บทที่ 3
วิธีการดำเนินการและเครื่องมือ

บทนี้จะกล่าวถึงวิธีการผสมผสานเทคนิคการรู้จำเสียงและการตรวจจับการเคลื่อนไหว โดยใช้กล้องคีเนค วิธีการทดสอบ โดยสร้างสถานการณ์การใช้เทคนิคควบคุมอินเตอร์เน็ตทีวีจำลอง สามเทคนิค ได้แก่ การรู้จำเสียง,การตรวจจับการเคลื่อนไหว และเทคนิคผสมผสานการรู้จำเสียง และการตรวจจับการเคลื่อนไหว และพัฒนาโปรแกรมส่วนควบคุมอินเตอร์เน็ตทีวี ดังต่อไปนี้

3.1 การวางแผนการดำเนินงาน

3.2 การศึกษาค้นคว้าหาข้อมูล

3.3 การออกแบบและพัฒนา Interface ส่วนควบคุมอินเทอร์เน็ตทีวีจำลอง

3.4 การพัฒนาระบบการรู้จำเสียง

3.5 การพัฒนาระบบการตรวจจับการเคลื่อนไหว

3.6 การพัฒนาระบบเทคนิคการผสมผสานการรู้จำเสียงและการตรวจจับการเคลื่อนไหว

3.7 การกำหนดแบบแผนการทดลองและกลุ่มผู้ทดสอบ

3.8 การวัดผลทั้งด้านความเร็วและความถูกต้องในการควบคุมอินเทอร์เน็ตทีวีจำลอง

3.9 เครื่องมือที่ใช้จัดทำระบบควบคุมอินเทอร์เน็ตทีวีจำลอง

3.1 การวางแผนการดำเนินงาน

การวางแผนการดำเนินงานเป็นขั้นตอนสำคัญ และเป็นขั้นตอนแรก ก่อนที่ผู้วิจัยจะเริ่มทำงานวิจัย เพราะในการวางแผนการดำเนินงานนั้นจะบอกถึงรายละเอียดงานที่ผู้วิจัยควรทำก่อนหลัง รวมไปถึงระยะเวลาในการทำงานในแต่ละขั้นตอนของงาน ทำให้ผู้วิจัยสามารถดำเนินงานได้เสร็จตรงตามเวลาที่กำหนดไว้ได้ โดยได้ผลของการวางแผนดำเนินงานตามภาพที่ 3.1

ตารางที่ 3.1 ระยะเวลาในการดำเนินการพัฒนาระบบ

		เดือน											
ที่	**แผนดำเนินงาน**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
1	การวางแผนการดำเนินงาน	⇔											
2	การศึกษาค้นคว้าหาข้อมูล	←	→	→									
3	การออกแบบและพัฒนา Interface ของโปรแกรมส่วนควบคุมอินเทอร์เน็ตทีวีจำลอง		←	→	→								
4	การพัฒนาระบบการรู้จำเสียง			←	→	→							
5	การพัฒนาระบบการตรวจจับการเคลื่อนไหว				←	→	→						
6	การพัฒนาระบบเทคนิคการผสมผสานการรู้จำเสียงและการตรวจจับการเคลื่อนไหว					←	→	→					
7	ทดสอบโปรแกรมและแก้ไขข้อผิดพลาดของโปรแกรม							←	→				
8	ทดสอบการติดตั้ง และใช้งานจริง								←	→			
9	การกำหนดแบบแผนการทดลองและกลุ่มผู้ทดสอบ										⇔		
10	การวัดผลทั้งด้านความเร็วและความถูกต้องในการควบคุมอินเทอร์เน็ตทีวีจำลอง										←	→	
11	จัดทำเอกสาร นำเสนอผลงานและส่งผลงาน												⇔

3.2 การศึกษาค้นคว้าหาข้อมูล

ผู้วิจัยได้ศึกษาและค้นคว้าหาข้อมูลจากแหล่งต่างๆ ที่จะสร้างการควบคุมอินเทอร์เน็ตทีวีจำลองด้วย ผสมผสานเทคนิคการรู้จำเสียงและการตรวจจับการเคลื่อนไหว ในหัวข้อต่อไปนี้

3.2.1 ศึกษาการใช้งานกล้องคีเนค

โดยผู้วิจัยได้ศึกษาหาข้อมูลเกี่ยวกับอุปกรณ์ฮาร์ดแวร์ต่างๆ ที่มีคุณสมบัติที่มีความสามารถในการรับรู้ภาพและเสียงได้ในเวลาเดียวกัน ซึ่งอุปกรณ์ Kinect ของไมโครซอฟท์ นั้นมีความสามารถในการรับรู้ภาพและเสียง โดยมีคุณสมบัติเด่นๆดังนี้

3.2.1.1 การรับรู้ภาพสี RBG มีความสามารถถ่ายภาพหรือบันทึกภาพเป็นวิดีโอได้

3.2.1.2 Depth sensor มีความสามารถในการแยกแยะภาพทางลึกของภาพได้

3.2.1.3 Multi-array microphone มีความสามารถในการรับรู้เสียง บันทึกเสียงได้ และสามารถแยกแยะทิศทางเสียงที่ได้เข้าสู่ตัวไมโครโฟน มีระบบตัดเสียงรบกวนผู้วิจัยจึงเลือกใช้กล้องคีเนคในงานวิจัย

3.2.2 ศึกษาวิธีสร้างการรู้จำเสียง

โดยผู้วิจัยได้ศึกษาตัวแบบของการรู้จำเสียง จากบทความที่ตีพิมพ์ในนิตยสารวิชาการต่างประเทศและในประเทศ รวมถึงบทความจากเว็บไซด์ต่างๆ หลากหลายตัวแบบผู้วิจัยจึงเลือกใช้โมดูลของไมโครซอฟท์ Speech SDK 5.1 ซึ่งมีคุณสมบัติรองรับภาษาต่างๆมากมาย ทั้งภาษาอังกฤษ จีน ญี่ปุ่น และสามารถทำงานร่วมกับอุปกรณ์กล้องคีเนคที่ผู้วิจัยใช้ได้เป็นอย่างดี โดยผู้วิจัยได้เลือกใช้วิธี Speech To Text (STT) ในการสร้างกระบวนการรู้จำเสียง

