ระบบช่วยเหลือคนตาบอดผ่านวีดีโอคอลในโทรศัพท์ร่วมกับการประกอบภาพทางไกล

นางสาวณัฐณิชา มณีแสง

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

บทคัดย่อและแฟ้มข้อมูลฉบับเต็มของวิทยานิพนธ์ตั้งแต่ปีการศึกษา 2554 ที่ให้บริการในคลังปัญญาจุฬาฯ (CUIR)
เป็นแฟ้มข้อมูลของนิสิตเจ้าของวิทยานิพนธ์ ที่ส่งผ่านทางบัณฑิตวิทยาลัย
The abstract and full text of theses from the academic year 2011 in Chulalongkorn University Intellectual Repository (CUIR)
are the thesis authors' files submitted through the University Graduate School.

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต
สาขาวิชาวิศวกรรมคอมพิวเตอร์ ภาควิชาวิศวกรรมคอมพิวเตอร์
คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2557
ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

Mobile video-based Tele-assistance system with remote scene construction for the blinds

Miss Nutnicha Maneesaeng

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Engineering Program in Computer Engineering
Department of Computer Engineering
Faculty of Engineering
Chulalongkorn University
Academic Year 2014

Copyright of Chulalongkorn University

หัวข้อวิทยานิพนธ์	ระบบช่วยเหลือคนตาบอดผ่านวีดีโอคอลในโทรศัพท์ร่วมกับการประกอบภาพทางไกล
โดย	นางสาวณัฐณิชา มณีแสง
สาขาวิชา	วิศวกรรมคอมพิวเตอร์
อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก	รองศาสตราจารย์ ดร.อติวงศ์ สุชาโต
อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม	ผู้ช่วยศาสตราจารย์ ดร.โปรดปราน บุณยพุกกณะ

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้นับวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญามหาบัณฑิต

.....คณบดีคณะวิศวกรรมศาสตร์

(ศาสตราจารย์ ดร. บัณฑิต เอื้ออาภรณ์)

คณะกรรมการสอบวิทยานิพนธ์

.....ประธานกรรมการ

(ผู้ช่วยศาสตราจารย์ ดร. ธนารัตน์ ชลิดาพงศ์)

.....อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก

(รองศาสตราจารย์ ดร.อติวงศ์ สุชาโต)

.....อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม

(ผู้ช่วยศาสตราจารย์ ดร.โปรดปราน บุณยพุกกณะ)

.....กรรมการภายนอกมหาวิทยาลัย

(ดร. ณัฐนันท์ ทัดพิทักษ์กุล)

ณัฐณิชา มณีแสง : ระบบช่วยเหลือคนตาบอดผ่านวีดีโอคอลในโทรศัพท์ร่วมกับการประกอบภาพทางไกล (Mobile video-based Tele-assistance system with remote scene construction for the blinds) อ.ที่ปรึกษาวิทยานิพนธ์หลัก: รศ. ดร.อติวงศ์ สุชาโต, อ.ที่ปรึกษาวิทยานิพนธ์ร่วม: ผศ. ดร.โปรดปราน บุณยพุกกณะ, 71 หน้า.

โปรแกรมวิดีโอคอลบนโทรศัพท์มือถือเป็นช่องทางสำคัญในการสื่อสารสำหรับคนตาบอดในสถานการณ์ที่คนตาบอดอยู่ตามลำพังและจำเป็นต้องอาศัยความช่วยเหลือจากผู้ช่วย แต่อย่างไรก็ตาม โปรแกรมที่มีอยู่ในปัจจุบันไม่สามารถใช้ช่วยเหลือคนตาบอดได้ ทั้งนี้เกิดจากหลายสาเหตุตัวอย่างเช่น ปัญหาด้านมุมมองภาพที่แสดงเพียงภาพวิดีโอปัจจุบันที่ได้รับจากคนตาบอด ซึ่งผู้ช่วยเหลือจำเป็นต้องบอกคนตาบอดให้หมุนกล้องไปมา จนกว่าจะถึงตำแหน่งที่ถูกต้อง ซึ่งทำได้ยากเนื่องจากคนตาบอดไม่สามารถมองเห็นสิ่งต่าง ๆ ที่กำลังถ่ายอยู่ได้ ดังนั้นในวิทยานิพนธ์นี้จึงให้ความสนใจไปที่การนำวิดีโอที่ได้รับจากคนตาบอดมาสร้างเป็นภาพขนาดใหญ่โดยใช้เทคนิคการต่อภาพ เพื่อให้ผู้ช่วยเหลือสามารถใช้งานช่วยเหลือคนตาบอดผ่านโปรแกรมวิดีโอคอลได้อย่างมีประสิทธิภาพ

อัลกอริทึมของการสร้างภาพทางไกลที่นำเสนอในวิทยานิพนธ์นี้รวมอยู่ในระบบต่อแสงสุข ซึ่งประกอบด้วย 3 ส่วน ส่วนแรก คือ โปรแกรมสำหรับคนตาบอด โดยเป็นโปรแกรมวิดีโอคอลบนโทรศัพท์ที่ใช้ระบบปฏิบัติการแอนดรอยด์ ซึ่งทำหน้าที่ในการถ่ายและส่งวิดีโอจากคนตาบอดไปยังผู้ช่วยเหลือ ที่ออกแบบโดยอ้างอิงจาก WCAG ส่วนที่สองคือส่วนการให้บริการการสื่อสารระหว่างผู้ใช้ ซึ่งเป็นการสื่อสารผ่าน WebRTC และส่วนสุดท้ายคือโปรแกรมสำหรับผู้ช่วยเหลือ ซึ่งรวมการทำงานของอัลกอริทึมของการสร้างภาพทางไกลที่มีพื้นฐานบนอัลกอริทึม SIFT โดยการประมวลผลนี้ได้นำการประมวลผลแบบเทรดมาใช้สำหรับการประมวลผลในกรณีภาพไม่ต่อเนื่อง โดยรองรับการทำงานสูงสุด 5 เทรด นอกจากนี้ผู้ใช้ยังสามารถขยายภาพที่ตำแหน่งต่าง ๆ และใช้โทนสีในการบอกตำแหน่งปัจจุบันแก่ผู้ช่วยเหลือด้วยเช่นกัน

การทดสอบการทำงานของระบบต่อแสงสุข ทำโดยผู้ทดลองที่เป็นคนตาบอดจำนวน 1 คน และผู้ช่วยเหลือที่เป็นบุคคลทั่วไปจำนวน 2 คน กับกิจกรรมจำนวน 2 กิจกรรม โดยเป็นการใช้งานระบบต่อแสงสุขที่รวมการทำงานของการสร้างภาพทางไกล และการทำงานที่ไม่รวมการสร้างภาพทางไกลเพื่อทำการเปรียบเทียบ ซึ่งการทดลองออกแบบให้สลับลำดับการใช้งานเพื่อป้องกันการเรียนรู้ของผู้ทดลอง โดยสุดท้ายจากการวัดผลการทดลองในรูปแบบของเวลาที่ใช้ในแต่ละกิจกรรมพบว่า การใช้งานระบบต่อแสงสุขร่วมกับการสร้างภาพทางไกลทำให้การช่วยเหลือเร็วขึ้นประมาณ 31% หรือ 1.75 นาที

ภาควิชา	วิศวกรรมคอมพิวเตอร์	ลายมือชื่อนิสิต .....
สาขาวิชา	วิศวกรรมคอมพิวเตอร์	ลายมือชื่อ อ.ที่ปรึกษาหลัก .....
ปีการศึกษา	2557	ลายมือชื่อ อ.ที่ปรึกษาร่วม .....

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

5770411821 : MAJOR COMPUTER ENGINEERING
KEYWORDS: ASSISTIVE TECHNOLOGY / IMAGE PROCESSING

NUTNICHA MANEESAENG: Mobile video-based Tele-assistance system with remote scene construction for the blinds. ADVISOR: ASSOC. PROF. ATIWONG SUCHATO, Ph.D., CO-ADVISOR: ASST. PROF. PROADPRAN PUNYABUKKANA, Ph.D., 71 pp.

For people with visual disabilities, the use of video-call application on mobile devices can be an essential means to remotely communicate with their assistants whenever tasks require help from sighted person. However, such applications on the market were not readily usable by blind users. One major drawback is that the assistant can only view current video stream. When the blind fails to take necessary scenes, the assistant must direct the blind to re-shoot the video until required task can be accomplished. Apparently, it is highly challenging for the blinds to capture the videos at the right spot. Thereby, this thesis focuses on constructing large scene by stitching images real-time from video streams taken by the users.

A new system, “Tor Sang Suk”, was developed to include our proposed algorithm. It comprises three components. The first component, for the blind users, is a WCAG complied video-call application on Android that simply captures and sends videos. The second component is the communication between Android device the PC, enabled through WebRTC. The third component for the assistants contains our proposed remote scene construction algorithm based on SIFT (Scale-Invariant Feature Transform) that allows up to five threads of disconnected video streams. It also features real-time zooming function and graying of older portion of images.

We evaluated Tor Sang Suk by one blind user, two assistants, and two scenarios. For each of the scenarios, the blind user performed the same task twice per one assistant, once using Tor Sang Suk with remote scene construction feature, and another without. For the two times the tasks were performed, it was done with different assistant to avoid prior knowledge from the assistant. We measured the time to complete each task and found that Tor Sang Suk with remote scene construction showed 31% time reduction, or an average of 1.75 minutes.

Department:	Computer Engineering	Student's Signature	.....
Field of Study:	Computer Engineering	Advisor's Signature	.....
Academic Year:	2014	Co-Advisor's Signature	.....

กิตติกรรมประกาศ

ข้าพเจ้าขอขอบพระคุณอาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก รศ. ดร.อติวงศ์ สุชาโต และอาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม ผศ. ดร.โปรดปราน บุณยพุกกณะ ที่ให้ความช่วยเหลือ ความรู้ ข้อคิด และคำแนะนำที่เป็นประโยชน์ต่อการทำวิทยานิพนธ์นี้ให้สำเร็จลุล่วงไปด้วยดี ขอขอบคุณ ผศ. ดร.ธนารัตน์ ชลิดาพงศ์ และ ดร.ณัฐนันท์ ทัดพิทักษ์กุล ที่ให้คำแนะนำ และแนวคิดที่เป็นประโยชน์ต่อการทำวิทยานิพนธ์นี้

ขอขอบคุณ คุณธมนิฏฐ์ ชาญจรัสพงศ์ สำหรับปัญหาของคนตาบอดซึ่งเป็นแนวคิดเริ่มต้นของวิทยานิพนธ์ และ น.ส.สโรชา กิตติสิริพันธุ์ สำหรับความร่วมมือในการทดสอบต่าง ๆ ตลอดการทำวิทยานิพนธ์ อีกทั้งขอบคุณเพื่อนที่ร่วมโครงการต่อแสงสุข และเพื่อนนิสิต หลักสูตรวิศวกรรมศาสตร์มหาบัณฑิต ที่ช่วยให้คำแนะนำในการพัฒนาโปรแกรม

นอกจากนี้ข้าพเจ้าขอขอบคุณ พี่ ๆ สมาชิกห้องปฏิบัติการระบบภาษาพูด ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ที่ให้กำลังใจ และความช่วยเหลือจนกระทั่งวิทยานิพนธ์นี้สำเร็จ

ขอขอบพระคุณครอบครัวของข้าพเจ้า ที่เข้าใจ มอบกำลังใจและให้ความสนับสนุนแก่ข้าพเจ้าเสมอมา

ท้ายที่สุดนี้ ขอขอบพระคุณภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ที่ให้ทุนการศึกษาอัจฉริยะคืนรัง รวมถึงบัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย ที่ให้ทุนสนับสนุนการนำเสนอผลงานวิจัยในต่างประเทศแก่ข้าพเจ้า

สารบัญ

หน้า

บทคัดย่อภาษาไทย ..... ง
บทคัดย่อภาษาอังกฤษ ..... จ
กิตติกรรมประกาศ ..... ฉ
สารบัญ ..... ช
สารบัญภาพ ..... ฌ
สารบัญตาราง ..... ฎ
บทที่ 1 บทนำ 1
ที่มาและความสำคัญของปัญหา 1
วัตถุประสงค์ของการวิจัย 4
ขอบเขตการวิจัย 4
ขั้นตอนการดำเนินงาน 5
ประโยชน์ที่คาดว่าจะได้รับ 6
ผลงานตีพิมพ์ 6
โครงสร้างวิทยานิพนธ์ 7
บทที่ 2 ทฤษฎี และงานวิจัยที่เกี่ยวข้อง 8
ทฤษฎีที่เกี่ยวข้อง 8
1. การสื่อสารในเวลาจริง (Real-Time Communication) 8
2. อัลกอริทึมการต่อภาพ (Image Stitching Algorithm) 10
3. แนวทางการเข้าถึงเนื้อหาของเว็บ (Web Content Accessibility Guidelines) 19
งานวิจัยที่เกี่ยวข้อง 20
1. ระบบช่วยเหลือทางไกล (Tele-Assistance System) 20
2. การประยุกต์ใช้ WebRTC ในระบบการสื่อสารต่าง ๆ 22

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

หน้า

3. อัลกอริทึมที่ใช้ในการต่อภาพ 23
บทที่ 3 วิธีดำเนินการวิจัย 24
เครื่องมือที่ใช้ในการวิจัย 24
ขั้นตอนการดำเนินงานวิจัย 24
1. การทดสอบการใช้งานโปรแกรมวิดีโอคอลที่มีอยู่ในปัจจุบันในระบบช่วยเหลือทางไกล 25
2. การออกแบบระบบและรูปแบบการทำงาน 26
3. การแก้ไขปัญหาเรื่องส่วนต่อประสานผู้ใช้สำหรับคนตาบอด 28
4. การพัฒนาส่วนการประกอบภาพทางไกล 31
5. การพัฒนาต้นแบบของโปรแกรมวิดีโอคอลที่ใช้งานในระบบช่วยเหลือคนตาบอด 39
บทที่ 4 ผลการวิจัย 46
ผลการวิจัย 46
1. ผลการวิเคราะห์ปัญหาในการสื่อสารผ่านโปรแกรมวิดีโอคอลที่มีอยู่ในปัจจุบันในระบบช่วยเหลือทางไกลระหว่างคนตาบอดและผู้ช่วยเหลือ 46
2. ผลการเสนอวิธีการจัดการรูปภาพที่ได้จากวิดีโอคอล เพื่ออำนวยความสะดวกให้แก่ผู้ช่วยเหลือ ในการใช้งานในระบบช่วยเหลือทางไกล 52
3. ผลการพัฒนาตัวอย่างโปรแกรมวิดีโอคอลเพื่อใช้งานในระบบช่วยเหลือคนตาบอดโดยบุคคลทั่วไป 55
บทที่ 5 สรุปผล 62
สรุปผลการวิจัย 62
ข้อเสนอแนะ 65
รายการอ้างอิง 66
ประวัติผู้เขียนวิทยานิพนธ์ 71

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

สารบัญภาพ

รูปที่ 1.1 ตัวอย่างการนำทางโดยใช้ RFID 2
รูปที่ 2.1 การให้บริการต่าง ๆ ของ WebRTC 9
รูปที่ 2.2 สถาปัตยกรรมของ RTCPeerConnection 9
รูปที่ 2.3 ขั้นตอนในการทำงานของอัลกอริทึม SIFT 11
รูปที่ 2.4 ตัวอย่างปริภูมิในมิติของขนาดและระยะทางในอัลกอริทึม SURF 14
รูปที่ 2.5 ตัวอย่างการกำหนดทิศทางของจุดคุณลักษณะด้วยอัลกอริทึม SURF 15
รูปที่ 2.6 ตัวอย่างการเปรียบเทียบของอัลกอริทึม SURF 17
รูปที่ 2.7 ตัวอย่างการปรับตำแหน่งโดยใช้ Homography matrix 18
รูปที่ 2.8 ตัวอย่างหน้าจอของโอเปอเรเตอร์ในงานวิจัยของ P. Baranski 21
รูปที่ 2.9 แนวคิดของระบบช่วยเหลือทางไกลสำหรับคนตาบอดโดย P. Baranski 21
รูปที่ 3.1 ขั้นตอนการดำเนินงาน 24
รูปที่ 3.2 ภาพรวมของการใช้งานในการสื่อสารทางไกล 27
รูปที่ 3.3 ตัวอย่างการอ้างอิงตำแหน่งภาพส่วนล่าสุดโดยใช้ตัวชี้ (pointer) 31
รูปที่ 3.4 ตัวอย่างภาพมุมกว้างที่แสดงแก่ผู้ช่วยเหลือ 32
รูปที่ 3.5 ขั้นตอนในการทำงานส่วนการประกอบภาพทางไกล 34
รูปที่ 3.6 ขั้นตอนในการทำงานของมอดูลของส่วนการทำงานต่อรูปภาพ 35
รูปที่ 3.7 ขั้นตอนในการทำงานของมอดูลการประมวลผลสำหรับภาพไม่ต่อเนื่อง 37
รูปที่ 3.8 แผนภาพแสดงการทำงานของผู้ใช้ระบบ 39
รูปที่ 3.9 ขั้นตอนการทำงานของส่วนการให้บริการสื่อสารระหว่างผู้ใช้ทั้งสองประเภท 42
รูปที่ 3.10 ขั้นตอนการแปลงค่าข้อมูลวิดีโอที่ได้รับ 45
รูปที่ 4.1 ตัวอย่างการต่อรูปภาพโดยใช้ชุดข้อมูลนำเข้าจากวิดีโอคอลร่วมกับอัลกอริทึม SURF และอัลกอริทึม SIFT 53

รูปที่ 4.2 หน้าจอแสดงผลของโปรแกรมวิดีโอคอลสำหรับคนตาบอด..... 56

รูปที่ 4.3 ตัวอย่างหน้าจอของผู้ช่วยเหลือขณะใช้งานวิดีโอคอลร่วมกับการประกอบภาพทางไกล 57

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

สารบัญตาราง

ตารางที่ 1.1 รายละเอียดและระยะเวลาในการดำเนินงาน 5

ตารางที่ 2.1 การเปรียบเทียบผลลัพธ์ในการต่อรูปภาพโดยใช้ SIFT และ SURF 23

ตารางที่ 3.1 ปัญหาและแนวทางแก้ไข 26

ตารางที่ 3.2 แสดงรายละเอียดการทำงานต่าง ๆ ในส่วนการแสดงภาพมุมกว้าง 32

ตารางที่ 3.3 รายละเอียดของการเลือกประเภทผู้ใช้งาน 40

ตารางที่ 3.4 รายละเอียดของการใช้งานวิดีโอคอล 40

ตารางที่ 3.5 รายละเอียดของการเชื่อมต่อใหม่ 40

ตารางที่ 3.6 รายละเอียดของการแสดงภาพมุมกว้าง 41

ตารางที่ 4.1 ตารางแสดงความหมายและหน่วยของตัวชี้วัดที่ใช้ในการทดลองเพื่อวิเคราะห์ส่วนต่อประสานผู้ใช้สำหรับคนตาบอด 47

ตารางที่ 4.2 ตารางแสดงค่า M1 และ M2 ที่ได้จากการทดลองใช้งานโดยคนตาบอด (การทดลอง ก.) 49

ตารางที่ 4.3 ตารางแสดงค่า M3 M4 และ M5 ที่ได้จากการทดลองใช้งานโดยผู้วิจัย (การทดลอง ข.) 49

ตารางที่ 4.4 ตารางแสดงหลักการการออกแบบส่วนต่อประสานผู้ใช้สำหรับโปรแกรมที่ใช้งานโดยคนตาบอดบนโทรศัพท์มือถือ 50

ตารางที่ 4.5 ตารางแสดงการเปรียบเทียบเวลาที่ใช้ในกิจกรรมต่าง ๆ (นาที) ของเวลาที่ดีที่สุดใน 6 โปรแกรมที่มีอยู่ในปัจจุบันและโปรแกรมที่ออกแบบตามหลักการการออกแบบที่นำเสนอ 51

ตารางที่ 4.6 สถิติของเวลาการประมวลผลโดยอัลกอริทึม SURF และ SIFT ในหน่วยวินาที 52

ตารางที่ 4.7 ตารางแสดงผลลัพธ์ของแบบสอบถามความพึงพอใจต่อการใช้งานวิดีโอคอลร่วมกับส่วนการแสดงภาพมุมกว้าง 54

ตารางที่ 4.8 เวลาที่ใช้ในการประมวลผลโดยอัลกอริทึมที่นำเสนอในหน่วยนาที 55

ตารางที่ 4.9 ตารางเปรียบเทียบแบนด์วิดท์ที่ใช้ในการส่งข้อมูลในรูปแบบต่าง ๆ ของโปรแกรมสำหรับผู้ช่วยเหลือ 58

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตารางที่ 4.10 ตารางแสดงรายละเอียดกิจกรรม ผู้ทดลอง และรูปแบบโปรแกรมที่ใช้งาน..... 59

ตารางที่ 4.11 ตารางแสดงผลการทดลองใช้งานในรูปแบบของเวลาที่ใช้ และจำนวนครั้งในการเคลื่อนกล้องกลับไปยังตำแหน่งเดิม..... 60

ตารางที่ 4.12 ผลการวิเคราะห์เวลาที่ใช้ในการทดลอง 61

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

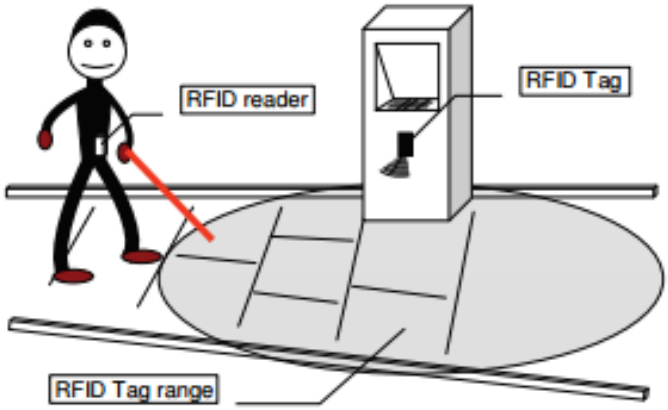
บทที่ 1
บทนำ

ที่มาและความสำคัญของปัญหา

ตามความหมายของแผนการจัดการศึกษาเฉพาะบุคคลกระทรวงศึกษาธิการ [1] คนตาบอดเป็นประเภทหนึ่งของบุคคลที่มีความบกพร่องทางการมองเห็น โดยหมายถึงบุคคลที่สูญเสียการมองเห็นมาก เมื่อตรวจวัดความชัดของสายตาข้างดีเมื่อแก้ไขแล้วอยู่ในระดับ 20/200 ฟุตลงมาจนถึงบอดสนิท สามารถมองเห็นได้น้อยกว่า 30 องศา นั่นคือ บุคคลปกติจะสามารถมองเห็นวัตถุใดวัตถุหนึ่งในระยะ 200 ฟุตได้อย่างชัดเจน แต่คนตาบอดจะเห็นวัตถุนั้นในระยะไม่ถึง 20 ฟุต ซึ่งข้อมูลจากสำนักงานส่งเสริมและพัฒนาคุณภาพชีวิตคนพิการแห่งชาติ [2] พบว่าในปี พ.ศ. 2558 มีจำนวนคนพิการทางการมองเห็นที่ที่มีบัตรประจำตัวคนพิการเป็นจำนวนประมาณ 180,000 คน ซึ่งคิดเป็น 12% ของคนพิการทางการมองเห็นทั้งหมดของประเทศไทย ตัวเลขเหล่านี้แสดงให้เห็นว่าสังคมในปัจจุบันประกอบไปด้วยคนตาบอดจำนวนมาก ซึ่งรูปแบบการใช้ชีวิตประจำวันของบุคคลเหล่านี้แตกต่างจากบุคคลทั่วไป ทั้งนี้เกิดจากปัจจัยสำคัญคือการที่คนตาบอดไม่สามารถมองเห็นสิ่งต่าง ๆ ที่อยู่รอบตัวได้ และหากมองไปที่การใช้ชีวิตประจำวันของคนตาบอด การเดินทางไปยังสถานที่ต่าง ๆ ของคนตาบอดมือ 4 วิธี [3] คือการเดินทางกับผู้นำทาง (Sighted Guide), การเดินทางกับสุนัขนำทาง (Dog Guide), การเดินทางโดยใช้ไม้เท้าขาว (White Cane) และการเดินทางโดยใช้เครื่องมืออิเล็กทรอนิกส์ (Electronic Aids) ทั้งนี้การเดินทางกับผู้นำทาง และการเดินทางโดยใช้ไม้เท้าขาวเป็นวิธีที่เหมาะสมกับสภาพเศรษฐกิจ และสภาพสังคมของประเทศไทยมากที่สุด [4]

การเดินทางโดยใช้เครื่องมืออิเล็กทรอนิกส์ (Electronic Aids) อยู่ในรูปแบบของการนำทางคนตาบอดในสถานที่ต่าง ๆ โดยสามารถแบ่งออกได้เป็นการนำทางในสถานที่ภายในอาคารและสถานที่ภายนอกอาคาร โดยการนำทางภายในอาคารเป็นการนำทางโดยใช้โครงสร้างพื้นฐานที่เป็นระบบฝังตัว (Embedded infrastructure) ซึ่งเป็นการฝังอุปกรณ์อิเล็กทรอนิกส์ไว้กับวัตถุที่ต้องการ เช่น ป้ายบอกทาง หรือทางเข้า พร้อมกันนี้แต่ละวัตถุจะมีรหัสประจำตัวซึ่งสามารถประมวลผลได้จากตัวอ่านสัญญาณ สำหรับในยุคแรกนั้นใช้หุ่นยนต์ในการรับสัญญาณและแปลค่า [5] ซึ่งในภายหลังมีการพัฒนาตัวอ่าน RFID ให้เป็นอุปกรณ์ที่สามารถพกพาได้โดยคนตาบอด [6] ตามแสดงในรูปที่ 1.1

นอกจากนั้นยังมีวิธีการอื่นอีก เช่น ใช้สัญญาณไวไฟ (Wi-Fi) ภายในตัวอาคาร [7] หรือเทคนิคการประมวลผลรูปภาพที่ถ่ายจากกล้องโทรศัพท์มือถือ [8] ซึ่งข้อจำกัดของงานวิจัยประเภทนี้คือราคาของอุปกรณ์ต่าง ๆ มีราคาค่อนข้างสูง และมีรูปแบบของการใช้งานที่ถูกออกแบบมาสำหรับใช้ในพื้นที่หนึ่ง ๆ เท่านั้น



รูปที่ 1.1 ตัวอย่างการนำทางโดยใช้ RFID คัดลอกจาก [9]

การเดินทางโดยใช้เครื่องมืออิเล็กทรอนิกส์ในสถานที่ภายนอกอาคารส่วนมากมักใช้ร่วมกับอุปกรณ์โทรศัพท์มือถือ และมีการออกแบบให้เจาะจงกับรูปแบบการใช้งานหนึ่ง ๆ ยกตัวอย่างเช่น ใช้ GPS ในการบอกตำแหน่งปัจจุบันของรถประจำทาง เพื่อช่วยอำนวยความสะดวกแก่คนตาบอดในการโดยสารรถประจำทาง [9] หรือใช้หลักการของ Crowdsourcing ในการให้ข้อมูลเกี่ยวกับป้ายรถสาธารณะแก่คนตาบอด [10] ซึ่งเห็นได้ว่าข้อจำกัดของงานวิจัยประเภทนี้คือสามารถใช้งานได้เฉพาะรูปแบบที่ถูกออกแบบไว้เท่านั้น หรืออาจใช้งานไม่ได้ในสถานที่ซึ่งไม่มีสัญญาณ GPS นอกจากนี้การเดินทางโดยใช้เครื่องมืออิเล็กทรอนิกส์ของคนตาบอด ยังอยู่ในรูปแบบของการช่วยให้คนตาบอดเดินหลบสิ่งกีดขวางต่าง ๆ ได้ โดยใช้การวัดระยะห่างของผู้ใช้กับวัตถุด้วยระยะโฟกัสของกล้อง [11] หรือการใช้สัญญาณอัลตราโซนิคเพื่อตรวจจับสิ่งกีดขวางต่าง ๆ และแจ้งไปยังคนตาบอดด้วยเสียงหรือการสั่น [12] แม้ว่าการใช้งานไม่เจาะจงในสถานที่ใดสถานที่หนึ่ง แต่เป็นการประมวลผลซึ่งไม่สามารถบอกรายละเอียดของสิ่งต่าง ๆ ที่อยู่รอบตัวคนตาบอด หรือสิ่งที่คนตาบอดต้องการทราบได้

เทคโนโลยีช่วยเหลือ (Assistive technology) ยกตัวอย่างเช่น อักษรเบรลล์, ภาพนูน, หนังสือเสียงระบบ DAISY, เครื่องจดบันทึกอักษรเบรลล์ หรือเครื่องแสดงผลอักษรเบรลล์ (Braille Display) [13] เป็นอุปกรณ์ที่ออกแบบมาโดยเจาะจงไปในการช่วยเหลือให้ผู้พิการสามารถมีชีวิตอิสระและสามารถดำเนินกิจกรรมต่าง ๆ ในชีวิตประจำวันที่อาจยากเมื่อทำตามลำพังได้ [14] โดยเทคโนโลยีอ่านหน้าจอ (Screen reader) [15] เป็นหนึ่งในเทคโนโลยีช่วยเหลือที่ทำหน้าที่ในการอ่านสิ่งต่าง ๆ ที่ปรากฏอยู่บนหน้าจอออกมาเป็นเสียง โดยใช้หลักการของการสังเคราะห์เสียงจากข้อความ (text-to-speech) เพื่อช่วยเหลือการใช้งานของคนตาบอด ให้สามารถโต้ตอบกับอุปกรณ์ได้ง่ายขึ้น เช่น โปรแกรมเปล่งเสียงสังเคราะห์ (ตาทิพย์ [16] และ Jaws for windows [17]), Apple VoiceOver [18], Google TalkBack [19]

