลงไปเพิ่มได้ โดยส่วนแบ่งข้อมูลสามารถแสดงได้อยู่ในรูปแบบต้นไม้โดยให้จุดยอดแม่แทนข้อมูลก่อนถูกตัดแบ่ง และมีจุดยอดลูกแปดจุดแทนส่วนที่ตัดได้แต่ละส่วน



รูปที่ 2.6 การแทนฉากสามมิติด้วยต้นไม้แบบอัฐภาค

### 2.1.9 การฉายแบบทัศนมิติ

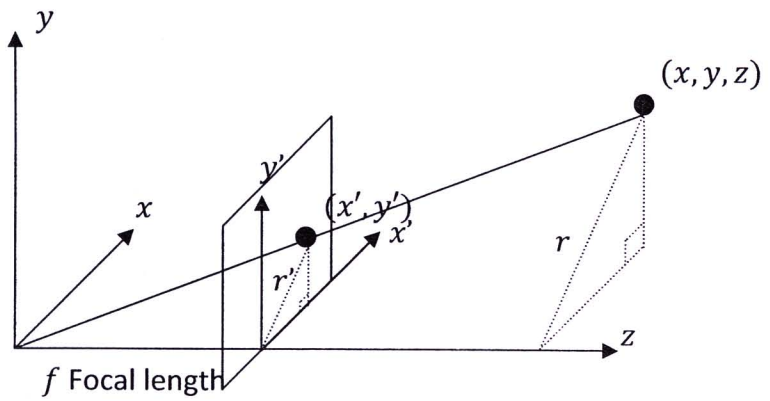
การฉายแบบทัศนมิติ (Perspective Projection) เป็นการฉายภาพแบบหนึ่ง โดยที่แสงที่เข้ามาในกล้องไม่ได้เป็นแสงขนาน ซึ่งการฉายภาพประเภทนี้พบทั่วไปในลักษณะของการถ่ายภาพต่างๆ โดยมีความสัมพันธ์ระหว่างพิกัดภาพ $(x', y')$ และพิกัดจริง $(x, y, z)$ เมื่อ $r$ และ $r'$ คือระยะทางจากแนวแกนความลึกไปยัง พิกัดจริงและพิกัดภาพ

โดยตำแหน่งของวัตถุจากการฉายแบบทัศนมิติจะเปลี่ยนไป ซึ่งสามารถแสดงได้จากสมการจากคุณสมบัติของสามเหลี่ยมคล้าย ดังสมการต่อไปนี้

$$\frac{\mathrm{r}}{\mathrm{r}'} = \frac{z}{f}$$
$$\frac{x}{x'} = \frac{y}{y'} = \frac{r}{r'}$$

จากสมการข้างต้น เราจะได้ความสัมพันธ์ระหว่างพิกัดจริง และพิกัดภาพ ดังสมการต่อไปนี้

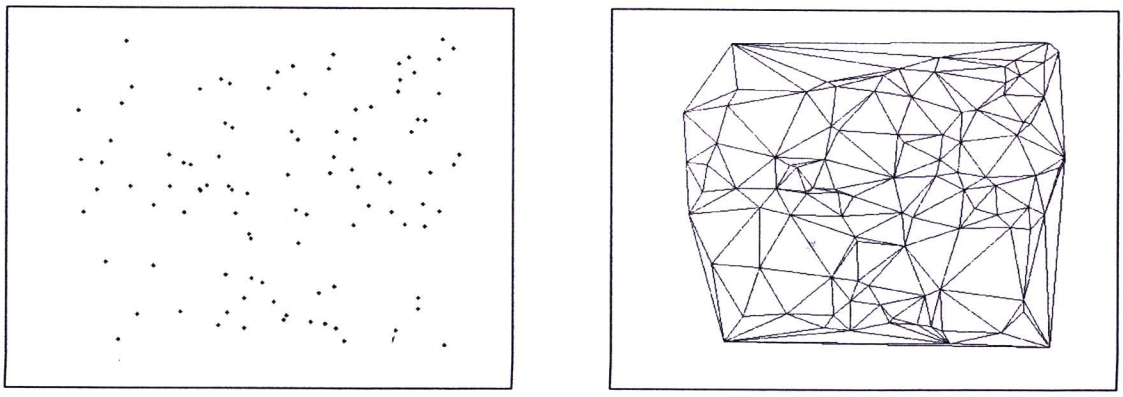
$$\mathrm{x}' = \frac{\mathrm{f}}{\mathrm{z}}\mathrm{x}$$
$$\mathrm{y}' = \frac{\mathrm{f}}{\mathrm{z}}\mathrm{y}$$