3.2.3 ศึกษาวิธีสร้างการตรวจจับการเคลื่อนไหว

ผู้วิจัยได้เลือกใช้ไลบรารี่ Kinect SDK Beta 1 มาเป็นเครื่องมือในการสร้างการตรวจจับการเคลื่อนไหว โดยใช้ภาษา C# โดยศึกษาวิธีการใช้งานจาก ตำราภาษาต่างประเทศ และบทความต่างๆทั้งในวารสารหรือบทความที่อยู่ในเว็บไซด์ในอินเทอร์เน็ต เพื่อนำความรู้ที่ได้มาสร้างกระบวนการตรวจจับการเคลื่อนไหว

3.2.4 ศึกษาวิธีสร้าง User Interface สำหรับติดต่อผู้ใช้

ผู้วิจัยได้เลือกใช้ Windows Presentation Foundation (WPF) ในการเขียนโปรแกรมเพื่อสร้าง User Interface สำหรับอินเทอร์เน็ตทีวีจำลอง โดยศึกษาวิธีการใช้งานจาก ตำราภาษาต่างประเทศ และบทความต่างๆทั้งในวารสารหรือบทความที่อยู่ในเว็บไซด์ในอินเทอร์เน็ต เพื่อนำความรู้ที่ได้มาสร้าง User Interface สำหรับติดต่อผู้ใช้

3.3 การออกแบบและพัฒนา Interface ส่วนควบคุมอินเทอร์เน็ตทีวีจำลอง

ในการออกแบบหน้าตาของอินเตอร์เฟส จะใช้ Windows Presentation Foundation (WPF) และใช้ XAML ในการออกแบบ Interface โดยจะมีรูปภาพบอกถึงสถานการณ์เข้าในงานโหมด การรู้จำเสียง และในขณะที่เปิดโหมดการทำงานนี้เมื่อผู้ใช้งานสั่งงานอินเทอร์เน็ตทีวีใดๆก็จะถูกแสดงผลในรูปแบบของ ข้อความที่หน้าจอ เพื่อบอกว่าผู้ใช้ได้ออกเสียงเพื่อสั่งงานอะไรไป หรือ Curser มือที่จะใช้สำหรับการการตรวจจับการเคลื่อนไหว

ภาพที่ 3.1 แสดงอินเทอร์เฟสของการระบบควบคุมอินเทอร์เน็ตทีวีจำลอง

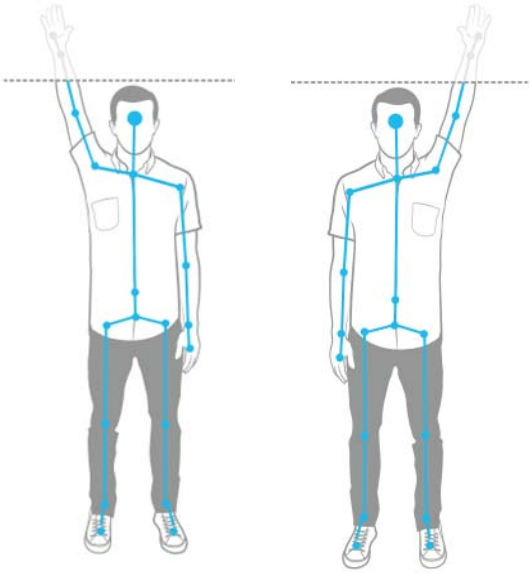
1. แสดงผลอินเตอร์เน็ตทีวีช่องที่กำลังใช้งาน
2. แสดงสถานการณ์เปิดหรือปิดระบบการรู้จำเสียง
3. แสดงข้อความที่ผู้ใช้ออกเสียงในการเปลี่ยนช่องอินเทอร์เน็ตทีวี
4. ระดับเสียงของอินเตอร์เน็ตทีวี
5. แสดงผังรายการให้ผู้ใช้
6. แสดงภาพของผู้ใช้ในกรณีที่ใช้โหมดการตรวจสอบการเคลื่อนไหว
7. แสดงรูป Curser มือใช้ในกรณีการตรวจจับการเคลื่อนไหวของมือขวาของผู้ใช้

3.4 การพัฒนาระบบการรู้จำเสียง

ในการตรวจสอบการรู้จำเสียงอินพุทที่เข้ามาผ่านไมโครโฟนบนตัวของกล้อง Kinect ได้นำไลบรารีของ Microsoft Speech Platform Runtime 11 โดยจะรองรับภาษา En-Us เขียนโปรแกรมผ่าน Visual Studio 2010 เพื่อนำอินพุทเสียงที่ได้แปลงเป็นตัวอักษรโดยใช้เทคนิค Speech to Text (STT) จากนั้นนำคำที่ได้ไปเปรียบเทียบกับคำที่ระบุไว้หากไม่ตรงตามที่ระบุไว้ก็จะ Reject ทิ้ง หากตรงตามเงื่อนไขก็จะ Accepted

รูปแบบของการ Capture Audio โดยจะให้คุณภาพของสียงอยู่ที่ อัตราการสุ่มเสียง 16000 Sampling Rate ค่าความลึกของเสียง 16 bit/sec ช่องสัญญาณเสียง 1 Channel อัตราการ Stream ข้อมูลเสียง 32000 byte/sec