ระบบช่วยเหลือทางไกล (Tele-assistance system) เป็นแนวคิดของการช่วยเหลือคนตาบอดให้สามารถเดินทางไปยังสถานที่ต่าง ๆ หรือทำกิจกรรมต่าง ๆ ตามลำพังได้ โดยอาศัยการสื่อสารกับผู้ช่วยเหลือส่วนตัวที่อยู่ทางไกล แนวคิดนี้เกิดขึ้นในปี ค.ศ. 2003 [20] ที่มหาวิทยาลัย Brunel และถูกใช้เป็นแนวคิดต้นแบบสำหรับงานวิจัยต่าง ๆ [21], [22] ที่พัฒนาระบบช่วยเหลือนี้ขึ้นเพื่ออำนวยความสะดวกให้กับคนตาบอด ซึ่งอยู่ในรูปแบบของการติดต่อกับผู้ช่วยเหลือส่วนตัวที่ใช้งานบนเครื่องคอมพิวเตอร์ แต่การใช้งานระบบช่วยเหลือทางไกลเหล่านี้ ยังมีข้อจำกัดในเรื่องของอุปกรณ์ที่ผู้ใช้งานซึ่งเป็นคนตาบอดต้องจัดหาเป็นพิเศษ เมื่อพิจารณาร่วมกับความก้าวหน้าของเทคโนโลยีในปัจจุบัน การสื่อสารผ่านโปรแกรมวิดีโอคอลเป็นทางเลือกหนึ่งที่เหมาะสมที่สำหรับการนำมาประยุกต์ใช้งาน เนื่องจากมีการส่งข้อมูลภาพวิดีโอร่วมกับเสียงแบบทันกาล ซึ่งเป็นปัจจัยสำคัญสำหรับการสื่อสารกับผู้ช่วยเหลือที่ไม่ได้อยู่ร่วมกับคนตาบอดในสถานที่นั้น ๆ

การใช้งานโปรแกรมวิดีโอคอลเป็นช่องทางในการสื่อสารระหว่างคนตาบอดกับผู้ช่วยเหลือในระบบช่วยเหลือทางไกล ยกตัวอย่างเช่นโปรแกรม BeMyEyes [23] มีจุดประสงค์เพื่อให้ผู้ช่วยเหลือที่ทำหน้าที่เปรียบเสมือนดวงตาแก่คนตาบอดมองสิ่งต่าง ๆ ที่อยู่รอบตัวคนตาบอด อธิบายและโต้ตอบกับคนตาบอดได้อย่างทันที อย่างไรก็ตามมุมมองภาพที่เห็นโดยทั่วไปเป็นเพียงมุมมองปัจจุบันที่ถูกถ่ายจากกล้องของคนตาบอดเท่านั้น ซึ่งในบางสถานการณ์ภาพเหล่านั้นอาจไม่เพียงพอสำหรับการช่วยเหลือ ที่ควรได้ข้อมูลที่มากกว่าภาพปัจจุบันเพียงอย่างเดียว ตัวอย่างเช่นการช่วยคนตาบอดหาสิ่งของต่าง ๆ หรือ การบรรยายสภาพแวดล้อมต่าง ๆ รอบตัวคนตาบอด ซึ่งในการส่งภาพวิดีโอ ภาพ

วิดีโอที่ตำแหน่งนั้นจะถูกแสดงแก่ผู้ช่วยเหลือในระยะเวลาหนึ่ง ๆ และเปลี่ยนไปเมื่อคนตาบอดเปลี่ยนมุมมองของกล้องไปยังตำแหน่งอื่น

ดังนั้นในงานวิจัยนี้จึงนำเสนอระบบช่วยเหลือคนตาบอดที่ใช้งานวิดีโอคอลร่วมกับการประกอบภาพทางไกล เพื่อใช้งานในการช่วยเหลือคนตาบอดให้ไปยังสถานที่ต่าง ๆ หรือดำเนินกิจกรรมต่าง ๆ ตามลำพังได้ โดยเห็นถึงความสำคัญของการจัดการภาพจากวิดีโอที่ถูกส่งจากคนตาบอดมายังผู้ช่วยเหลือให้สามารถนำมาใช้ประโยชน์ได้มากขึ้น โดยนอกจากจะเป็นการอำนวยความสะดวกในการใช้งานสำหรับผู้ช่วยเหลือแล้ว ยังสามารถลดการเคลื่อนที่ไปยังตำแหน่งต่าง ๆ ของคนตาบอดได้ด้วย นอกจากนี้ยังมีการเพิ่มเติมสำหรับส่วนต่อประสานผู้ใช้ที่รองรับการใช้งานบนโทรศัพท์มือถือของคนตาบอดด้วยเช่นกัน

วัตถุประสงค์ของการวิจัย

1. ศึกษาปัญหาในการสื่อสารผ่านโปรแกรมวิดีโอคอลที่มีอยู่ในปัจจุบันในระบบช่วยเหลือทางไกลระหว่างคนตาบอดและผู้ช่วยเหลือ
2. เสนอวิธีการจัดการรูปภาพที่ได้จากวิดีโอคอล เพื่ออำนวยความสะดวกให้แก่ผู้ช่วยเหลือ ในการใช้งานในระบบช่วยเหลือทางไกล
3. พัฒนาโปรแกรมวิดีโอคอลที่สามารถใช้งานในระบบช่วยเหลือทางไกลสำหรับคนตาบอดโดยบุคคลทั่วไป

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ขอบเขตการวิจัย

1. การจัดการรูปภาพที่ได้จากการสื่อสารผ่านวิดีโอคอลที่นำเสนอ จะใช้ในกรณีคนตาบอดอยู่นิ่งอยู่กับที่และหมุนกล้องไปรอบตัวเท่านั้น
2. โปรแกรมวิดีโอคอลที่พัฒนาขึ้นเพื่อใช้งานร่วมกับเทคนิคการประกอบภาพทางไกลอยู่ในรูปแบบของโปรแกรมต้นแบบเท่านั้น
3. โปรแกรมวิดีโอคอลที่ใช้งานโดยคนตาบอดจะใช้งานผ่าน Google Chrome บนโทรศัพท์มือถือที่ใช้ระบบปฏิบัติการแอนดรอยด์ 4.0 ขึ้นไปเท่านั้น
4. โปรแกรมวิดีโอคอลต้นแบบที่ใช้งานโดยผู้ช่วยเหลือ จะใช้งานผ่าน Google Chrome บนเครื่องคอมพิวเตอร์เท่านั้น

ขั้นตอนการดำเนินงาน

1. รวบรวมความรู้ที่เกี่ยวข้อง
2. รวบรวมความต้องการของระบบ
3. ออกแบบระบบและรูปแบบการทำงานของโปรแกรม
4. พัฒนาโปรแกรม
5. ทดสอบการทำงานของโปรแกรม
6. สรุปผลการวิจัย

กิจกรรม	10-57	11-57	12-57	1-58	2-58	3-58	4-58	5-58	6-58	7-58
รวบรวมความรู้ที่เกี่ยวข้อง										
• ทบทวนวรรณกรรมที่เกี่ยวข้อง										
• กำหนดขอบเขตและวัตถุประสงค์ของงานวิจัย										
ค้นหาปัญหา										
• ออกแบบการทดลองเพื่อค้นปัญหาในการใช้งาน										
• ทำการทดลองเพื่อค้นปัญหาในการใช้งาน										
ออกแบบระบบและรูปแบบการทำงานของโปรแกรม										
• สรุปผลการทดลอง และกำหนดการทำงานพื้นฐานของโปรแกรม										
• จัดทำโครงร่างวิจัย										
• ออกแบบการทดสอบการใช้งานโปรแกรมวีดีโอคอลของคนตาบอด										
• ทดสอบการใช้งานโปรแกรมวีดีโอคอลของคนตาบอด										
• วิเคราะห์ปัญหา และสรุปเป็นหลักการสำหรับการออกแบบ										
• ทบทวนวรรณกรรมที่เกี่ยวข้องเพิ่มเติม										
• กำหนดรูปแบบการจัดการข้อมูลจากวีดีโอคอลสำหรับผู้ช่วยเหลือ										
พัฒนาโปรแกรม										
• ศึกษาเนื้อหาในเชิงเทคนิค										
• พัฒนาส่วนวีดีโอคอล										
• พัฒนาส่วนการนำเสนอข้อมูลจากวีดีโอคอล										
• พัฒนาส่วนการแปลงค่าข้อมูลที่ได้จากวีดีโอคอล										
• รวมการทำงานส่วนวีดีโอคอลและส่วนการนำเสนอข้อมูลเข้าด้วยกัน										
• สอบโครงร่างวิจัย										
ทดสอบการทำงานของโปรแกรม										
• ออกแบบการทดสอบการทำงานของโปรแกรม										
• ทดสอบการทำงานของโปรแกรม										
• แก้ไขปัญหาและพัฒนาโปรแกรมเพิ่มเติม										
สรุปผลการวิจัย										
• ออกแบบการทดลองเพื่อสรุปผล										
• ทำการทดลองเพื่อสรุปผล										
• สรุปผลวิจัย										
• นำเสนอต่อคณะกรรมการ										
• จัดทำรูปเล่มฉบับสมบูรณ์										

ตารางที่ 1.1 รายละเอียดและระยะเวลาในการดำเนินงาน

ประโยชน์ที่คาดว่าจะได้รับ

1. ทราบถึงปัญหาในการสื่อสารผ่านโปรแกรมวิดีโอคอลในระบบช่วยเหลือทางไกลระหว่างคนตาบอดและผู้ช่วยเหลือ
2. มีวิธีในการจัดการข้อมูลที่ได้จากการสื่อสารผ่านวิดีโอคอล เพื่ออำนวยความสะดวกให้แก่ผู้ช่วยเหลือ ในการใช้งานในระบบช่วยเหลือทางไกล
3. มีตัวอย่างโปรแกรมวิดีโอคอลเพื่อใช้งานในการช่วยเหลือคนตาบอดโดยบุคคลทั่วไป ที่สามารถแก้ไขปัญหาที่เกิดขึ้นได้
4. เป็นแนวทางสำหรับงานวิจัยในการพัฒนาโปรแกรมเพื่อช่วยเหลือคนตาบอดในอนาคต

ผลงานตีพิมพ์

งานวิทยานิพนธ์นี้ถูกแบ่งเป็น 2 ส่วนเพื่อตีพิมพ์เป็นบทความทางวิชาการ ส่วนแรกเป็นเนื้อหาของการศึกษาส่วนต่อประสานผู้ใช้ของโปรแกรมวิดีโอคอลบนโทรศัพท์มือถือ โดยมีหัวข้อเรื่องคือ “Acessible Video-Call Application on Android for the Blind” จัดทำโดย “Nutnicha Maneesaeng, Proadpran Punyabukkana, Atiwong Suchato” ถูกนำเสนอในงานประชุมวิชาการ “2015 4th International Conference on Humanity, History and Society: ICHHS 2015” ที่จัดขึ้นในวันที่ 8-9 มีนาคม 2558 ณ กรุงโซล สาธารณรัฐเกาหลี และตีพิมพ์ลงใน Lecture Notes on Software Engineering (LNSE) Vol. 4, No. 2

ส่วนที่ 2 เป็นเนื้อหาของการนำเทคนิคประมวลผลภาพมาใช้ในการนำเสนอรูปภาพที่ได้จากวิดีโอคอล โดยมีหัวข้อเรื่องคือ “Tele-Assistance System for the Blinds using Video-Call with Remote Scene Construction” ถูกนำเสนอในงานประชุมวิชาการ “The Forth International Conference on Informatics, Electronics & Vision: ICIEV 2015” ที่จัดขึ้นในวันที่ 15-18 มิถุนายน 2558 ณ เมืองฟุกุโอกะ ประเทศญี่ปุ่น

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

โครงสร้างวิทยานิพนธ์

เนื้อหาในวิทยานิพนธ์ฉบับนี้แบ่งเนื้อหาในการนำเสนอออกเป็น 5 ส่วน โดยในบทที่ 2 กล่าวถึงทฤษฎีและงานวิจัยที่เกี่ยวข้องซึ่งประกอบด้วย งานวิจัยที่เกี่ยวข้องกับระบบช่วยเหลือทางไกล (Tele-Assistance System), ทฤษฎีการสื่อสารในเวลาจริง (Real-Time Communication) และการประยุกต์ใช้ในระบบสื่อสารต่าง ๆ และอัลกอริทึมในการต่อภาพ (Image Stitching Algorithm)

บทที่ 3 เป็นวิธีดำเนินการวิจัย โดยจะกล่าวถึงเครื่องมือที่ใช้ในการวิจัย และการดำเนินงานวิจัยในขั้นตอนต่าง ๆ ได้แก่ การทดสอบการใช้งานโปรแกรมวิดีโอคอลที่มีอยู่ในปัจจุบันในระบบช่วยเหลือทางไกล การออกแบบระบบและรูปแบบการทำงาน การวิเคราะห์ส่วนต่อประสานผู้ใช้สำหรับคนตาบอด และการพัฒนาโปรแกรม บทที่ 4 อธิบายผลการวิจัยโดยวิเคราะห์ตามวัตถุประสงค์ของการวิจัย และสุดท้ายบทที่ 5 เป็นสรุปผลการวิจัย และข้อเสนอแนะ

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

บทที่ 2
ทฤษฎี และงานวิจัยที่เกี่ยวข้อง

ทฤษฎีที่เกี่ยวข้อง

ทฤษฎีที่เกี่ยวข้องกับงานวิจัยแบ่งออกเป็นสามส่วน ส่วนแรกคือทฤษฎีการสื่อสารในเวลาจริง (Real-Time Communication) ซึ่งเกี่ยวข้องกับการทำวิดีโอคอล ส่วนที่สองคืออัลกอริทึมในการต่อภาพ (Image Stitching Algorithm) ซึ่งเกี่ยวข้องกับการนำเสนอรูปภาพจากวิดีโอคอลแก่ผู้ช่วยเหลือ และส่วนที่สามคือหลักการการออกแบบเว็บไซต์เพื่อให้เข้าถึงได้ (Web Content Accessibility Guidelines) ซึ่งนำมาอ้างอิงสำหรับการออกแบบส่วนต่อประสานผู้ใช้แก่คนตาบอด

1. การสื่อสารในเวลาจริง (Real-Time Communication)

การสื่อสารในเวลาจริง (Real-Time Communications หรือ RTC) เป็นการสื่อสารระหว่างบุคคลซึ่งสามารถโต้ตอบกันได้ทันทีผ่านเครือข่ายต่าง ๆ ของการสื่อสาร โดยสามารถสื่อสารด้วยข้อความ ภาพเคลื่อนไหว หรือเสียง เช่นการแชท (Chat) หรือ การสื่อสารทางเสียงผ่านโครงข่ายอินเทอร์เน็ต (Voice over Internet Protocol หรือ VoIP)

1.1 การสื่อสารในเวลาจริงผ่านเว็บ (Web Real-Time Communication)

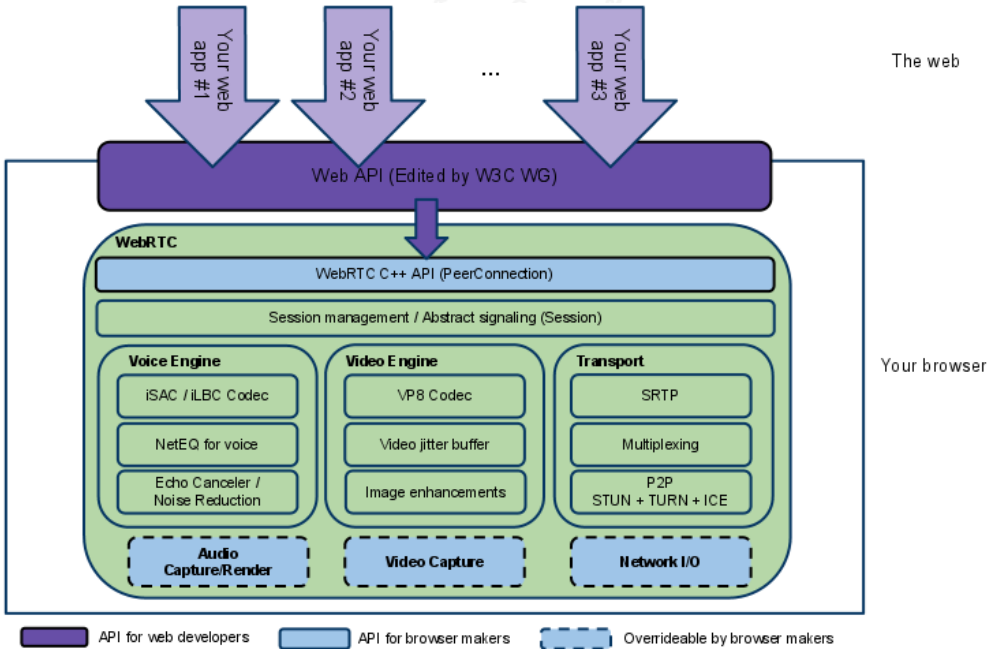
การสื่อสารในเวลาจริงผ่านเว็บ (Web Real-Time Communication หรือ WebRTC) เป็น API ที่ให้บริการในการสื่อสารระหว่างผู้ใช้งานในสถาปัตยกรรมแบบเพียร์ทูเพียร์ (Peer-to-peer) ที่อยู่ภายใต้การพัฒนาของ W3C [24] และ IETF [25] ซึ่งมีจุดมุ่งหมายเพื่อให้นักพัฒนาเบราว์เซอร์ต่าง ๆ เพิ่มเติมการทำงานของการแลกเปลี่ยนข้อมูลได้ง่ายและสะดวกมากขึ้น

WebRTC ใช้งานร่วมกับ Node.js ซึ่งเป็นเว็บเซิร์ฟเวอร์ที่ทำงานอยู่บนพื้นฐานของจาวาสคริปต์(JavaScript) และทำหน้าที่ช่วยเหลือให้ผู้ใช้งานต่าง ๆ เชื่อมต่อระหว่างกันได้ [26] ซึ่ง Google Chrome และ Mozilla Firefox ได้นำ WebRTC มาให้บริการการสื่อสารแบบทันกาลผ่านเว็บเบราว์เซอร์ (Web browser)

การให้บริการต่าง ๆ ของ WebRTC สามารถแบ่งได้เป็นสามส่วนดังแสดงในรูปที่ 2.1 โดยส่วนแรกเรียกว่า Media Stream คือส่วนที่ทำให้ผู้พัฒนาสามารถเรียกข้อมูลของกล้องและไมโครโฟนจากเครื่องผู้ใช้ได้โดยตรงส่วนที่สองคือ Peer Connection ทำให้ผู้พัฒนาสามารถสร้างการสื่อสารด้วยภาพและเสียงได้โดยใช้การเชื่อมต่อในสถาปัตยกรรมเพียร์ทูเพียร์ตามสถาปัตยกรรมที่แสดงในรูปที่ 2.2 และส่วนสุดท้ายคือ Data Channel ที่ไว้สำหรับส่งผ่านไฟล์ทุกประเภทที่เบราว์เซอร์รองรับ โดยในงานวิจัยนี้ก็จะนำสองส่วนแรกคือ Media Stream และ Peer Connection มาใช้ในการสร้างการสื่อสารระหว่างคนตาบอดและผู้ช่วยเหลือ โดยขั้นตอนการทำงานในการสื่อสารสามารถแบ่งได้ออกเป็น 3 ขั้นตอน [27] ได้แก่

1. เบราว์เซอร์เข้าไปยังเว็บไซต์ที่ให้บริการการใช้งานของ WebRTC
2. เซิร์ฟเวอร์ของเว็บไซต์ส่งหมายเลขประจำตัวของเบราว์เซอร์ให้กับผู้ใช้งาน
3. ผู้ใช้งานติดต่อกับเบราว์เซอร์อื่น ๆ ที่เจาะจงโดยใช้ส่งหมายเลขประจำตัวที่ได้รับ

รูปที่ 2.1 การให้บริการต่าง ๆ ของ WebRTC



รูปที่ 2.2 สถาปัตยกรรมของ RTCPeerConnection คัดลอกจาก [22]

2. อัลกอริทึมการต่อภาพ (Image Stitching Algorithm)

การต่อภาพ (Image Stitching) เป็นหนึ่งอัลกอริทึมในเทคโนโลยีการประมวลผลภาพ (Image processing) โดยเป็นการเชื่อมต่อภาพที่มีส่วนทับซ้อนกันบางส่วน ตั้งแต่สองภาพขึ้นไปให้กลายเป็นภาพเดียวกัน ทั้งนี้ผลจากการต่อภาพด้วยวิธีดังกล่าวเรียกว่าภาพโมเสก (Mosaic image) ซึ่งถูกนำไปประยุกต์ใช้ในหลายรูปแบบ เช่น การสร้างภาพพาโนรามา (Panoramic image)

ขั้นตอนในการต่อภาพแบ่งเป็นสองส่วน [28] ได้แก่การลงทะเบียนภาพ (Image registration) และการรวมภาพ (Image fusion)

2.1 การลงทะเบียนภาพ (Image registration)

เป็นขั้นตอนสำคัญในการต่อภาพ เนื่องจากการต่อภาพเป็นการประมวลผลจากข้อมูลภายในภาพมากกว่าหนึ่งภาพที่มีความแตกต่างกัน การลงทะเบียนภาพเป็นการแทนค่าคุณลักษณะต่าง ๆ ที่ปรากฏในภาพด้วยข้อมูลชุดหนึ่ง และทำการเปรียบเทียบกับภาพอื่นที่ต้องการเชื่อมต่อเพื่อหาส่วนทับซ้อน (Overlap) ซึ่งถูกนำไปใช้ในขั้นตอนการรวมภาพ (Image fusion) โดนขั้นตอนในการลงเบียนภาพสามารถจำแนกได้เป็น 3 ประเภท [29] ได้แก่

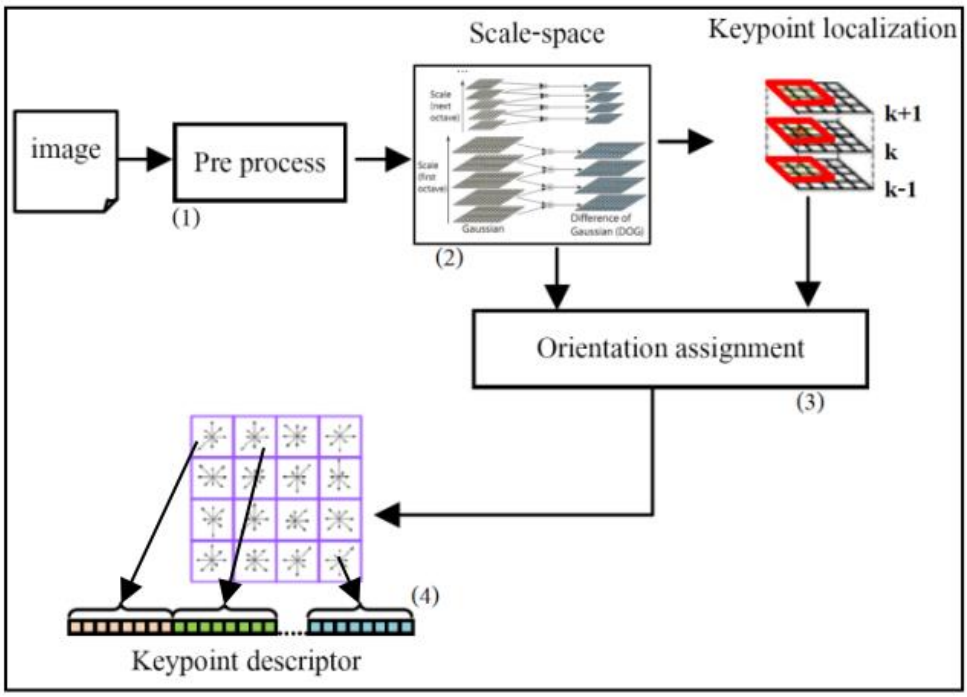
1. วิธีอิงคุณลักษณะ (Feature-based method)
2. วิธีข้อมูลระดับเทา (Gray-Scale Information method)
3. วิธีแปลงโดยอิงโดเมน (Transform-domain-based method)

ในงานวิจัยนี้ประยุกต์ใช้วิธีอิงคุณลักษณะในขั้นตอนการลงทะเบียนภาพ ซึ่งเป็นการสร้างความสัมพันธ์ของจุดคุณลักษณะต่าง ๆ ของภาพในรูปแบบของตัวเลข โดยจุดคุณลักษณะดังกล่าวถูกคำนวณโดยใช้ค่าความเข้มของสีในภาพ ความหนาแน่น รูปทรง และพื้นผิว ทั้งนี้แต่ละจุดคุณลักษณะจะถูกคำนวณและคัดเลือกโดยอัลกอริทึมต่าง ๆ เช่น SIFT (Scale-Invariant Feature Transform) หรือ SURF (Speeded-Up Robust Feature) เพื่อใช้เป็นตัวแทนข้อมูลที่จะนำไปเปรียบเทียบกับภาพอื่น

2.1.1 SIFT (Scale-Invariant Feature Transform)

Scale-Invariant Feature Transform (SIFT) [30] เป็นอัลกอริทึมพื้นฐานที่ใช้ในการตรวจสอบและอธิบายคุณลักษณะต่าง ๆ ในรูปภาพ ที่คิดค้นโดยนาย David G. Lowe ในปี ค.ศ. 1993 [31] โดยรูปที่ 2.3 แสดงขั้นตอนในการทำงานของอัลกอริทึม SIFT ทั้ง 4 ขั้นตอน [31] ได้แก่

1. การหาปริภูมิในมิติของขนาดและระยะทาง (Scale-space extrema detection)
2. การกำหนดตำแหน่งจุดสนใจ (Keypoint localization)
3. การกำหนดทิศทางของจุดสนใจ (Orientation assignment)
4. การสร้างคำอธิบายลักษณะเด่นของภาพ (Keypoint descriptor)



รูปที่ 2.3 ขั้นตอนในการทำงานของอัลกอริทึม SIFT คัดลอกจาก [32]

2.1.1.1 การหาปริภูมิในมิติของขนาดและระยะทาง (Scale-space Extrema Detection)

การหาปริภูมิในมิติของขนาดและระยะทาง เป็นการกำหนดปริภูมิในมิติของขนาดและระยะทาง (Scale-space) ซึ่งนิยามโดยค่า $L(x, y, \sigma)$ ดังแสดงในสมการ (2.1) โดยเป็นการคอนโวลูชันภาพ $I(x, y)$ ด้วยฟังก์ชันตัวกรองแบบเกาส์เซียนที่มีขนาด σ ตามสมการ (2.2) ซึ่งเมื่อผ่านการกรองแล้วจะนำมาคำนวณหาความแตกต่างกับข้อมูลชั้นที่อยู่ติดกันโดยใช้ DoG (Difference of Gaussian) ตามสมการ (2.3)

$$L(x, y, \sigma) = G(x, y, k\sigma) * I(x, y) \tag{2.1}$$

$$\mathrm{G(x, y, \sigma)} = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2} \tag{2.2}$$

$$\mathrm{DoG(x, y, \sigma)} = L(x, y, \sigma_n) - L(x, y, \sigma_{n-1}) \tag{2.3}$$

2.1.1.2 การกำหนดตำแหน่งจุดสนใจ (Keypoint localization)

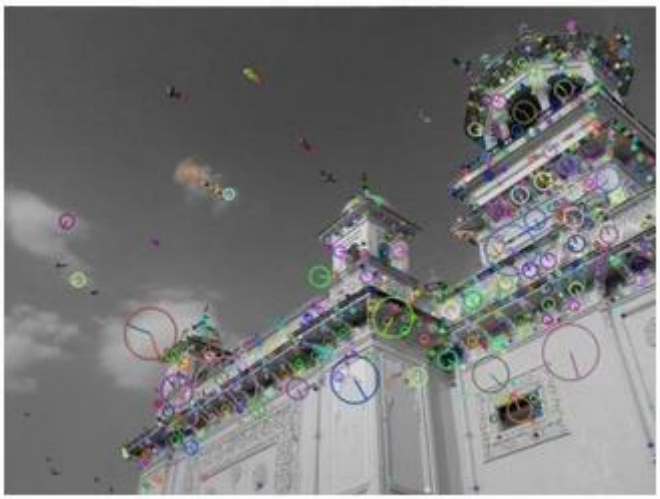
การกำหนดตำแหน่งจุดสนใจ (Keypoint localization) เป็นการเลือกจุดสนใจในภาพโดยใช้การคำนวณจุดที่มีค่าสูงสุดและต่ำสุดจากจุดเพื่อนบ้านที่ใกล้เคียง 8 จุด (Nearest neighbor) ในปริภูมิ DoG โดยเมื่อได้จุดที่สนใจจำนวน 2 จุดแล้ว จะถูกนำมาจับคู่กับจุดสนใจอื่นที่เป็นจุดเพื่อนบ้านใกล้เคียง 8 จุดถัดไปเพื่อหาจุดสนใจอื่นในภาพสำหรับใช้คำนวณในขั้นตอนถัดไป

2.1.1.3 การกำหนดทิศทางของจุดสนใจ (Orientation assignment)