รูปที่ 2.7 การฉายภาพแบบทัศนมิติ

### 2.1.10 การสร้างสามเหลี่ยมแบบเดลาอูไน

การสร้างสามเหลี่ยมเป็นวิธีการในการสร้างสามเหลี่ยมจากกลุ่มจุด [13] โดยที่จะพยายามสร้างให้สามหลี่ยมที่มีมุมภายในที่น้อยที่สุดให้มีค่ามากที่สุด โดยการสร้างสามเหลี่ยมแบบเดลาอูไน ทำการสร้างสามเหลี่ยมโดยการใช้หลักการของแผนภาพโวโรนอย (Voronoi diagram) และการสร้างวงกลม (Circumcircle) ซึ่งการสร้างวงกลมคือการสร้างวงกลมโดยให้จุดสามจุดนั้นอยู่บนทรงกลมที่สร้าง ซึ่งวิธีการนี้จะอาศัยหลักโดยการให้ทรงกลมที่สร้างนั้นจะไม่ล้อมรอบเซตของกลุ่มจุด



ก) ข)

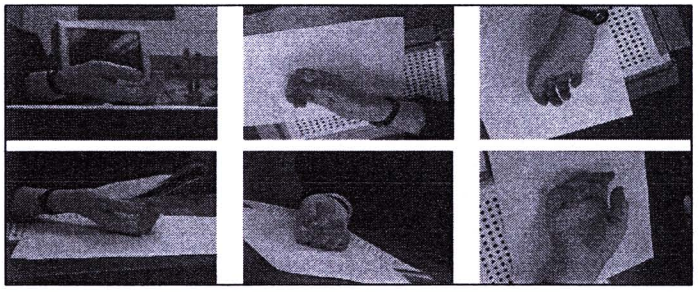
รูปที่ 2.8 การสร้างสามเหลี่ยมแบบเดลาอูไน ก) กลุ่มจุดนำเข้า ข) ผลลัพธ์การสร้างสามเหลี่ยมจากกลุ่มจุด

## 2.2 งานวิจัยที่เกี่ยวข้อง

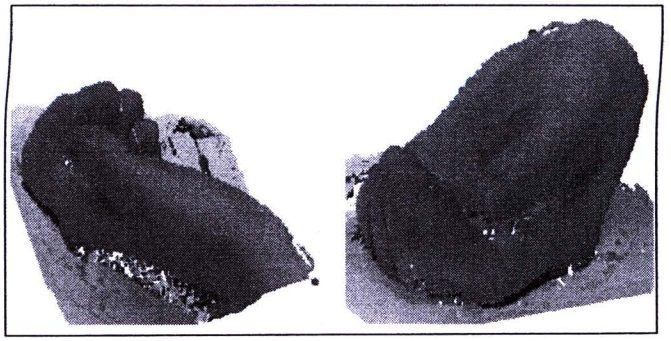
งานวิจัยนี้มุ่งเน้นในการขึ้นรูปสามมิติจากรูปภาพ (Image-based Modeling) โดยงานวิจัยที่มุ่งเน้นในการขึ้นรูปสามมิติจากภาพ สามารถแบ่งแยกได้ตามจำนวนของรูปภาพนำเข้าเป็นสองกลุ่ม ดังนี้

### 2.2.1 การขึ้นรูปแบบจำลองสามมิติจากหลายภาพ

การขึ้นรูปแบบจำลองสามมิติจากหลายภาพจะอาศัยการนำเข้าของรูปภาพที่มีความเกี่ยวเนื่องกัน เช่นด้านตำแหน่งของภาพ หรือแสงเงา ตัวอย่างงานวิจัยในลักษณะนี้ได้แก่ งานวิจัยของคูทูลาโคส [14] โดยอาศัยภาพจากหลายมุมมอง มาเป็นข้อมูลฐานในการสร้างภาพสามมิติ นอกจากนี้แนวคิดในการขึ้นรูปแบบจำลองสามมิติจากภาพหลายมุมมองยังถูกนำไปใช้โดยการอาศัยมุมกล้องขนาน (Stereo Images) โดยอาศัยภาพสองภาพที่คู่ขนานกันในการประมาณค่าความลึกของตำแหน่งภายในจุดภาพ ซึ่งสามารถเห็นได้ในการถ่ายรูปสามมิติในปัจจุบัน



ก)