โดยการสั่งงานด้วยเสียงผู้วิจัยจะออกแบบให้มีกระบวนการเปิดและปิดระบบการรู้จำเสียงดังภาพที่ 3.4 เพื่อป้องกันเสียงรบกวนในขณะที่ผู้ใช้งานยังไม่ได้มีความต้องการในการใช้คำสั่งจะกำหนดให้ผู้ใช้ยกมือขวาขึ้นเหนือศีรษะจะเป็นการเปิดระบบการรู้จำเสียงและหากต้องการปิดให้ผู้ใช้ยกมือซ้ายขึ้นเหนือศีรษะจะเป็นการปิดระบบเสียง โดยจะอาศัยตำแหน่งของข้อต่อที่ได้รับจาก API และแต่ละตำแหน่งของข้อต่อจะมีค่า X,Y,Z ในหน่วยมิลิเมตร จึงได้นำตำแหน่ง Joint Hand Right และ Joint Hand Left ทำการเปรียบเทียบตำแหน่ง Position ด้านแกน Y กับ Joint Head ดังภาพที่ 3.3



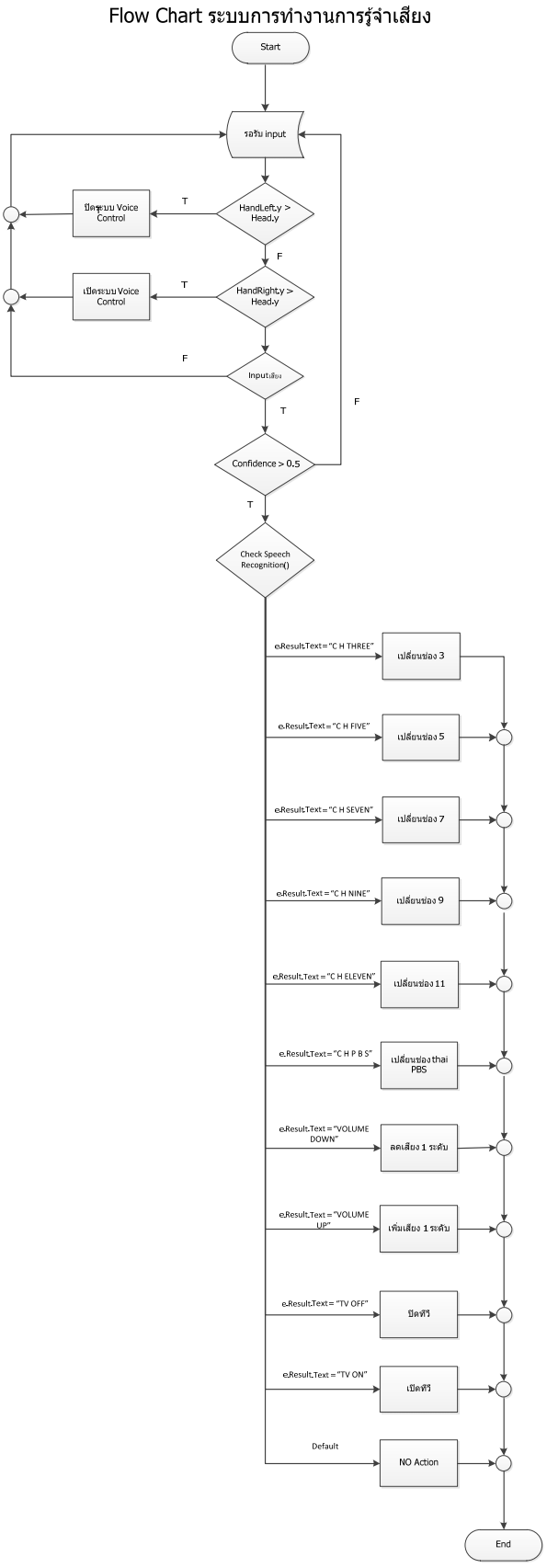
ภาพที่ 3.2 แสดงวิธีการเปิดปิดระบบการสั่งการด้วยเสียง

ที่มา: (http://video.ch9.ms/teched/2012/eu/DEV330.pptx,2012)

ภาพที่ 3.3 แสดง Interface ระบบการสั่งงานด้วยเสียง

ที่มา: (http://video.ch9.ms/teched/2012/eu/DEV330.pptx,2012)

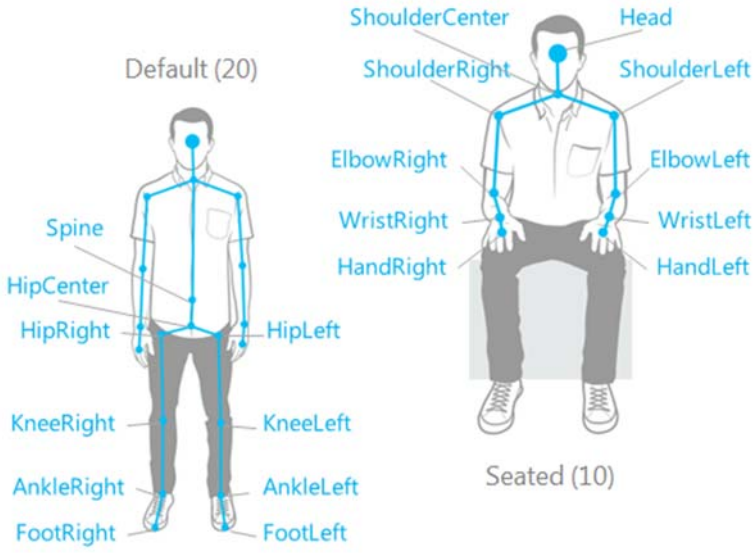
ในระบบการรู้จำเสียงของ Input ที่เข้ามา โดยจะทำการเปลี่ยนสัญญาณเสียงที่เข้าเพื่อนำไปเปรียบเทียบกับคำที่ระบุไว้จะ มีฟิลด์ Confidence แสดงสถานะ การตรวจจับของเสียงนั้นๆ โดยจะมีค่าระหว่าง 0 – 1 ซึ่ง 1 หมายถึง สามารถตรวจจับได้ และตรงกับเงื่อนไขของคำที่ระบุไว้ โดยจะระบุ ถ้ามากกว่า 0.5 ก็จะเป็นการ Accepted คำนั้นๆเพื่อนนำไปตรวจสอบ หากน้อยกว่า 0.5 ก็จะ Reject คำนั้นๆทิ้งไป