การกำหนดทิศทางของจุดที่สนใจ (Orientation assignment) เป็นขั้นตอนการคำนวณขนาด $m(x, y)$ และทิศทางของจุดที่สนใจ $\theta(x, y)$ ดังสมการที่ (2.4) และ (2.5) ตามลำดับ โดยขนาดและทิศทางของจุดที่สนใจเป็นการคำนวณจากคอนโวลูชันของฟังก์ชันตัวกรองแบบเกาส์เซียน $L(x, y)$ ซึ่งจะถูกนำไปคำนวณเป็นลักษณะเด่นของภาพ (Keypoint) ดังแสดงในรูปที่ 2.4 เพื่อใช้ในการเปรียบเทียบในขั้นตอนการจับคู่ค่าคุณลักษณะ (Feature matching)

$$m(x, y) = \sqrt{L(x+1, y) - L(x-1, y)^2 + L(x, y+1) - L(x, -1y)^2} \tag{2.4}$$

$$\theta(x, y) = \tan^{-1}\left(\frac{L(x,y+1) - L(x,y-1)}{L(x+1,y) - L(x-1,y)}\right) \tag{2.5}$$



รูปที่ 2.4 ตัวอย่างจุดสนใจในภาพที่ถูกคำนวณด้วยอัลกอริทึม SIFT คัดลอกจาก [32]

2.1.1.4 การสร้างคำอธิบายลักษณะเด่นของภาพ (Keypoint descriptor)

หลังจากการคำนวณลักษณะเด่นของภาพ (Keypoint) จากขนาดและระยะทางจะสร้างคำอธิบายลักษณะเด่นของภาพ (Keypoint descriptor) ในรูปแบบของเวกเตอร์ โดยสร้างพื้นที่ล้อมรอบจุดที่สนใจจำนวน 16 ช่องในขนาด 4x4 เพื่อเก็บข้อมูลทิศทางรอบจุดสนใจ โดยในแต่ละช่องจะสร้างภาพแท่งความถี่ (Histogram) ขนาด 8 แกน

หลังจากสร้างคำอธิบายลักษณะเด่นของภาพ ถือว่าเป็นการสิ้นสุดการทำงานในขั้นตอนการลงทะเบียนภาพ (Image registration) โดยการจับคู่ค่าคุณลักษณะ (Feature matching) ในอัลกอริทึม SIFT เรียกว่าการจับคู่ลักษณะเด่นของภาพ (Keypoint matching) ซึ่งจะกล่าวในหัวข้อ 2.2.1

2.1.2 SURF (Speeded-Up Robust Feature)

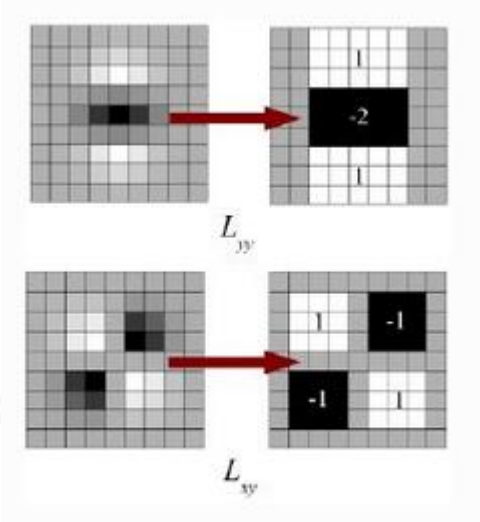
Speeded-Up Robust Feature (SURF) เป็นอัลกอริทึมที่คิดค้นโดย Bay Herbert และคณะในปีค.ศ. 2006 [33] โดยเป็นการพัฒนาต่อยอดจากอัลกอริทึม SIFT ในด้านการลดเวลาในการประมวลผล

ขั้นตอนในการทำงานของอัลกอริทึม SURF มี 4 ขั้นตอน [34] ได้แก่

1. การหาปริภูมิในมิติของขนาดและระยะทาง (Scale-space extrema detection)
2. การกำหนดตำแหน่งจุดคุณลักษณะ (Feature point localization)
3. การกำหนดทิศทางของจุดคุณลักษณะ (Orientation assignment)
4. การสร้างคำอธิบายลักษณะเด่นของภาพ (Feature descriptor)

2.1.2.1 การหาปริภูมิในมิติของขนาดและระยะทาง (Scale-space extrema detection)

การหาปริภูมิในมิติของขนาดและระยะทาง เป็นการกำหนดปริภูมิในมิติของขนาดและระยะทาง (Scale-space) ซึ่งใช้การประมาณค่าโดยใช้ตัวกรองแบบกล่อง (Box filter) ซึ่งเป็นการคอนโวลูชันภาพด้วยฟังก์ชันตัวกรองแบบเกาส์เซียนดังแสดงในรูปที่ 2.4



รูปที่ 2.4 ตัวอย่างปริภูมิในมิติของขนาดและระยะทางในอัลกอริทึม SURF คัดลอกจาก [35]

2.1.2.2 การกำหนดตำแหน่งจุดคุณลักษณะ (Feature point localization)

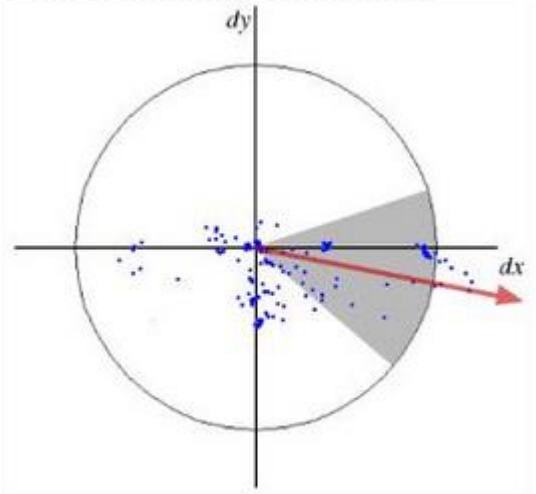
การกำหนดตำแหน่งจุดคุณลักษณะ (Feature point localization) ในอัลกอริทึม SURF อ้างอิงจาก Hession matrix $\mathbf{H}(\mathbf{x}, \sigma)$ ซึ่งคำนวณจากปริภูมิในมิติของขนาดและระยะทางดังแสดงในสมการที่ (2.6)

$$\mathrm{H}(\mathrm{x}, \sigma) = \begin{bmatrix} L_{xx}(x, \sigma) & L_{xy}(x, \sigma) \\ L_{xy}(x, \sigma) & L_{yy}(x, \sigma) \end{bmatrix} \quad (2.6)$$

การพิจารณาตำแหน่งจุดคุณลักษณะที่จะนำมาใช้ในการจับคู่คุณลักษณะ (Feature matching) ใช้การตรวจจับ $\mathbf{H}(\mathbf{x}, \sigma)$ โดยใช้ค่าขีดแบ่ง (Threshold) ที่กำหนดขึ้น โดยหลังจากคัดเลือกจุดคุณลักษณะที่สนใจจากค่าขีดแบ่งแล้ว จะทำการคัดเลือกอีกครั้งโดยใช้การเปรียบเทียบกับจุดเพื่อนบ้านที่ใกล้ที่สุด 8 จุด เพื่อคัดเลือกเป็นจุดคุณลักษณะที่จะใช้ในการจับคู่คุณลักษณะ

2.1.2.3 การกำหนดทิศทางของจุดคุณลักษณะ (Orientation assignment)

การกำหนดทิศทางของจุดคุณลักษณะ (Orientation assignment) ใช้การตอบสนองเวฟเล็ตในแนวแกนตั้งและแกนนอนต่อจุดเพื่อนบ้าน และทำการถ่วงน้ำหนักด้วยสมการเกาส์เซียน จากนั้นจะทำการวางจุดดังกล่าวลงไปในแนวแกน x และ y ดังแสดงในรูปที่ 2.5 โดยทิศทางของจุดคุณลักษณะถูกกำหนดโดยการคำนวณผลรวมของการตอบสนองภายในมุมขนาด 60 องศา



รูปที่ 2.5 ตัวอย่างการกำหนดทิศทางของจุดคุณลักษณะด้วยอัลกอริทึม SURF คัดลอกจาก [35]

2.1.2.4 การสร้างคำอธิบายลักษณะเด่นของภาพ (Feature Descriptor)

การสร้างคำอธิบายลักษณะเด่นของภาพ (Feature Descriptor) ใช้ HAAR Wavelet ในการกำหนดทิศทางในแนวแกน x และแกน y ซึ่งเขียนอยู่ในรูปของเวกเตอร์ดังแสดงในสมการ (2.7) โดยหลังจากขั้นตอนนี้จะเป็นขั้นตอนการจับคู่คุณลักษณะซึ่งจะกล่าวถึงในหัวข้อ 2.2.2

$$\mathrm{v} = (\sum dx, \sum dy, \sum |dx|, \sum |dy|) \qquad (2.7)$$

2.2 การรวมภาพ (Image fusion)

กระบวนการรวมภาพ (Image fusion) เป็นขั้นตอนหลังจากกระบวนการลงทะเบียนภาพ (Image registration) โดยเป็นการรวมภาพที่มีความสัมพันธ์กันตั้งแต่สองรูปขึ้นไปให้กลายเป็นภาพเดียวกัน ขั้นตอนสำคัญในกระบวนการรวมภาพ คือการจับคู่คุณลักษณะ (Feature Matching) ซึ่งเป็นการหาตำแหน่งของภาพที่ทับซ้อนกันในภาพทั้งสอง โดยวิธีการในการจับคู่คุณลักษณะจะแตกต่างกันออกไปตามรูปแบบของผลลัพธ์ที่ได้จากกระบวนการลงทะเบียนภาพ ซึ่งในที่นี้จะกล่าวถึงการจับคู่คุณลักษณะสำหรับผลลัพธ์ที่ได้จากการลงทะเบียนภาพที่เรียกว่าการจับคู่ลักษณะเด่นของภาพ (Keypoint Matching) โดยใช้อัลกอริทึม SIFT และ SURF

2.2.1 การจับคู่ลักษณะเด่นของภาพ (Keypoint Matching)

การจับคู่ลักษณะเด่นของภาพ (Keypoint Matching) เป็นการจับคู่โดยระบุจุดเพื่อนบ้านที่ใกล้ที่สุดโดยใช้การคำนวณการขจัดของเวคเตอร์ที่เป็นคำอธิบายลักษณะเด่นของภาพ (Feature Descriptor) ด้วยระยะทางยูคลิด (Euclidean distance) ซึ่งหลังจากการคำนวณการขจัด จะจับคู่ลักษณะเด่นของภาพที่มีค่าขจัดน้อยที่สุดโดยจะทำการคำนวณเพื่อหาการขจัดทั้งหมดสองครั้ง เพื่อคำนวณค่าอัตราส่วนของการขจัดที่น้อยที่สุดจากทั้งสองครั้ง (R) ตามสมการที่ (2.8) ซึ่งจะนำอัตราส่วนนี้ไปเทียบกับค่าขีดจำกัด หากมีค่าอัตราส่วนมากกว่า 0.8 ถือว่ามีความคลาดเคลื่อนสูง ซึ่งจะไม่ยอมรับการจับคู่ลักษณะเด่นของภาพในจุดนั้น

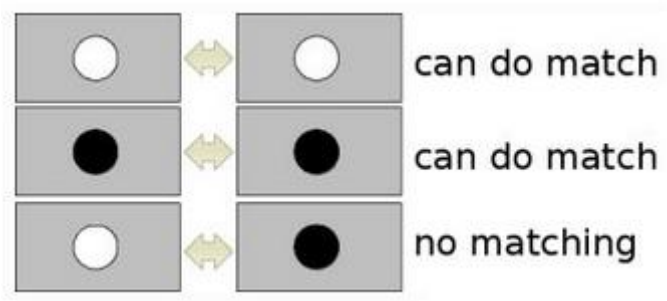
$$\mathrm{R} = \frac{First\ closet\ distance}{Second\ closet\ distance} \qquad (2.8)$$

หลังจากการจับคู่ลักษณะเด่นของภาพจะทำการรวมภาพทั้งสองเข้าด้วยกันโดยใช้ Homography matrix ซึ่งจะกล่าวในหัวข้อ 2.2.3

2.2.2 การจับคู่คุณลักษณะสำหรับผลลัพธ์ที่ได้จากการลงทะเบียนภาพโดยอัลกอริทึม SURF

การจับคู่คุณลักษณะสำหรับผลลัพธ์ที่ได้จากการลงทะเบียนภาพโดยอัลกอริทึม SURF แตกต่างจากวิธีการของอัลกอริทึม SIFT โดยเน้นไปที่การนำข้อมูลที่คำนวณได้จากขั้นตอนการลงทะเบียนภาพมาใช้ในการจับคู่คุณลักษณะแทนการประมวลผลเพิ่มเติม

การจับคู่คุณลักษณะสำหรับอัลกอริทึม SURF เปรียบเทียบจากความเข้มของคุณลักษณะโดยใช้วิธีการ RANSAC ซึ่งเป็นวิธีในการเลือกใช้เซตที่เล็กที่สุดของจุดคุณลักษณะที่สอดคล้องกัน โดยการจับคู่จะเกิดขึ้นก็ต่อเมื่อภาพทั้งสองมีความเข้มของคุณลักษณะในรูปแบบเดียวกันดังแสดงในรูปที่ 2.6 เท่านั้น



Chulalongkorn University

รูปที่ 2.6 ตัวอย่างการเปรียบเทียบของอัลกอริทึม SURF คัดลอกจาก [35]

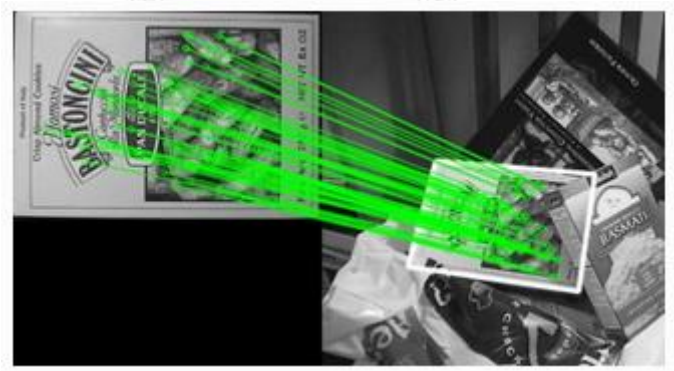
2.2.3 การรวมภาพโดยใช้ Homography matrix

Homography matrix (M) เป็นเมทริกซ์การแปลง (Transform matrix) ขนาด 3 x 3 ดังแสดงในสมการ (2.9) ซึ่งเป็นการแสดงความสัมพันธ์ของตำแหน่งที่เกี่ยวข้องกันระหว่างภาพสองภาพ โดยการคำนวณ Homography matrix จะอ้างอิงจากความสัมพันธ์ตามสมการ (2.10) ซึ่งจะพิจารณาจุดที่นำเข้าจำนวนสองจุด ได้แก่ X1 ในแนวแกน x y และ z (u1, v1, 1) และจุด X2 ในแนวแกน x y และ z (u2, v2, 1)

$$\mathrm{M} = \begin{bmatrix} m1 & m3 & m3 \\ m4 & m5 & m6 \\ m7 & m8 & m9 \end{bmatrix} \tag{2.9}$$

$$\begin{bmatrix} u1 \\ v1 \\ 1 \end{bmatrix} = \begin{bmatrix} m1 & m3 & m3 \\ m4 & m5 & m6 \\ m7 & m8 & m9 \end{bmatrix} \begin{bmatrix} u2 \\ v2 \\ 1 \end{bmatrix} \tag{2.10}$$

ภาพนำเข้าที่ต้องการเชื่อมต่อกับภาพอื่นจะถูกปรับให้อยู่ในตำแหน่งที่สอดคล้องกับภาพที่ต้องการเชื่อมต่อโดยใช้ Homography matrix ในการเปลี่ยนมุมของจุดคุณลักษณะต่าง ๆ ในภาพทีละจุด ดังแสดงในรูปที่ 2.7



รูปที่ 2.7 ตัวอย่างการปรับตำแหน่งโดยใช้ Homography matrix คัดลอกจาก [36]

3. แนวทางการเข้าถึงเนื้อหาของเว็บ (Web Content Accessibility Guidelines)

งานวิจัยนี้มีในส่วนการพัฒนาโปรแกรมวิดีโอต้นแบบ มีส่วนเกี่ยวข้องกับการออกแบบส่วนต่อประสานผู้ใช้สำหรับคนตาบอดให้สามารถเข้าถึงฟังก์ชันต่าง ๆ ที่มีในโปรแกรมได้ โดยหลักการสำหรับการออกแบบส่วนต่อประสานผู้ใช้สำหรับคนตาบอดถูกกล่าวถึงในแนวทางการเข้าถึงเนื้อหาของเว็บ (Web Content Accessibility Guidelines หรือ WCAG) ดังนั้นในงานวิจัยนี้จึงนำบางส่วนของแนวทางการเข้าถึงเนื้อหาของเว็บมาประยุกต์ใช้ในการออกแบบส่วนต่อประสานผู้ใช้สำหรับคนตาบอด

แนวทางการเข้าถึงเนื้อหาของเว็บ (Web Content Accessibility Guidelines หรือ WCAG) [37] ถูกพัฒนาขึ้นผ่านกระบวนการของ W3C ภายใต้ความร่วมมือของบุคคลและองค์กรต่าง ๆ ทั่วโลกโดยมีเป้าหมายเพื่อเป็นมาตรฐานที่ใช้งานร่วมกันในการสร้างการเข้าถึงเนื้อหาของเว็บที่ตรงกับความต้องการของบุคคล องค์กร และรัฐบาลในระดับสากล

WCAG เป็นหลักการสำหรับการออกแบบเนื้อหาของเว็บซึ่งในปัจจุบันไม่ได้มีการทำเป็นเอกสารชัดเจนสำหรับการออกแบบบนโทรศัพท์มือถือ โดยเอกสารของ WCAG เป็นการอธิบายวิธีการที่ทำให้เนื้อหาของเว็บสามารถเข้าถึงได้สำหรับบุคคลทุกประเภท ซึ่งมีการกล่าวถึงการออกแบบสำหรับผู้พิการเพิ่มเติม โดย WCAG 1.0 ถูกประกาศใช้เป็นครั้งแรกเมื่อปี ค.ศ. 1999 [38] และภายหลังจากนั้นมีการปรับเนื้อหาภายใน และประกาศเป็น WCAG 2.0 ในปี ค.ศ. 2008 [39] ซึ่งประกอบด้วย 4 หลักการ ได้แก่

1) สามารถรับรู้ได้ (Perceivable)

- ข้อมูลและส่วนประกอบของส่วนต่อประสานผู้ใช้จะต้องแสดงในรูปแบบที่สามารถรับรู้ได้

2) สามารถปฏิบัติการได้ (Operable)

- ต้องสามารถปฏิบัติการส่วนต่าง ๆ ของส่วนต่อประสานผู้ใช้ และการนำทางต่าง ๆ ได้

3) สามารถเข้าใจได้ (Understandable)

- ต้องสามารถเข้าใจข้อมูลและการดำเนินงานต่าง ๆ ของส่วนต่อประสานผู้ใช้ได้

4) ทนทาน (Robust)

- สามารถเข้ากันได้กับเทคโนโลยีที่มีอยู่ในปัจจุบันและอนาคต รวมถึงเทคโนโลยีช่วยเหลือ

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

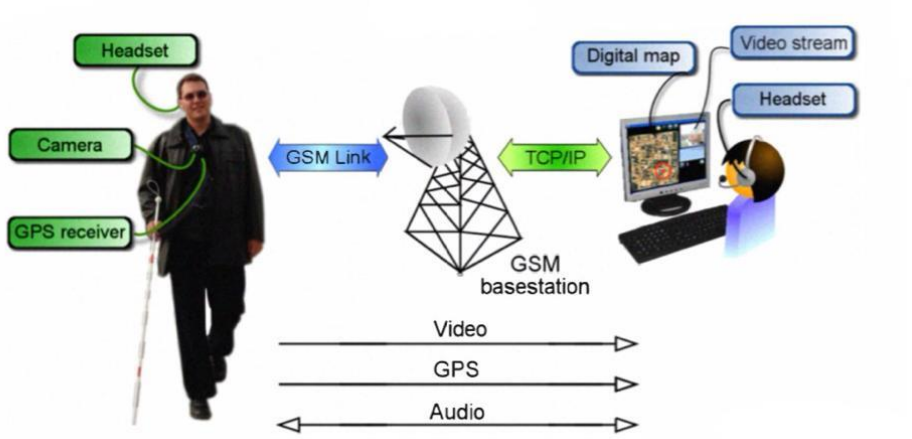
งานวิจัยที่เกี่ยวข้อง

1. ระบบช่วยเหลือทางไกล (Tele-Assistance System)

ระบบช่วยเหลือทางไกล (Tele-assistance system) เป็นแนวคิดของการช่วยเหลือคนตาบอดให้สามารถเดินทางไปยังสถานที่ต่าง ๆ หรือทำกิจกรรมต่าง ๆ ตามลำพังได้ โดยในตอนเริ่มต้นระบบช่วยเหลือทางไกลสำหรับคนตาบอดเป็นการประมวลผลโดยใช้ค่า GPS ที่ได้จากคนตาบอดร่วมกับการใช้การรู้จำเสียง (Speech recognition) ที่คนตาบอดถาม หรือการพิมพ์คำถามผ่านแป้นพิมพ์ของคนตาบอด [40] นอกจากนี้มีการใช้การนำทางโดยอาศัยการวางแผนล่วงหน้า [41] เมื่อคนตาบอดเดินทางตามลำพังก็ใช้การติดต่อกับฐานข้อมูลที่เก็บข้อมูลที่ได้วางแผนไว้ แต่จะเห็นได้ว่ารูปแบบการใช้งานต้องใช้งานผ่านอุปกรณ์เฉพาะที่มีขนาดใหญ่ รวมไปถึงการโต้ตอบที่ใช้การรู้จำเสียงทำให้มีรูปแบบของคำถามและการโต้ตอบที่ตายตัว ทั้งนี้อาจเนื่องมาจากข้อจำกัดของความก้าวหน้าของเทคโนโลยีในสมัยนั้น

ในปีค.ศ. 2003 [20] มีการเสนอแนวคิดสำหรับระบบช่วยเหลือทางไกลที่เป็นการสื่อสารระหว่างคนตาบอดและผู้ช่วยเหลือ แต่ไม่ได้มีการรายงานผลจากการวิจัยในครั้งนี้ จนกระทั่ง 3 ปีต่อมา Technical University of Lodz ได้พัฒนาโปรแกรมต้นแบบของแนวคิดนี้ขึ้นโดยมีแนวคิดว่าผู้ช่วยเหลือที่เป็นโอเปอเรเตอร์รับภาพวิดีโอ และ GPS จากคนตาบอด โดยผู้ช่วยเหลือพูดโต้ตอบกลับไป ในระบบดังกล่าวใช้อุปกรณ์ 2 อย่าง ได้แก่ อุปกรณ์เซนเซอร์ที่ถือโดยคนตาบอด และ เครื่องคอมพิวเตอร์สำหรับผู้ช่วยเหลือที่หน้าจอของโปรแกรมประกอบไปด้วยรูปภาพที่ได้รับจากคนตาบอดในมุมซ้าย ส่วนด้านขวาจะเป็นแผนที่ซึ่งใช้แผนที่กูเกิ้ล (Google Maps) ตามแสดงในรูปที่ 2.8 นอกจากนี้ในรูปที่ 2.9 ยังได้แสดงแนวคิดของระบบช่วยเหลือทางไกลสำหรับคนตาบอดซึ่งถูกนำเสนอในงานวิจัยดังกล่าวด้วยเช่นกัน

รูปที่ 2.8 ตัวอย่างหน้าจอของโอเปอเรเตอร์ในงานวิจัยของ P. Baranski [18], [19]



รูปที่ 2.9 แนวคิดของระบบช่วยเหลือทางไกลสำหรับคนตาบอดโดย P. Baranski [18], [19]

โปรแกรมต้นแบบของแนวคิดดังกล่าวสำเร็จและนำไปทดสอบระบบกับอาสาสมัครที่เป็นคนตาบอดทั้งหมดสามคนเมื่อปี ค.ศ. 2010 [21], [22] โดยเมื่อเปรียบเทียบผลลัพธ์ที่ได้จากการทดสอบกับการเดินทางของคนตาบอดตามลำพังพบว่า การสื่อสารกับผู้ช่วยเหลือผ่านระบบช่วยเหลือทางไกลช่วยลดข้อผิดพลาดต่าง ๆ ที่เกิดขึ้นได้มาก และลดระยะเวลาในการเดินทางของคนตาบอดด้วยเช่นกัน นอกจากนี้ผู้ทดสอบยังกล่าวว่าการเดินทางกับผู้ช่วยเหลือนี้ทำให้รู้สึกปลอดภัยมากขึ้น แต่อย่างไรก็ตามการใช้งานร่วมกับ GPS ของคนตาบอดมีข้อจำกัดคือ ไม่สามารถใช้งานได้ในพื้นที่ที่ไม่มีสัญญาณ GPS รวมไปถึงข้อจำกัดด้านอุปกรณ์ที่คนตาบอดจำเป็นต้องจัดหาเพิ่มเติมเพื่อใช้งานด้วยเช่นกัน

การสื่อสารของคนตาบอดกับผู้ช่วยเหลือในระบบช่วยเหลือทางไกลนี้มีข้อดีคือ สามารถให้ข้อมูลต่าง ๆ แก่คนตาบอดได้อย่างทันกาล และมีรูปแบบของการใช้งานที่หลากหลาย ซึ่งเมื่อพิจารณาร่วมกับเทคโนโลยีในปัจจุบัน การสื่อสารผ่านวิดีโอคอลสามารถนำมาเป็นช่องทางในการสื่อสารระหว่างคนตาบอดและผู้ช่วยเหลือได้ นอกจากนี้เพื่อลดข้อจำกัดในเรื่องการจัดหาอุปกรณ์สำหรับคนตาบอด ในงานวิจัยนี้จึงเลือกใช้การสื่อสารผ่านวิดีโอคอลที่คนตาบอดสามารถใช้งานผ่านโทรศัพท์มือถือได้ ในขณะที่ผู้ช่วยเหลือก็สามารถใช้งานได้ในอุปกรณ์เครื่องคอมพิวเตอร์ทั่วไปได้เช่นกัน โดยประเด็นหลักที่เน้นในงานวิจัยนี้ อยู่ที่การใช้งานของผู้ช่วยเหลือซึ่งทำหน้าที่เปรียบเสมือนดวงตาของคนตาบอดที่มีปัญหาเรื่องมุมมองภาพที่จำกัด ซึ่งเป็นไปตามมาตรฐานของการส่งไฟล์วิดีโอผ่านโปรแกรมวิดีโอคอลโดยทั่วไป ทำให้ภาพที่ได้รับอาจไม่เพียงพอที่จะใช้ในการช่วยเหลือ หรือทำให้ผู้ช่วยเหลือไม่สามารถมองเห็นสิ่งต่าง ๆ ที่อยู่รอบตัวคนตาบอดได้

2. การประยุกต์ใช้ WebRTC ในระบบการสื่อสารต่าง ๆ

เมื่อพิจารณาการสื่อสารในระบบช่วยเหลือทางไกลร่วมกับเทคโนโลยีที่มีอยู่ในปัจจุบัน วิดีโอคอลเป็นทางเลือกที่เหมาะสมสำหรับการสื่อสารระหว่างคนตาบอดและผู้ช่วยเหลือเนื่องจากสามารถส่งข้อมูลภาพและเสียงได้พร้อมกัน และอยู่ในรูปของการสื่อสารแบบทันกาล โดย WebRTC เป็นรูปแบบหนึ่งของการใช้งานวิดีโอคอล โดยใช้งานผ่านเว็บเบราว์เซอร์

WebRTC ถูกนำไปใช้ในระบบต่าง ๆ มากมาย เช่น ระบบการประชุมทางไกล [42] ระบบช่วยเหลือสำหรับผู้สูงอายุ [43], [44] และระบบที่ให้บริการด้านสุขภาพทางไกล [45] โดยการใช้งานผ่าน WebRTC นี้ผู้ใช้สามารถใช้งานได้โดยไม่จำเป็นต้องดาวน์โหลดโปรแกรมใด ๆ ในขณะที่ผู้พัฒนาก็สามารถพัฒนาต่อได้จากโอเพนซอร์ส (Open source) ที่มีอยู่ในปัจจุบัน ซึ่งจะเห็นได้ว่าเป็นช่องทางที่สะดวกและเป็นอีกทางเลือกหนึ่งสำหรับการสื่อสารแบบทันกาล ที่ผู้ใช้สามารถใช้งานได้ผ่านอุปกรณ์ได้หลากหลาย นอกจากนี้การพัฒนา WebRTC สามารถใช้ EasyRTC APIs ซึ่งเป็นโอเพนซอร์สแบบเต็มรูปแบบ (Full stack open source) ของ WebRTC ที่เปิดให้ผู้พัฒนานำไปพัฒนาต่อได้ในหลากหลายรูปแบบ เช่น การสร้างห้องเพื่อพูดคุยระหว่างเครื่องสองเครื่อง การสนทนาด้วยข้อความหรือการพูดคุยผ่านวิดีโอคอล

ตามที่กล่าวไปบางส่วนในทฤษฎีที่เกี่ยวข้อง ในงานนี้จะใช้ API ในส่วนของ MediaStream และ RTCPeerConnection ซึ่งในส่วนของการส่งภาพและเสียงผ่าน PeerConnection จะมีการจัดการเพิ่มเติมในเรื่องการส่งข้อมูลจากผู้ช่วยเหลือไปยังคนตาบอดที่ไม่จำเป็นต้องใช้ภาพด้วยเช่นกัน

