บทที่ 3

แนวคิดการแก้ไขปัญหาและขั้นตอนในการแก้ปัญหา

3.1. แนวคิดการแก้ไขปัญหา

จากทฤษฎีและงานวิจัยที่เกี่ยวข้องที่ได้ศึกษาเกี่ยวกับการพิจารณาความสัมพันธ์ระหว่างพิกัดบนสองภาพ ซึ่งเป็นพื้นฐานในการนำไปประยุกต์ใช้ในงานการประมวลผลภาพมาร์คเกอร์เพื่อให้มุมมองหรือลักษณะทางกายภาพของมาร์คเกอร์ในฉากจริงและมาร์คเกอร์ต้นแบบที่มีความสัมพันธ์ในมุมมองเดียวกัน โดยทั่วไปเทคโนโลยีความจริงเสริมนั้นมีการเลือกใช้รูปแบบมาร์คเกอร์ที่มีลักษณะเฉพาะที่ง่ายต่อการสกัดคุณลักษณะเด่นเพื่อความแม่นยำในการตรวจจับและทนทานต่อการเปลี่ยนแปลงมุมมองต่างๆ สำหรับมาร์คเกอร์ที่มีรูปแบบเป็นลายมือเขียนซึ่งเกิดจากประกอบด้วยตัวอักษรแต่ละตัวนำมาประกอบกันเป็นคำศัพท์ซึ่งมีแบบการเขียนที่หลากหลายและแตกต่างกันออกไปตามลายมือของผู้เขียน อีกทั้งการมีรูปแบบเป็นลายเส้นที่ไม่แน่นอนทำให้ยากต่อการสกัดคุณลักษณะเด่นเพื่อใช้ในการเชื่อมต่อมุมมองของสองรูปภาพ เพื่อที่จะระบุตำแหน่งมาร์คเกอร์แบบลายมือเขียนได้ถูกต้องจึงศึกษาการระบุตำแหน่งข้อความบนรูปภาพเพื่อที่จะได้นำพื้นที่ของมาร์คเกอร์แบบคำศัพท์มาทำการสกัดคุณลักษณะเด่นที่ใช้สำหรับระบุตำแหน่งที่สัมพันธ์ของสองรูป เนื่องจากขั้นตอนการคำนวนค่าตำแหน่งเชิง 3 มิติ ของมาร์คเกอร์เมื่อเทียบกับฉากจริงนั้นต้องการจุดตำแหน่งใดๆที่เป็นจุดตำแหน่งเดียวกันระหว่างสองรูปภาพเพื่อใช้ในการคำนวนค่าโฮโมกราฟฟีเมทริกซ์ที่สามารถระบุความสัมพันธ์ของจุดในตำแหน่งเดียวกันของทั้งสองรูปได้

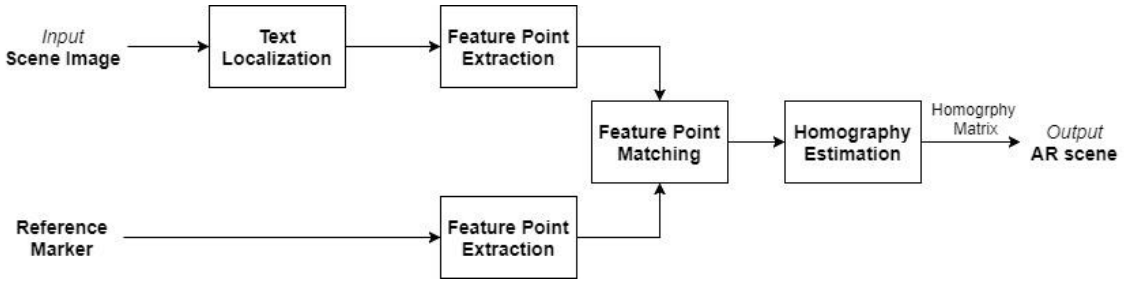
ดังนั้นวิทยานิพนธ์นี้จึงมุ่งเน้นในการพัฒนาวิธีการพิจารณาความสัมพันธ์ระหว่างพิกัดบนสองภาพ ที่มีความสัมพันธ์ในมุมมองเดียวกัน โดยในงานวิจัยนี้จะพัฒนาการสกัดคุณลักษณะเด่นของสองรูปภาพที่ต่างมุมมองกัน และทำพิจารณาคุณลักษณะเด่นเหล่านั้นเพื่อใช้ในการคำนวนหาความสัมพันธ์ของจุดตำแหน่งใดๆที่เป็นจุดตำแหน่งเดียวกันระหว่างสองรูปได้ โดยวิธีการสกัดคุณลักษณะเด่นนั้นมีความทนทานต่อการเปลี่ยนแปลงมุมมองต่างๆ เช่น การหมุน การเปลี่ยนขนาด การย้ายตำแหน่ง และมุมมองจากเปลี่ยนมุมต่างๆกับพื้นระนาบ ซึ่งมาร์คเกอร์แบบลายมือเขียนนั้นจะมาจากลายมือของผู้เขียนที่แตกต่างกัน โดยการที่จะสกัดคุณลักษณะเด่นของมาร์คเกอร์แบบลายมือเขียนนั้นจะอยู่พื้นฐานของการระบุตำแหน่งข้อความด้วยวิธีการ MSER ที่สามารถพิจารณาพื้นที่ที่มีความเข้มสีต่อเนื่องกันและแตกต่างจากความเข้มสีของขอบของพื้นที่นั้นๆ