ภาพที่ 3.4 Flow Chart แสดงการทำงานระบบรู้จำเสียง

3.5 การพัฒนาระบบการตรวจจับการเคลื่อนไหว

ในการตรวจสอบการเคลื่อนไหวภาพผ่านหน้ากล้อง Kinectได้นำไลบรารีของ Kinect SDK โดยจะเขียนโปรแกรมผ่านโปรแกรม Visual Studio 2010 โดยทำการตรวจจับรูปร่างของคน แล้วจะนำ Skeleton ไปซ้อนทับบนภาพ โดยท่ายืนจะกำหนดเป็น 20 Joint และท่านั่งจะกำหนดเป็น 10 Joint จากนั้นก็ทำการเขียนโปรแกรมเก็บค่า เฉพาะ Joint Hand Right และ Joint Hand Left เพื่อนำค่า Position ของตำแหน่งมือไปเปรียบเทียบกับตำแหน่ง Cursor มือบนจอภาพ เพื่อให้ได้ตำแหน่งและพิกัด สำหรับการควบคุมอินเทอร์เน็ตทีวี โดยจะแบ่งขั้นตอนออกเป็นดังนี้

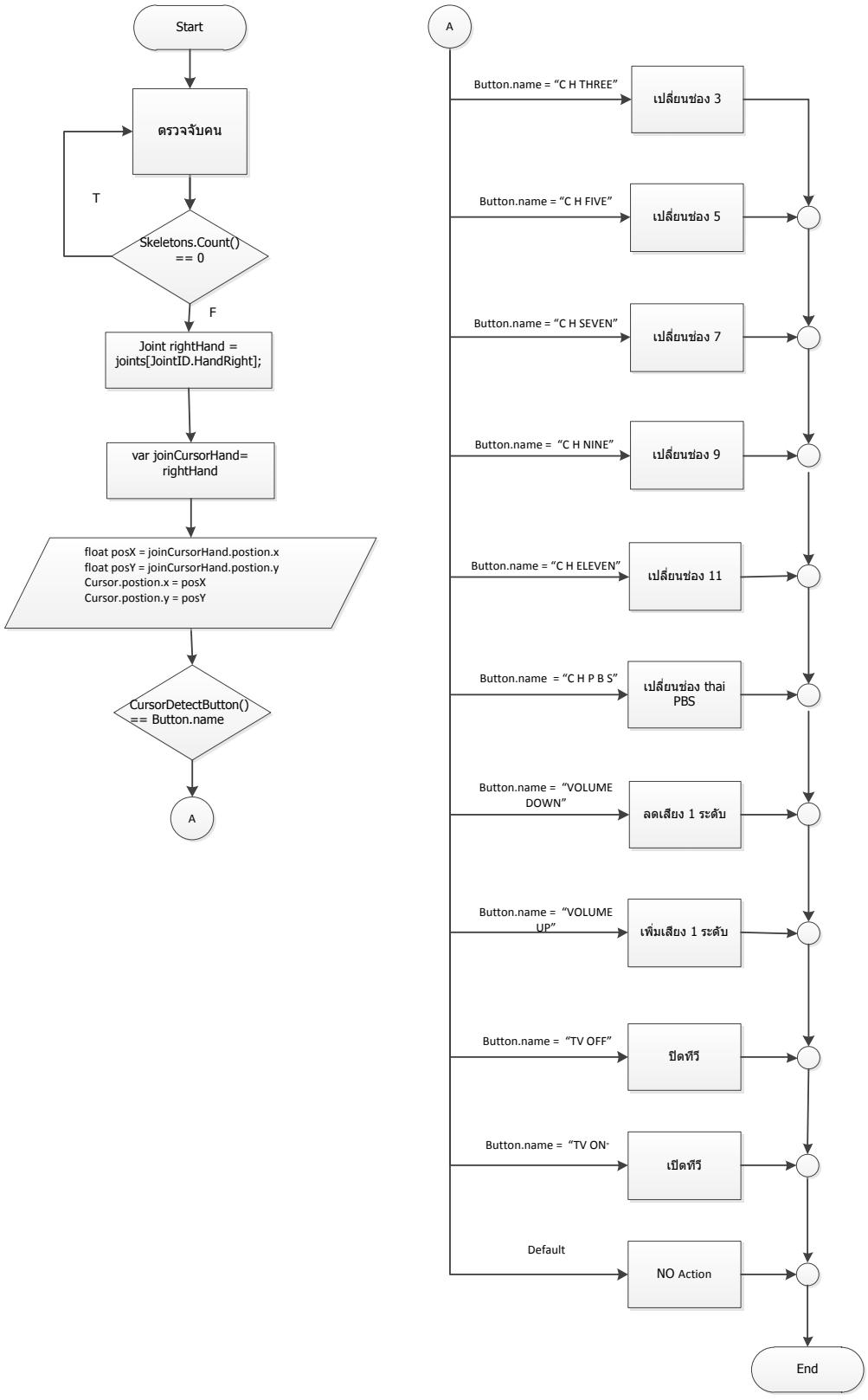


ภาพที่ 3.5 แสดงจุดข้อต่อที่ API ส่งค่ามา

ที่มา: (http://video.ch9.ms/teched/2012/eu/DEV330.pptx,2012)