3. อัลกอริทึมที่ใช้ในการต่อภาพ

ตามที่ได้กล่าวไปในทฤษฎีที่เกี่ยวข้อง ในการต่อภาพในงานวิจัยนี้เลือกใช้วิธีอิงคุณลักษณะ (Feature-based method) ในกระบวนการลงทะเบียนภาพ โดยอัลกอริทึมที่เป็นที่นิยมสำหรับ วิธีอิงคุณลักษณะในปัจจุบันได้แก่ SIFT และ SURF

ในปี ค.ศ. 2013 P. M. Panchal และคณะ [46] ได้เปรียบเทียบอัลกอริทึมทั้งสองในการจับคู่คุณลักษณะที่เหมือนกันในรูปภาพสองรูป โดยผลลัพธ์ในการศึกษาในตารางที่ 2.1 แสดงให้เห็นว่า SIFT สามารถตรวจจับคุณลักษณะได้มากกว่าเมื่อเทียบกับ SURF แต่การประมวลผลก็ใช้เวลานานกว่า SURF เช่นกัน นอกจากนี้ผลลัพธ์ในการต่อรูปภาพแสดงให้เห็นว่า SURF สามารถต่อรูปภาพได้มีประสิทธิภาพดีเช่นเดียวกับ SIFT

อัลกอริทึม	จำนวนจุดคุณลักษณะที่ตรวจจับได้		จำนวนจุดคุณลักษณะที่จับคู่กันได้	เวลาที่ใช้ในการประมวลผล
	ภาพที่ 1	ภาพที่ 2		
SIFT	892	934	41	1.543 วินาที
SURF	281	245	28	0.546 วินาที

ตารางที่ 2.1 การเปรียบเทียบผลลัพธ์ในการต่อรูปภาพโดยใช้ SIFT และ SURF [46]

เมื่อพิจารณาขั้นตอนการประมวลผลโดยอัลกอริทึม SIFT และ SURF ร่วมกับผลการเปรียบเทียบ SIFT และ SURF โดย P. M. Panchal และคณะ [46] ดังนั้นแล้วในงานวิจัยนี้จะเลือกใช้อัลกอริทึมของ SURF ในการประมวลผลรูปภาพ เนื่องจากใช้เวลาในการประมวลผลน้อยกว่า ในขณะที่ผลลัพธ์ของทั้งสองอัลกอริทึมมีประสิทธิภาพใกล้เคียงกัน

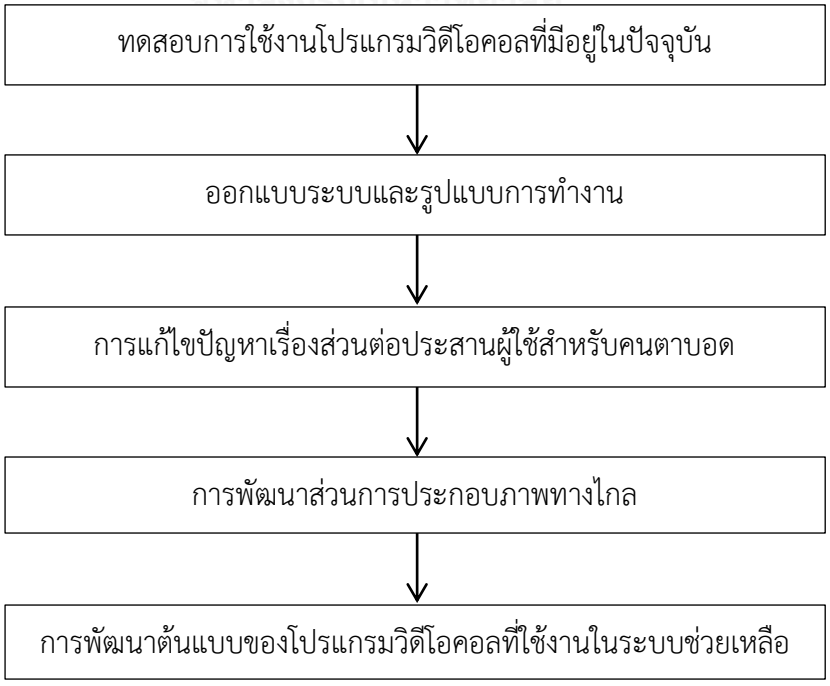
บทที่ 3
วิธีดำเนินการวิจัย

เครื่องมือที่ใช้ในการวิจัย

1. ไลบรารี่สำหรับการประมวลผลรูปภาพ, OpenCV [47]
2. เครื่องมือพัฒนาโปรแกรมภาษา C++, Microsoft Visual Studio 2010
3. โอเพนซอร์สแบบเต็มรูปแบบ (Full stack open source) ของ WebRTC, EasyRTC [48]

ขั้นตอนการดำเนินงานวิจัย

การดำเนินงานสามารถแบ่งได้เป็น 5 ขั้นตอนดังแสดงในรูปที่ 3.1 ขั้นตอนแรกเป็นการทดสอบการใช้งานโปรแกรมวิดีโอคอลที่มีอยู่ในปัจจุบันในระบบช่วยเหลือทางไกล ซึ่งมีการวิเคราะห์ปัญหาจากผลการทดลอง และสรุปเป็นแนวคิดเบื้องต้นในการพัฒนาโปรแกรม ต่อมาเป็นการออกแบบระบบและรูปแบบการทำงาน จากนั้นเป็นการแก้ไขปัญหาที่เกิดขึ้น โดยเริ่มจากการแก้ไขปัญหาเรื่องส่วนต่อประสานผู้ใช้สำหรับคนตาบอด จากนั้นเป็นการพัฒนาส่วนการประกอบภาพทางไกล และสุดท้ายเป็นการพัฒนาต้นแบบของโปรแกรมวิดีโอคอลที่ใช้งานในระบบช่วยเหลือคนตาบอดที่สามารถแก้ไขปัญหาที่เกิดขึ้นได้



รูปที่ 3.1 ขั้นตอนการดำเนินงาน

1. การทดสอบการใช้งานโปรแกรมวิดีโอคอลที่มีอยู่ในปัจจุบันในระบบช่วยเหลือทางไกล

ขั้นตอนแรกเป็นการทดลองเพื่อค้นหาปัญหาในการใช้งานโปรแกรมวิดีโอคอลในการช่วยเหลือทางไกลสำหรับคนตาบอดโดยบุคคลทั่วไป โดยทดลองกับกลุ่มตัวอย่างที่เป็นคนตาบอดจำนวน 3 คนที่แตกต่างกัน ได้แก่ ผู้ที่ตาบอดภายหลังและใช้ชีวิตร่วมกับผู้ช่วยเหลือจำนวน 2 คน และผู้ที่ตาบอดแต่กำเนิดและใช้ชีวิตประจำวันตามลำพังจำนวน 1 คน โดยในการทดลองได้จับคู่คนตาบอดคู่กับบุคคลทั่วไปจำนวน 3 คนเพื่อทำกิจกรรมต่าง ๆ ตามสถานการณ์ที่กำหนดไว้ ได้แก่

1. การนำช่วยคนตาบอดอ่านรายการเครื่องดื่มที่มีในร้านกาแฟ
2. การช่วยคนตาบอดเลือกซื้อน้ำในร้านสะดวกซื้อ
3. การช่วยคนตาบอดสอบถามยอดเงินคงเหลือในบัญชีผ่านตู้ ATM
4. การช่วยคนตาบอดให้เดินทางไปยังห้องที่กำหนดไว้
5. การช่วยคนตาบอดหาของที่ตกอยู่ที่พื้น

ในการทดลองคนตาบอดใช้งานวิดีโอคอลผ่านโปรแกรมไลน์ (LINE) บนเครือข่าย 3G ในโทรศัพท์มือถือซัมซุงกาแลกซี่เอส 3 (Samsung Galaxy S3) ที่ใช้ระบบปฏิบัติการแอนดรอยด์ 4.3 พร้อมกับหูฟัง ในขณะที่ผู้ช่วยเหลือใช้งานวิดีโอคอลผ่านโปรแกรมไลน์ (LINE) บนสัญญาณไวไฟ (Wi-Fi) ในเครื่องคอมพิวเตอร์

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

<u>การวิเคราะห์ผลการทดลองเบื้องต้น</u>

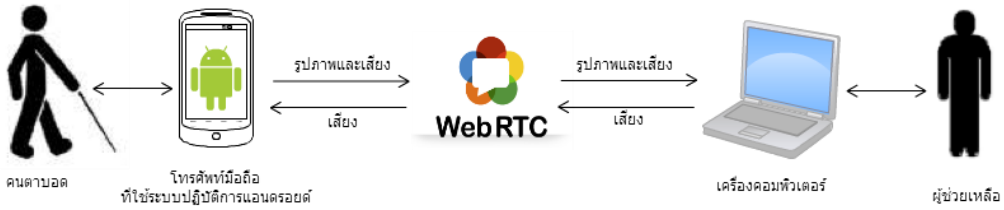
การวัดผลการทดลองอยู่ในรูปของเวลาที่ใช้ในการทำกิจกรรมต่าง ๆ ตามสถานการณ์ที่กำหนดไว้ ร่วมกับการสังเกตโดยตรงจากผู้วิจัยขณะทำการทดลอง ทั้งนี้สามารถสรุปปัญหาที่เกิดขึ้นและแนวทางในการแก้ไขได้ทั้งหมด 4 ข้อ ตามตารางที่ 3.1

ลำดับ	ปัญหา	แนวทางแก้ไข
P1	การบอกให้คนตาบอดเลื่อนกล้องไปยังทิศทางต่าง ๆ ทำได้ยาก	ออกแบบฟังก์ชันที่ช่วยอำนวยความสะดวกเพื่อทดแทนการเลื่อนกล้องไปยังจุดต่าง ๆ
P2	ภาพที่มองเห็นในปัจจุบันไม่เพียงพอสำหรับการใช้งานในการช่วยเหลือ ซึ่งบางครั้งต้องใช้การมองเห็นสิ่งที่อยู่รอบ ๆ ตัวคนตาบอด	หาเทคนิคอื่น ๆ มาประยุกต์ใช้เพื่อให้ผู้ช่วยเหลือสามารถมองเห็นรอบตัวคนตาบอดได้
P3	การส่งข้อมูลภาพและเสียงอาจขาดหายไปในพื้นที่ที่ไม่มีสัญญาณ หรือสัญญาณอ่อน ซึ่งคนตาบอด ไม่สามารถรับรู้ และแก้ไขได้ ในขณะที่ผู้ช่วยเหลือก็อาจได้รับข้อมูลที่ขาดหายหรือไม่เพียงพอสำหรับการช่วยเหลือ	เก็บสำรอง หรือบันทึกข้อมูลภาพวิดีโอที่ถูกส่งมาจากคนตาบอดสำหรับผู้ช่วยเหลือให้สามารถดูภายหลังได้
P4	ส่วนต่อประสานผู้ใช้ของโปรแกรมวิดีโอคอลที่มีอยู่ในปัจจุบันยังไม่รองรับกับการใช้งานของคนตาบอด	ออกแบบส่วนต่อประสานผู้ใช้ให้สอดคล้องกับการทำงานของเทคโนโลยีอ่านหน้าจอบนโทรศัพท์มือถือสำหรับคนตาบอด

ตารางที่ 3.1 ปัญหาและแนวทางแก้ไข

2. การออกแบบระบบและรูปแบบการทำงาน

ภาพรวมของการใช้งานในการสื่อสารจะเป็นไปตามรูปที่ 3.2 ซึ่งเห็นได้ว่าระบบประกอบด้วย 3 ส่วน ได้แก่ส่วนโปรแกรมที่ใช้งานในระบบจำนวน 2 ส่วน คือส่วนโปรแกรมสำหรับคนตาบอดที่ใช้งานผ่านเว็บเบราว์เซอร์บนโทรศัพท์มือถือที่ใช้ระบบปฏิบัติการแอนดรอย์ และส่วนโปรแกรมสำหรับผู้ช่วยเหลือที่ใช้งานผ่านเว็บเบราว์เซอร์บนเครื่องคอมพิวเตอร์ ที่ทำงานร่วมกับส่วนประกอบภาพทางไกล และส่วนที่ 3 เป็นส่วนการทำงานของการสื่อสารระหว่างผู้ใช้ทั้งสองประเภท โดยหลังจากเชื่อมต่อสำเร็จแล้วผู้ใช้ทั้งสองจะสื่อสารกันผ่าน WebRTC ในสถาปัตยกรรมแบบเพียร์ทูเพียร์ (peer-to-peer) โดยคนตาบอดจะส่งภาพวิดีโอและเสียงให้กับผู้ช่วยเหลือ ในขณะที่ผู้ช่วยเหลือจะส่งเพียงเสียงพูดกลับไปให้คนตาบอดเท่านั้น



รูปที่ 3.2 ภาพรวมของการใช้งานในการสื่อสารทางไกล

ในการพัฒนาโปรแกรมที่ใช้งานในระบบทั้ง 2 ส่วน มีแนวคิดจากปัญหาที่พบในการทดลองเพื่อค้นหาปัญหาในการใช้งาน โดยส่วนโปรแกรมที่ใช้งานโดยคนตาบอดมุ่งเน้นไปที่การแก้ไขปัญหาส่วนต่อประสานผู้ใช้เพื่อให้คนตาบอดสามารถใช้งานโปรแกรมต้นแบบได้ ในขณะที่ส่วนที่สอง ซึ่งเป็นส่วนโปรแกรมที่ใช้งานโดยผู้ช่วยเหลือ มุ่งเน้นไปที่การแก้ปัญหาโดยประยุกต์ใช้เทคนิคทางคอมพิวเตอร์ ซึ่งเป็นประเด็นสำคัญสำหรับงานวิจัยนี้

จากตารางที่ 3.1 พบว่าปัญหาในการใช้งานของผู้ช่วยเหลือ ซึ่งเป็นปัญหาในลำดับ P1-P3 มีสาเหตุจากสิ่งเดียวกันคือรูปภาพที่ได้จากคนตาบอดไม่เพียงพอสำหรับการช่วยเหลือ ทั้งนี้อาจเกิดจากการเคลื่อนมุมกล้องของคนตาบอดไปตลอดเวลาซึ่งอาจผ่านตำแหน่ง หรือรูปภาพที่สำคัญไปโดยผู้ช่วยเหลือไม่ทันสังเกต หรืออาจเกิดจากปัญหาในการส่งข้อมูลที่บ้างครั้งสัญญาณอ่อนเกินไปจนไม่สามารถส่งภาพวิดีโอได้ ซึ่งในการสื่อสารผ่านวิดีโอคอลนั้น ภาพวิดีโอที่ถูกส่งมาไม่ได้มีการจัดการใด ๆ เพิ่มเติม ดังนั้นแนวคิดสำหรับการออกแบบส่วนโปรแกรมที่ใช้งานโดยผู้ช่วยเหลือ จึงออกแบบให้ใช้งานร่วมกับส่วนประกอบภาพทางไกลที่นำรูปภาพที่ถูกส่งจากคนตาบอด ณ เวลาต่าง ๆ มาประมวลผลโดยประยุกต์ใช้เทคนิคการต่อภาพ (Image stitching) และนำมาแสดงให้แก่ผู้ช่วยเหลือสามารถดูภาพที่ผ่านมาได้ โดยไม่จำเป็นต้องบอกให้คนตาบอดกลับไปถ่ายภาพที่ตำแหน่งนั้น

เทคนิคการต่อภาพ (Image stitching) เป็นเทคนิคที่ใช้ในการรวมรูปภาพตั้งแต่สองรูปขึ้นไป โดยมีจุดประสงค์เพื่อแสดงรายละเอียดของภาพได้ในมุมที่กว้างขึ้น เมื่อนำมาพิจารณาร่วมกับการแสดงรูปภาพที่เก็บไว้ ณ เวลาต่าง ๆ ในกรณีที่ถูกถ่ายเมื่อคนตาบอดหยุดนิ่งอยู่กับที่ หรือ มีการเคลื่อนกล้องในแนว 2 มิติ การแสดงภาพมุมกว้างเป็นทางเลือกที่เหมาะสมสำหรับการแสดงรูปภาพที่บันทึกไว้ใน ณ เวลาใด ๆ ซึ่งเมื่อแสดงรูปภาพให้อยู่ในระนาบ 2 มิติได้แล้ว ผู้ช่วยเหลือก็จะมีมุมมองที่มากขึ้น และสามารถดูยังจุดต่าง ๆ ที่ผ่านมาได้ โดยไม่จำเป็นต้องบอกให้คนตาบอดหันกลับไปยังจุดต้องการได้ด้วยเช่นกัน

3. การแก้ไขปัญหาเรื่องส่วนต่อประสานผู้ใช้สำหรับคนตาบอด

ปัญหาหลักสำหรับโปรแกรมที่ใช้งานโดยคนตาบอดคือส่วนต่อประสานผู้ใช้ที่ไม่รองรับกับการทำงานของเทคโนโลยีอ่านหน้าจอบนโทรศัพท์มือถือ ดังนั้นในงานวิจัยนี้จึงมีการทำการทดลองเพิ่มเติมอีก 2 ครั้งเพื่อศึกษารูปแบบการใช้งานของคนตาบอด และนำปัญหามาวิเคราะห์เพื่อสรุปเป็นหลักการสำหรับการออกแบบส่วนต่อประสานผู้ใช้ที่เหมาะสม

การแก้ไขปัญหาเรื่องส่วนต่อประสานผู้ใช้สำหรับคนตาบอด แบ่งวิธีการดำเนินงานเป็น 2 ส่วนใหญ่ ได้แก่ทดสอบการใช้งานโปรแกรมวิดีโอคอลของคนตาบอด และการสรุปหลักการสำหรับการออกแบบส่วนต่อประสานผู้ใช้สำหรับคนตาบอดโดยอ้างอิงและเพิ่มเติมจาก WCAG 2.0

ส่วนการทดสอบการใช้งานโปรแกรมวิดีโอคอลของคนตาบอดประกอบด้วยขั้นตอนย่อย 3 ขั้นตอน ได้แก่

1. การกำหนดกิจกรรมที่ผู้ใช้จะต้องทำ
2. การออกแบบวิธีการวัดผลทดลองและการวัดผล
3. การทดลองและการวัดผล

3.1 การทดสอบการใช้งานโปรแกรมวิดีโอคอลของคนตาบอด

การทดสอบการใช้งานโปรแกรมวิดีโอคอลของคนตาบอดในงานนี้แบ่งออกเป็น 2 การทดลอง ซึ่งเป็นการทดสอบการใช้งานโปรแกรมวิดีโอคอลบนโทรศัพท์มือถือที่ใช้ระบบปฏิบัติการแอนดรอยด์ จำนวน 6 โปรแกรม ได้แก่ Line, Skype, Tango, WeChat, Hangout และ Viber เพื่อค้นหาปัญหาในการใช้งาน

3.1.1 การกำหนดกิจกรรมที่ผู้ใช้จะต้องทำ

การทดลองทั้ง 2 รูปแบบเป็นการจำลองสถานการณ์การใช้งานโปรแกรมวิดีโอคอลของคนตาบอด ซึ่งในงานวิจัยนี้ได้กำหนดกิจกรรมทั้งหมดออกเป็น 7 กิจกรรมที่เกี่ยวข้องกับการใช้งานวิดีโอคอลของคนตาบอด เริ่มตั้งแต่การสมัครสมาชิก, การเพิ่มรายชื่อเพื่อน, กระบวนการต่าง ๆ ที่เกิดขึ้นในการใช้งานในการพูดคุย ได้แก่ การรับสายเรียกเข้า, การโทรออกหาเพื่อน, การสลับกล้องที่ใช้ในการแสดงผลภาพ, การเชื่อมต่อใหม่หากการเชื่อมต่อหลุด และการวางสาย

3.1.2 ออกแบบวิธีการวัดผลทดลองและการวัดผล

การทดลองทั้ง 2 รูปแบบเป็นการทดลองที่มีพื้นฐานอยู่บนหลักการของ Cognitive Walkthrough โดยเป็นการทดลองใช้งานโปรแกรมวิดีโอคอลตามกิจกรรมทั้ง 7 กิจกรรมที่กำหนดขึ้น ซึ่งมีความแตกต่างกันที่ตัวผู้ทำการทดลอง และการวัดผล

การทดลอง ก. เป็นการทดลองใช้งานโดยคนตาบอด โดยใช้ผู้ทดลองที่เป็นคนตาบอดจำนวน 1 คน ซึ่งเป็นผู้ที่คุ้นเคยกับการใช้งานโปรแกรมบนโทรศัพท์มือถือด้วยเทคโนโลยีอ่านหน้าจอ โดยในการทดลองนี้ ผู้วิจัยทำหน้าที่เป็นผู้สังเกตการทดลอง และ คอยช่วยเหลือคนตาบอดหากเกิดข้อผิดพลาด การทดลองนี้วัดประสิทธิภาพในรูปแบบของเวลาที่คนตาบอดใช้ในแต่ละกิจกรรม และวัดประสิทธิผลในรูปแบบของจำนวนของความผิดพลาด (Error) ที่เกิดขึ้นจากการทำงานที่ไม่เป็นไปตามรูปแบบที่กำหนดไว้ โดยนิยามของข้อผิดพลาดนี้ คือข้อผิดพลาดที่เกิดขึ้นจากผู้ใช้งานที่โดยปกติแล้วกิจกรรมนั้นสามารถทำสำเร็จได้ ในขณะที่ข้อผิดพลาดที่เกิดจากตัวโปรแกรมที่ไม่สามารถรองรับการทำงานของเทคโนโลยีอ่านหน้าจอได้จะถือว่าเป็นกิจกรรมที่ล้มเหลวซึ่งจะสิ้นสุดการทดสอบกิจกรรมนั้น และ ผู้สังเกตการณ์จะช่วยเหลือจนถึงจุดเริ่มต้นของกิจกรรมถัดไป และให้คนตาบอดทำการทดลองในกิจกรรมต่อไป

การทดลอง ข. เป็นการทดลองเพื่อวิเคราะห์ปัญหาการใช้งานของคนตาบอด โดยเป็นการทดลองที่ทำโดยผู้วิจัยที่ใช้งานเช่นเดียวกับคนตาบอดร่วมกับการวิเคราะห์ส่วนต่าง ๆ ของหน้าจอ การทดลองนี้จะทำกิจกรรมต่าง ๆ ให้สำเร็จโดยไม่สนใจความล้มเหลว และ ความผิดพลาดที่เกิดขึ้น โดยสำหรับการวัดผล จะบันทึกผลเป็นจำนวนของการกระทำที่เป็นใจความหลักของกิจกรรม (X1) และการกระทำที่เกี่ยวข้องกับการทำกิจกรรมทั้งหมดใน 1 กิจกรรม (X2) จากนั้นจะคำนวณค่าการกระทำที่เสียเปล่า (Y) จาก X1 และ X2 ตามสมการที่ 3.1

$$Y = \frac{(X2-X1)}{X1} \tag{3.1}$$

การนับจำนวนของการกระทำที่เกี่ยวข้องกับการกระทำที่เป็นใจความหลักของกิจกรรม (X1) นับจากการกระทำที่เป็นหัวใจสำคัญของกิจกรรมนั้น ๆ ซึ่งประกอบด้วยการกดปุ่มต่าง ๆ รวมไปถึงส่วนที่เป็นอินพุตจากผู้ใช้งาน ในขณะที่การกระทำที่เกี่ยวข้องกับการทำกิจกรรมทั้งหมดใน 1 กิจกรรม (X2) เป็นการใช้งานทั้งหมดของคนตาบอดที่ใช้ในการนำทางไปยังการกระทำที่เกี่ยวข้องกับการทำกิจกรรมต่าง ๆ โดยในทุกครั้งที่มีการรับอินพุตจะถือว่าเป็น 1 การโต้ตอบกับโปรแกรมเสมอ

3.1.3 การทดลองและการวัดผล

การทดลองใช้งานโดยคนตาบอดให้ผู้ทดลองใช้โปรแกรมทั้ง 6 โปรแกรม ทำกิจกรรมต่าง ๆ ตามลำดับ ได้แก่ สมัครสมาชิก, เพิ่มเพื่อนหนึ่งคน และเลือกเพื่อนคนนั้นเพื่อโทรวิดีโอคอล จากนั้นจะให้ผู้ทดลองสลับกล้องที่ใช้แสดงผล และวางสายวิดีโอคอล จากนั้นผู้สังเกตการณ์จะเปิดโปรแกรมไปที่หน้าหลัก และให้ผู้ทดสอบรับสายเรียกเข้า จากนั้นจะตัดสัญญาณอินเทอร์เน็ตที่ใช้ เพื่อให้ผู้ทดสอบทำกิจกรรมเชื่อมต่อใหม่หากกรณีสัญญาณหลุด ซึ่งหากผู้ทดสอบไม่สามารถทำกิจกรรมก่อนหน้าสำเร็จได้ ผู้สังเกตจะเป็นผู้ช่วยเหลือให้สามารถไปยังกิจกรรมต่อไปได้ โดยนับว่ากิจกรรมนั้นเป็นกิจกรรมที่ล้มเหลว

การทดลองเพื่อวิเคราะห์ปัญหาการใช้งานของคนตาบอดจะเป็นการทดลองกับผู้วิจัยซึ่งใช้งานโปรแกรมโดยทำงานร่วมกับเทคโนโลยีอ่านหน้าจอ โดยในการทดลองนี้จะไม่สนใจความล้มเหลวที่เกิดขึ้นเพื่อทำกิจกรรมนั้นให้สำเร็จ

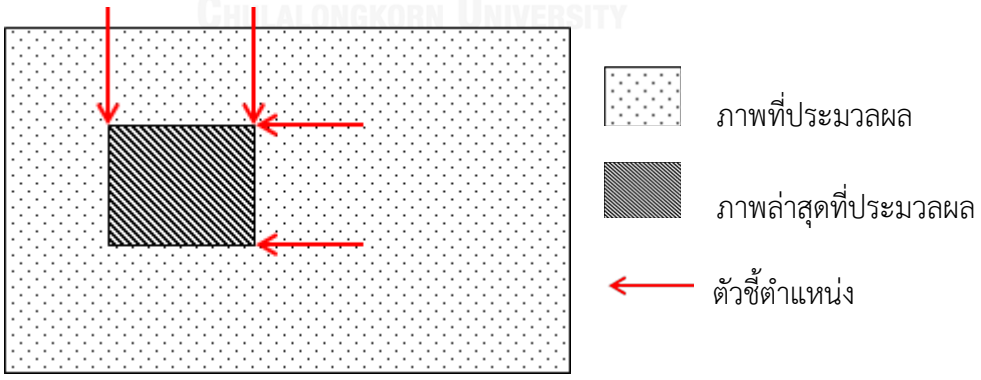
จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

การสรุปผลการทดลองบันทึกอยู่ในรูปแบบของตาราง และเอกสารที่ได้จากการสังเกต โดยบันทึกปัญหาต่าง ๆ ที่พบไว้ เพื่อนำไปวิเคราะห์ในขั้นตอนถัดไป

4. การพัฒนาส่วนการประกอบภาพทางไกล

ส่วนการประกอบภาพทางไกลพัฒนาโดยใช้ไลบรารี่ (Library) ของ OpenCV ในการต่อรูปภาพ (Image stitching) โดยมีข้อมูลนำเข้าเป็นเฟรมวิดีโอล่าสุดของคนตาบอดที่อยู่ในรูปแบบของภาพนิ่ง โดยส่วนการประกอบภาพทางไกลนี้มีความแตกต่างจากการต่อรูปภาพโดยทั่วไป เนื่องจากการต่อรูปภาพโดยทั่วไปตามหลักการของ OpenCV ใช้การหมุนรูปภาพที่นำเข้ามาใหม่ให้มีตำแหน่งและทิศทางตรงกับรูปภาพเดิม และทำการวางรูปภาพเดิมทับลงไปบนรูปภาพใหม่ที่ทำการหมุนแล้ว ซึ่งไม่สามารถนำมาใช้ได้กับงานวิจัยนี้ที่บางส่วนของรูปภาพจะถูกอัพเดทด้วยเฟรมล่าสุดที่ได้รับจากคนตาบอดที่ควรนำมาแสดงแก่ผู้ช่วยเหลือ นอกจากนั้นภาพนิ่งที่นำมาประมวลผลอาจไม่สามารถเชื่อมต่อกับภาพเก่าได้ด้วยเช่นกัน

ในการต่อรูปภาพได้เลือกใช้อัลกอริทึม SIFT ในการหาคุณลักษณะ (Feature) ต่าง ๆ ที่ปรากฏอยู่ในภาพ และการจับคู่คุณลักษณะ (Feature matching) ซึ่งให้ความแม่นยำสูงแต่ในขณะเดียวกันก็ใช้เวลาประมวลผลค่อนข้างมาก และแปรผันตามขนาดรูปภาพที่เพิ่มขึ้น จึงแก้ไขขั้นตอนในการจับคู่คุณลักษณะ โดยจับคู่รูปภาพนำเข้าใหม่กับรูปภาพที่เป็นส่วนล่าสุดที่ประมวลผลไปก่อนหน้าเพื่อลดจำนวนของจุดคุณลักษณะในการจับคู่ ทั้งนี้การอ้างอิงตำแหน่งภาพส่วนล่าสุดที่ประมวลผลจะใช้ตัวชี้ (pointer) ซึ่งเก็บค่า 4 ค่า ได้แก่ จุดเริ่มต้นในแนวแกน x, จุดสิ้นสุดในแนวแกน x, จุดเริ่มต้นในแนวแกน y และจุดสิ้นสุดในแนวแกน y ตามตัวอย่างที่แสดงในรูปที่ 3.3



รูปที่ 3.3 ตัวอย่างการอ้างอิงตำแหน่งภาพส่วนล่าสุดโดยใช้ตัวชี้ (pointer)

รูปที่ 3.4 แสดงตัวอย่างภาพมุมกว้างที่แสดงแก่ผู้ช่วยเหลือ ซึ่งได้ออกแบบการทำงานเพิ่มเติม โดยมีรายละเอียดของการทำงานต่าง ๆ แสดงในตารางที่ 3.2