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

3.2. แนวทางการแก้ไขปัญหา

ขั้นตอนวิธีการการที่จะสกัดคุณลักษณะเด่นของมาร์คเกอร์แบบลายมือเขียนนั้นจะอยู่พื้นฐานของวิธีการระบุตำแหน่งข้อความ ซึ่งในวิจัยนี้ปรับใช้ขั้นตอนวิธีจากการนำเสนอในงานวิจัยของ Intaratat และคณะ [21] โดยขั้นตอนในการแก้ปัญหานั้นได้เลือกใช้วิธี MSER มาช่วยในการระบุตำแหน่งมาร์คเกอร์แบบลายมือเขียนซึ่งเป็นเทคนิคที่สามารถทนทานต่อการย้ายตำแหน่ง การเปลี่ยนมุมมองและการขนาด โดยจะทำการปรับค่าพารามิเตอร์ที่เป็นลักษณะเฉพาะของคุณลักษณะของพื้นที่ตัวอักษรเพื่อระบุตำแหน่งของคำศัพท์ก่อนแล้วจึงปรับค่าพารามิเตอร์อีกครั้งเพื่อให้เหมาะสมสำหรับการค้นหาตัวอักษรบางตัวที่หายไป เมื่อได้พื้นที่ของคำศัพท์ทั้งหมดจึงทำการหาคุณลักษณะเด่นของคำนั้นๆเพื่อการเชื่อมต่อมุมมองของภาพเสริมกับมุมมองของฉากจริง โดยขั้นตอนของระบบการทำงานทั้งหมดสามารถแสดงได้ดังรูปที่ 3.1 จากข้อมูลนำเข้าซึ่งเป็นรูปจากฉากจริงนำเข้าสู่ขั้นตอนการระบุตำแหน่งของข้อความ (Text Localization) ซึ่งเป็นคุณลักษณะของมาร์คเกอร์แบบลายมือโดยจะทำการสกัดพื้นที่ของตัวอักษรที่เป็นไปได้ในรูปภาพจากการสกัดพื้นที่ด้วย MSER (MSER Extraction) จากนั้นสามารถคัดกรองพื้นที่ที่มีความเป็นไปไม่ได้ที่จะเป็นพื้นที่ของตัวอักษรด้วยคุณลักษณะเฉพาะประกอบด้วย ความทึบ (Solidity) ความยาวแกนหลัก (Major Axis Length) และความยาวแกนรอง (Minor Axis Length) หลังจากนั้นจะทำการจัดกลุ่มของตัวอักษรจากตำแหน่งจุดศูนย์กลาง (Centroid) ขนาด (Size) และความเข้มสี (Color Intensity) สุดท้ายจะทำการตรวจสอบพื้นที่ด้านข้างของกลุ่มพื้นที่นั้นด้วยคุณลักษณะโดยรวมของกลุ่มพื้นที่ตัวอักษรในพื้นที่นั้นๆเพื่อป้องกันความผิดพลาดในการสกัดพื้นที่ตัวอักษรที่ไม่สามารถผ่านเกณฑ์การสกัดพื้นที่ในครั้งแรกได้ เมื่อสามารถระบุตำแหน่งของมาร์คเกอร์ได้แล้ว จะนำพื้นที่ไปสกัดจุดคุณลักษณะเด่น (Feature Point Extraction) ด้วยวิธีการคอนเวกซ์ฮัลล์ ทำให้ได้เซตของจุดที่ตั้งอยู่บนขอบของคำศัพท์ ซึ่งสามารถลดจุดที่ไม่สำคัญออกไปได้ด้วยวิธีการลดรูปหลายเหลี่ยมให้อยู่ในรูปอย่างง่าย ทำให้สามารถตัดจุดที่ไม่สำคัญออกไปเหลือเพียงจุดคุณลักษณะเด่นของพื้นที่คำศัพท์นั้น จากนั้นทำการจับคู่คุณลักษณะเด่น (Feature Point Matching) ระหว่างมาร์คเกอร์จากฉากจริง และมาร์คเกอร์ต้นแบบซึ่งรูปแบบการจับคู่จะถูกคัดกรองด้วยระยะทางจากตำแหน่งจุดศูนย์กลางของพื้นที่ และมุมที่กระทำกับจุดศูนย์กลาง ซึ่งรูปแบบที่มีคุณลักษณะดังกล่าวใกล้เคียงกับรูปแบบจากมาร์คเกอร์ต้นแบบจะถูกลำดับให้ไปทำการสกัดคุณลักษณะเด่นเฉพาะที่ของแต่ละจุดคุณลักษณะเด่นด้วยวิธีการ Histogram of the oriented gradient (HOG) โดยวัดความแตกต่างของคุณลักษณะจาก HOG ที่น้อยที่สุดจะเป็นรูปแบบการจับคู่ของจุดที่เป็นตำแหน่งเดียวกันระหว่างสองรูป เมื่อได้จุดที่ตรงกันระหว่างสองรูปจะทำให้สามารถประมาณค่าโฮโมกราฟฟี (Homography Estimation) เพื่อเชื่อมต่อมุมมองของมาร์คเกอร์ทั้งสองให้ไปในทิศทางเดียวกันได้

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved



รูปที่ 3.1 ขั้นตอนระบบการทำงาน

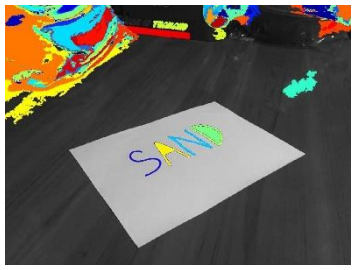
3.2.1. ขั้นตอนการระบุตำแหน่งข้อความ (Text Localization)