Hand Tracking โดยจะเป็นขั้นตอนของการเชื่อมโยงข้อมูลที่ได้รับจากข้อต่อ API ในส่วนของ Hand Right กับรูปภาพ Curser มือบนจอภาพ โดยอันดับแรกให้โปรแกรมทำการเรียกข้อมูลข้อต่อสำหรับ คน 1 คนขึ้นมา จากนั้นเมื่อได้ข้อมูลชุดข้อต่อจากที่สามารถเรียกได้ผ่านตัวแปร first Person แล้วก็ทำการนำข้อมูลข้อต่อของโหนด Hand Right หาค่า Position X, Y ออกมา และนำไปเปรียบเทียบกับตำแหน่งของ Cursor Hand บนจอภาพ

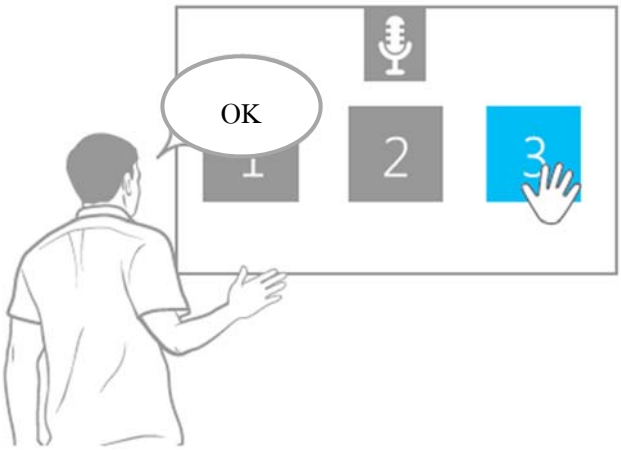
Flow Chart ระบบการทำงานการตรวจจับการเคลื่อนไหว



ภาพที่ 3.6 Flow Chart แสดงการทำงานระบบการตรวจจับการเคลื่อนไหว

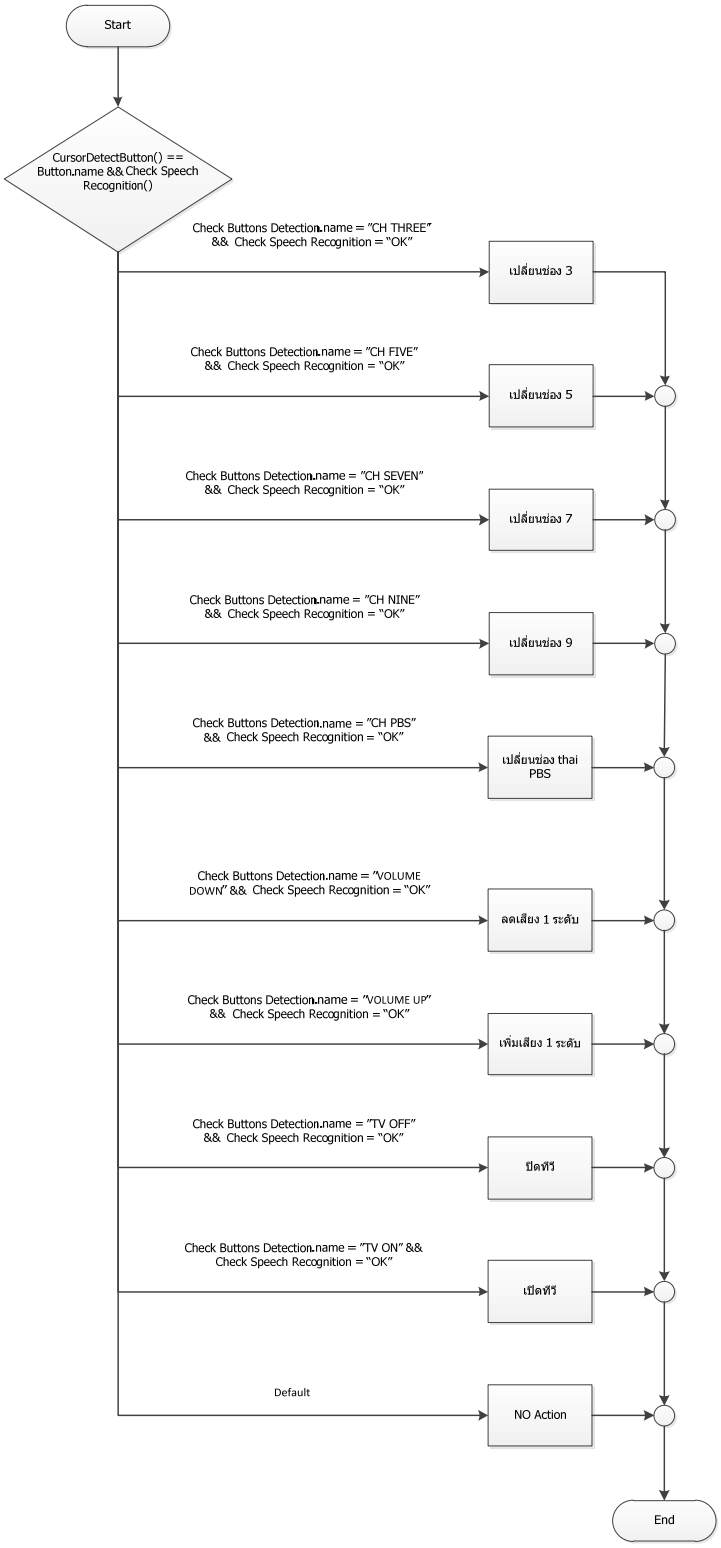
3.6 การพัฒนาระบบเทคนิคการผสมผสานการรู้จำเสียงและการตรวจจับการเคลื่อนไหว