รูปที่ 3.4 ตัวอย่างภาพมุมกว้างที่แสดงแก่ผู้ช่วยเหลือ

ลำดับ	รายละเอียดการทำงาน
1	สร้างกรอบสี่เหลี่ยมล้อมรอบภาพที่เป็นส่วนล่าสุดที่ประมวลผล
2	แสดงภาพขยาย โดยอ้างอิงจากการคลิกเมาส์ในตำแหน่งที่ต้องการของผู้ช่วยเหลือ
3	แสดงภาพด้วยภาพล่าสุดที่ได้จากคนตาบอดในตำแหน่งนั้นเสมอ
4	แสดงข้อมูลเวลาของภาพโดยใช้โทนสี

ตารางที่ 3.2 แสดงรายละเอียดการทำงานต่าง ๆ ในส่วนการแสดงภาพมุมกว้าง

ณ ตำแหน่งต่าง ๆ ของภาพที่แสดงแก่ผู้ช่วยเหลือนั้น ผู้ใช้สามารถขยายภาพในส่วนที่ต้องการได้ โดยใช้การคลิกเมาส์ นอกจากนี้ได้อ้างอิงตัวชี้ทั้ง 4 ตัว ในการสร้างกรอบสี่เหลี่ยมเพื่อแสดงตำแหน่งของภาพล่าสุดที่ประมวลผลแก่ผู้ช่วยเหลือด้วยเช่นกัน

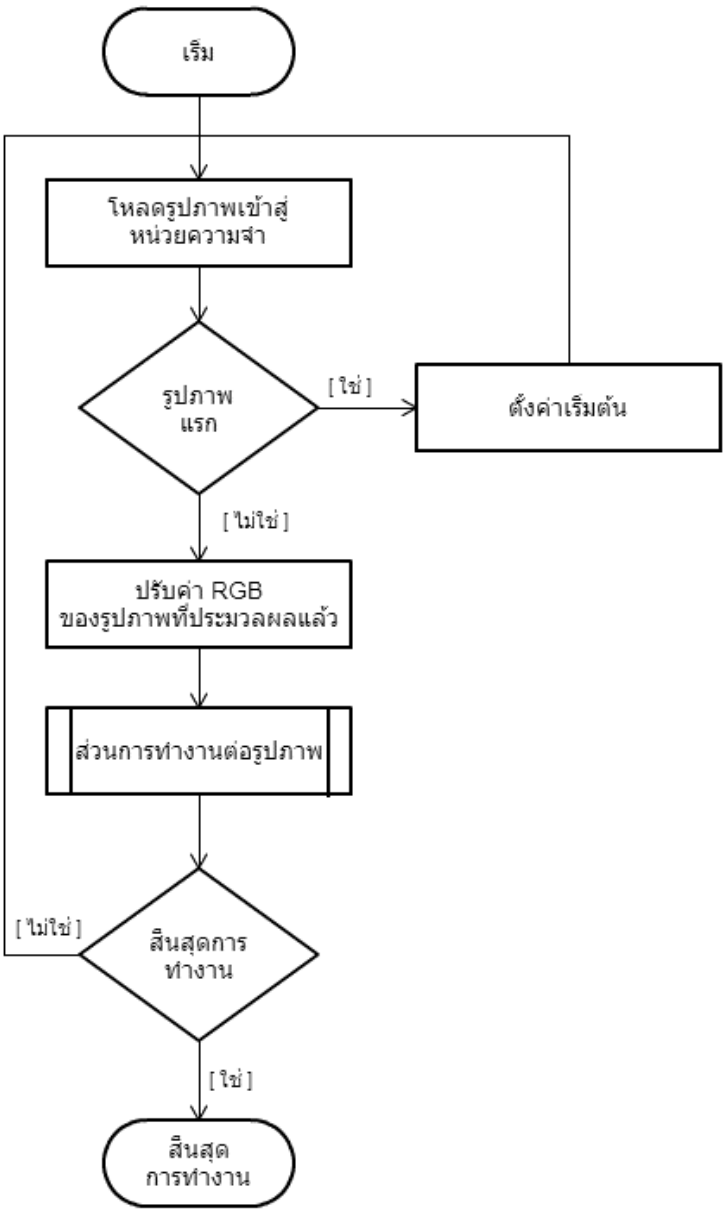
การแสดงข้อมูลเวลาในรูปแบบของโทนสีใช้การคำนวณและปรับค่า HSV ของรูปภาพที่ประมวลผลแล้วเพื่อปรับให้อยู่ในโทนสีขาว-ดำ โดยเป็นการปรับค่า S Channel ของแต่ละ Channel ในรูปภาพ ซึ่งสามารถคำนวณจากค่า Mean ที่ได้จากค่าสูงสุดและต่ำสุดของช่อง (channel) ต่าง ๆ ในช่อง RGB ตามสมการที่ (3.2) โดยในแต่ละช่องใน RGB (v3b[i]) ถูกปรับด้วยค่า Mean ร่วมกับค่าแอลฟาซึ่งเป็นค่าคงที่ตามสมการที่ (3.3)

$$\text{Mean} = \frac{Max\,(H,S,V) - Min\,(H,S,V)}{2} \tag{3.2}$$

$$v3b[i] = (v3b[i] - mean) \times \alpha + mean \;\; ; i = 0 - 2 \tag{3.3}$$

รูปที่ 3.5 แสดงขั้นตอนในการทำงานส่วนการประกอบภาพทางไกล โดยในครั้งแรกของการทำงานเป็นการตั้งค่าเริ่มต้นสำหรับการประมวลผล ประกอบไปด้วยการคัดลอกรูปภาพไปยังภาพที่ประมวลผลแล้ว การคัดลอกรูปภาพไปยังภาพล่าสุดที่ประมวลผล และการกำหนดค่าเริ่มต้นสำหรับตัวชี้ตำแหน่งล่าสุด หลังจากนั้นเมื่อมีรูปภาพใหม่เข้ามา จะทำการโหลดรูปภาพเข้าสู่หน่วยความจำของโปรแกรม ทำการปรับค่า RGB ของรูปภาพที่ประมวลผลแล้วเพื่อแสดงข้อมูลเวลาแก่ผู้ใช้ และเรียกใช้มอดูลของส่วนการทำงานต่อรูปภาพซึ่งเป็นการทำงานสำคัญของส่วนการประกอบภาพทางไกล

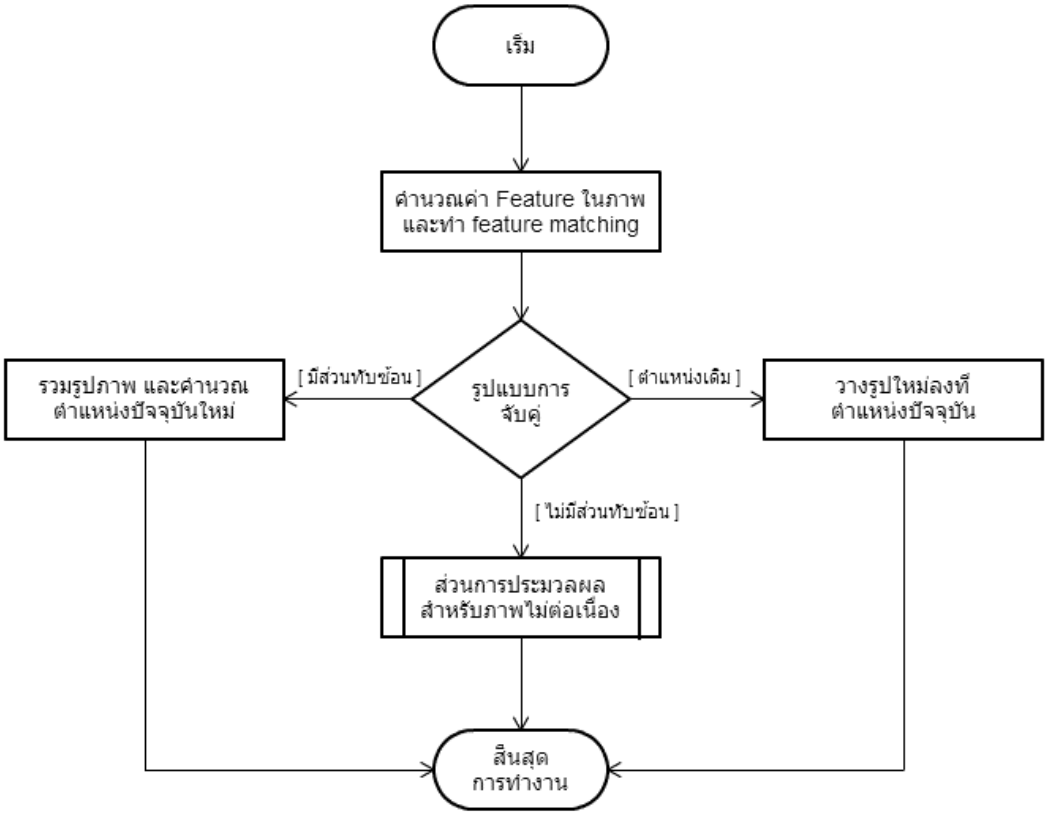
จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University



รูปที่ 3.5 ขั้นตอนในการทำงานส่วนการประกอบภาพทางไกล

รูปที่ 3.6 แสดงขั้นตอนในการทำงานของมอดูลของส่วนการทำงานต่อรูปภาพ ซึ่งมีหน้าที่หลักในการหาจุดคุณลักษณะ (Feature point) ในรูปภาพนำเข้าใหม่ และนำมาเปรียบเทียบกับส่วนภาพล่าสุดที่ประมวลผลเพื่อจับคู่คุณลักษณะ (Feature matching) โดยหลังจากจับคู่จุดคุณลักษณะที่มีในทั้งสองภาพได้แล้วจะคำนวณความคลาดเคลื่อนของตำแหน่งของจุดในสองภาพ ซึ่งสามารถแบ่งผลลัพธ์การคำนวณได้เป็น 3 กรณีได้แก่

1. ภาพใหม่อยู่ในตำแหน่งเดิม
2. ภาพใหม่มีบางส่วนทับซ้อนกับภาพที่ประมวลผลแล้ว
3. ภาพใหม่ไม่มีส่วนทับซ้อนกับภาพที่ประมวลผลแล้ว



รูปที่ 3.6 ขั้นตอนในการทำงานของมอดูลของส่วนการทำงานต่อรูปภาพ

4.1 ภาพใหม่อยู่ในตำแหน่งเดิม

กรณีคำนวณแล้วพบว่าภาพใหม่อยู่ในตำแหน่งเดิมจะทำการวางรูปภาพใหม่ลงไปในตำแหน่งเดิมและปรับค่าส่วนภาพล่าสุดที่ประมวลผลใหม่เพื่อรอการประมวลผลกับภาพนำเข้าถัดไป

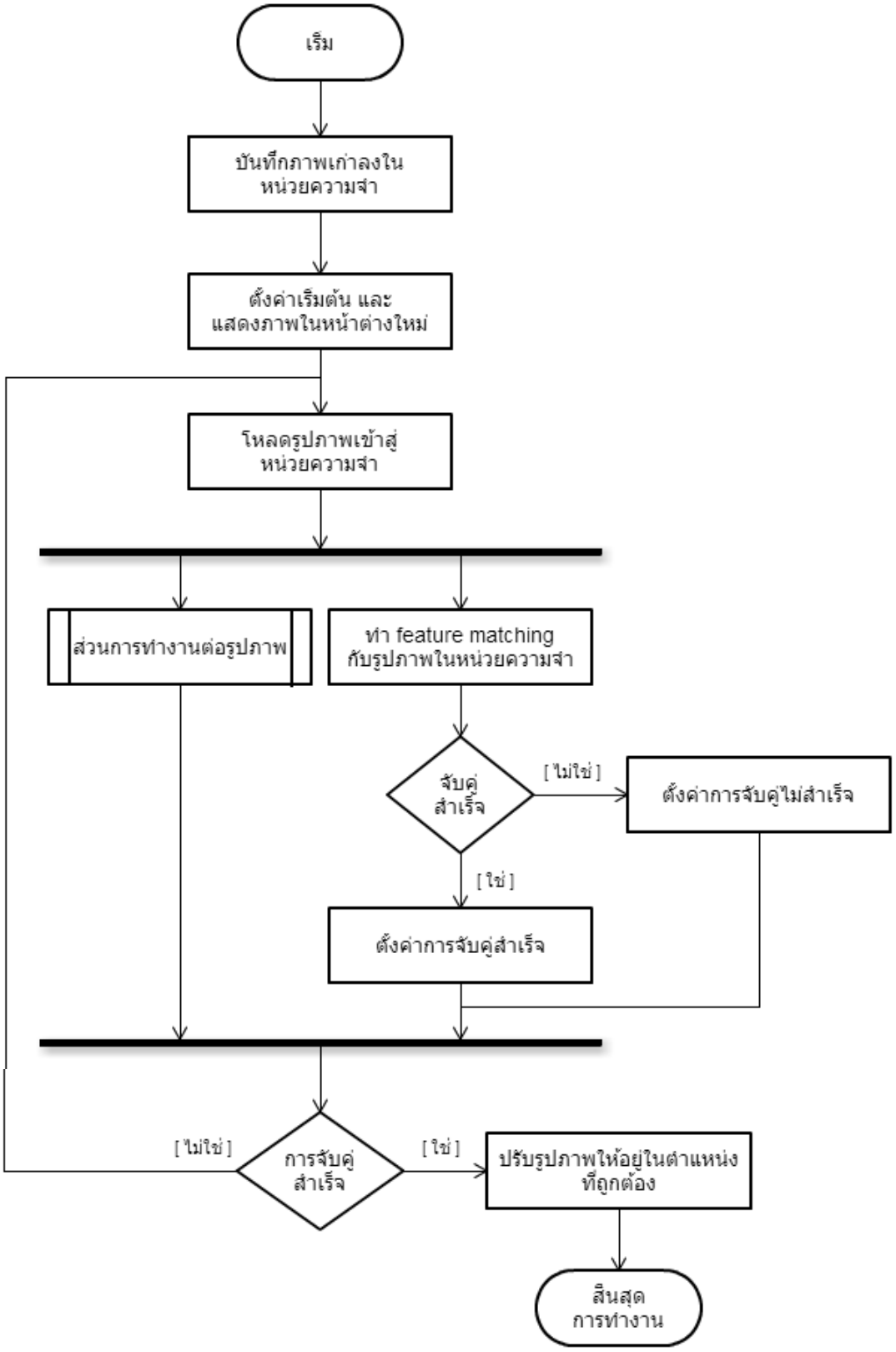
4.2 ภาพใหม่มีบางส่วนทับซ้อนกับภาพที่ประมวลผลแล้ว

กรณีคำนวณแล้วพบว่าภาพใหม่มีบางส่วนทับซ้อนกับภาพที่ประมวลผลแล้วจะทำการรวมรูปภาพโดยคำนวณ Homography matrix และทำการหมุนภาพใหม่ด้วย Homography matrix เพื่อวางผลไปในภาพที่ประมวลผลแล้ว จากนั้นทำการปรับค่าส่วนภาพล่าสุดที่ประมวลผล และคำนวณการเปลี่ยนแปลงในแนวแกน x (Δx) และการเปลี่ยนแปลงในแนวแกน y (Δy) บันทึกไว้ในหน่วยความจำของโปรแกรม

4.3 ภาพใหม่ไม่มีส่วนทับซ้อนกับภาพที่ประมวลผลแล้ว

กรณีคำนวณแล้วพบว่าภาพใหม่ไม่มีส่วนทับซ้อนกับภาพที่ประมวลผลแล้ว จะเรียกใช้มอดูลการประมวลผลสำหรับภาพไม่ต่อเนื่อง ซึ่งมีการนำการประมวลผลแบบเทรด (Thread) เข้ามาใช้งานตามแสดงในรูปที่ 3.7

การประมวลผลสำหรับภาพไม่ต่อเนื่องเริ่มต้นด้วยการคัดลอกข้อมูลการประมวลผลเก่าทั้งหมด ได้แก่ ภาพที่ประมวลผล ภาพล่าสุดที่ประมวลผล และ ตัวชี้ตำแหน่ง ลงในหน่วยความจำของโปรแกรม จากนั้นจะเริ่มต้นการทำงานใหม่โดยถือว่ารูปที่ไม่สามารถเชื่อมต่อกับภาพเก่าได้ เป็นภาพแรกสุดสำหรับการประมวลผล โดยเมื่อมีรูปภาพใหม่เข้ามาก็จะแบ่งการทำงานออกเป็น 2 เทรด โดยเทรดแรกเป็นเทรดหลักที่ทำหน้าที่เชื่อมต่อภาพโดยเรียกใช้มอดูลของส่วนการทำงานต่อรูปภาพและแสดงผลแก่ผู้ใช้ ในขณะที่เทรดที่สองจะทำงานเปรียบเสมือนการทำงานเบื้องหลัง มีหน้าที่ในการพยายามเชื่อมต่อรูปภาพที่นำเข้ากับส่วนภาพล่าสุดที่ประมวลผลเดิมที่เก็บไว้ในหน่วยความจำ ซึ่งเมื่อการทำงานของทั้งสองเทรดจบลง จะตรวจสอบว่าภาพที่นำเข้าใหม่สามารถเชื่อมต่อกับส่วนภาพล่าสุดที่ประมวลผลเดิมที่เก็บไว้ในหน่วยความจำได้หรือไม่ หากเชื่อมต่อได้จะทำการวางรูปภาพทั้งหมดในตำแหน่งที่ถูกต้อง และสิ้นสุดการทำงาน แต่หากไม่สามารถเชื่อมต่อได้ก็จะรับรูปต่อไปมาประมวลผลกับทั้งสองเทรดดังเดิม



รูปที่ 3.7 ขั้นตอนในการทำงานของมอดูลการประมวลผลสำหรับภาพไม่ต่อเนื่อง

การปรับรูปภาพให้อยู่ในตำแหน่งที่ถูกต้องใช้การอ้างอิงจากตัวชี้ตำแหน่งที่เก็บไว้ในหน่วยความจำ และผลลัพธ์การคำนวณตำแหน่งของรูปภาพมาเป็นหลักในการรวมทั้งสองภาพเข้าด้วยกันแทนการคำนวณจุดคุณลักษณะในภาพทั้งสอง และทำการจับคู่จุดคุณลักษณะ เพื่อลดเวลาในการประมวลผล โดยใช้การเปรียบเทียบตัวชี้ตำแหน่งของภาพล่าสุดที่ประมวลผลในภาพทั้งสอง ซึ่งหากปรับให้มีค่าเท่ากันได้ ก็สามารถรวมภาพทั้งสองได้โดยไม่จำเป็นต้องคำนวณเพิ่มเติม ซึ่งการปรับค่าตัวชี้ตำแหน่งนี้ จะคำนวณความคลาดเคลื่อนในแนวแกน x (Δx) ตามสมการที่ 3.4 และความคลาดเคลื่อนในแนวแกน y (Δy) ตามสมการที่ 3.5

$$\Delta x = \begin{cases} \mathrm{X1}_{\text{เก่า}} - \mathrm{X1}_{\text{ปัจจุบัน}} & \mathrm{X1}_{\text{เก่า}} > \mathrm{X1}_{\text{ปัจจุบัน}} \\ \mathrm{X2}_{\text{ปัจจุบัน}} - \mathrm{X2}_{\text{เก่า}} & \mathrm{X1}_{\text{เก่า}} < \mathrm{X1}_{\text{ปัจจุบัน}} \end{cases} \tag{3.4}$$

$$\Delta y = \begin{cases} \mathrm{Y1}_{\text{เก่า}} - \mathrm{Y1}_{\text{ปัจจุบัน}} & \mathrm{Y1}_{\text{เก่า}} > \mathrm{Y1}_{\text{ปัจจุบัน}} \\ \mathrm{Y2}_{\text{ปัจจุบัน}} - \mathrm{Y2}_{\text{เก่า}} & \mathrm{Y1}_{\text{เก่า}} < \mathrm{Y1}_{\text{ปัจจุบัน}} \end{cases} \tag{3.5}$$

โดย X1 = พิกัดของจุดเริ่มต้นของภาพล่าสุดที่ประมวลผลในแนวแกน X

X2 = พิกัดของจุดสิ้นสุดของภาพล่าสุดที่ประมวลผลในแนวแกน X

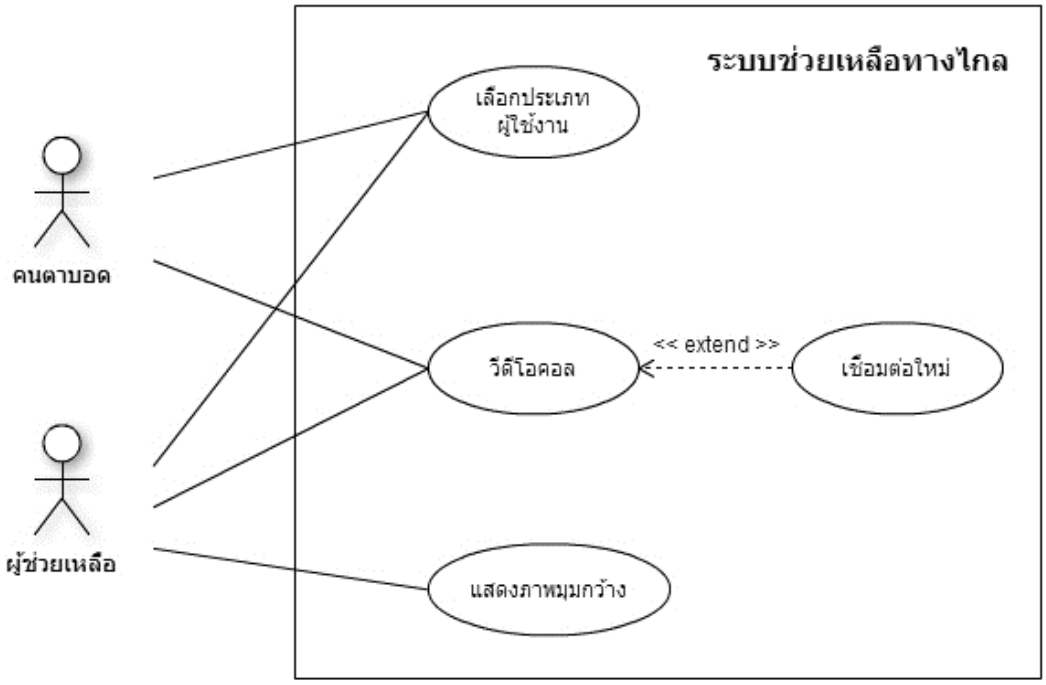
Y1 = พิกัดของจุดเริ่มต้นของภาพล่าสุดที่ประมวลผลในแนวแกน Y

Y2 =พิกัดของจุดสิ้นสุดของภาพล่าสุดที่ประมวลผลในแนวแกน Y

หลังจากการคำนวณความคลาดเคลื่อนในแนวแกน x (Δx) และความคลาดเคลื่อนในแนวแกน y (Δy) แล้วจะทำการปรับขนาดภาพที่ประมวลผล โดยปรับให้ภาพล่าสุดที่ประมวลผลอยู่ในตำแหน่งพิกัดเดียวกันกับพิกัดของตัวชี้ของภาพเก่า จากนั้นจะทำการรวมภาพทั้งสองเข้าด้วยกันโดยใช้การอ้างอิงจากตัวชี้ที่ปรับค่าแล้ว

5. การพัฒนาต้นแบบของโปรแกรมวิดีโอคอลที่ใช้งานในระบบช่วยเหลือคนตาบอด

โปรแกรมที่พัฒนาขึ้นอยู่ในรูปแบบของโปรแกรมต้นแบบที่สามารถแก้ไขปัญหาที่เกิดขึ้นจากการใช้งานในระบบช่วยเหลือคนตาบอดได้ โดยโปรแกรมที่พัฒนาขึ้นแบ่งเป็น 2 รูปแบบตามประเภทของผู้ใช้งาน ซึ่งรูปที่ 3.8 แสดงแผนภาพแสดงการทำงานของผู้ใช้ระบบที่ประกอบไปด้วยการเลือกประเภทผู้ใช้งาน การใช้งานวิดีโอคอล และการแสดงภาพมุมกว้าง โดยรายละเอียดของการทำงานต่าง ๆ อธิบายในตารางที่ 3.3 – 3.6



รูปที่ 3.8 แผนภาพแสดงการทำงานของผู้ใช้ระบบ

1. เลือกประเภทผู้ใช้งาน		
Actors:	คนตาบอด และผู้ช่วยเหลือ	
Descriptions:	ผู้ใช้เลือกประเภทผู้ใช้งานสำหรับการเข้าสู่ระบบ	
Preconditions:	-	
Postconditions:	-	
Normal Flow:	1	ผู้ใช้เข้าสู่หน้าหลักของเว็บไซต์
	2	ผู้ใช้เลือกรายการ “ฉันเป็นผู้ช่วยเหลือ”
	3	ระบบบันทึกประเภทผู้ใช้งานและทำการเชื่อมต่อเพื่อรับรหัสที่ใช้ในการเชื่อมต่อ
Alternative Flow:	2a	ผู้ใช้เลือกรายการ “ฉันเป็นคนตาบอด”

ตารางที่ 3.3 รายละเอียดของการเลือกประเภทผู้ใช้งาน

2. วิดีโอคอล		
Actors:	คนตาบอด และผู้ช่วยเหลือ	
Descriptions:	ผู้ใช้ใช้งานวิดีโอคอล	
Preconditions:	ผู้ใช้เลือกประเภทผู้ใช้งานสำหรับการเข้าสู่ระบบแล้ว	
Postconditions:	-	
Normal Flow:	1	ระบบจับคู่ผู้ใช้ประเภทตรงข้ามเพื่อทำการเชื่อมต่อ
	2	เริ่มต้นการใช้งานวิดีโอคอล
Alternative Flow:	2a	หากสัญญาณในการเชื่อมต่อหลุด เริ่มการทำงานของการเชื่อมต่อใหม่

ตารางที่ 3.4 รายละเอียดของการใช้งานวิดีโอคอล

3. เชื่อมต่อใหม่		
Actors:	คนตาบอด และผู้ช่วยเหลือ	
Descriptions:	เชื่อมต่อใหม่หากสัญญาณหลุด	
Preconditions:	ผู้ใช้อยู่ระหว่างการใช้งานวิดีโอคอล และขาดการเชื่อมต่อ	
Postconditions:	-	
Normal Flow:	1	ระบบตรวจสอบสัญญาณ และทำการเชื่อมต่อใหม่
	2	เริ่มต้นการใช้งานวิดีโอคอล
Alternative Flow:	1a	แจ้งเตือนแก่ผู้ใช้หากไม่มีสัญญาณอินเตอร์เน็ท

ตารางที่ 3.5 รายละเอียดของการเชื่อมต่อใหม่

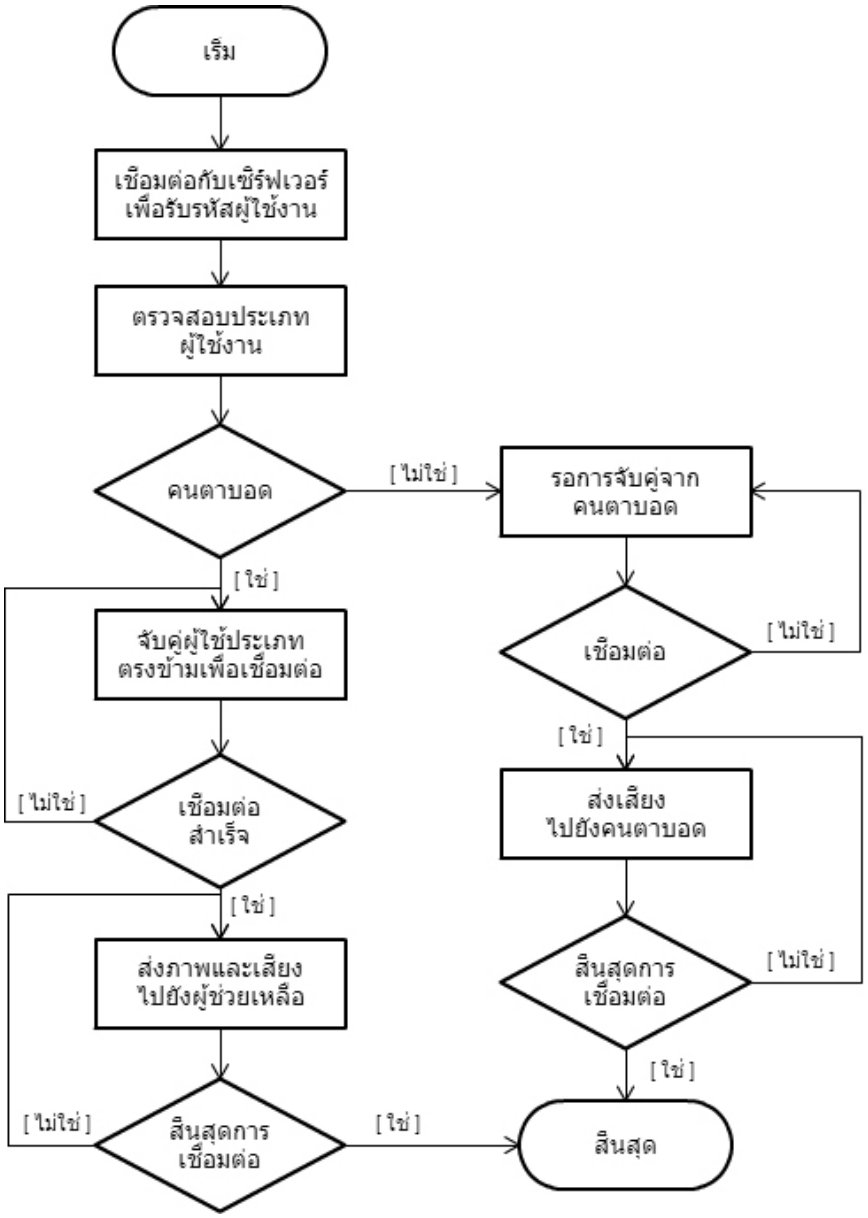
4. แสดงภาพมุมกว้าง		
Actors:	ผู้ช่วยเหลือ	
Descriptions:	ผู้ใช้ใช้งานวิดีโอคอล	
Preconditions:	ผู้ใช้อยู่ระหว่างการใช้งานวิดีโอคอล	
Postconditions:	-	
Normal Flow:	1	ผู้ใช้เลือกรายการ “แสดงภาพมุมกว้าง”
	2	ระบบจับภาพวิดีโอ และประมวลผลภาพมุมกว้างเพื่อแสดงแก่ผู้ใช้
Alternative Flow:	-	