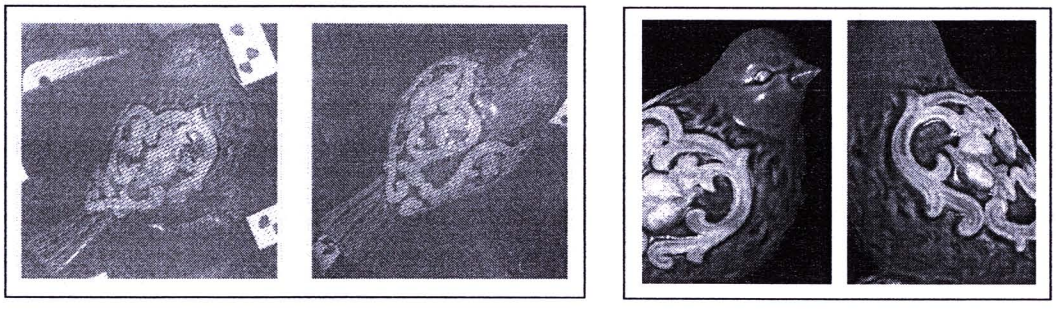


ข)

รูปที่ 2.9 ผลลัพธ์จากงานวิจัยของคูทูลาโคส ก) ส่วนหนึ่งของรูปภาพนำเข้า ข) ผลลัพธ์การประมาณแบบจำลองสามมิติ [14]

นอกจากนี้อาศัยมุมมองของภาพที่กล่าวมาแล้ว ยังสามารถขึ้นรูปโดยการใช้รูปแบบของแสงเงาที่ต่างกันภายในมุมมองเดียวกัน โดยหลักการนี้ถูกนำมาใช้ในการสร้างเครื่องสแกนเนอร์สามมิติ (3-d Scanner) โดยตัวอย่างของงานวิจัยนี้ เช่นงานวิจัยของฮอลรอยด์ [15] โดยจะทำการยิงแสงเป็นแถบในหลายรูปแบบ เพื่อหาส่วนเว้าโค้งของวัตถุที่ถูกถ่าย โดยในงานวิจัยนี้นอกจากจะได้รูปทรงของวัตถุแล้วยังได้ข้อมูลของลักษณะพื้นผิวอีกด้วย

แต่อย่างไรก็ตามงานวิจัยในลักษณะนี้ยังมีข้อจำกัดอยู่ โดยส่วนมากจะต้องการอุปกรณ์พิเศษในการสร้าง อีกทั้งรูปนำเข้าที่ใช้นั้นจะต้องอยู่ภายใต้สภาวะแวดล้อมเดียวกัน โดยไม่มีความคงทนต่อแสงเงาที่เปลี่ยนไปภายในภาพ



ก) ข)

รูปที่ 2.10 ผลลัพธ์จากงานวิจัยของฮอลรอยด์ ก) ส่วนหนึ่งของภาพนำเข้า ข) ผลลัพธ์การประมาณแบบจำลองสามมิติ พร้อมข้อมูลลักษณะพื้นผิว [15]

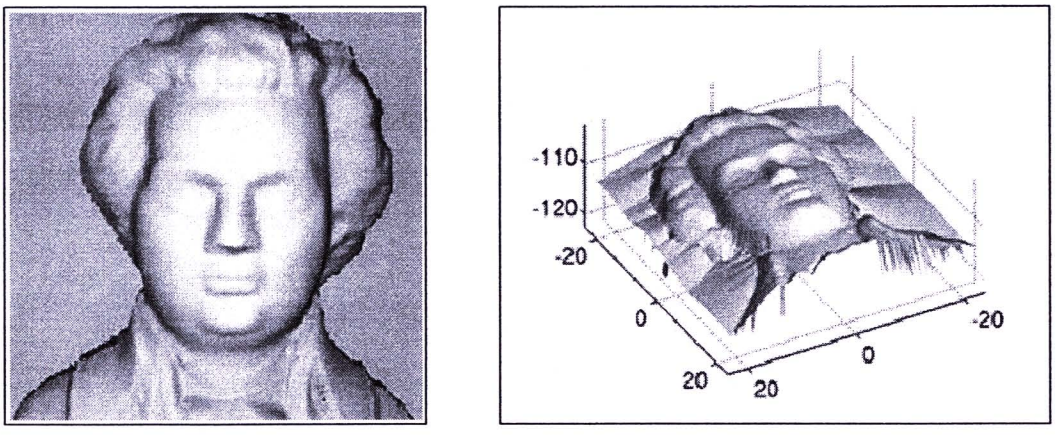
### 2.2.2 การขึ้นรูปแบบจำลองสามมิติจากภาพภาพเดียว