เมื่อได้วิธีการรู้จำเสียงและการตรวจจับการเคลื่อนไหวแล้ว จึงได้ออกแบบวิธีการผสมผสานเทคนิคการรู้จำเสียงและการตรวจสอบการเคลื่อนไหวโดยกำหนดให้ เมื่อ Position ของ Cursor บนหน้าจออินเทอร์เน็ตทีวี ไป Detect กับปุ่มใดๆบนอินเตอร์เฟสบนหน้าพร้อมทั้งออกเสียง “OK” ก็จะเป็นการสั่งงานกับเมนูนั้นๆ ดังภาพที่ 3.8 แต่ถ้าหาก Cursor ไป Detect กับปุ่มใดๆบนอินเตอร์เฟสบนหน้าจอแล้วมีการออกเสียงอื่นโปรแกรมก็จะไม่มีการทำงาน หรือถ้าออกเสียงอย่างเดียวแล้วไม่ได้มีการนำ Cursor ไป Detect กับปุ่มใดๆ โปรแกรมก็จะไม่มีการทำงานอีกเช่นกัน



ภาพที่ 3.7 แสดงรูปแบบการทำงานของวิธีการผสมผสานเทคนิคการรู้จำเสียงและการตรวจสอบการเคลื่อนไหว

Flow Chart ระบบการทำงานการผสมผสานการรู้จำเสียง
และการตรวจจับการเคลื่อนไหว



ภาพที่ 3.8 Flow Chart แสดงการทำงานระบบการผสมผสานรู้จำเสียงและการตรวจจับการเคลื่อนไหว

3.7 การกำหนดแบบแผนการทดลองและกลุ่มผู้ทดสอบ

งานวิจัยนี้ทำการทดสอบประสิทธิภาพโดยการวัดความเร็วและความถูกต้องของวิธีการที่นำเสนอและได้ทำการทดลองทั้งหมด 2 ครั้ง โดยครั้งที่ 1 จะเป็นการทดสอบเบื้องต้น(ตารางที่ 3.2) และการทดสอบครั้งที่ 2 (ตารางที่ 3.3) จะมีการปรับโจทย์การทดสอบให้เหมือนกับการใช้คำสั่งในชีวิตประจำวันมากขึ้น โดยการพัฒนาโปรแกรมทั้งสามระบบ คือ 1.ระบบการรู้จำเสียงสำหรับการควบคุมอินเทอร์เน็ตทีวี 2.ระบบการตรวจสอบการเคลื่อนไหวสำหรับการควบคุมอินเทอร์เน็ตทีวี 3.ระบบผสมผสานระหว่างการรู้จำเสียงและการตรวจจับการเคลื่อนไหว ขึ้นมาด้วยภาษา C# ใช้ Kinect SDK เวอร์ชัน Beta 1 และสร้างสถานการณ์สภาพแวดล้อมในห้องขึ้นมาโดยใช้ห้องประชุมขนาดประมาณ 4x8 เมตร โดยให้สภาพแวดล้อมแรกเป็นสภาพแวดล้อมที่ไม่มีเสียงเข้ามารบกวนโดยที่ค่าเฉลี่ยเสียงภายในห้องที่ทดสอบอยู่ที่เฉลี่ย 46 dB สภาพแวดล้อมที่สองคือสภาพแวดล้อมที่มีเสียงรบกวนโดยกำหนดให้เสียงรบกวนอยู่ในระดับความดังที่ผู้ทดสอบจะสามารถได้ยินรับรู้เสียงที่ออกมาจากอินเทอร์เน็ตทีวีค่าเฉลี่ยของเสียงในห้องอยู่ที่เฉลี่ย 70 dB ในการตรวจจับตั้งกล้องอยู่บนระนาบที่ขนานกับระนาบพื้นที่ทำการตรวจจับสูงจากระนาบพื้น 0.6 เมตร ห่างออกไป 1.5 เมตร โดยกำหนดโจทย์ที่ใช้ทดสอบครั้งที่หนึ่ง มีรายละเอียดดังต่อไปนี้ ปิด TV, เลือกช่อง (เปิด TV), เปลี่ยนช่อง, เพิ่มเสียง, ลดเสียง

ตารางที่ 3.2 แสดงถึงงานที่ผู้ทดสอบครั้งที่ 1 ต้องทำในแต่ละชุดการทดสอบในท่ายืน

โจทย์สำหรับใช้ทดสอบ		
ชุดทดสอบ A	ชุดทดสอบ B	ชุดทดสอบ C
เลือกช่อง 3	เลือกช่อง 5	เลือกช่อง 9
เปลี่ยนช่อง 7	เพิ่มเสียงขึ้น 1 ระดับ	เพิ่มเสียงขึ้น 1 ระดับ
ลดเสียงลง 1 ระดับ	เปลี่ยนช่อง 9	ปิด TV
เพิ่มเสียงขึ้น 1 ระดับ	ลดเสียงลง 1 ระดับ	เปลี่ยนช่อง 11
ปิด TV	ปิด TV	ลดเสียงลง 1 ระดับ

โดยจะใช้ผู้ทดสอบจำนวน 60 คนนักศึกษาชั้นปี 1 – 3 ช่วงอายุประมาณ 19 – 21 ปี สถาบันเทคโนโลยีไทย-ญี่ปุ่น ซึ่งไม่มีประสบการณ์การใช้ระบบเหล่านี้มาก่อน โดยจะแบ่งการทดสอบเป็น 2 กลุ่ม กลุ่มละ 30 คนโดยให้กลุ่มแรกทำการทดสอบในสภาพแวดล้อมที่ไม่มีเสียงเข้า

มารบกวนค่าเฉลี่ยเสียงภายในห้องที่ทดสอบอยู่ที่เฉลี่ย 46 dB และกลุ่มที่สองอยู่ในระดับความดังที่ผู้ทดสอบจะสามารถได้ยินรับรู้เสียงที่ออกมาจากอินเทอร์เน็ตทีวีค่าเฉลี่ยของเสียงในห้องอยู่ที่เฉลี่ย 70 dB โดยให้แต่กลุ่มทำการทดสอบดังนี้