ตารางที่ 3.6 รายละเอียดของการแสดงภาพมุมกว้าง

โปรแกรมสำหรับผู้ใช้ทั้งสองประเภทนี้ใช้การสื่อสารผ่านส่วนการให้บริการสื่อสารเดียวกัน แต่อย่างไรก็ตามโปรแกรมทั้งสองรูปแบบมีความแตกต่างกันในระดับการทำงานย่อยภายใน ได้แก่ส่วนต่อประสานผู้ใช้สำหรับคนตาบอด และส่วนการประกอบภาพทางไกลสำหรับผู้ช่วยเหลือ โดยการเลือกรูปแบบการทำงานของโปรแกรมจะเกิดจากการระบุประเภทผู้ใช้งานจากผู้ใช้ในตอนเริ่มต้นของโปรแกรม โดยสำหรับผู้ช่วยเหลือจะมีการทำงานเพิ่มเติมในส่วนของการแสดงภาพมุมกว้าง

การสื่อสารระหว่างผู้ใช้ทั้งสองประเภท เป็นการใช้งานวิดีโอคอลผ่านเว็บเบราว์เซอร์ ซึ่งพัฒนาโดยใช้โอเพนซอร์ส (Open source) ของ EasyRTC ซึ่งเป็นโอเพนซอร์สแบบเต็มรูปแบบ (Full stack open source) ของ WebRTC โดยใช้ HTML ร่วมกับจาวาสคริปต์ (JavaScript) ส่วนการให้บริการการสื่อสารระหว่างผู้ใช้นี้แบ่งออกเป็น 2 ส่วน ได้แก่ ส่วนผู้ให้บริการ (Server) ซึ่งจะถูกรันด้วย Node.js เพื่อจัดการการติดต่อผ่านซ็อกเก็ต (Socket) และส่วนผู้รับบริการ (Client) ที่รันผ่านเว็บเบราว์เซอร์ โดยการทำงานมีขั้นตอนดังแสดงในรูปที่ 3.9

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University



รูปที่ 3.9 ขั้นตอนการทำงานของส่วนการให้บริการสื่อสารระหว่างผู้ใช้ทั้งสองประเภท

การทำงานส่วนการให้บริการการสื่อสารระหว่างผู้ใช้ ส่วนผู้รับบริการ (Client) ที่เป็นเว็บเบราว์เซอร์จะรับอินพุตจากผู้ใช้ซึ่งเป็นการระบุประเภทผู้ใช้งาน จากนั้นจะทำการเชื่อมต่อกับส่วนผู้ให้บริการ (Server) เพื่อขอรับรหัสผู้ใช้งาน (EasyRTC ID) ซึ่งเป็นรหัสผู้ใช้แบบครั้งเดียวสำหรับการแทนตัวตนในระบบ จากนั้นเป็นส่วนของการจับคู่ผู้ใช้งาน ซึ่งเป็นการทำงานโดยอัตโนมัติ เริ่มต้นจากการตรวจสอบประเภทผู้ใช้งาน หากเข้าใช้งานในรูปแบบคนตาบอดจะทำการจับคู่กับผู้ใช้อื่นที่เข้าใช้งานในรูปแบบของผู้ช่วยเหลือ ซึ่งหากเชื่อมต่อสำเร็จก็จะเริ่มการทำงานในส่วนของวิดีโอคอล โดยใช้บริการของ EasyRTC ที่ส่งภาพวิดีโอและเสียงไปยังผู้ช่วยเหลือ ในทางตรงกันข้ามหากเข้าใช้งานในรูปแบบของผู้ช่วยเหลือ จะรอจนได้รับสัญญาณการขอเชื่อมต่อจากคนตาบอด ซึ่งหากเชื่อมต่อสำเร็จก็จะเริ่มการทำงานในส่วนของวิดีโอคอลเช่นกัน แต่ต่างจากคนตาบอดในด้านที่การส่งข้อมูล ซึ่งจะส่งเพียงเสียงกลับไปเท่านั้น

5.1 การพัฒนาโปรแกรมสำหรับคนตาบอด

โปรแกรมที่ใช้งานโดยคนตาบอดประกอบไปด้วยการทำงานพื้นฐานต่าง ๆ ที่เกี่ยวข้องกับการใช้งานวิดีโอคอลเป็นหลัก ซึ่งมีการเพิ่มเติมในด้านส่วนต่อประสานผู้ใช้งานเพื่อแก้ไขปัญหาที่เกิดขึ้นในการใช้งานของคนตาบอด

การทำงานในโปรแกรมสำหรับคนตาบอดเป็นการทำงานโดยอัตโนมัติของโปรแกรมหลังจากผู้ใช้ระบุประเภทผู้ใช้งานเป็นคนตาบอดที่หน้าแรกของเว็บไซต์ ซึ่งขั้นตอนการทำงานเป็นไปตามที่ได้กำหนดไว้ในส่วนการให้บริการการสื่อสารระหว่างผู้ใช้ โดยการทำงานทั้งหมดถูกควบคุมโดย JavaScript และใช้ CSS ในการซ่อนหรือแสดงส่วนประกอบต่างๆ บนหน้าจอเพื่อให้โปรแกรมสามารถทำงานร่วมกับเทคโนโลยีอ่านหน้าจอ (Screen reader) บนโทรศัพท์มือถือได้อย่างถูกต้อง

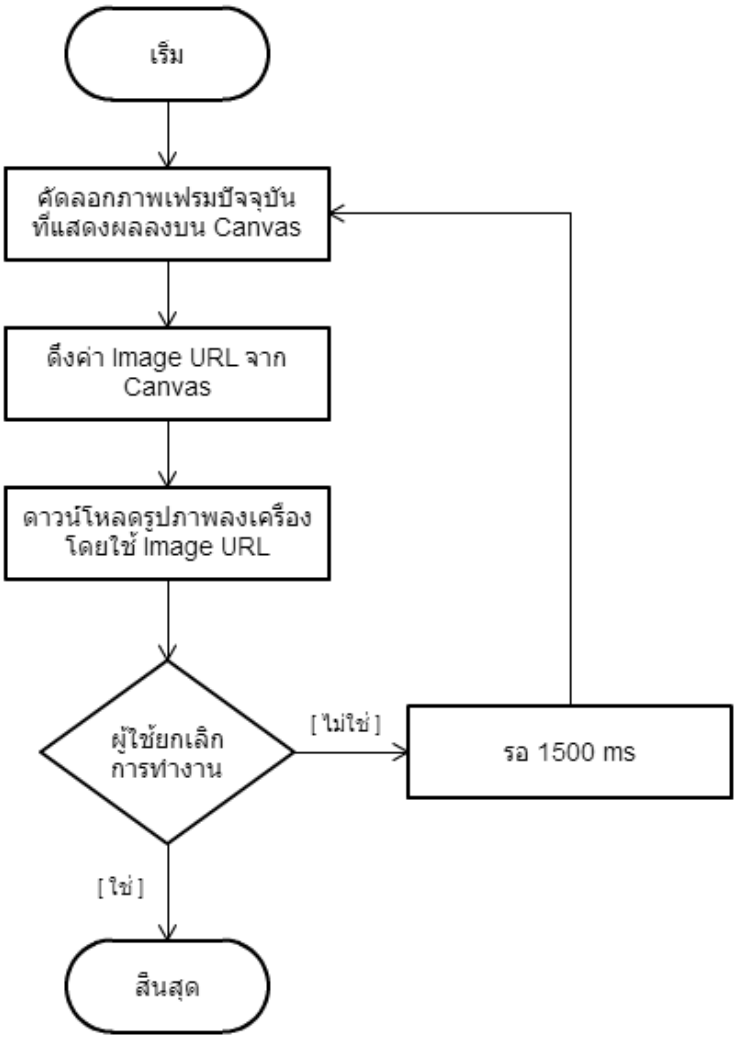
5.2 การพัฒนาโปรแกรมสำหรับผู้ช่วยเหลือ

โปรแกรมสำหรับผู้ช่วยเหลือเป็นโปรแกรมวิดีโอคอลที่ใช้งานบนเครื่องคอมพิวเตอร์ ซึ่งสามารถแบ่งได้เป็น 3 ส่วนหลัก โดยส่วนแรกเป็นส่วนของโปรแกรมวิดีโคอล ที่เป็นไปตามที่ได้กำหนดไว้ในส่วนการให้บริการการสื่อสารระหว่างผู้ใช้ ส่วนที่สองคือการทำงานที่เกี่ยวข้องกับการแปลงค่าข้อมูลวิดีโอที่ได้รับ และส่วนสุดท้ายคือส่วนประมวลผลภาพทางไกลซึ่งมีรายละเอียดตามที่กล่าวไปในหัวข้อที่ 4 ซึ่งการทำงานส่วนนี้เริ่มต้นพร้อมกับการเปิดหน้าเว็บไซต์เพื่อใช้งานวิดีโอคอลของผู้ใช้ แต่

จะอยู่ในรูปแบบการทำงานเบื้องหลัง และถูกเรียกใช้เมื่อมีการเปลี่ยนแปลงของไฟล์รูปภาพเกิดขึ้นในตำแหน่งที่เก็บภาพที่กำหนดไว้

กิจกรรมของผู้ช่วยเหลือเริ่มต้นที่การได้รับการแจ้งเตือนจากคนตาบอด ซึ่งเมื่อตอบรับการช่วยเหลือแล้วก็จะใช้งานวิดีโอคอลที่เป็นไปตามที่ได้กำหนดไว้ในส่วนการให้บริการการสื่อสารระหว่างผู้ใช้ ซึ่งระหว่างการใช้งานวิดีโอคอลผู้ใช้สามารถเลือกเปิดใช้งานภาพมุมกว้างได้ โดยการทำงานของโปรแกรมสำหรับผู้ช่วยเหลือถูกควบคุมด้วย JavaScript และมีการใช้ CSS ในการซ่อนหรือแสดงส่วนประกอบต่างๆ บนหน้าจอเช่นเดียวกันกับโปรแกรมสำหรับคนตาบอด นอกจากนี้โปรแกรมสำหรับผู้ช่วยเหลือได้เพิ่ม JavaScript สำหรับแปลงค่าข้อมูลวิดีโอที่ได้รับเพื่อใช้ในการประมวลผลภาพมุมกว้างเข้าไปด้วยเช่นกัน

เนื่องจากการส่งข้อมูลภาพผ่านโอเพนซอร์สของ EasyRTC ใช้การเข้ารหัส (Encode) ข้อมูลวิดีโอในรูปแบบของ BLOB (Binary Large Object) ซึ่งไม่สามารถดึงข้อมูลเฟรมวิดีโอมาได้โดยตรงผ่านทาง ไลบรารี่ (Library) ของ OpenCV จึงมีการเพิ่มเติมส่วนการแปลงค่าข้อมูลวิดีโอที่ได้รับให้อยู่ในรูปของภาพนิ่งเพื่อส่งไปประมวลผล โดยขั้นตอนการทำงานแสดงอยู่ในรูปที่ 3.10 โดยเมื่อผู้ใช้เปิดการใช้งานภาพมุมกว้าง โปรแกรมฝั่งผู้ช่วยเหลือจะใช้ JavaScript ในการจับรูปภาพเฟรมวิดีโอปัจจุบันที่ปรากฏอยู่บนหน้าเว็บเบราว์เซอร์ และวาดรูปภาพลงไปใน Canvas element ซึ่งเป็นคุณสมบัติทางด้านกราฟฟิกที่มาพร้อมกับ HTML5 โดยเป็นพื้นที่สี่เหลี่ยมที่ไว้สำหรับการแสดงผลกราฟฟิก หลังจากนั้น JavaScript จะทำการดึงค่า URL ของรูปภาพดังกล่าวและทำการบันทึกลงหน่วยความจำสำรองในเครื่องผู้ใช้



Chulalongkorn University

รูปที่ 3.10 ขั้นตอนการแปลงค่าข้อมูลวิดีโอที่ได้รับ

การทำงานในส่วนนี้เกี่ยวข้องกับคำสั่งดาวน์โหลดบนเบราว์เซอร์ ซึ่งมีข้อจำกัดในด้านของเวลาในการประมวลผลภาพที่มีค่าเฉลี่ยของเวลาที่ใช้ในการประมวลผลประมาณ 1400 ms ดังนั้นการทำงานในส่วนนี้จึงอยู่ในรูปแบบของการทำซ้ำ (Loop) ทุก ๆ 1500 ms นั่นเอง

บทที่ 4
ผลการวิจัย

ผลการวิจัย

การวิเคราะห์ผลการวิจัยนี้เป็นไปตามวัตถุประสงค์ของการวิจัย ซึ่งแบ่งได้เป็น 3 ข้อได้แก่

1. ผลการวิเคราะห์ปัญหาในการสื่อสารผ่านโปรแกรมวิดีโอคอลที่มีอยู่ในปัจจุบันในระบบช่วยเหลือทางไกลระหว่างคนตาบอดและผู้ช่วยเหลือ
2. ผลการจัดการรูปภาพที่ได้จากวิดีโอคอล เพื่ออำนวยความสะดวกให้แก่ผู้ช่วยเหลือ ในการใช้งานในระบบช่วยเหลือทางไกล
3. ผลการพัฒนาตัวอย่างโปรแกรมวิดีโอคอลเพื่อใช้งานในระบบช่วยเหลือคนตาบอดโดยบุคคลทั่วไป

1. ผลการวิเคราะห์ปัญหาในการสื่อสารผ่านโปรแกรมวิดีโอคอลที่มีอยู่ในปัจจุบันในระบบช่วยเหลือทางไกลระหว่างคนตาบอดและผู้ช่วยเหลือ

จากขั้นตอนการทดสอบการใช้งานโปรแกรมวิดีโอคอลที่มีอยู่ในปัจจุบันในระบบช่วยเหลือทางไกลเพื่อค้นหาปัญหาในการใช้งานโปรแกรมวิดีโอคอลที่มีอยู่โดยคนตาบอดและผู้ช่วยเหลือโดยใช้โปรแกรมไลน์ (LINE) บนโทรศัพท์มือถือซัมซุงกาแลกซี่เอส 3 (Samsung Galaxy S3) ที่ใช้ระบบปฏิบัติการแอนดรอยด์ 4.3 พร้อมกับหูฟัง ในขณะที่ผู้ช่วยเหลือใช้โปรแกรมวิดีโอคอลผ่านสัญญาณไวไฟ (Wi-Fi) บนเครื่องคอมพิวเตอร์ตามที่ได้กล่าวไปในหัวข้อ 3.1 นั้น สามารถสรุปปัญหาที่เกิดขึ้นได้ 4 ข้อดังแสดงอยู่ในตารางที่ 3.1 โดยปัญหาในลำดับ P1 – P3 เป็นปัญหาที่เกิดขึ้นซึ่งส่งผลกระทบต่อการใช้งานของผู้ช่วยเหลือในด้านของมุมมองภาพที่ได้รับ และการบอกให้คนตาบอดซึ่งไม่สามารถมองเห็นเคลื่อนกล้องไปยังตำแหน่งที่ต้องการ ซึ่งทำให้การช่วยเหลือล่าช้า และนอกจากนี้ยังเป็นการสิ้นเปลืองการส่งข้อมูลซ้ำหลายครั้งโดยไม่จำเป็นด้วยเช่นกัน ซึ่งโปรแกรมวิดีโอคอลที่มีอยู่ในปัจจุบัน ทำได้เพียงแสดงภาพวิดีโอปัจจุบันที่ถูกส่งมาจากคนตาบอด โดย ณ ตำแหน่งที่ผ่านไปอาจเป็นภาพที่จำเป็นต่อการใช้งานในการช่วยเหลือที่ไม่สามารถย้อนกลับ หรือบันทึกข้อมูลเก็บไว้ได้

ปัญหา P4 เป็นปัญหาที่เกิดขึ้นในการใช้งานโดยคนตาบอด ซึ่งเป็นเรื่องส่วนต่อประสานผู้ใช้ของโปรแกรมไลน์ที่ไม่รองรับกับการใช้งานร่วมกับเทคโนโลยีอ่านหน้าจอบนโทรศัพท์มือถือได้ทั้งหมด โดยในงานวิจัยนี้ได้ทำการศึกษาเพิ่มเติมเกี่ยวกับส่วนต่อประสานผู้ใช้สำหรับคนตาบอดของโปรแกรมวิดีโอคอลที่มีอยู่ในปัจจุบันเพื่อวิเคราะห์ปัญหาดังกล่าวเพิ่มเติมในขั้นตอนของการออกแบบส่วนต่อประสานผู้ใช้สำหรับคนตาบอดในวิธีการดำเนินงาน

ตารางที่ 4.1 แสดงตัววัดชี้ใช้ในการทดลองทั้งหมด 5 ค่า โดยการทดลอง ก. ซึ่งเป็นการทดลองโดยคนตาบอดเป็นการวัดประสิทธิภาพในรูปแบบของเวลาที่คนตาบอดใช้ในแต่ละกิจกรรม (M1) และ วัดประสิทธิผลในรูปแบบของจำนวนของความผิดพลาด (Error) ที่เกิดขึ้นจากการทำงานที่ไม่เป็นไปตามรูปแบบที่กำหนดไว้ (M2) ส่วนตัวชี้วัดอีก 3 ค่า เป็นการบันทึกค่าการกระทำที่เป็นใจความหลักของกิจกรรม (M3) ค่าการกระทำที่เกี่ยวข้องกับการทำกิจกรรมทั้งหมดใน 1 กิจกรรม (M4) และค่าการกระทำที่เสียเปล่า (M5) ซึ่งคำนวณจาก M3 และ M4

ชื่อ	ความหมาย	หน่วย	หมายเหตุ
M1	เวลาที่คนตาบอดใช้ในแต่ละกิจกรรม	นาที	-
M2	จำนวนของความผิดพลาด (Error) ที่เกิดขึ้นจากการทำงานที่ไม่เป็นไปตามรูปแบบที่กำหนดไว้	ครั้ง	-
M3	การกระทำที่เป็นใจความหลักของกิจกรรม	ครั้ง	-
M4	กระทำที่เกี่ยวข้องกับการทำกิจกรรมทั้งหมดใน 1 กิจกรรม	ครั้ง	-
M5	ค่าการกระทำที่เสียเปล่า	-	$M5 = \frac{(M4 - M3)}{M3}$

ตารางที่ 4.1 ตารางแสดงความหมายและหน่วยของตัวชี้วัดที่ใช้ในการทดลองเพื่อวิเคราะห์ส่วนต่อประสานผู้ใช้สำหรับคนตาบอด

จากการทดสอบส่วนต่อประสานผู้ใช้ของโปรแกรมวิดีโอคอลที่มีอยู่ในปัจจุบันจำนวน 6 โปรแกรม ได้แก่ Line, Skype, Tango, WeChat, Hangout และ Viber ผลลัพธ์ที่ได้จากการทดลองใช้งานโดยคนตาบอดแสดงในตารางที่ 4.2 การทดลองใช้งานโปรแกรมวิดีโอคอลที่มีในปัจจุบันโดยคนตาบอดแสดงให้เห็นว่าไม่มีโปรแกรมใดเลยที่สามารถใช้งานได้อย่างสมบูรณ์แบบโดยคนตาบอด โปรแกรมแทงโก (Tango) สามารถทำกิจกรรมสำเร็จได้มากที่สุดคือ 6 จาก 7 กิจกรรมที่กำหนดไว้ ในขณะที่โปรแกรม Hangouts มีประสิทธิภาพ (M1) มากที่สุดเนื่องจากเวลาที่ใช้ในแต่ละกิจกรรมมี

ค่าน้อยที่สุดรวมทั้งสิ้น 3 กิจกรรมจาก 5 กิจกรรมที่สามารถทำได้สำเร็จ นอกจากนี้ในด้านประสิทธิผล (M2) พบว่าทั้ง 6 โปรแกรมมีค่าไม่ต่างกันมากนัก ทั้งนี้เนื่องจากทุกโปรแกรมมีข้อผิดพลาดเกิดขึ้นจากผู้ทดลองที่เป็นคนตาบอดจำนวนหนึ่งครั้งซึ่งเกิดขึ้นในกิจกรรมที่แตกต่างกันออกไป นอกจากนี้กิจกรรมเชื่อมต่อใหม่เป็นกิจกรรมที่มีข้อผิดพลาดเกิดขึ้นในทุกโปรแกรมที่สามารถทำกิจกรรมจนเสร็จสิ้นได้ เนื่องจากผู้ใช้งานไม่สามารถรับรู้ได้เมื่อขาดการติดต่อ

ตารางที่ 4.3 เป็นผลการทดลองใช้งานโดยผู้วิจัย ซึ่งใช้งานโปรแกรมวิดีโอคอลเช่นเดียวกับคนตาบอดโดยในภาพรวมจะเห็นว่าการกระทำที่เกี่ยวข้องกับการทำกิจกรรมทั้งหมดใน 1 กิจกรรม (M4) มีค่ามากกว่าการกระทำที่เป็นใจความหลักของกิจกรรม (M3) โดยเฉลี่ยแล้วประมาณ 5 เท่า โดยโปรแกรม Skype เป็นโปรแกรมที่มีค่าเฉลี่ยของการกระทำที่เสียเปล่า (M5) น้อยที่สุด ในขณะที่โปรแกรม Line มีค่าเฉลี่ยของการกระทำที่เสียเปล่าสูงที่สุด เนื่องจากการสมัครสมาชิกที่ผู้ใช้ต้องปัดหน้าจอมากกว่า 200 ครั้งเพื่ออ่านข้อตกลงและเงื่อนไขในการใช้งาน

การทดลองทั้ง 2 ครั้งแสดงให้เห็นว่าส่วนต่อประสานผู้ใช้งานของโปรแกรมวิดีโอคอลที่มีอยู่ในปัจจุบันยังไม่สามารถทำงานร่วมกับเทคโนโลยีช่วยเหลือได้ดีมากนัก เนื่องจากมีการกระทำที่เสียเปล่าค่อนข้างมาก นอกจากนี้ยังมีบางกิจกรรมที่ไม่สามารถกระทำได้เนื่องจากโปรแกรมถูกออกแบบให้แสดงภาพวิดีโอแบบเต็มหน้าจอ ซึ่งจะซ่อนเมนูต่าง ๆ ไว้ และไม่สามารถเรียกกลับมาได้เมื่อใช้งานร่วมกับเทคโนโลยีช่วยเหลือ อีกทั้งปุ่มต่าง ๆ ที่เป็นรูปภาพซึ่งไม่ได้ถูกระบุไว้ด้วยข้อความที่สื่อความหมายชัดเจน ทำให้ผู้ใช้เกิดความสับสน และปัญหาของการออกแบบการโต้ตอบที่ใช้การสัมผัสที่ปุ่มและลากไปยังตำแหน่งต่าง ๆ ที่ถูกกำหนดไว้บนหน้าจอ ซึ่งเป็นรูปแบบการโต้ตอบที่ไม่รองรับกับการทำงานของเทคโนโลยีอ่านหน้าจอบนโทรศัพท์มือถือ

โปรแกรม	สมัครสมาชิก		เพิ่มเพื่อนใหม่		โทรออก		รับสายเรียกเข้า		สลับกล้องที่ใช้งาน		เชื่อมต่อใหม่		วางสาย	
	M1	M2	M1	M2	M1	M2	M1	M2	M1	M2	M1	M2	M1	M2
LINE	6.15	0	1.37	0	0.58	0	0.08	0	-	-	1.41	1	-	-
SKYPE	-	-	-	-	2.19	1	0.05	0	1.14	0	2.41	1	0.13	0
HANGOUTS	-	-	1.52	1	0.38	0	-	-	0.16	0	1.04	1	0.21	0
TANGO	7.38	0	-	-	1.53	0	0.06	0	0.21	0	2.01	1	0.06	0
WECHAT	6.42	1	3.03	0	1.42	0	0.05	0	-	-	2.18	1	-	-
VIBER	4.36	0	3.55	1	1.04	0	-	-	-	-	1.34	1	-	-
ค่าเฉลี่ย	6.0775	-	2.3675	-	1.19	-	0.06	-	0.5033	-	1.7317	-	0.1333	-

ตารางที่ 4.2 ตารางแสดงค่า M1 และ M2 ที่ได้จากการทดลองใช้งานโดยคนตาบอด (การทดลอง ก.)

โปรแกรม	สมัครสมาชิก			เพิ่มเพื่อนใหม่			โทรออก			รับสายเรียกเข้า			สลับกล้องที่ใช้			เชื่อมต่อใหม่			วางสาย			รวม		
	M3	M4	M5	M3	M4	M5	M3	M4	M5	M3	M4	M5	M3	M4	M5	M3	M4	M5	M3	M4	M5	M3	M4	M5
LINE	6	275	45	6	21	2.5	3	23	6.7	1	5	4	1	1	0	4	25	5.3	1	4	3	22	354	15
SKYPE	9	22	1.4	6	13	1.2	4	15	2.8	1	2	1	2	3	0.5	3	11	2.7	1	5	4	26	71	1.7
HANGOUTS	13	63	3.8	3	4	0.3	3	10	2.3	1	2	1	1	4	3	2	8	3	1	3	2	24	94	2.9
TANGO	6	20	2.3	4	24	5	4	17	3.3	2	9	3.5	1	12	11	1	3	2	1	11	10	19	96	4.1
WECHAT	6	20	2.3	6	20	2.3	6	38	5.3	1	3	2	1	2	1	2	18	8	1	3	2	23	104	3.5
VIBER	4	16	3	6	29	3.8	4	14	2.5	1	2	1	1	2	1	2	7	2.5	1	6	5	19	76	3
ค่าเฉลี่ย	7.3	69	9.6	5.2	19	2.5	4	20	3.8	1.2	3.8	2.1	1.2	4	2.8	2.3	12	3.9	1	5.3	4.3	22	133	5.1

ตารางที่ 4.3 ตารางแสดงค่า M3 M4 และ M5 ที่ได้จากการทดลองใช้งานโดยผู้วิจัย (การทดลอง ข.)