การขึ้นรูปแบบจำลองสามมิติจากภาพภาพเดียวอาศัยรูปภาพนำเข้าเพียงภาพเดียว โดยรายละเอียดของแบบจำลองที่ได้จะลดลงเนื่องจากข้อมูลที่ขาดหายไปจากการลดมิติ แต่เนื่องจากมีความยืดหยุ่นในการใช้งานมากกว่า โดยมีงานวิจัยในลักษณะมีหลายด้านเช่น การขึ้นรูปสามมิติจากแสงเงา, การค้นหาโมเดลสามมิติจากฐานข้อมูล และการขึ้นรูปสามมิติจากขั้นตอนวิธีเรียนรู้ของคอมพิวเตอร์

*2.2.2.1 การขึ้นรูปสามมิติจากแสงเงา*

การขึ้นรูปสามมิติจากแสงเงา [16] (Shape from Shading) เป็นการสร้างแบบจำลองสามมิติโดยการอาศัยส่วนกลับของสมการการสะท้อนแสง เพื่อหาค่าเวกเตอร์ปกติของตำแหน่งในแต่ละจุดภาพ ได้มีผู้เสนอวิธีการแก้ปัญหาดังกล่าว โดยการใช้เทคนิคการหาค่าความเหมาะสม เช่นงานวิจัยของไซด์และชาร์ [17] ที่ใช้การแก้ไขปัญหาแบบเชิงเส้นในการแก้สมการ หรืองานของลียงส์ [18] ที่ใช้สมการความหนืดเข้าช่วย

แต่การขึ้นรูปสามมิติโดยการใช้วิธีนี้สามารถใช้ได้ในสภาวะแวดล้อมที่ถูกจัดไว้กับวัตถุที่มีขนาดเล็ก และวัตถุที่ใช้ต้องเป็นวัตถุที่เป็นเนื้อเดียว อีกทั้งไม่สามารถใช้กับวัสดุที่มีความมันวาวได้ การใช้วิธีดังกล่าวจึงไม่สามารถใช้ได้กับรูปภาพทั่วไปได้

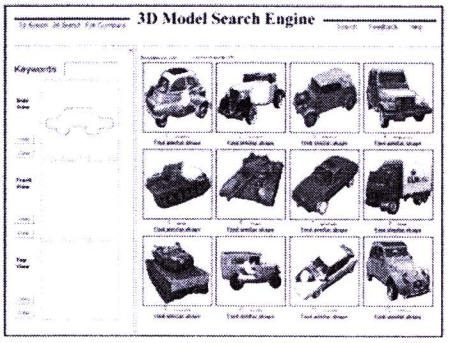



ก) ข)

รูปที่ 2.11 ผลลัพธ์จากการสร้างฉากสามมิติโดยใช้แสงเงา ก) ภาพนำเข้า ข) ผลลัพธ์การประมาณแบบจำลองสามมิติ [16]

*2.2.2.2 การค้นหาแบบจำลองสามมิติจากภาพ*

การค้นหาแบบจำลอง [19] จากภาพเป็นวิธีการในการขึ้นรูปสามมิติจากภาพรูปแบบหนึ่ง โดยการทำการเก็บข้อมูลแบบจำลองสามมิติในฐานข้อมูล แล้วทำการหาความแบบจำลองสามมิติที่เหมือนกับภาพนำเข้า โดยวิธีการนี้สามารถค้นหาแบบจำลองสามมิติได้อย่างมีประสิทธิภาพ รวมถึงแบบจำลองสามมิติที่ได้ไม่มีการขาดหายเนื่องจากใช้ข้อมูลที่มีความสมบูรณ์ในทุกมุมมองต่างจากวิธีอื่น แต่วิธีนี้ประสิทธิภาพจะขึ้นโดยตรงกับข้อมูลที่เก็บอยู่ในฐานข้อมูล ซึ่งหากไม่มีข้อมูลที่ใกล้เคียงกับภาพนำเข้าจะไม่สามารถทำงานได้




รูปที่ 2.12 การค้นหาแบบจำลองสามมิติจากภาพ [19]