1.ให้ผู้ทดสอบ สุ่มเลือก เทคนิคสำหรับควบคุมอินเทอร์เน็ตทีวี จาก 3 วิธี มา 1 วิธี พร้อมทั้งนำหัวข้อเทคนิคที่ผู้ใช้เลือกได้ออกจากกองสุ่มเลือก

2.จากนั้นให้ผู้ทดสอบ สุ่มเลือก ชุดทดสอบ จาก 3 ชุดทดสอบ มา 1 ชุดทดสอบ พร้อมทั้งนำหัวข้อชุดทดสอบ ที่ผู้ใช้เลือกได้ออกจากกองสุ่มเลือก

3.ให้ผู้ทดสอบ เข้าอยู่ในบริเวณที่จะทดสอบ หากผู้ทดสอบเตรียมตัวพร้อมผู้ประเมินจะเริ่มทำการจับเวลาในการทดสอบ

4.เมื่อผู้ทดสอบ ทำตามเงื่อนไขจากชุดทดสอบจนเสร็จ ผู้ประเมินจะทำการหยุดเวลาพร้อมจดบันทึกเวลาที่ผู้ทดสอบได้ใช้ไป

5.ให้ผู้ทดสอบทำการ สุ่มเลือก เทคนิคสำหรับควบคุมอินเทอร์เน็ตทีวี จาก 2 วิธีที่เหลือมา 1 วิธี พร้อมทั้งนำหัวข้อเทคนิคที่ผู้ใช้เลือกได้ออกจากกองสุ่มเลือก

6.จากนั้นให้ผู้ทดสอบ สุ่มเลือก ชุดทดสอบ จาก 2 ชุดทดสอบ มา 1 ชุดทดสอบ พร้อมทั้งนำหัวข้อชุดทดสอบ ที่ผู้ใช้เลือกได้ออกจากกองสุ่มเลือก

7.ให้ผู้ทดสอบ เข้าอยู่ในบริเวณที่จะทดสอบ หากผู้ทดสอบเตรียมตัวพร้อมผู้ประเมินจะเริ่มทำการจับเวลาในการทดสอบ

8.เมื่อผู้ทดสอบ ทำตามเงื่อนไขจากชุดทดสอบจนเสร็จ ผู้ประเมินจะทำการหยุดเวลาพร้อมจดบันทึกเวลาที่ผู้ทดสอบได้ใช้ไป

9.จากนั้นให้ผู้ทดสอบทำการทดสอบ เทคนิคสำหรับควบคุมอินเทอร์เน็ตทีวีและชุดทดสอบ ที่เหลืออยู่

10.ให้ผู้ทดสอบ เข้าอยู่ในบริเวณที่จะทดสอบ หากผู้ทดสอบเตรียมตัวพร้อมผู้ประเมินจะเริ่มทำการจับเวลาในการทดสอบ

11.เมื่อผู้ทดสอบ ทำตามเงื่อนไขจากชุดทดสอบจนเสร็จ ผู้ประเมินจะทำการหยุดเวลาพร้อมจดบันทึกเวลาที่ผู้ทดสอบได้ใช้ไป

12.นำข้อมูลที่ได้จากทั้ง 3 กลุ่ม มาวิเคราะห์เปรียบเทียบความแตกต่างในเรื่องของความเร็วและความถูกต้องของการทำงาน

โจทย์ที่ใช้ทดสอบครั้งที่ 2 มีรายละเอียดดังต่อไปนี้ เปิด TV, ปิด TV, เลือกช่อง, เปลี่ยนช่อง, เพิ่มเสียง, ลดเสียง

ตารางที่ 3.3 แสดงถึงงานที่ผู้ทดสอบครั้งที่ 2 ต้องทำในชุดการทดสอบในท่ายืน

ลำดับที่	โจทย์สำหรับใช้ทดสอบ
1	เปิด TV (TV ON)
2	เปลี่ยนช่อง 5 (CH FIVE)
3	เปลี่ยนช่อง 9 (CH NINE)
4	เพิ่มเสียง 1 ระดับ (Volume UP)
5	เปลี่ยนช่อง 3 (CH THREE)
6	เปลี่ยนช่อง 7 (CH SEVEN)
7	ลดเสียง 1 ระดับ (Volume DOWN)
8	เพิ่มสียง 1 ระดับ (Volume Up)
9	ปิด TV (TV OFF)

โดยจะใช้ผู้ทดสอบจำนวน 30 คนนักศึกษาชั้นปี 1 – 3 ช่วงอายุประมาณ 19 – 21 ปี สถาบันเทคโนโลยีไทย-ญี่ปุ่น ซึ่งไม่มีประสบการณ์การใช้ระบบเหล่านี้มาก่อน โดยจะแบ่งการโดยให้ทำการทดสอบในสภาพแวดล้อมที่มีเสียงเข้ามารบกวนอยู่ในระดับความดังที่ผู้ทดสอบจะสามารถได้ยินรับรู้เสียงที่ออกมาจากอินเทอร์เน็ตทีวีค่าเฉลี่ยของเสียงในห้องอยู่ที่เฉลี่ย 70 dB โดยให้แต่กลุ่มทำการทดสอบดังนี้

1. ใส่ลำดับเทคนิคสำหรับควบคุมอินเตอร์เน็ตทีวี ทั้ง 6 ลำดับ ลำดับล่ะ 5 ชุด

ตารางที่ 3.4 ตารางแสดงลำดับการควบคุมในแต่ล่ะเทคนิคการควบคุมอินเทอร์เน็ตทีวี

ลำดับการควบคุม	**1 : Voice Only**	**2 : Motion Detection Only**	**3 : A Combination**
A	1	2	3
B	1	3	2

C	2	1	3
D	2	3	1
E	3	1	2
F	3	2	1

2. ผู้เข้าทดสอบจะสุ่มหยิบลำดับการควบคุม 1 ลำดับการควบคุมจากกองสุ่มเลือก (1 คน ต้องทำทั้ง 3 เทคนิค)