หลังจากการทดสอบส่วนต่อประสานผู้ใช้ได้ทำการศึกษาเพิ่มเติม โดยวิเคราะห์ปัญหาที่เกิดขึ้นเพื่อสรุปเป็นหลักการสำหรับการออกแบบส่วนต่อประสานผู้ใช้ที่อ้างอิงและเพิ่มเติมจาก WCAG 2.0 ดังแสดงในตารางที่ 4.4

ลำดับ	หลักการการออกแบบ	ความสอดคล้องกับ WCAG
1	ปุ่มทุกปุ่มและฟังก์ชั่นต่าง ๆ ต้องปรากฏอยู่บนจอและสามารถเข้าถึงได้จากเทคโนโลยีอ่านหน้าจอด้วยการใช้งานแบบปัด	WCAG 2.0 ข้อ 4.1
2	ควรมีคำอธิบายหรือการแจ้งเตือนด้วยเสียงเมื่อเกิดข้อผิดพลาดหรือมีการเปลี่ยนแปลงต่าง ๆ เกิดขึ้น	WCAG 2.0 ข้อ 3.3
3	ใช้ข้อความที่สามารถเข้าใจได้ และเพิ่มคำอธิบายสำหรับรูปภาพ	WCAG 2.0 ข้อ 1.1
4	ปุ่มหรือฟังก์ชั่นการใช้งานที่มักเกิดขึ้นบ่อยควรเรียงลำดับไว้ให้เข้าถึงได้ในลำดับต้น ๆ และการทำงานของเทคโนโลยีอ่านหน้าจอจะสอดคล้องกับลำดับ ของการวางที่ตำแหน่งต่างๆ และทำงานโดยเรียงลำดับจากซ้ายไปขวา และบนลงล่างเสมอ	-
5	ควรมีตัวเลือกสำหรับเปิดปิดคำอธิบายที่เป็นข้อความ	-

ตารางที่ 4.4 ตารางแสดงหลักการการออกแบบส่วนต่อประสานผู้ใช้สำหรับโปรแกรมที่ใช้งานโดยคนตาบอดบนโทรศัพท์มือถือ

หลักการสำหรับการออกแบบที่สรุปได้จากการทดลองอ้างอิงจาก WCAG 2.0 จำนวน 3 ข้อ ได้แก่ สามารถรับรู้ได้ สามารถเข้าใจได้ และมีความทนทาน นอกจากนี้หลักการสำหรับการออกแบบอีก 2 ข้อเป็นการสรุปที่เพิ่มเติมจาก WCAG 2.0 โดยลำดับที่ 4 มีสาเหตุเนื่องจากการทำงานของเทคโนโลยีอ่านหน้าจอบนโทรศัพท์มือถือมีการทำงานแบบเรียงลำดับจากซ้ายไปขวา และบนลงล่างเสมอ และลำดับที่ 5 เป็นการสร้างทางเลือกสำหรับคำอธิบายต่าง ๆ โดยทั้งสองข้อที่เพิ่มขึ้นมา มีวัตถุประสงค์เพื่อลดจำนวนการปัดหน้าจอที่ไม่เกิดผลสำหรับคนตาบอดให้เหลือน้อยที่สุดนั่นเอง

หลังจากสรุปหลักการสำหรับการออกแบบ ได้ทำการทดลองอีกครั้งกับคนตาบอดเพื่อวัดผลส่วนต่อประสานผู้ใช้ของโปรแกรมวิดีโอคอลที่ออกแบบ โดยทดลองใช้งานส่วนต่อประสานผู้ใช้ที่ออกแบบจากหลักการ วัดผลในรูปของเวลาที่ใช้ในแต่ละกิจกรรมและนำมาเปรียบเทียบกับผลลัพธ์ที่ดีที่สุดจากโปรแกรมที่มีอยู่ในปัจจุบันทั้ง 6 โปรแกรม ทั้งนี้ผลลัพธ์จากการทดลองแสดงในตารางที่ 4.5

กิจกรรม	เวลาที่ดีที่สุดของโปรแกรมวิดีโอคอล 6 โปรแกรมที่มีในปัจจุบัน	โปรแกรมที่ออกแบบตามหลักการการออกแบบที่นำเสนอ	%
สมัครสมาชิก	4.36	1.56	64.22
เพิ่มเพื่อนใหม่	1.37	1.05	23.36
โทรออก	0.58	0.29	50.00
รับสาย	0.05	0.06	-20.00
สลับกล้องที่ใช้งาน	0.16	0.07	56.25
เชื่อมต่อใหม่	1.04	0.46	55.77
วางสาย	0.06	0.08	-33.33

ตารางที่ 4.5 ตารางแสดงการเปรียบเทียบเวลาที่ใช้ในกิจกรรมต่าง ๆ (นาที) ของเวลาที่ดีที่สุดใน 6 โปรแกรมที่มีอยู่ในปัจจุบันและโปรแกรมที่ออกแบบตามหลักการการออกแบบที่นำเสนอ

ผลลัพธ์จากการทดลองแสดงให้เห็นว่า เวลาที่ใช้ในแต่ละกิจกรรมของโปรแกรมที่ออกแบบตามหลักการการออกแบบที่นำเสนอในภาพรวมแล้วน้อยกว่าเวลาที่ดีที่สุดของโปรแกรมวิดีโอคอล 6 โปรแกรมที่มีในปัจจุบัน สำหรับกิจกรรมการรับสาย และวางสายที่ใช้เวลาในการทำกิจกรรมมากกว่าจะเห็นได้ว่าเวลาที่แตกต่างอยู่ในหน่วยวินาที อีกทั้งเวลาที่ใช้ก็อยู่ในหน่วยวินาทีด้วยเช่นกัน

2. ผลการเสนอวิธีการจัดการรูปภาพที่ได้จากวิดีโอคอล เพื่ออำนวยความสะดวกให้แก่ผู้ช่วยเหลือในการใช้งานในระบบช่วยเหลือทางไกล

การจัดการรูปภาพที่ได้จากวิดีโอคอลเพื่ออำนวยความสะดวกให้แก่ผู้ช่วยเหลือในการใช้งานในระบบช่วยเหลือทางไกลที่นำเสนอในงานวิจัยนี้ คือการนำภาพวิดีโอที่ถูกส่งมาจากคนตาบอดในกรณีที่คนตาบอดหยุดนิ่งอยู่กับที่และหมุนกล้องไปรอบ ๆ มาประมวลผลและแสดงเป็นภาพมุมกว้างร่วมกับการแสดงตำแหน่งล่าสุดที่ถูกประมวลผล และเวลาของแต่ละตำแหน่งโดยการใช้โทนสีดังแสดงในรูปที่ 3.5

ในการเลือกอัลกอริทึมที่ใช้ในกระบวนการลงทะเบียนภาพ และการรวมภาพ ได้เลือกเปรียบเทียบอัลกอริทึมที่ใช้ในวิธีอิงคุณลักษณะ ซึ่งได้แก่ SIFT และ SURF โดยทดลองใช้สองอัลกอริทึมในการประมวลผล ตารางที่ 4.6 แสดงเวลาในหน่วยวินาทีที่ใช้ในอัลกอริทึมทั้งสองในการประมวลผลกับภาพตัวอย่างจำนวน 100 ภาพที่ตัดมาจากรูปภาพนิ่งขนาดใหญ่ บนเครื่องคอมพิวเตอร์ที่มีคุณสมบัติคือ Intel Core i7 (Dual-core) ความเร็ว 1.7 GHz ซึ่งแต่ละรูปภาพที่นำเข้ามีขนาด 300 x 300 จุดภาพ (pixel) และรูปภาพมุมกว้างมีขนาด 1500 x 1200 จุดภาพ ผลลัพธ์จากการประมวลผลแสดงให้เห็นว่าการประมวลผลด้วยอัลกอริทึม SURF ใช้เวลาในการประมวลผลน้อยกว่าโดยมีค่าเฉลี่ยของเวลาที่ใช้อยู่ที่ 0.44 วินาที ในขณะที่การประมวลผลโดยใช้อัลกอริทึม SIFT มีค่าเฉลี่ยของเวลาที่ใช้อยู่ที่ 0.63 วินาที ซึ่งทางสถิติถือว่ามีความแตกต่างกันอย่างมีนัยสำคัญที่ระดับความเชื่อมั่น 95% แต่อย่างไรก็ตาม เวลาการประมวลผลของทั้งสองอัลกอริทึมก็สามารถนำไปประยุกต์ใช้ในการประมวลผลแบบเกือบทันกาล (nearly real-time) ได้

ค่าสถิติ	SURF	SIFT
ค่าเฉลี่ย	0.43827	0.62564
ความแปรปรวน	0.001479896	0.003249849
ค่าสูงสุด	0.506	0.78
ค่าต่ำสุด	0.289	0.507
P (T <= t) one-tail	3.31594813955403E-69	

ตารางที่ 4.6 สถิติของเวลาการประมวลผลโดยอัลกอริทึม SURF และ SIFT ในหน่วยวินาที

รูปภาพที่ 4.1 แสดงผลลัพธ์ และการเปรียบเทียบการประมวลผลทั้ง 4 รูปแบบกับชุดข้อมูลนำเข้าซึ่งมาจากภาพวิดีโอที่จำลองสถานการณ์การใช้งานจริงที่มีลำดับภาพเป็นภาพที่ต่อเนื่องทั้งหมด โดยจะเห็นว่าการประมวลผลโดยใช้อัลกอริทึม SURF มีข้อผิดพลาดเกิดขึ้นในขณะที่การประมวลผลโดยใช้อัลกอริทึม SIFT สามารถประมวลผลได้ ทั้งนี้เนื่องจากอัลกอริทึม SIFT สามารถตรวจจับค่าคุณลักษณะในภาพได้มากกว่า SURF ทำให้สามารถจับคู่ภาพได้แม่นยำและมีประสิทธิภาพมากกว่านั่นเอง ดังนั้นในงานวิจัยนี้จึงเลือกใช้อัลกอริทึม SIFT ในการคำนวณ และจับคู่จุดคุณลักษณะในภาพ

(ก) (ข)

รูปที่ 4.1 ตัวอย่างการต่อรูปภาพโดยใช้ชุดข้อมูลนำเข้าจากวิดีโอคอลร่วมกับ
(ก) อัลกอริทึม SURF (ข) อัลกอริทึม SIFT

การทดลองใช้งานในรอบที่ 1 เป็นการทดลองกับกลุ่มตัวอย่างที่เป็นบุคคลทั่วไปจำนวน 6 คน โดยจำลองสถานการณ์ในการช่วยคนตาบอดเลือกซื้อเครื่องดื่มจากเครื่องขายน้ำดื่มอัตโนมัติ โดยผลลัพธ์ในการทดลองแสดงอยู่ในตารางที่ 4.7

รายละเอียด	ค่าเฉลี่ยของการดูภาพผ่านวิดีโอ	ค่าเฉลี่ยของการดูภาพผ่านภาพมุมกว้าง
1. สามารถช่วยคนตาบอดหาของตามที่กำหนดได้	2.67	4.33
2. สามารถเข้าใจสภาพแวดล้อมต่าง ๆ คนตาบอดได้	2.67	4.50
3. สามารถบอกให้คนตาบอดหมุนกล้องไปยังตำแหน่งที่ต้องการได้	2.67	3.50
4. สามารถดูภาพที่ตำแหน่งต่าง ๆ ตามที่ต้องการได้	2.50	4.33
5. ความพึงพอใจโดยรวม	2.83	4.17
6. ภาพมุมกว้างที่เห็นมีคุณภาพ มีความคมชัดเพียงพอสำหรับการช่วยเหลือ	-	4.00
7. ภาพมุมกว้างที่เห็นสามารถเข้าใจได้ถึงตำแหน่งปัจจุบันของคนตาบอด และเวลาที่ภาพในตำแหน่งต่าง ๆ ถูกถ่ายไว้ได้	-	4.33
8. การนำภาพวิดีโอคอลมาประมวลผลเป็นภาพมุมกว้างทำให้การช่วยเหลือสะดวกมากขึ้น	-	4.33
9. สามารถเก็บรักษาเฟรมต่าง ๆ ที่ได้จากวิดีโอคอลไว้ใช้ประโยชน์ได้	-	4.67

ตารางที่ 4.7 ตารางแสดงผลลัพธ์ของแบบสอบถามความพึงพอใจต่อการใช้งานวิดีโอคอลร่วมกับส่วนการแสดงภาพมุมกว้าง

จุฬาลงกรณ์มหาวิทยาลัย

ผลลัพธ์จากการทอลองอยู่ในรูปแบบของความพึงพอใจในการใช้งานโดยเปรียบเทียบการช่วยเหลือผ่านโปรแกรมวิดีโอคอลที่พัฒนาขึ้นที่ไม่รวมการทำงานส่วนการประมวลผลภาพเข้าไป กับการใช้งานร่วมกันกับส่วนการประมวลผลภาพ โดยผู้ทดลองให้คะแนนตามสเกลลิเคิร์ท (5 = เห็นด้วยมาก, 4 = เห็นด้วย, 3 = เฉย ๆ, 2 = ค่อนข้างไม่เห็นด้วย, 1 = ไม่เห็นด้วย) ซึ่งคะแนนเฉลี่ยของการดูภาพผ่านวิดีโอคอลทั่วไปคือ 2.67 ในขณะที่ค่าเฉลี่ยของการดูภาพผ่านภาพมุมกว้างคือ 4.17 นอกจากนี้ในส่วนการเพิ่มเติมของการแสดงภาพมุมกว้าง สิ่งที่ผู้ทดลองพึงพอใจมากที่สุดคือการที่สามารถเก็บรักษาเฟรมต่าง ๆ ที่ได้จากวิดีโอคอลมาใช้ประโยชน์ได้ อย่างไรก็ตามจากการทดลองใช้งานในรอบที่ 1 พบว่ามีปัญหาในการใช้งานการประกอบภาพทางไกล เนื่องจากภาพที่ได้รับมาประมวลผลมีความคลาดเคลื่อน โดยในบางครั้งไม่สามารถนำภาพทั้งสองมาต่อกันได้จึงมีการเพิ่มเติมการประมวลผลสำหรับกรณีภาพไม่ต่อเนื่องเพิ่มเข้าไป ซึ่งภายหลังการปรับปรุงอัลกอริทึม และทำการ

ทดลองกับภาพตัวอย่างจำนวน 50 ภาพที่ตัดมาจากรูปภาพนิ่งขนาดใหญ่พบว่าเวลาที่ใช้ในการประมวลผลมีค่าสูงขึ้นโดยเฉลี่ยประมาณ 2 เท่าดังแสดงในตารางที่ 4.8

ค่าสถิติ	การประมวลผล กรณีปกติ	การประมวลผล กรณีภาพไม่ต่อเนื่อง
ค่าสูงสุด	0.858	1.648
ค่าต่ำสุด	0.639	1.404
ค่าเฉลี่ย	0.729	1.497

ตารางที่ 4.8 เวลาที่ใช้ในการประมวลผลโดยอัลกอริทึมที่นำเสนอในหน่วยนาที

3. ผลการพัฒนาตัวอย่างโปรแกรมวิดีโอคอลเพื่อใช้งานในระบบช่วยเหลือคนตาบอดโดยบุคคลทั่วไป

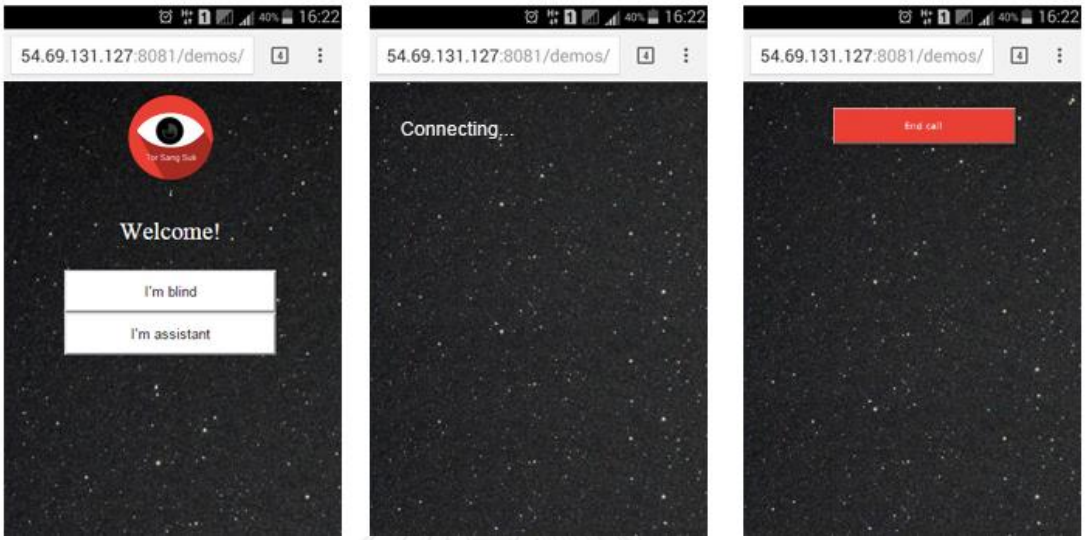
ตัวอย่างโปรแกรมวิดีโอคอลเพื่อใช้งานในระบบช่วยเหลือคนตาบอดร่วมกับการประกอบภาพทางไกลที่นำเสนอในงานวิจัยนี้แบ่งเป็น 2 รูปแบบ ได้แก่

1. โปรแกรมวิดีโอคอลสำหรับคนตาบอด
2. โปรแกรมวิดีโอคอลสำหรับผู้ช่วยเหลือ

โปรแกรมวิดีโอคอลสำหรับคนตาบอด เป็นการพัฒนาโดยคำนึงถึงปัญหาด้านส่วนต่อประสานผู้ใช้ (User interface) ซึ่งออกแบบโดยคำนึงร่วมกับหลักการสำหรับการออกแบบเว็บไซต์เพื่อให้เข้าถึงได้ WCAG 2.0 แต่อย่างไรก็ตามได้นำหลักการดังกล่าวมาพิจารณาร่วมกับประสบการณ์ผู้ใช้ (User experience) ที่ได้จากการทดลองเพื่อวิเคราะห์ปัญหาที่เกิดขึ้นในการใช้งานของคนตาบอด โดยพิจารณาจาก 6 โปรแกรมวิดีโอคอลที่ใช้งานในปัจจุบัน ซึ่งจากการวิเคราะห์ได้ลดจำนวนปุ่มภายในหน้าจอ และออกแบบให้เป็นการทำงานโดยอัตโนมัติ เพื่อลดปัญหาที่เกิดจากการโต้ตอบระหว่างคนตาบอดและอุปกรณ์ที่อาจเกิดขึ้น

ตัวอย่างหน้าจอแสดงผลของโปรแกรมวิดีโอคอลสำหรับคนตาบอดซึ่งใช้งานผ่านเบราว์เซอร์แสดงอยู่ในรูปที่ 4.2 โดยตอนเริ่มของโปรแกรมในรูป (ก) ผู้ใช้จะเลือกประเภทของผู้ใช้งานโดยสามารถเข้าถึงได้จากการปัดหน้าจอจำนวน 3 ครั้ง จากนั้นโปรแกรมจะแจ้งเตือนผู้ใช้ถึงสถานะการเชื่อมต่อกับผู้ช่วยเหลือดังรูป (ข) ซึ่งในหน้านี้ไม่มีส่วนรับอินพุตใด ๆ จากคนตาบอด โดยเมื่อสามารถ

จับคู่ผู้ช่วยเหลือและทำการใช้งานวิดีโอคอล ที่หน้าจอสำหรับการใช้งานวิดีโอคอล (ค) ผู้ใช้สามารถเข้าถึงปุ่มวางสายได้โดยการปัดหน้าจอจำนวน 1 ครั้ง



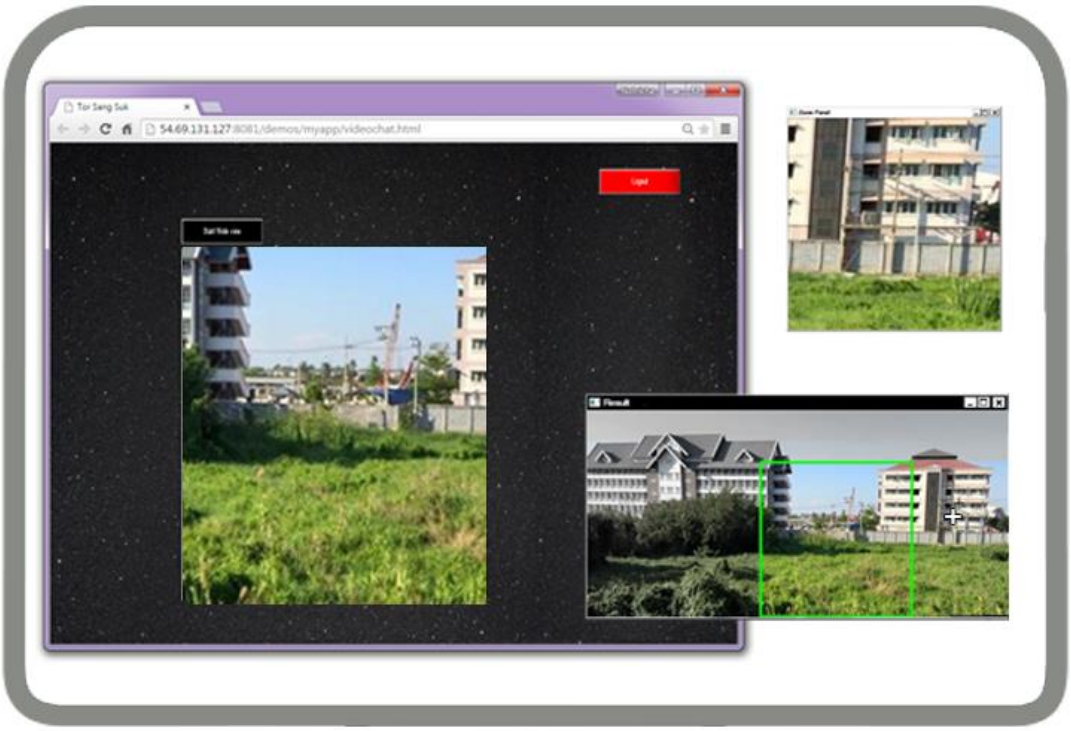
(ก) (ข) (ค)

รูปที่ 4.2 หน้าจอแสดงผลของโปรแกรมวิดีโอคอลสำหรับคนตาบอด

การทำงานของโปรแกรมสำหรับคนตาบอดถูกควบคุมโดย JavaScript และใช้ CSS ในการซ่อนและแสดงส่วนประกอบต่าง ๆ บนหน้าจอ โดยเมื่อคนตาบอดเข้าใช้งานโดยการกดปุ่ม “I’m blind” ที่หน้าแรกของโปรแกรม จะทำการเชื่อมต่อกับ EasyRTC โดยอัตโนมัติ นอกจากนั้นจะตั้งค่ารูปแบบ (style) ของส่วนประกอบต่าง ๆ บนหน้าจอให้กลายเป็นไม่ปรากฏ (invisible) และแสดงข้อความใหม่แก่คนตาบอดเพื่อแจ้งถึงสถานะการเชื่อมต่อกับผู้ช่วยเหลือ ในส่วนนี้ JavaScript จะทำการตรวจสอบการจับคู่ ซึ่งหากจับคู่กับผู้ช่วยเหลือได้แล้ว จะเริ่มการใช้งานวิดีโอคอล พร้อมกับตั้งค่ารูปแบบ (style) ของส่วนประกอบต่าง ๆ บนหน้าจอให้กลายเป็นไม่ปรากฏ (invisible) และตั้งค่าส่วนประกอบที่เป็นปุ่มสำหรับวางสายแสดงแก่คนตาบอด

โปรแกรมสำหรับผู้ช่วยเหลือ เป็นโปรแกรมบนเครื่องคอมพิวเตอร์อยู่ในรูปแบบของโปรแกรม .exe ที่พัฒนาโดยใช้ไลบรารี่ของ OpenCV โดยโปรแกรมเป็นผู้เปิดหน้าเว็บเบราว์เซอร์เพื่อใช้งานในส่วนของวีดีโอคอล และทำงานเป็นการทำงานเบื้องหลังเพื่อรอการประมวลผลภาพมุมกว้าง

รูปที่ 4.3 แสดงตัวอย่างหน้าจอของผู้ช่วยเหลือขณะใช้งานวิดีโอคอลร่วมกับการประกอบภาพทางไกล ซึ่งประกอบด้วยหน้าต่างเว็บเบราว์เซอร์ที่อยู่ในด้านซ้ายของภาพ โดยในหน้าต่างเว็บเบราว์เซอร์ขณะใช้งานวิดีโอคอล จะมีปุ่มสั่งสำหรับเริ่มต้น หรือสิ้นสุดการทำงานส่วนประกอบภาพทางไกล ส่วนด้านขวาของรูปเป็นส่วนแสดงผลลัพธ์ของการประมวลผลภาพแสดง ซึ่งเป็นหน้าต่างใหม่ที่สร้างโดยโปรแกรมที่ประมวลผลภาพทางไกล



รูปที่ 4.3 ตัวอย่างหน้าจอของผู้ช่วยเหลือขณะใช้งานวิดีโอคอลร่วมกับการประกอบภาพทางไกล

โปรแกรมฝั่งผู้ช่วยเหลือที่มีการเพิ่มเติมในส่วนของการส่งเพียงเสียงกลับไปยังคนตาบอด พบว่าในสามารถลดแบนด์วิดท์ (bandwidth) ของโปรแกรมสำหรับผู้ช่วยเหลือได้ โดยตารางที่ 4.9 แสดงรายละเอียดแบนด์วิดท์ที่ใช้ในการส่งข้อมูลของโปรแกรมสำหรับผู้ช่วยเหลือ โดยเปรียบเทียบการส่งข้อมูลเสียงร่วมกับภาพวิดีโอ และการส่งข้อมูลเสียงเพียงอย่างเดียว โดยทำการวัดแบนด์วิดท์ด้วยโปรแกรม BitMeter II [49] บนเครื่องคอมพิวเตอร์ด้วยข้อมูลนำเข้าที่เป็นคลิปวิดีโอเดียวกัน โดยจากการวัดพบว่าการส่งข้อมูลเสียงเพียงอย่างเดียวสามารถลดจำนวนการส่งข้อมูลได้ประมาณ 96% ในขณะที่ปริมาณการดาวน์โหลดมีค่าคงที่เนื่องจากเป็นการรับข้อมูลภาพและเสียงจากคนตาบอดเช่นเดียวกันทั้งสองครั้ง

รูปแบบ	ปริมาณการดาวน์โหลด (เมกกะไบต์)	ปริมาณการอัพโหลด (เมกกะไบต์)
ส่งภาพและเสียง	78.48	97.54
ส่งเสียงอย่างเดียว	80.72	3.35

ตารางที่ 4.9 ตารางเปรียบเทียบแบนด์วิดท์ที่ใช้ในการส่งข้อมูลในรูปแบบต่าง ๆ ของโปรแกรมสำหรับผู้ช่วยเหลือ

การทดลองใช้งานโปรแกรมที่พัฒนาขึ้น เป็นการทดลองกับกลุ่มตัวอย่างจำนวน 3 คน ซึ่งเป็นคนตาบอดแต่กำเนิดจำนวน 1 คน และบุคคลทั่วไปจำนวน 2 คน โดยจับคู่คนตาบอดกับผู้ช่วยเหลือเพื่อทำสถานการณ์ต่าง ๆ ตามที่กำหนดไว้ทั้งหมด 2 กิจกรรมดังแสดงในตารางที่ 4.10 ซึ่งรูปแบบการทดลองเป็นการทดลองใช้งานต้นแบบของโปรแกรมที่พัฒนาขึ้น โดยในแต่ละกิจกรรมเริ่มต้นจากการทดลองกับผู้ช่วยเหลือคนที่ 1 ซึ่งใช้งานโปรแกรมที่ไม่รวมส่วนการแสดงภาพมุมกว้าง จากนั้นทดลองซ้ำโดยเปลี่ยนสิ้นค้าที่เลือกซื้อด้วยโปรแกรมที่รวมส่วนการแสดงภาพมุมกว้าง จากนั้นทดลองซ้ำกับผู้ช่วยเหลือคนที่ 2 โดยเริ่มจากการใช้งานร่วมกับส่วนการแสดงภาพมุมกว้าง และทดลองอีกครั้งโดยเลือกซื้อสิ้นค้าที่แตกต่างด้วยโปรแกรมที่ไม่ร่วมส่วนการแสดงภาพมุมกว้าง

ลำดับ	กิจกรรม	ผู้ทดลอง	รูปแบบโปรแกรมที่ใช้งาน
1	ซื้อน้ำอัดลมจากเครื่องขายน้ำดื่มอัตโนมัติเครื่องที่ 1	คนตาบอดและผู้ช่วยคนที่ 1	โปรแกรมที่ไม่รวมส่วนการแสดงภาพมุมกว้าง
2	ซื้อน้ำผลไม้กระป๋องจากเครื่องขายน้ำดื่มอัตโนมัติเครื่องที่ 2	คนตาบอดและผู้ช่วยคนที่ 1	โปรแกรมที่รวมส่วนการแสดงภาพมุมกว้าง
3	ซื้อน้ำเปล่าจากเครื่องขายน้ำดื่มอัตโนมัติเครื่องที่ 1	คนตาบอดและผู้ช่วยคนที่ 2	โปรแกรมที่รวมส่วนการแสดงภาพมุมกว้าง
4	ซื้อน้ำผลไม้แบบขวดจากเครื่องขายน้ำดื่มอัตโนมัติเครื่องที่ 2	คนตาบอดและผู้ช่วยคนที่ 2	โปรแกรมที่ไม่รวมส่วนการแสดงภาพมุมกว้าง
5	ดูผลงานชิ้นที่ 1 และ 2 ในนิทรรศการ	คนตาบอดและผู้ช่วยคนที่ 2	โปรแกรมที่ไม่รวมส่วนการแสดงภาพมุมกว้าง
6	ดูผลงานชิ้นที่ 3 และ 4 ในนิทรรศการ	คนตาบอดและผู้ช่วยคนที่ 2	โปรแกรมที่รวมส่วนการแสดงภาพมุมกว้าง
7	ดูผลงานชิ้นที่ 5 และ 6 ในนิทรรศการ	คนตาบอดและผู้ช่วยคนที่ 1	โปรแกรมที่รวมส่วนการแสดงภาพมุมกว้าง
8	ดูผลงานชิ้นที่ 7 และ 8 ในนิทรรศการ	คนตาบอดและผู้ช่วยคนที่ 1	โปรแกรมที่ไม่รวมส่วนการแสดงภาพมุมกว้าง

ตารางที่ 4.10 ตารางแสดงรายละเอียดกิจกรรม ผู้ทดลอง และรูปแบบโปรแกรมที่ใช้งาน

จากการทดลองซึ่งแบ่งออกเป็น 2 กิจกรรมหลัก ได้แก่การเลือกซื้อเครื่องดื่มจากเครื่องขายน้ำดื่มอัตโนมัติในลำดับที่ 1-4 และการนำทางในงานนิทรรศการเพื่อดูผลงานต่าง ๆ ในลำดับที่ 5-8 โดยการช่วยซื้อเครื่องดื่มในลำดับที่ 1-4 พบว่าการช่วยเหลือผ่านโปรแกรมที่ไม่รวมส่วนการแสดงภาพมุมกว้าง ผู้ช่วยเหลือจำเป็นต้องบอกให้คนตาบอดเคลื่อนกล้องกลับไปมาเพื่อระบุตำแหน่งของเครื่องดื่ม และทำการช่วยเหลือ ในขณะที่การใช้งานโปรแกรมร่วมกับส่วนการแสดงภาพมุมกว้างผู้ช่วยเหลือเริ่มต้นจากการให้คนตาบอดเคลื่อนกล้องไปรอบ ๆ ในตอนเริ่มต้น โดยหลังจากนั้นสามารถช่วยเหลือคนตาบอด และระบุตำแหน่งของเครื่องดื่มที่ต้องการได้โดยไม่จำเป็นต้องให้คนตาบอดเคลื่อนที่กลับไปยังตำแหน่งดังกล่าวซ้ำ ในขณะที่การนำทางในงานนิทรรศการในลำดับที่ 5-8 พบว่ามี