*2.2.2.3 การขึ้นรูปสามมิติจากขั้นตอนวิธีเรียนรู้ของคอมพิวเตอร์*

การขึ้นรูปสามมิติจากขั้นตอนวิธีเรียนรู้ของคอมพิวเตอร์จะทำงานโดยการพยายามสร้างรูปสามมิติโดยการหาข้อมูลสำคัญที่ได้จากส่วนย่อยของรูปซึ่งทำการหาโดยวิธีการทางขั้นตอนวิธีเรียนรู้ของคอมพิวเตอร์ เช่น การแบ่งกลุ่ม โดยวิธีการนี้มีสามารถแก้ปัญหาได้ในขอบเขตที่ค่อนข้างกว้าง โดยสามารถทำได้ในรูปที่มีลักษณะแตกต่างกันไป และออกแบบมาสำหรับการสร้างฉากสามมิติที่ประกอบด้วยวัตถุต่างๆ ซึ่งงานในลักษณะนี้เหมาะสมกับงานที่ไม่ต้องการรายละเอียดภายในแต่ละวัตถุมากนัก โดยวิธีการนี้มีขั้นตอนทั้งหมดสามขั้นตอนได้แก่ การแบ่งภาพ, การแบ่งกลุ่ม และการสร้างฉากสามมิติ

ในปี 2005 โฮเอ็มได้เสนองานวิจัยการผุดขึ้นของภาพแบบอัตโนมัติ [2,20-22] ซึ่งทำการการสร้างฉากแบบผุดขึ้น โดยมีสมมติฐานที่ว่าวัตถุนั้นตั้งอยู่บนระนาบเพียงระนาบเดียว และอยู่ในลักษณะตั้งฉากเท่านั้น โดยจะทำการการหาเลเบิล (Label) ของแต่ละซูเปอร์พิกเซล โดยแบ่งประเภทของซูเปอร์พิกเซลออกเป็นสามอย่างคือ พื้นดิน ส่วนตั้งฉาก และท้องฟ้า ซึ่งจะมีขั้นตอนการทำงานดังต่อไปนี้

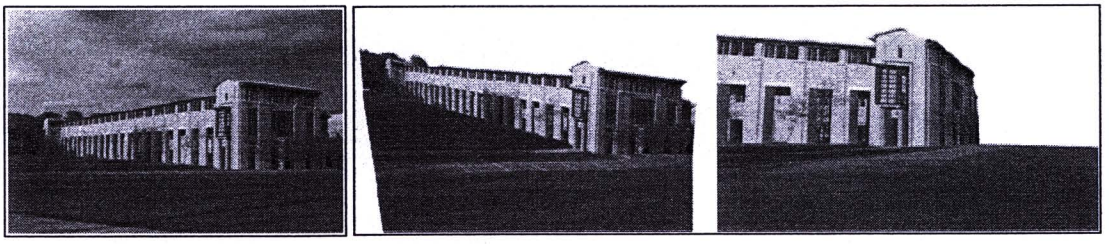
ทำการแบ่งรูปออกเป็นส่วนย่อย เรียกว่าซูเปอร์พิกเซลโดยใช้วิธีการแบ่งย่อยในเชิงกราฟอย่างมีประสิทธิภาพโดยเมื่อได้ซูเปอร์พิกเซลแล้ว จากนั้นจะทำการจัดกลุ่มของซูเปอร์พิกเซลที่น่าจะมีเลเบิลที่เหมือนกัน แล้วเลือกกลุ่มที่ดีที่สุดมาเป็นผลลัพธ์

การทำการจัดกลุ่มภาพนั้นจะใช้วิธีขั้นตอนวิธีเรียนรู้ของคอมพิวเตอร์ ซึ่งจำเป็นจะต้องสกัดค่าจุดเด่นบางอย่างออกมา เพื่อใช้ในอัลกอริทึม ในงานวิจัยนี้ได้จัดสกัดค่าจุดเด่นของ

แต่ละซูเปอร์พิกเซลโดยแบ่งออกเป็น 4 กลุ่ม ได้แก่ สี พื้นผิว ตำแหน่ง และลักษณะเชิงเรขาคณิต โดยเรียนรู้ผ่านแบบจำลองทางความน่าจะเป็นของอาดาบูสต์ [23]

การสร้างฉากสามมิติในงานวิจัยนี้ได้เสนอการขึ้นรูปแบบ ตัดและพับ (Cut and fold) โดยการตัดส่วนท้องฟ้าทิ้งไป แล้วตั้งส่วนตั้งฉากขึ้น ซึ่งจะสามารถจำลองรูปสามมิติที่มีความสมจริง