เริ่มด้วยจำกัดพื้นที่ต่างๆในภาพที่มีความน่าจะเป็นพื้นที่ของคำศัพท์ด้วยวิธีการใช้เทคนิค MSER ในการสกัดพื้นที่โดยรูปภาพนำเข้าจากฉากจริงจะถูกปรับช่องสัญญาณสีเป็นแบบโทนขาวดำ (Gray-scale) โดยผลลัพธ์จากการสกัด MSER จะเป็นกลุ่มพิกเซลที่ติดกันซึ่งจำนวนและลักษณะของแต่ละกลุ่มพิกเซลมาจากการปรับพารามิเตอร์ที่สำคัญต่างๆคือ ค่าเกณฑ์ความเปลี่ยนแปลง (Threshold Delta) และ ค่าช่วงของขนาดพื้นที่ (Region Area Range) โดยการตั้งค่าเกณฑ์ความเปลี่ยนแปลงในระดับที่สูงนั้นจะทำให้สกัดพื้นที่ที่มีโอกาสเป็นคำศัพท์สูงแต่อาจสูญเสียพื้นที่ของตัวอักษรบางตัวไปได้ ซึ่งถ้าหากปรับค่าเกณฑ์ความเปลี่ยนแปลงในระดับที่ต่ำไปนั้นจะทำให้สกัดพื้นที่ที่ไม่จำเป็นที่ไม่ใช่พื้นที่ของตัวอักษรได้มากทำให้มีการคำนวณที่ใช้เวลานาน ดังนั้นในการสกัดหาพื้นที่ของคำศัพท์ในครั้งแรกจึงใช้ค่าพารามิเตอร์ที่สูงเพื่อลดการสกัดพื้นที่อื่นๆที่ไม่ใช่พื้นที่ของคำศัพท์

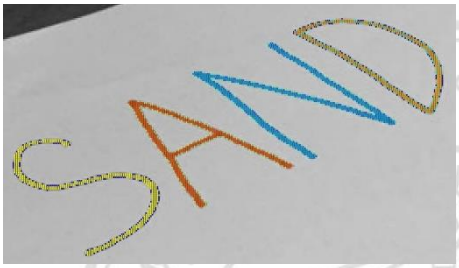
ถึงแม้ว่าผลลัพธ์ที่ได้จากเทคนิค MSER ที่ปรับค่าพารามิเตอร์ต่างๆให้อยู่ในเกณฑ์ที่สูงแต่ก็อาจจะมีพื้นที่อื่นๆที่ไม่ใช่พื้นที่ของคำศัพท์ได้ จึงนำพื้นที่ผลลัพธ์จาก MSER มาสกัดหาคุณลักษณะต่างๆ เช่น ค่าความทึบ (Solidity) ความแกนหลัก (Major Axis Length) ความแกนรอง (Minor Axis Length) อัตราส่วนแกน (Rate Axis) และค่าความเข้มสี (Intensity) [4] เพื่อใช้ในการคัดเลือกพื้นที่เป็นพื้นที่ของตัวอักษรและรวมกันเป็นพื้นที่ของคำศัพท์ หลังจากการผ่านเงื่อนไขเกณฑ์คุณลักษณะที่กำหนดดังกล่าวระบบจะสามารถระบุตำแหน่งพื้นที่ของคำศัพท์และจัดการตัดพื้นที่อื่นๆออกหมดได้ ซึ่งตำแหน่งพื้นที่คำศัพท์นั้นจะเป็นตำแหน่งของมาร์คเกอร์แบบลายมือเขียนเพื่อใช้ในการสกัดคุณลักษณะเด่นต่อไป โดยสามารถแสดงตัวอย่างในรูปที่ 3.2 เป็นตัวอย่างการระบุตำแหน่งของคำศัพท์ในภาพมาร์คเกอร์ ผลลัพธ์สุดท้ายจะถูกแทนที่ด้วยสีอื่นๆแทนสีโทนขาวดำโดยพื้นที่แต่ละพื้นที่สกัดได้จะมีสีที่แตกต่างกันไป

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

(ก)



(ข)



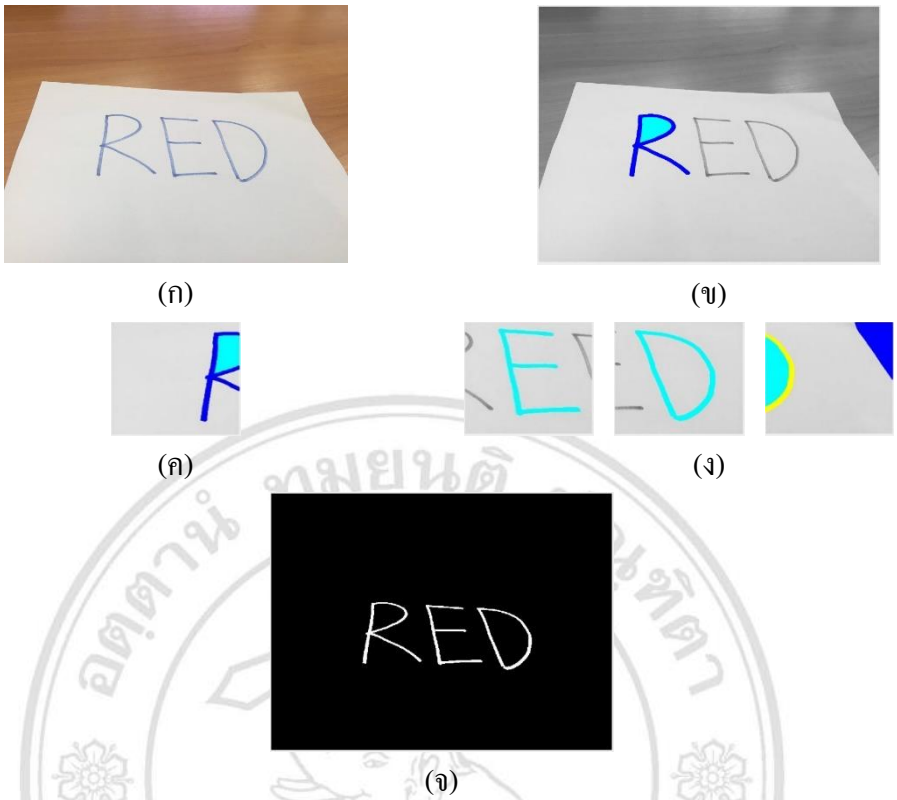
(ค)