การย้อนหลับเพื่อดูแผนที่ของผลงานต่าง ๆ เนื่องจากผู้ช่วยเหลือไม่สามารถจำตำแหน่งของผลงานได้ ในขณะที่การช่วยเหลือผ่านโปรแกรมที่รวมส่วนการแสดงภาพมุมกว้าง ทำให้ผู้ช่วยเหลือสามารถเก็บภาพแผนที่ขนาดใหญ่ที่แสดงอยู่ตรงบริเวณทางเข้า และนำทางคนตาบอดไปยังผลงานต่าง ๆ ได้โดยไม่จำเป็นต้องกลับมาดูแผนที่ใหม่ โดยการวัดผลอยู่ในรูปแบบของเวลาที่ใช้ และจำนวนครั้งในการบอกให้คนตาบอดเคลื่อนกล้องกลับไปยังตำแหน่งที่เคยผ่านมาแล้วดังแสดงในตารางที่ 4.11

ลำดับการทดลอง	เวลาที่ใช้ (นาที)	การย้อนกลับตำแหน่งเก่า (ครั้ง)
1	6.27	9
2	4.30	3
3	4.14	1
4	6.09	8
5	4.58	4
6	2.89	0
7	3.94	1
8	5.31	3

ตารางที่ 4.11 ตารางแสดงผลการทดลองใช้งานในรูปแบบของเวลาที่ใช้ และจำนวนครั้งในการเคลื่อนกล้องกลับไปยังตำแหน่งเดิม

การวิเคราะห์เวลาที่ใช้ในการทดลองแสดงอยู่ในตารางที่ 4.12 โดยจะเห็นได้ว่าเมื่อจับคู่การทดลองในลำดับที่ 1 และ 4 เพื่อหาค่าเฉลี่ยของเวลาที่ใช้ในการทำกิจกรรมที่ 1 โดยโปรแกรมที่ไม่มีส่วนการแสดงภาพมุมกว้าง และนำมาเปรียบเทียบกับค่าเฉลี่ยของเวลาในการทำกิจกรรมที่ 1 โดยโปรแกรมที่มีส่วนการแสดงภาพมุมกว้างในลำดับที่ 2 และ 3 พบว่าการใช้งานโปรแกรมร่วมกับส่วนการแสดงภาพมุมกว้างทำให้การช่วยเหลือเร็วขึ้นโดยเฉลี่ย 31.75% ในขณะเดียวกันเมื่อ จับคู่การทดลองในลำดับที่ 5 และ 8 เพื่อหาค่าเฉลี่ยของเวลาที่ใช้ในการทำกิจกรรมที่ 2 โดยโปรแกรมที่ไม่มีส่วนการแสดงภาพมุมกว้าง และนำมาเปรียบเทียบกับค่าเฉลี่ยของเวลาในการทำกิจกรรมที่ 2 โดยโปรแกรมที่มีส่วนการแสดงภาพมุมกว้างในลำดับที่ 6 และ 7 พบว่าการใช้งานโปรแกรมร่วมกับส่วนการแสดงภาพมุมกว้างทำให้การช่วยเหลือเร็วขึ้นโดยเฉลี่ย 30.94% นอกจากนี้ยังพบว่าสามารถลดจำนวนครั้งในการบอกให้คนตาบอดเคลื่อนที่กลับไปยังตำแหน่งที่ผ่านมาแล้วได้ด้วยเช่นกัน

กิจกรรมที่	ค่าเฉลี่ยเวลาที่ใช้โปรแกรมที่ไม่มีส่วนการแสดงภาพมุมกว้าง (นาที)	ค่าเฉลี่ยเวลาที่ใช้โปรแกรมที่มีส่วนการแสดงภาพมุมกว้าง (นาที)	ค่าเฉลี่ยของเวลาที่ลดลง (นาที)	%
1	6.18	4.22	1.96	31.75
2	4.945	3.415	1.53	30.94

ตารางที่ 4.12 ผลการวิเคราะห์เวลาที่ใช้ในการทดลอง

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

บทที่ 5
สรุปผล

สรุปผลการวิจัย

ระบบช่วยเหลือคนตาบอดผ่านโทรศัพท์ร่วมกับการประกอบภาพทางไกลที่นำเสนอนี้ เป็นระบบที่พัฒนาขึ้นโดยเน้นการแก้ไขปัญหาที่เกิดขึ้นกับผู้ช่วยเหลือที่มีข้อจำกัดในด้านมุมมองของภาพที่ได้รับจากคนตาบอด ซึ่งอาจไม่เพียงพอสำหรับการช่วยเหลือ โดยผู้วิจัยมีแนวคิดว่าภาพวิดีโอที่ถูกส่งมาจากคนตาบอดในเวลาหนึ่ง ๆ อาจเป็นภาพวิดีโอที่จำเป็นต่อการใช้งาน และถูกส่งมาแล้วแต่ไม่มีการเก็บไว้หรือนำมาประมวลผลเพื่อแสดงแก่ผู้ช่วยเหลือในเวลาต่อมา นอกจากการใช้งานโปรแกรมวิดีโอคอลของคนตาบอดในปัจจุบันยังคงมีปัญหา ซึ่งเกิดจากความไม่สอดคล้องกันของการทำงานต่าง ๆ ในโปรแกรมกับเทคโนโลยีอ่านหน้าจอ

วิทยานิพนธ์นี้เริ่มจากการศึกษาปัญหาที่เกิดขึ้นในการใช้งานโปรแกรมวิดีโอคอลที่มีอยู่ในปัจจุบันในการช่วยเหลือทางไกลสำหรับคนตาบอดโดยบุคคลทั่วไป ซึ่งสามารถสรุปปัญหาที่เกิดขึ้นได้ทั้งหมด 2 ประเด็นหลัก ได้แก่ ปัญหาด้านมุมมองภาพที่ไม่เพียงพอต่อการช่วยเหลือคนตาบอดของผู้ช่วยเหลือ และปัญหาด้านส่วนต่อประสานผู้ใช้ของโปรแกรมที่ไม่รองรับการทำงานร่วมกับเทคโนโลยีอ่านหน้าจอบนโทรศัพท์มือถือของคนตาบอด

การแก้ปัญหาด้านมุมมองภาพ ในงานวิจัยนี้ได้นำเสนอวิธีการจัดการรูปภาพที่ได้จากการสื่อสารผ่านวิดีโอคอล โดยประยุกต์ใช้เทคนิคการประมวลผลรูปภาพเพื่อนำภาพวิดีโอที่ถูกส่งมาจากคนตาบอดมาประมวลผลและแสดงแก่ผู้ช่วยเหลือ ซึ่งจากการศึกษาและทดลอง จึงเลือกใช้อัลกอริทึม SIFT มาใช้ในการประมวลผลรูปภาพนำเข้าใหม่ที่ได้จากวิดีโอคอลร่วมกับการเปรียบเทียบกับบางส่วนของรูปภาพที่ประมวลผลแล้วเพื่อลดเวลาที่ใช้ในการประมวลผล อีกทั้งนำเทคนิคการปรับสีในช่อง RGB ของภาพ มาประยุกต์ใช้ในการแสดงข้อมูลเวลาของรูปภาพในตำแหน่งต่าง ๆ

เนื่องจากการประมวลผลภาพถ่ายมีรูปภาพนำเข้าเป็นภาพที่ถูกถ่ายโดยคนตาบอด และได้จากการสตรีมผ่านอินเทอร์เน็ต ซึ่งอาจเกิดความไม่ต่อเนื่องของรูปภาพนำเข้าเนื่องจากปัญหาด้านสัญญาณที่ขาดหาย หรือการเคลื่อนกล้องที่เร็วเกินไปของคนตาบอด ซึ่งเมื่อนำมาประมวลผลกับ

อัลกอริทึมที่พัฒนาในงานนี้พบว่ามีปัญหาในด้านการจับคู่คุณลักษณะในภาพ (Feature matching) ที่ส่งผลให้เกิดการประมวลผลที่ผิดพลาดขึ้นได้ ดังนั้นจึงเพิ่มส่วนของการทำนายตำแหน่งภาพเพื่อแก้ไขปัญหาดังกล่าว โดยใช้การอ้างอิงจากการเปลี่ยนแปลงครั้งก่อนหน้าที่เกิดขึ้น และแบ่งการประมวลผลออกเป็น 2 ส่วน ได้แก่ส่วนการประมวลผลหลักที่ทำหน้าที่ประมวลผลภาพใหม่แสดงแก่ผู้ใช้โดยไม่คำนึงถึงข้อผิดพลาดที่เกิดขึ้นจากการจับคู่คุณลักษณะ และส่วนการทำงานเบื้องหลังที่ทำหน้าที่ในการตรวจสอบการจับคู่กับภาพเก่าที่เคยประมวลผล เพื่อรวมเป็นภาพเดียวกันและแสดงแก่ผู้ช่วยเหลือ

ในท้ายที่สุดได้นำวิธีการจัดการรูปภาพมารวมกับโปรแกรมวิดีโอคอลที่พัฒนาขึ้นเพื่อใช้งานผ่านเว็บเบราว์เซอร์ ซึ่งอยู่ในรูปของโปรแกรมต้นแบบที่ใช้งานในระบบช่วยเหลือทางไกลสำหรับคนตาบอด โดยในการใช้งานโปรแกรมวิดีโอคอลที่พัฒนาขึ้นนั้น คนตาบอดสามารถใช้งานผ่านโทรศัพท์มือถือที่ใช้ระบบปฏิบัติการแอนดรอยด์ 4.0 ขึ้นไป ซึ่งมีการออกแบบส่วนต่อประสานผู้ใช้ในลักษณะการลดการโต้ตอบกับอุปกรณ์ของคนตาบอดให้เหลือน้อยที่สุด โดยเมื่อคนตาบอดเข้าสู่โปรแกรมซึ่งระบุประเภทผู้ใช้งานเป็นคนตาบอด ระบบก็จะทำการจับคู่คนตาบอดกับผู้ช่วยเหลือให้อัตโนมัติ ในขณะที่ผู้ช่วยเหลือสามารถใช้งานบนเครื่องคอมพิวเตอร์เท่านั้น ทั้งนี้เนื่องจากเป็นการใช้งานควบคู่กับโปรแกรมสำหรับการประมวลผลภาพที่พัฒนาขึ้นที่ทำงานเป็นฉากหลัง (Background process) ที่จะถูกเรียกขึ้นมาทำงานเมื่อผู้ใช้เปิดการทำงานของการประมวลผลภาพมุมกว้างนั่นเอง

งานวิจัยนี้มีการทดลองทั้งหมด 2 การทดลอง โดยการทดลองแรกเป็นการทดลองเพื่อเปรียบเทียบการใช้งานตัวอย่างโปรแกรมที่พัฒนาขึ้น ซึ่งใช้งานในการช่วยเหลือคนตาบอดเลือกซื้อน้ำดื่มจากเครื่องขายน้ำดื่มอัตโนมัติ โดยทำการทดลองกับกลุ่มตัวอย่างทั้งหมด 6 คน คนละ 2 ครั้ง ได้แก่ การช่วยเหลือในรูปแบบวิดีโอคอลปกติ และการช่วยเหลือโดยใช้งานร่วมกับการประกอบภาพทางไกลซึ่งไม่ได้รวมส่วนการประมวลผลสำหรับภาพไม่ต่อเนื่องเข้าไป โดยผลลัพธ์ในการทดสอบอยู่ในรูปแบบของความพึงพอใจของผู้ใช้งานตามสเกลลิเคิร์ท ซึ่งแสดงให้เห็นว่า ผู้ทดลองพึงพอใจต่อการใช้งานช่วยเหลือร่วมกับการใช้งานส่วนการประมวลผลภาพมุมกว้างมากกว่าโดยเฉลี่ยแล้วสูงกว่าการใช้งานทั่วไปประมาณ 1.5 เท่า โดยสิ่งที่ผู้ทดลองพึงพอใจมากที่สุดคือการที่สามารถเก็บรักษาเฟรมต่าง ๆ ที่ได้จากวิดีโอคอลมาใช้ประโยชน์ได้ และค่าเฉลี่ยต่อความพึงพอใจต่อเทคนิคการแสดงภาพมุมกว้างคือ 4.33

ภายหลังจากการทดลองครั้งที่ 1 ได้มีการเพิ่มเติมการทำงานสำหรับภาพที่ไม่ต่อเนื่อง และทำการทดลองเพิ่มเติม โดยทดลองกับคนตาบอดจำนวน 1 คน และผู้ช่วยเหลือซึ่งเป็นบุคคลทั่วไปที่คุ้นเคยกับคนตาบอดจำนวน 2 คน โดยกำหนดสถานการณ์จำลอง 2 สถานการณ์ ได้แก่ การซื้อเครื่องดื่มจากเครื่องขายของอัตโนมัติ และการนำทางในงานนิทรรศการ โดยทำการทดลองสถานการณ์ละ 4 ครั้ง โดยปรับเปลี่ยนรายละเอียดภายในบางส่วน ได้แก่รูปแบบของเครื่องดื่มที่เลือกซื้อ และ ผลงานที่เลือก นอกจากนี้ได้สลับลำดับการใช้งานโปรแกรม และผู้ช่วยเหลือเพื่อป้องกันการเรียนรู้ของผู้ใช้ด้วยเช่นกัน

การวัดผลการทดลองในรูปแบบของเวลาที่ใช้ในการช่วยเหลือคนตาบอด และจำนวนครั้งที่ผู้ช่วยเหลือบอกให้คนตาบอดย้อนกลับไปถ่ายยังตำแหน่งที่เคยผ่านมาแล้ว โดยผลลัพธ์จากการทดลองพบว่าการใช้งานโปรแกรมวิดีโอคอลร่วมกับการประกอบภาพทางไกลใช้เวลาในการช่วยเหลือคนตาบอดน้อยกว่าการใช้งานที่มีแค่โปรแกรมวิดีโอคอลประมาณ 31% นอกจากนี้การใช้งานร่วมกับการประกอบภาพทางไกล ผู้ช่วยเหลือสามารถเก็บภาพในส่วนที่ต้องการในรูปแบบของภาพมุมกว้างได้ และไม่มีการบอกให้คนตาบอดย้อนกลับไปยังตำแหน่งที่ผ่านมาแล้วซึ่งมีภาพเก็บไว้แล้วด้วยเช่นกัน

โดยสรุปของโปรแกรมที่พัฒนาขึ้นซึ่งถึงแม้ว่า WebRTC รองรับการใช้งานผ่าน Google Chrome, Mozilla Firefox และ Opera บนระบบปฏิบัติการแอนดรอยด์ แต่อย่างไรก็ตามมีข้อจำกัดของการใช้งานของคนตาบอดที่รองรับบนโทรศัพท์ที่ใช้ระบบปฏิบัติการแอนดรอยด์ 4.0 ขึ้นไป ทั้งนี้เกิดจาก 2 ประเด็น ได้แก่

1. การทำงานของเทคโนโลยีอ่านหน้าจอบนโทรศัพท์มือถือที่รองรับบนโทรศัพท์ที่ใช้ระบบปฏิบัติการแอนดรอยด์ 4.0 ขึ้นไป
2. การใช้บริการของ WebRTC ที่ต้องใช้งานผ่านเบราว์เซอร์ที่รองรับเท่านั้น โดยการใช้งาน WebRTC บนโทรศัพท์มือถือปัจจุบันรองรับแค่ระบบปฏิบัติการแอนดรอยด์ ผ่าน Google Chrome Mozilla Firefox และ Opera เท่านั้น ไม่สามารถใช้งานผ่านเบราว์เซอร์ที่รองรับ รวมไปถึง Safari บนระบบปฏิบัติการ iOS ได้

จากการทดสอบการใช้งาน ได้กำหนดให้โปรแกรมสำหรับผู้ช่วยเหลือรองรับการทำงานบนเครื่องคอมพิวเตอร์ที่มีคุณสมบัติตั้งแต่ Intel Core i7 (Dual-core) ความเร็ว 1.7 GHz เป็นต้นไป

Chulalongkorn University

โดยส่วนการแสดงภาพมุมกว้างที่พัฒนารองรับสำหรับภาพนำเข้าที่ไม่มีความผิดพลาดจากปัญหาในเรื่องการเชื่อมต่อเท่านั้น โดยหากภาพนำเข้าเป็นภาพสีดำเนื่องจากสัญญาณขาดหาย หรือภาพเบลอจากการสั่น จะส่งผลให้ส่วนการแสดงภาพมุมกว้างไม่สามารถประมวลผลได้ ดังนั้นจึงไม่นำภาพดังกล่าวมาประมวลผลในอัลกอริทึมที่พัฒนาขึ้นในงานวิจัยนี้

ข้อเสนอแนะ

โปรแกรมวิดีโอคอลต้นแบบที่พัฒนามีพื้นฐานอยู่บน EasyRTC ซึ่งเป็นโอเพนซอร์สแบบเต็มรูปแบบของ WebRTC ที่ให้บริการการสื่อสารผ่านวิดีโอคอลผ่านเว็บเบราว์เซอร์ ซึ่งทำให้เกิดข้อจำกัดกับระบบในหลายประเด็น ประเด็นแรกเป็นเรื่องการใช้งานผ่านเว็บเบราว์เซอร์ของคนตาบอดซึ่งถึงแม้ฟังก์ชั่นและส่วนต่อประสานผู้ใช้ต่าง ๆ ของโปรแกรมถูกออกแบบให้รองรับกับการใช้งานของเทคโนโลยีอ่านหน้าจอบนโทรศัพท์มือถือ แต่การเข้าถึงเบราว์เซอร์บนโทรศัพท์มือถือ และการใช้งานผ่านเบราว์เซอร์ของคนตาบอดร่วมกับเทคโนโลยีอ่านหน้าจอมีโอกาสเกิดข้อผิดพลาดในการใช้งานได้มากกว่าใช้งานในรูปแบบของโปรแกรมทั่วไปบนโทรศัพท์มือถือ ซึ่งสามารถเขียนโปรแกรมให้แสดงหน้าเว็บได้โดยใช้ WebView ซึ่งเป็นส่วนสำหรับแสดงผลหน้าเว็บไซต์สำหรับโปรแกรมในระบบปฏิบัติการแอนดรอยด์ ทั้งนี้การใช้งาน WebRTC ผ่าน WebView สามารถใช้ได้กับ WebView v36 ซึ่งปัจจุบันกำลังอยู่ระหว่างการพัฒนา [50] โดยในอนาคตสามารถพัฒนาโปรแกรมสำหรับคนตาบอดเพื่อใช้งานในรูปแบบของโปรแกรมบนโทรศัพท์ที่ใช้ระบบปฏิบัติการแอนดรอย์ได้ นอกจากนี้สามารถอ้างอิงหลักการออกแบบส่วนต่อประสานผู้ใช้สำหรับคนตาบอดซึ่งเป็นส่วนหนึ่งของผลลัพธ์ของการวิจัยในครั้งนี้ได้ด้วยเช่นกัน

ในส่วนการประมวลผลภาพมุมกว้าง ที่มีข้อมูลนำเข้าเป็นภาพนิ่งที่ถูกบันทึกลงเครื่องผู้ใช้ของเบราว์เซอร์ มีข้อกำจัดในด้านความเร็วในการบันทึกภาพที่เป็นข้อจำกัดจากตัวเบราว์เซอร์ จึงมีข้อเสนอแนะให้ทำการรวมส่วนการทำงานให้เป็นส่วนเดียวกัน เพื่อทำให้ส่วนการประมวลผลภาพสามารถรับภาพที่ถูกส่งมาจากคนตาบอดได้โดยตรงเพื่อลดปัญหาคอขวดในการบันทึกภาพ

การสื่อสารผ่านวิดีโอคอลโดยใช้บริการของ WebRTC ซึ่งเป็นการจัดการการสื่อสารและรับส่งข้อมูลต่าง ๆ ผ่านผู้ใช้ทั้งสองผ่านอินเทอร์เน็ต สามารถเพิ่มเติมในการรับส่งข้อมูลต่าง ๆ เพื่ออำนวยความสะดวกในการใช้งานได้ เช่น การส่งข้อมูล GPS จากโทรศัพท์ของคนตาบอดเป็นต้น

รายการอ้างอิง

1. กระทรวงศึกษาธิการ, ความพิการ *9* ประเภท. 2011: [Online]. Avaliable: http://disabilities.nfe.go.th/popup.php?name=knowledge&file=p_readknowledge&id=32.
2. สำนักงานปลัดกระทรวงการพัฒนาสังคมและความมั่นคงของมนุษย์, สถิติการจดทะเบียนคนพิการในประเทศไทย. 2015.
3. Jacobson, W.H., *The Art and Science of Teaching Orientation and Mobility to Persons with Visual Impairments*. 1993: AFB Press.
4. เฉลิมพล สมบัติยานุชิต, สถาปัตยกรรมบำบัด "กรณีศึกษาสำหรับคนตาบอด. 2011, กรุงเทพฯ: สาขาวิชาสถาปัตยกรรม ภาควิชาสถาปัตยกรรม มหาวิทยาลัยศิลปากร.
5. Kulyukin, V., et al. *RFID in robot-assisted indoor navigation for the visually impaired*. in *Intelligent Robots and Systems, 2004. (IROS 2004). Proceedings. 2004 IEEE/RSJ International Conference on*. 2004.
6. Ganz, A., et al. *PERCEPT: Indoor navigation for the blind and visually impaired*. in *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*. 2011.
7. Rajamki, J., et al., *LaureaPOP indoor navigation service for the visually impaired in a WLAN environment*, in *Proceedings of the 6th WSEAS International Conference on Electronics, Hardware, Wireless and Optical Communications*. 2007, World Scientific and Engineering Academy and Society (WSEAS): Corfu Island, Greece. p. 96-101.
8. Elloumi, W., et al. *Indoor navigation assistance with a Smartphone camera based on vanishing points*. in *Indoor Positioning and Indoor Navigation (IPIN), 2013 International Conference on*. 2013.
9. Harrington, N., L. Antuna, and Y. Coady, *ABLE Transit: A Mobile Application for Visually Impaired Users to Navigate Public Transit*, in *Proceedings of the 2012 Seventh International Conference on Broadband, Wireless Computing, Communication and Applications*. 2012, IEEE Computer Society. p. 402-407.

10. Prasain, S., *StopFinder: improving the experience of blind public transit riders with crowdsourcing*, in *The proceedings of the 13th international ACM SIGACCESS conference on Computers and accessibility*. 2011, ACM: Dundee, Scotland, UK. p. 323-324.
11. Ueda, T. and L. de Araújo, *Virtual Walking Stick: Mobile Application to Assist Visually Impaired People to Walking Safely*, in *Universal Access in Human-Computer Interaction. Aging and Assistive Environments*, C. Stephanidis and M. Antona, Editors. 2014, Springer International Publishing. p. 803-813.
12. Vorapatratorn, S. and K. Nambunmee, *iSonar: an obstacle warning device for the blind*, in *Proceedings of the 7th International Convention on Rehabilitation Engineering and Assistive Technology*. 2013, Singapore Therapeutic, Assistive & Rehabilitative Technologies (START) Centre: Gyeonggi-do, South Korea. p. 1-4.
13. ณุชนาฎ โต๊ะดี และสิริพร พลายแสง, ความรู้เกี่ยวกับคนพิการทางการมองเห็น. 2009.
14. Framingham, S., *Assistive Technologies*. 2014, Website by Enervision Media Inc.
15. Adams, D., et al., *Interviewing blind photographers: design insights for a smartphone application*, in *Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility*. 2013, ACM: Bellevue, Washington. p. 1-2.
16. สมาคมคนตาบอดแห่งประเทศไทย. โปรแกรมตาทิพย์.
17. *JAWS: The World's Most Popular Windows Screen Reader*. [Online]. Avaliable: http://www.freedomscientific.com/Products/Blindness/JAWS.
18. Apple Inc. *VoiceOver for iOS*.
19. Google Inc. *Google Talkback*.
20. Garaj, V., et al., *A system for remote sighted guidance of visually impaired pedestrians.* British Journal of Visual Impairment May 2003: p. 9.
21. Baranski, P., M. Polanczyk, and P. Strumillo. *A remote guidance system for the blind*. in *e-Health Networking Applications and Services (Healthcom), 2010 12th IEEE International Conference on*. 2010.
22. Bujacz, M., et al. *Remote mobility and navigation aid for the visually disabled*. in *7th Int. Conf. on Disability, Virtual Reality and Assoc Technologies with Art ArtAbilitation*. Maia, Portugal.

23. Wiberg, H.J. *BeMyEyes*. 2015.
24. Bergkvist, A., et al., *Webrtc 1.0: Real-time communication between browsers*. 2013, [Online]. Available: http://dev.w3.org/2011/webrtc/editor/webrtc.html.
25. IETF, *Rtcweb status pages*. 2013, [Online]. Available: http://tools.ietf.org/wg/rtcweb/charters.
26. Johansson, D. and M. Holmgren. *Towards Implementing Web-Based Adaptive Application Mobility Using Web Real-Time Communications*. in *Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS), 2014 Eighth International Conference on*. 2014.
27. Rhinow, F., et al. *P2P live video streaming in WebRTC*. in *Computer Applications and Information Systems (WCCAIS), 2014 World Congress on*. 2014.
28. Chen, C.Y. and R. Klette, *Image Stitching - Comparisons and New Techniques.* Computer Analysis of Images and Patterns, 1998: p. 615-622.
29. Goshtasby, A.A., *2-D and 3-D Image Registration: for Medical, Remote Sensing, and Industrial Applications*. 2005: Wiley-Interscience.
30. Brown, M. and D. Lowe, *Automatic Panoramic Image Stitching using Invariant Features.* International Journal of Computer Vision, 2007. 74(1): p. 59-73.
31. Lowe, D., *Distinctive Image Features from Scale-Invariant Keypoints.* International Journal of Computer Vision, 2004. 60(2): p. 91-110.
32. Mordvintsev, A. and A. K., *Introduction to SIFT (Scale-Invariant Feature Transform).* [Online]. Available: http://opencv-python-tutroals.readthedocs.org/en/latest/py_tutorials/py_feature2d/py_sift_intro/py_sift_intro.html#sift-intro.
33. Bay, H., T. Tuytelaars, and L. Van Gool, *SURF: Speeded Up Robust Features*, in *Computer Vision – ECCV 2006*, A. Leonardis, H. Bischof, and A. Pinz, Editors. 2006, Springer Berlin Heidelberg. p. 404-417.
34. Ying, Z., Y. Lei, and W. Zhujun. *Research on Video Image Stitching Technology Based on SURF*. in *Computational Intelligence and Design (ISCID), 2012 Fifth International Symposium on*. 2012.
35. Mordvintsev, A. and A. K., *Introduction to SURF (Speeded-Up Robust Features).* [Online]. Available: http://opencv-python-

tutroals.readthedocs.org/en/latest/py_tutorials/py_feature2d/py_surf_intro/py_surf_intro.html#surf.

36. Mordvintsev, A. and A. K., *Feature Matching + Homography to find Objects.* [Online]. Available: http://opencv-python-tutroals.readthedocs.org/en/latest/py_tutorials/py_feature2d/py_feature_homography/py_feature_homography.html#feature-homography.
37. Henry, S.L., *Web Content Accessibility Guidelines (WCAG) Overview*. (Online). http://www.w3.org/WAI/intro/wcag.
38. Chisholm, W., G. Vanderheiden, and I. Jacobs, *Web Content Accessibility Guidelines 1.0*. (Online). http://www.w3.org/TR/1999/WAI-WEBCONTENT-19990505/.
39. Caldwell, B., et al., *Web Content Accessibility Guidelines (WCAG) 2.0*. (Online). http://www.w3.org/TR/WCAG20/.
40. Loomis, J.M., et al., *Personal Guidance System for People with Visual Impairment: A Comparison of Spatial Displays for Route Guidance.* Journal of visual impairment & blindness, 2005. 99(4): p. 219-232.
41. Strothotte, T., et al., *Development of dialogue systems for a mobility aid for blind people: initial design and usability testing*, in *Proceedings of the second annual ACM conference on Assistive technologies*. 1996, ACM: Vancouver, British Columbia, Canada. p. 139-144.
42. Amirante, A., et al., *On the seamless interaction between webRTC browsers and SIP-based conferencing systems.* Communications Magazine, IEEE, 2013. 51(4): p. 42-47.
43. Hongo, N., H. Yamamoto, and K. Yamazaki. *Web shopping support system for elderly people using WebRTC*. in *Advanced Communication Technology (ICACT), 2014 16th International Conference on*. 2014.
44. Chuan-Yen, C., et al. *A Video Conferencing System Based on WebRTC for Seniors*. in *Trustworthy Systems and their Applications (TSA), 2014 International Conference on*. 2014.
45. Jang-Jaccard, J., et al., *WebRTC-based video conferencing service for telehealth.* Computing, 2014: p. 1-25.

46. Panchal, P.M., S.R. Panchal, and S.K. Shah, *A comparison sift and surf.* International Journal of Innovative Research in Computer and Communication Engineering, April 2013. Vol. 1: p. 323-327.
47. Bradski, G., *The OpenCV Library.* Dr. Dobb's Journal of Software Tools, 2000.
48. Priologic Software Inc. *EasyRTC*.
49. Codebox Software, *BitMeter II*. [Online]. Avaliable: http://codebox.org.uk/pages/bitmeter2.
50. CC-By 3.0 license, *WebView for Android.* [Online]. Available: https://developer.chrome.com/multidevice/webview/overview.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ประวัติผู้เขียนวิทยานิพนธ์

นางสาวณัฐณิชา มณีแสง เกิดเมื่อวันที่ 29 พฤษภาคม พ.ศ. 2535 สำเร็จการศึกษาระดับมัธยมที่โรงเรียนราชินีบน กรุงเทพมหานคร จากนั้นได้เข้าศึกษาระดับปริญญาบัณฑิต ที่ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย และสำเร็จการศึกษาในปี พ.ศ. 2557

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University