งานวิจัยนี้ก็ยังมีข้อจำกัดบางประการ เนื่องจากสมมติฐานที่ได้กล่าวมาแล้ว ทำให้วิธีการนี้สามารถใช้ได้กับฉากที่มีระนาบปรากฏได้เพียงระนาบเดียว รวมถึงไม่สามารถใช้กับรูปภาพที่มีวัตถุเป็นจำนวนมากได้ เนื่องจากวิธีนี้จะมองกลุ่มของวัตถุทั้งหมดเป็นชิ้นเดียว และวิธีการขึ้นรูปสามมิติเป็นเพียงการขึ้นรูปแบบคร่าวๆเท่านั้น



ก) ข)

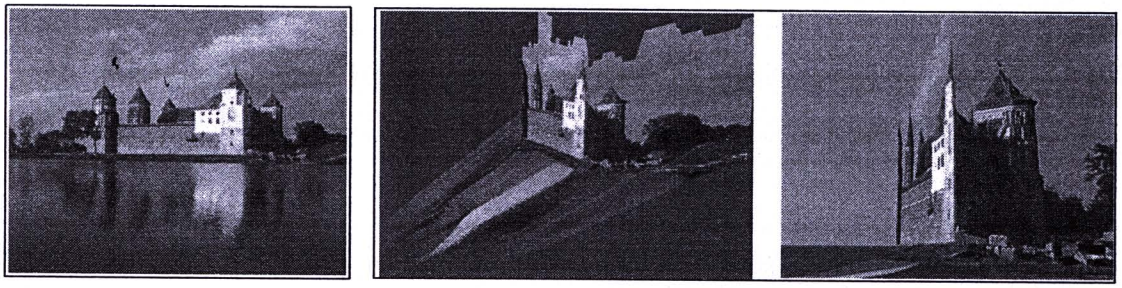
รูปที่ 2.13 ผลลัพธ์จากงานวิจัยการผุดขึ้นของภาพแบบอัตโนมัติ ก) ภาพนำเข้า ข) ผลลัพธ์การประมาณแบบจำลองสามมิติ [2]

นอกจากนี้ยังมีในปี 2007 แซกซีนาได้เสนองานวิจัยการเรียนรู้โครงสร้างฉากสามมิติจากรูปภาพเดี่ยว [3-5] โดยใช้มาร์คอฟแรนดอมฟิลด์ (Markov Random Field) ซึ่งเป็นโมเดลทางสถิติเพื่อใช้ในการหาตำแหน่งของแต่ละส่วนประกอบของรูปภาพ โดยได้เสนอมาร์คอฟแรนดอมฟิลด์ไว้สองรูปแบบ คือ แบบพารามิเตอร์ระนาบ (Plane parameter) และแบบจุด (point-wise)

แบบพารามิเตอร์ระนาบนั้น จะทำการแบ่งภาพเป็นซูเปอร์พิกเซล โดยที่ตั้งสมมติฐานว่าซูเปอร์พิกเซลนั้นมีลักษณะเป็นระนาบ โดยจะทำการหาตำแหน่งและทิศทางของแต่ละระนาบเพื่อนำมาประกอบเป็นรูป ส่วนในแบบจุดนั้นจะใช้จุดพิกเซลของภาพแทนการใช้ซูเปอร์พิกเซล แล้วหาตำแหน่งของแต่ละจุดพิกเซล โดยใช้ค่าจุดเด่นสองประเภท คือค่าจุดเด่น

ของซูเปอร์พิกเซลเดียว และค่าจุดเด่นที่เกิดจากความแตกต่างระหว่างคู่ซูเปอร์พิกเซล จากผลการทดลองพบว่ามาร์คอฟแรนดอมฟแลด์แบบพารามิเตอร์ระนาบนั้นให้ผลลัพธ์ที่ดีกว่า

โดยฉากที่ได้จากวิธีการนี้จะมีลักษณะเป็นแผ่นเดียวทั้งฉาก ซึ่งสามารถสร้างฉากที่มีความแม่นยำ แต่วิธีการดังกล่าวยังมีข้อจำกัดเนื่องจากลักษณะของฉากที่ได้ ทำให้ไม่สามารถแสดงผลของฉากที่มีวัตถุจำนวนมากได้