(ง)

รูปที่ 3.2 ภาพตัวอย่างขั้นตอนในการสกัดตัวอักษร (ก) รูปนำเข้าจากกล้อง (ข) ผลลัพธ์จากการสกัดพื้นที่ด้วย MSER (ค) ผลลัพธ์พื้นที่ของคำศัพท์ และ (ง) รูปตำแหน่งพื้นที่ของคำศัพท์ตามตำแหน่งในภาพจากฉากจริง

แต่เนื่องจากการตั้งค่าเกณฑ์ความเปลี่ยนแปลงที่สูงจึงอาจทำให้พื้นที่คำศัพท์ที่สกัดได้มีบางส่วนที่ขาดหายไป จึงทำการตรวจสอบพื้นที่ด้านข้างทั้งซ้ายและขวาของพื้นที่คำศัพท์ที่สกัดได้ในครั้งแรก โดยการตรวจสอบพื้นที่ด้านข้างนั้นจะใช้วิธีการ MSER ในการสกัดพื้นที่เหมือเดิม แต่จะใช้เกณฑ์ค่าความเปลี่ยนแปลงและการจำกัดขนาดของพื้นที่ให้ต่ำลง เนื่องจากพื้นที่ของตัวอักษรที่ขาดหายไปไหนอาจมีความเข้มสีหรือขนาดของพื้นที่ที่ไม่สามารถผ่านเกณฑ์ในครั้งแรกได้ โดยพื้นที่ที่ทำการตรวจสอบจะถูกจำกัดขนาดของพื้นที่ให้เป็นไปตามขนาดโดยเฉลี่ยของกลุ่มพื้นที่ของคำศัพท์ที่สกัดได้ในครั้งแรก ซึ่งพื้นที่ที่สามารถสกัดได้ในครั้งนี้จะถูกคัดกรองจากคุณลักษณะต่างๆเช่นเดียวกับการสกัดครั้งแรกของกลุ่มพื้นที่ตัวอักษรโดยเฉลี่ยจากคุณลักษณะของผลลัพธ์จากการสกัดครั้งแรกทั้งหมด พื้นที่ผลลัพธ์จากการตรวจสอบจะถูกรวมเข้ากับกลุ่มผลลัพธ์ของพื้นที่คำศัพท์ในครั้งแรก โดยการตรวจสอบจะหยุดตรวจสอบเมื่อผลลัพธ์สุดท้ายในการค้นหาคำศัพท์นั้นไม่เปลี่ยนแปลง สามารถแสดงขั้นตอนการตรวจหาคำพื้นที่ของคำศัพท์ที่ขาดหายไปได้ดังตัวอย่างรูปที่ 3.3 เป็นตัวอย่างมาร์คเกอร์ลายมือเขียนที่มีของลายเส้นไม่ชัดเจนทำให้ไม่สามารถผ่านเกณฑ์ค่าการเปลี่ยนแปลงในครั้งแรกได้

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved



(ก) (ข)

(ค) (ง)

(จ)

รูปที่ 3.3 การค้นหาพื้นที่ตัวอักษรที่ขาดหาย (ก) รูปนำเข้าจากกล้อง (ข) ผลลัพธ์จากการสกัดพื้นที่ด้วย MSER (ค) ผลลัพธ์การตรวจสอบพื้นที่ตัวอักษรที่หายไปทางด้านขวา ง) ผลลัพธ์การตรวจสอบพื้นที่ตัวอักษรที่หายไปทางด้านซ้าย และ (จ) ตำแหน่งพื้นที่ของคำศัพท์ตามตำแหน่งในภาพนำเข้าจากฉากจริง

จากตัวอย่างในรูปที่ 3.3 (ข) ผลลัพธ์จากการสกัดพื้นที่ด้วย MSER นั้นสามารถสกัดพื้นคำศัพท์ได้เพียงตัว R ตัวเดียวเนื่องจากมีค่าผ่านเกณฑ์ความเปลี่ยนแปลงที่กำหนดในครั้งแรก ดังนั้นระบบจึงตรวจพื้นที่ด้านข้างของพื้นที่ตัวอักษร R โดยใช้เกณฑ์เฉลี่ยต่างๆของพื้นที่คำศัพท์แรกที่สามารถสกัดได้เป็นเกณฑ์ในการเลื่อนไปทั้งด้านซ้ายและขวาได้ผลลัพธ์ดังรูปที่ 3.3 (ค) เป็นการเลื่อนตรวจสอบทางด้านขวาและ (ง) เป็นการเลื่อนตรวจสอบทางด้านซ้าย โดยทุกครั้งที่มีการตรวจสอบจะนำผลลัพธ์ทั้งสองรูปผลลัพธ์คือผลลัพธ์ก่อนหน้าตรวจสอบกับผลลัพธ์ล่าสุดมาหาค่าความผลต่างกันให้อยู่ในช่วงที่กำหนด ซึ่งเป็นการตรวจความแตกต่างระหว่างผลลัพธ์พื้นที่คำศัพท์เดิมกับผลลัพธ์พื้นที่คำศัพท์ใหม่ ซึ่งถ้าการตรวจสอบพื้นที่ตัวอักษรจากด้านข้างไม่พบพื้นที่ที่เป็นตัวอักษรก็จะหยุดตรวจสอบและคืนผลลัพธ์สุดท้ายดังรูปที่ 3.3 (จ)