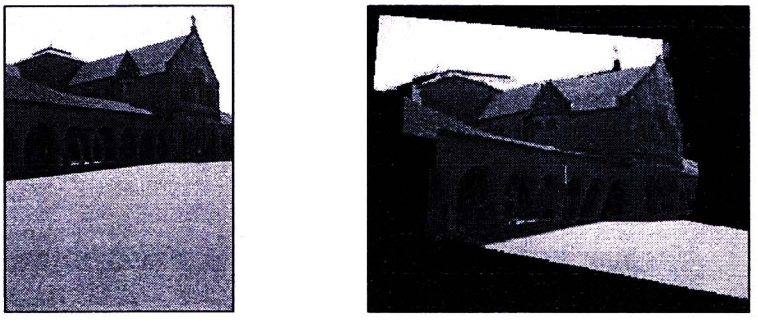
ก) ข)

รูปที่ 2.14 ผลลัพธ์จากงานวิจัยการเรียนรู้โครงสร้างฉากสามมิติจากรูปภาพเดี่ยว ก) ภาพนำเข้า ข) ผลลัพธ์การประมาณแบบจำลองสามมิติ [5]

ในปี 2010 ลิว และคณะ ได้เสนองานวิจัยการประมาณค่าความลึกของรูปเดี่ยวจากเลเบิล [24] โดยในงานวิจัยนี้เสนอการหาความสัมพันธ์ระหว่างเลเบิลของแต่ละจุดภาพกับค่าความลึก โดยนำความสัมพันธ์ที่ได้มาใช้ในการสร้างฉากสามมิติ

ในงานวิจัยนี้ทำการแบ่งคลาสของวัตถุภายในฉากไว้ทั้งหมดแปดกลุ่ม ได้แก่ ท้องฟ้า , ถนน, น้ำ, หญ้า, สิ่งก่อสร้าง, ภูเขา และวัตถุฉาก โดยทำการเรียนรู้หาความสัมพันธ์ของตำแหน่งในภาพกับความลึกในรูปแบบของลอการิทึมฐานสิบ แล้วใช้มาร์คอฟแรนดอมฟิลด์คำนวณหาค่าความลึกเช่นเดียวกับงานของแซกซีน่า

งานวิจัยนี้สามารถหาความลึกของภาพได้อย่างแม่นยำ แต่ในงานวิจัยชิ้นนี้ขึ้นอยู่กับตำแหน่งของกล้องมาก และไม่สามารถสร้างแบบจำลองที่มีความซับซ้อนมากได้



ก)　　ข)

รูปที่ 2.15 ผลลัพธ์จากงานวิจัยการประมาณค่าความลึกของรูปเดี่ยวจากเลเบิล เดี่ยว ก) ภาพนำเข้า ข) ผลลัพธ์การประมาณแบบจำลองสามมิติ [24]

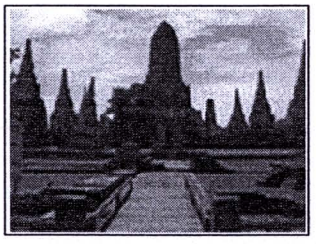
ในปี 2008 ฟูตระกูลได้เสนองานวิจัยการประมาณฉากสามมิติโดยการวิเคราะห์ความสัมพันธ์ระหว่างวัตถุ เพื่อสร้างฉากที่มีความซับซ้อน โดยใช้วิธีการทางการเรียนรู้ทางคอมพิวเตอร์ ในการจำแนกความสัมพันธ์ระหว่างคู่ซูเปอร์พิกเซลทั้งหมดหกรูปแบบ คือ วัตถุเดียวกัน บดบัง ถูกบดบัง อยู่ในระนาบเดียวกัน ตั้งอยู่บน และถูกทับ ซึ่งมีขั้นตอนการทำงานดังต่อไปนี้

วิธีการนี้จะทำการแบ่งซูเปอร์พิกเซลออกเป็นส่วนย่อย แล้วทำการปรับปรุงคุณภาพ และนำซูเปอร์พิกเซลทั้งหมดมาทำการหาความสัมพันธ์โดยผ่านอัลกอริทึมเพื่อนบ้านใกล้ที่สุด (Nearest neighbor) ซึ่งเมื่อได้ความสัมพันธ์ระหว่างทุกคู่พิกเซล จะทำการหาความลึกสัมพัทธ์ระหว่างซูเปอร์พิกเซลภาพ แล้วใช้ความรู้ด้านมุมกล้องแบบทัศนมิติ (Perspective) ในการหาความลึกสัมบูรณ์ของแต่ละจุดภาพ และนำมาสร้างฉากสามมิติ