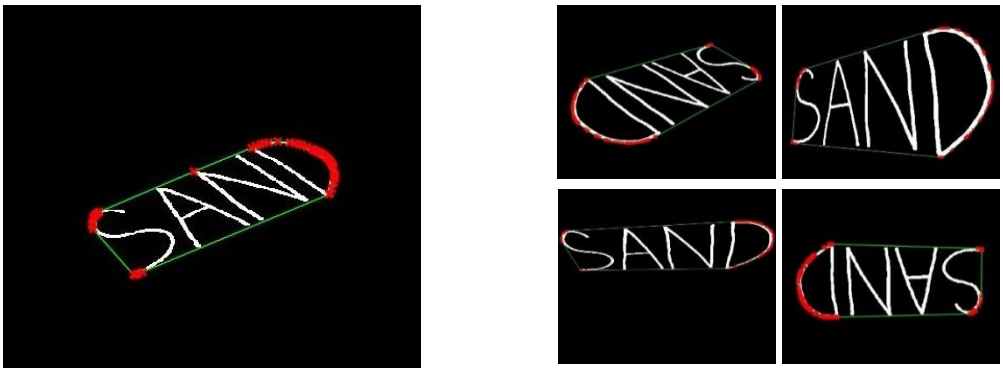
ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

3.2.2. ขั้นตอนสกัดลักษณะเด่นของพื้นที่คำศัพท์ (Feature Point Extraction)

หลังจากสามารถสกัดพื้นที่ของคำศัพท์ได้แล้วจึงนำพื้นที่ทั้งหมดมาสกัดจุดคุณลักษณะเด่น (Feature Point) โดยเริ่มจากการหาขอบรูปหลายเหลี่ยม (Polygon) ที่ล้อมรอบพื้นที่ของคำศัพท์ โดยใช้เทคนิค คอนเวกซ์ฮัลล์ ซึ่งเป็นเซตนูนของขอบที่เล็กที่สุดที่ครอบคลุมพื้นที่ของคำศัพท์คล้ายกับยางรัดล้อมรอบพื้นที่ของคำศัพท์ไว้ สามารถแสดงตัวอย่างได้ดังรูปที่ 3.4 (ก) ซึ่งแทนจุดในเซตของคอนเวกฮัลล์ด้วยจุดกากบาทสีแดงที่อยู่บนขอบของพื้นที่ โดยในมาร์คเกอร์คำศัพท์เดียวกันจะสามารถคาดหวังจุดในเซตของคอนเวกฮัลล์ที่มีความทนทานต่อมุมมองที่หลากหลายได้ซึ่งสามารถแสดงได้ดังตัวอย่างมุมมองอื่น ๆที่แตกต่างกันของรูปแบบคำศัพท์เดียวกันในรูปที่ 3.4 (ข) เนื่องจากเซตของจุดที่เป็นคอนเวกซ์ฮัลล์ของพื้นที่คำศัพท์นั้นมีจำนวนจุดเป็นจำนวนมากเพื่อลดเวลาใช้การดำเนินการจับคู่คุณลักษณะเด่นจึงนำเซตดังกล่าวไปทำการลดรูปของขอบรูปหลายเหลี่ยมให้เหลือจำนวนจุดครอบพื้นที่ที่สำคัญและมีความแปรผันตามมุมมองต่าง ๆเป็นคุณลักษณะเด่นที่ไม่เปลี่ยนไปของมาร์คเกอร์ ซึ่งทำการลดจุดมุมจากความสำคัญของแต่ละจุด โดยการพิจารณาจากมุมและระยะห่างระหว่างจุดโดยจุดที่มีค่ามุมที่น้อยและระยะห่างระหว่างจุดมากก็มีโอกาสที่จะโดนตัดออกน้อยสุด โดยจะทำการตัดจุดที่ไม่สำคัญออกจนเหลือจำนวนจุดตามที่กำหนด ซึ่งขั้นตอนนี้จะทำการสกัดคุณลักษณะเด่นในส่วนของรูปมาร์คเกอร์ต้นแบบด้วยเพื่อใช้ในการจับคู่จุดที่เป็นตำแหน่งเดียวกันระหว่างสองรูปในขั้นตอนต่อไป

สำหรับการประมาณค่ามุมระหว่างแกนหลักกับแกนแนวระนาบ และขนาดของคุณลักษณะเด่นจากจุดคุณลักษณะเด่นที่สกัดได้ ซึ่งแต่ละจุดคุณลักษณะเด่นนั้นจะมีคุณลักษณะเฉพาะของตัวเองจากพื้นพิกเซลโดยรอบจุดนั้นๆกระจายออกเป็นสี่เหลี่ยมจัสตุรัส โดยที่ค่ามุมระหว่างแกนหลักกับแกนแนวระนาบของทุกจุดคุณลักษณะเด่นจะถูกอ้างอิงจากค่ามุมระหว่างแกนหลักกับแกนแนวระนาบของพื้นที่คำศัพท์ทั้งหมดจากฉากจริงซึ่งแทนด้วย Φ อย่างไรก็ตามค่ามุมระหว่างแกนหลักกับแกนแนวระนาบของทุกจุดคุณลักษณะเด่นนั้นสามารถเป็นได้ทั้ง Φ หรือ $-\Phi$ เนื่องจากไม่สามารถรู้ได้ว่าพื้นที่คำศัพท์ที่สกัดได้นั้นเป็นมุมองที่กลับด้านหรือไม่ ส่วนขนาดของลักษณะเด่นเฉพาะของจุดลักษณะเด่นจะแทนด้วยขนาด σ ซึ่งจะแปรผันตรงกับความยาวของค่าแกนหลักของตัวอักษรที่จุดลักษณะเด่นนั้นตั้งอยู่ โดยตัวอย่างกำหนดลักษณะเด่นห้าจุดสุดท้ายของมาร์คเกอร์ลายมือจากฉากจริง และมาร์คเกอร์ต้นแบบ สามารถแสดงได้ดังรูปที่ 3.4 (ข)