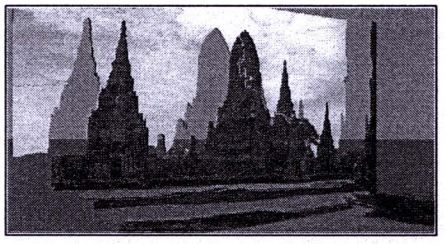
ซึ่งในงานวิจัยนี้ได้วิเคราะห์ค่าจุดเด่นโดยใช้ค่าจุดเด่นสองประเภท คือค่าจุดเด่นของตัวซูเปอร์พิกเซลเอง และค่าจุดเด่นระหว่างคู่ซูเปอร์พิกเซลที่ติดกัน และในแต่ละประเภทได้แบ่งออกเป็นสี่หมวดหมู่ คือ ด้านลายพื้นผิว ด้านเส้นขอบ ด้านเส้นตรง และด้านตำแหน่ง

แต่อย่างไรก็ตามเนื่องจากความสัมพันธ์ระหว่างซูเปอร์พิกเซลในงานวิจัยนี้มีความสัมพันธ์แบบเฉพาะส่วน ซึ่งไม่สามารถระบุถึงความสัมพันธ์ของทั้งฉากได้ ซึ่งทำให้เกิดข้อ

ขัดแย้งของความสัมพันธ์ของวัตถุ ทำให้การประมาณฉากสามมิติยังมีข้อบกพร่อง รวมถึงค่าจุดเด่นบางค่านั้นไม่มีความแม่นยำเพียงพอที่สามารถแยกความสัมพันธ์ได้อย่างมีประสิทธิภาพ



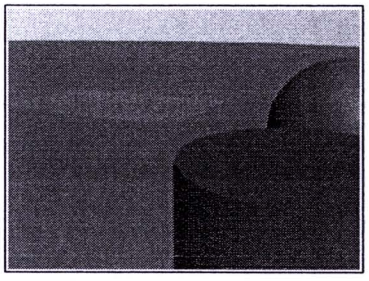
ก)



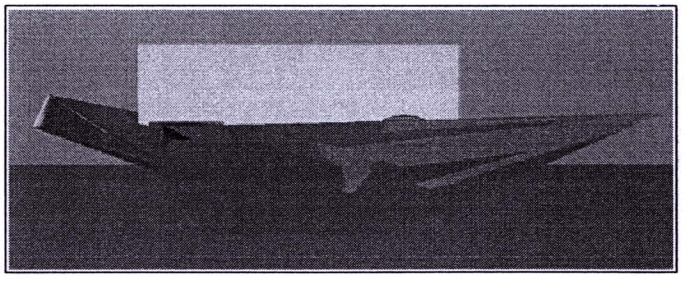
ข)

รูปที่ 2.16 ผลลัพธ์จากงานวิจัยการประมาณฉากสามมิติโดยการวิเคราะห์ความสัมพันธ์ระหว่างวัตถุ ก) ภาพนำเข้า ข) ผลลัพธ์การประมาณแบบจำลองสามมิติ [7]

โดยจากงานวิจัยที่ได้กล่าวมาสามารถสร้างฉากสามมิติจากรูปภาพเดี่ยวซึ่งประกอบไปด้วยวัตถุจำนวนมากได้ ถึงอย่างไรก็ดีการกำหนดความสัมพันธ์แบบเฉพาะส่วนไม่สามารถแก้ไขปัญหาการในการกำหนดความสัมพันธ์ได้อย่างสมบูรณ์ เนื่องจากความสัมพันธ์แบบเฉพาะส่วนนั้นไม่สามารถแสดงความสัมพันธ์ของซูเปอร์พิกเซลที่ห่างกันออกไปได้ รวมถึงยังไม่สามารถกำจัดข้อขัดแย้งระหว่างความสัมพันธ์ นอกจากนี้วิธีการในการสร้างภาพสามมิติยังคงประสบปัญหาความผิดพลาดในการสร้างฉาก ซึ่งส่งผลให้เกิดฉากที่มีความไม่สมเหตุสมผลขึ้น ซึ่งปัญหาที่ได้กล่าวมาสามารถแก้ไขได้โดยการเพิ่มความสัมพันธ์แบบวงกว้างสำหรับทุกวัตถุภายในฉาก รวมถึงการปรับปรุงขั้นตอนวิธีในการสร้างฉากสามมิติ



ก)



ข)

รูปที่ 2.17 การสร้างฉากสามมิติที่ผิดพลาดอันเนื่องมาจากความขัดแย้งของความสัมพันธ์ ก) ภาพนำเข้า ข) ผลลัพธ์การประมาณแบบจำลองสามมิติที่มีข้อผิดพลาด