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved



(ก) (ข)

(ค) (ง)

(จ) (ฉ)

(ช) (ซ)

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

รูปที่ 3.4 การสกัด Feature point (ก) รูปเซตของจุดจากคอนเวกซ์ฮัลล์ (ข) ตัวอย่างมุมมองต่างๆของ (ค) จุดคุณลักษณะเด่นของรูปมาร์คเกอร์จากฉากจริง (ง) จุดคุณลักษณะเด่นของรูปมาร์คเกอร์ต้นแบบ (จ) คุณลักษณะเด่นเฉพาะของจุดคุณลักษณะเด่นของรูปมาร์คเกอร์จากฉากจริง (ฉ) คุณลักษณะเด่นเฉพาะของจุดคุณลักษณะเด่นของรูปมาร์คเกอร์ต้นแบบ (ช) ตัวอักษรจากรูปมาร์คเกอร์จากฉากจริง (ซ) ตัวอักษรจากรูปมาร์คเกอร์ต้นแบบ

3.2.3. ขั้นตอนการจับคู่ลักษณะเด่น (Feature Point Matching)

เพื่อที่จะประมาณค่า โฮโมกราฟฟีเมทริกซ์ (Homography Matrix) จำเป็นจะต้องใช้จุดที่เป็นตำแหน่งเดียวกันระหว่างสองรูปอย่างน้อย 4 คู่ ดังนั้นกำหนดให้ n และ m เป็นจำนวนของจุดคุณลักษณะเด่นที่ถูกสกัดจากรูปจากฉากจริงและรูปมาร์คเกอร์ต้นแบบ ซึ่ง n และ m นั้นถูกกำหนดค่าไว้ในขั้นตอนการสกัดจุดคุณลักษณะเด่น ดังนั้นทำให้จุดคุณลักษณะเด่นซึ่งตั้งอยู่บนขอบของพื้นที่คำศัพท์จึงมีคุณสมบัติลำดับแบบวงกลมโดยที่ลำดับของจุดคุณลักษณะเด่นนั้นสำคัญ การจับคู่ลักษณะเด่นของทั้งสองรูปนั้นจะเป็นการจับคู่แบบตามลำดับไม่สามารถข้ามลำดับได้ โดยจะใช้รูปตั้งต้นเป็นหลักซึ่งจะเรียงตามลำดับที่กำหนดจำนวนลักษณะเด่นไว้ จากตัวอย่างในตารางที่ 1 กำหนดห้าจุดจึงเรียงตามลำดับ 1, 2, 3 และ 4 ซึ่งจะไปจับคู่ลักษณะเด่นของรูปจากฉากจริงที่สามารถสกัดมาได้เช่นกัน ซึ่งจากตัวอย่างมีห้าจุด การจับคู่จะตัดทีละจุดจากฝั่งภาพจากฉากจริงออก เช่น จุดที่ 1, 2 , 3 และ 4 ของรูปต้นแบบ จะจับคู่และวัดความถูกต้องจากระยะทางระหว่างลักษณะเด่นที่น้อยที่สุด เริ่มจากการตัดจุดที่ 1 จากรูปกล้องโดยจะได้คู่ที่จับคู่คือ {2, 3, 4, 5}, {3, 4, 5, 2}, {4, 5, 2, 3} และ{5, 2, 3, 4} จากนั้นทดลองตัดจุดถัดไปและจับคู่ไปเรื่อยๆ ทำให้เกิดจำนวนรูปแบบการจับคู่กันที่เกิดขึ้นได้ทั้งหมด $4\times\frac{n!}{(n-4)!4!}\times\frac{m!}{(m-4)!4!}$ รูปแบบ สามารถแสดงตัวอย่างการจับคู่ระหว่างรูปต้นแบบ n=5 จุด กับ รูปจากฉากจริง m=5 จุด ตามลำดับจากซ้ายไปขวา มีจำนวนรูปแบบทั้งหมด 100 รูปแบบดังตารางที่ 1

ตารางที่ 3.1 เซตของรูปแบบการจับคู่ n=5 และ m=5 ระหว่างรูปต้นแบบ และรูปจากฉากจริง

จุดจากรูปต้นแบบ	จุดจากรูปจากฉากจริง
{1,2,3,4}, {1,2,3,5}, {1,2,4,5}, {1,3,4,5}, {2,3,4,5}	{1,2,3,4}, {2,3,4,1}, {3,4,1,2} , {4,1,2,3}
	{5,1,2,3}, {1,2,3,5}, {2,3,5,1}, {3,5,1,2}
	{4,5,1,2}, {5,1,2,4}, {1,2,4,5}, {2,4,5,1}
	{3,4,5,1}, {4,5,1,3}, {5,1,3,4}, {1,3,4,5}
	{2,3,4,5}, {3,4,5,2}, {4,5,2,3}, {5,2,3,4}
ทั้งหมด **100** รูปแบบ	

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

โดยจะทำการเลือกหนึ่งรูปแบบจากทั้งหมดที่เป็นรูปแบบการจับคู่ของจุดที่เป็นตำแหน่งเดียวกันระหว่างสองรูป ซึ่งจากจำนวนรูปแบบทั้งหมดนั้นสามารถคัดเลือกรูปแบบการจับคู่ที่มีโอกาสจับคู่ลักษณะเด่นถูกต้องสูงได้ด้วยการเปรียบเทียบมุมและระยะทางระหว่างจุดศูนย์กลางกับจุดลักษณะเด่นจากรูปต้นแบบ และรูปจากฉากจริง ซึ่งจะทำให้สามารถคัดการจับคู่ที่ผิดพลาดออกไปได้ส่วนหนึ่ง โดยการกำหนดให้ $P=\{(p_i,q_i)\}_{i=1}^{4}$ เป็นรูปแบบของการจับคู่จุดลักษณะเด่นที่เป็นไปได้ โดยที่ p_i แทนจุดคุณลักษณะเด่นของรูปจากฉากจริง และ q_i แทนจุดคุณลักษณะเด่นของรูปต้นแบบที่คาดว่าจะเป็นจุดตำแหน่งเดียวกันกับจุด p_i ดังนั้นเพื่อที่จะหารูปแบบที่ดีที่สุดงานวิจัยนี้จึงได้แบ่งขั้นตอนการจับคู่จุดคุณลักษณะเด่นออกเป็นสองขั้นตอน

1. เพื่อที่จะหารูปแบบที่มีความน่าจะเป็นการจับคู่ของจุดที่สูงขึ้นจึงใช้การเปรียบเทียบตวามแตกต่างมุมและระยะทางระหว่างจุดศูนย์กลางกับจุดลักษณะเด่นของรูปต้นแบบ และรูปจากฉากจริง โดยสามารถคำนวนได้ดังสมการที่ 3.1

$$E_d(P)=\sum_{i=1}^{4}\left|d(p_i)-d(q_i)\right| \text{ and } E_\theta(P)=\min_{\pm\Phi}\left(\sum_{i=1}^{4}\left|\theta(p_i)-\theta(q_i)\pm\Phi\right|\right) \tag{3.1}$$

เมื่อ $d(\cdot)$ คือระยะทางจากจุดคุณลักษณะเด่นถึงจุดศูนย์กลางของคำศัพท์ ซึ่งจะถูกทำให้อยู่ในรูปมาตรฐานด้วยค่าความยาวแกนหลักของคำศัพท์ และ $\theta(\cdot)$ คือความแตกต่างระหว่างมุมของทิศทางจากจุดศูนย์กลางของพื้นที่คำศัพท์ ดังนั้นจึงรวม E_d และ E_θ เป็นฟังก์ชันเปรียบเทียบความต่างเป็นสมการในขั้นตอนที่ 1 แทนด้วย $E(P)$ แสดงได้ดังสมการที่ 3.2

$$E(P)=\frac{E_d(P)}{\max_P\left(E_d(P)\right)}+\frac{E_\theta(P)}{\max_P\left(E_\theta(P)\right)} \tag{3.2}$$

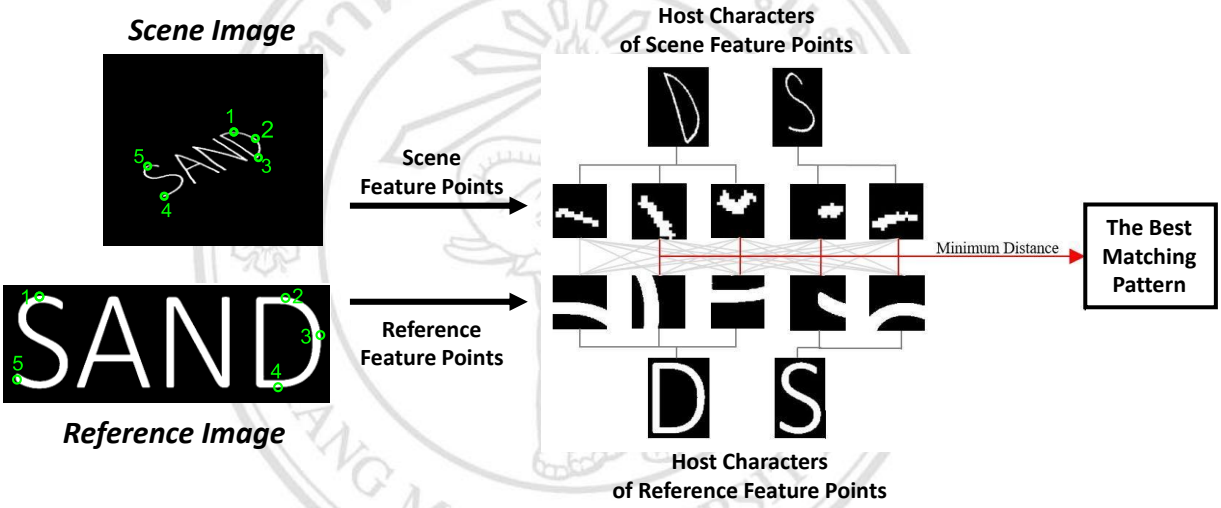
2. หลังจากที่รูปแบบการจับคู่ทั้งหมดถูกจัดลำดับด้วยฟังก์ชันเปรียบเทียบความต่างในขั้นตอนที่ 1 จึงสามารถเลือกลำดับการจับคู่ที่สูงที่สุด K รูปแบบ ในจำนวนรูปแบบที่ได้คัดเลือกมานั้นสามารถลดความผิดพลาดในการประมาณค่าโฮโมกราฟฟีเมทริกซ์ได้ด้วยการตรวจสอบการกระจายตัวของจุดคุณลักษณะเด่นในรูปแบบนั้นๆจะต้องกระจายตัวออกเป็นสี่ควอดรันต์โดยแบ่งจากจุดศูนย์กลางของพื้นที่ โดยรูปแบบที่มีการกระจายจะถูกพิจารณาสกัดคุณลักษณะเด่นเฉพาะของแต่ละจุดก่อน ในขั้นตอนที่สองนี่เลือกวิธี HOG มาเป็นวิธีที่ใช้ในการสกัดคุณลักษณะเด่นของภาพคุณลักษณะเด่นเฉพาะของแต่ละจุด และคุณลักษณะเด่นของตัวอักษรที่แต่ละจุดคุณลักษณะเด่นตั้งอยู่ ซึ่งกำหนดขนาดของพื้นที่กรอบสี่เหลี่ยมที่ล้อมรอบจุดคุณลักษณะเด่นด้วยขนาด σ ในแต่ละจุดคุณลักษณะเด่นซึ่งในงานวิจัยชิ้นนี้ได้ใช้ขนาด 0.25σ ในการทดลอง ในขณะที่ค่ามุมระหว่างแกนหลักกับแกนแนวระนาบของพื้นที่คำศัพท์ทั้งหมดจากฉากจริงที่มีค่า $\pm\Phi$ จะถูกเลือกจากค่า $E_\theta(P)$ ที่น้อยที่สุด

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

ท้ายที่สุดแล้วนำคุณลักษณะเด่นเฉพาะกับคุณลักษณะเด่นของตัวอักษรที่ตั้งไปทำการคำนวณระยะห่างในมิติของคุณลักษณะเด่นทั้งสองที่ได้ไปใช้ในการจับคู่แต่ละจุดจากการจับคู่ที่มีระยะทางรวมน้อยที่สุดระหว่างคุณลักษณะเด่นฉากจริงและรูปต้นแบบ สามารถแสดงได้ดังสมาการที่ 3.3 โดยที่ $h(\cdot)$ แทนคุณลักษณะเด่นจากการสกัดทั้งสองขั้นตอน

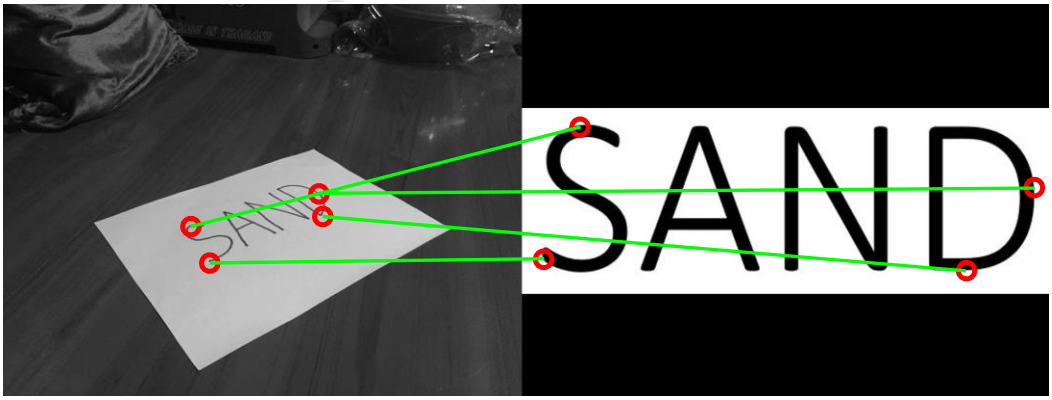
$$P_{\text{opt}} = \arg\min_{P}\left(\sum_{i=1}^{4}\left|h(p_i) - h(q_i)\right|\right) \tag{3.3}$$

โดยตัวอย่างการแสดงตัวอย่างการจับคู่รูปแบบจากคุณลักษณะเด่นเฉพาะ และคุณลักษณะเด่นของตัวอักษรที่จุดคุณลักษณะเด่นนั้นๆตั้งอยู่แสดงได้ดังรูปที่ 3.5



รูปที่ 3.5 ขั้นตอนการจับคู่ลักษณะเด่น

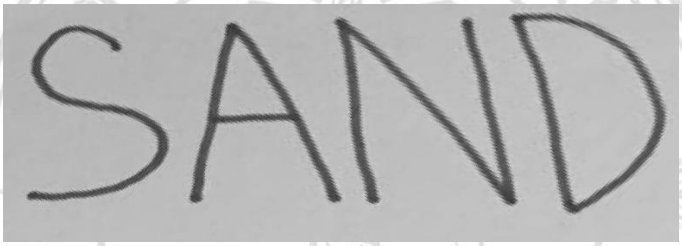
ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่



รูปที่ 3.6 ผลลัพธ์การจับคู่ลักษณะเด่น

3.2.4. การประมาณค่าโฮโมกราฟฟี (Homography Estimation)

ขั้นตอนการคำนวนค่าตำแหน่งเชิง 3 มิติ ของมาร์คเกอร์ตั้งต้นเมื่อเทียบกับมุมมองของมาร์คเกอร์จากฉากจริงซึ่งต้องการจุดที่เป็นตำแหน่งเดียวกันระหว่างสองรูปอย่างน้อย 4 จุด เพื่อเชื่อมต่อมุมมองระหว่างสองรูปได้ ซึ่งสามารถใช้ผลลัพธ์จากความจับคู่ 4 จุดคุณลักษณะเด่นในขั้นตอนที่แล้วได้ โดยจากการจับคู่จุดคุณลักษณะเด่นนั้นสามารถนำลักษณะเด่นที่จับคู่ได้ไปสร้างโฮโมกราฟฟีเมทริกซ์เพื่อใช้สำหรับประมาณตำแหน่งมุมของมาร์คเกอร์และตำแหน่งของโมเดลวัตถุสามมิติในโลกเสริมให้มีทิศทางไปในทางเดียวกันจากการคืนปรับเทียบทั้งสองรูปให้มีมุมมองภาพที่ตรงกับตามตัวอย่างรูปที่ 3.7 และสามารถเชื่อมต่อมุมมองเพื่อผสมผสานโลกเสริมลงไปบนฉากจริงได้ดังตัวอย่างในรูปที่ 3.8



รูปที่ 3.7 ผลลัพธ์จากการเชื่อมต่อมุมมองจากฉากจริงเทียบกับมุมมองต้นแบบ

รูปที่ 3.8 การเชื่อมต่อมุมมองของโลกเสริมกับฉากจริง