3. ผู้เข้าทดสอบหยิบโจทย์ที่ใช้ในการควบคุม (Task)

4. ผู้เข้าทดสอบ เข้ามายืนในจุด เริ่มการทดสอบ ผู้ควบคุมเวลาจะเริ่มต้นจับเวลา

5. ระหว่างที่ผู้ทดสอบทำโจทย์ทดสอบ ผู้ควบคุมเวลาจะคอยบันทึก ข้อผิดพลาดที่เกิดขึ้นในการควบคุม TV

6. เมื่อผู้ทดสอบทำ โจทย์ จนครบ ผู้ควบคุมเวลาจะบันทึกเวลาที่ใช้ในแต่ละวิธีในการควบคุม TV

7. ผู้ทดสอบคนต่อไปเข้ามาสุ่มหยิบลำดับและทำกระบวนการนี้ ซ้ำๆ จนครบทั้ง 30 คน

3.8 การวัดผลทั้งด้านความเร็วและความถูกต้องในการควบคุมอินเทอร์เน็ตทีวีจำลอง

ในส่วนของเครื่องมือในการวิจัยนั้น นอกจากการออกแบบเทคนิคการผสมผสานการรู้จำเสียงและการตรวจจับการเคลื่อนไหวในการควบคุมอินเตอร์เน็ตทีวีแล้ว ยังมีการวัดสถานะของระบบ ในเรื่องของความเร็วในการควบคุมอินเทอร์เน็ตทีวี และความถูกต้องในการสั่งการอินเทอร์เน็ตทีวี ทั้ง 3 วิธี คือ 1.ระบบการรู้จำเสียงสำหรับการควบคุมอินเทอร์เน็ตทีวี 2.ระบบการตรวจสอบการเคลื่อนไหวสำหรับการควบคุมอินเทอร์เน็ตทีวี 3.ระบบผสมผสานระหว่างการรู้จำเสียงและการตรวจจับการเคลื่อนไหว

3.8.1 การวัดผลด้านความเร็วในการควบคุมอินเทอร์เน็ตทีวีจำลอง

ในการวัดผลด้านความเร็วในการควบคุมคุมอินเทอร์เน็ตทีวีจำลองนั้น จะให้ผู้ทดสอบทำโจทย์การทดสอบ จนกว่าจะเสร็จสิ้นกระบวนการการสั่งการอินเทอร์เน็ตทีวีจำลอง ดังตารางที่ 3.2 จากนั้นจะนำค่าเวลาที่ได้ ของแต่ละคนทั้ง 30 คนมาหาค่าเฉลี่ยความเร็วในการควบคุมอินเทอร์เน็ตทีวีจำลอง ทั้ง 3 วิธี และ 2 สถานการณ์ คือ ในสภาพแวดล้อมที่ไม่มีเสียงเข้ามารบกวนค่าเฉลี่ยเสียงภายในห้องที่ทดสอบอยู่ที่เฉลี่ย 46 dB และกลุ่มที่สองอยู่ในระดับความดังที่ผู้ทดสอบจะสามารถได้ยินและรับรู้เสียงที่ออกมาจากอินเทอร์เน็ตทีวีค่าเฉลี่ยของเสียงในห้องอยู่ที่เฉลี่ย 70 dB ผู้วิจัยใช้สูตรดังนี้

$$\text{Average}_{(v,z)} = \frac{\sum N}{\text{Y}}$$

เมื่อ

Average = ค่าเฉลี่ยความเร็วในการควบคุมอินเทอร์เน็ตทีวีจำลองของแต่ละวิธี

V = รูปแบบที่ใช้ในการควบคุมอินเทอร์เน็ตทีวีจำลอง

Z = สภาพแวดล้อมในการทดสอบ

N = เวลาที่ได้จากการทำแบบทดสอบของแต่ละวิธี (วินาที)

Y = จำนวนผู้ทดลองทั้งหมดในแต่ละวิธี

3.8.2 การวัดผลด้านความถูกต้องในการควบคุมอินเทอร์เน็ตทีวีจำลอง

ในการวัดผลด้านความถูกต้องในการควบคุมอินเทอร์เน็ตทีวีจำลองนั้น จะให้ผู้ทดสอบทำโจทย์การทดสอบ จนกว่าจะเสร็จสิ้นกระบวนการการสั่งการอินเทอร์เน็ตทีวีจำลอง ดังตารางที่ 3.2 และ 3.3 แล้วนับจำนวนครั้งที่ผู้ทดสอบออกคำสั่งผิดพลาด หรือระบบทำงานผิดพลาด และ 2 สถานการณ์ คือ ในสภาพแวดล้อมที่ไม่มีเสียงเข้ามารบกวนค่าเฉลี่ยเสียงภายในห้องที่ทดสอบอยู่ที่เฉลี่ย 46 dB และกลุ่มที่สองอยู่ในระดับความดังที่ผู้ทดสอบจะสามารถได้ยินรับรู้เสียงที่ออกมาจากอินเทอร์เน็ตทีวีค่าเฉลี่ยของเสียงในห้องอยู่ที่เฉลี่ย 70 dB

3.9 เครื่องมือที่ใช้จัดทำระบบควบคุมอินเทอร์เน็ตทีวีจำลอง

ในการจัดทำระบบควบคุมอินเทอร์เน็ตทีวีจำลองนั้นใช้เครื่องมือประกอบด้วยส่วนของฮาร์ดแวร์ (Hardware) และส่วนของซอฟต์แวร์ (Software) สรุปได้ดังนี้

3.9.1 ฮาร์ดแวร์ (Hardware)

ในส่วนฮาร์ดแวร์นั้นผู้วิจัยได้ใช้เครื่องคอมพิวเตอร์แบบพกพา (Notebook Computer) ในการจัดทำระบบควบคุมอินเทอร์เน็ตทีวีจำลองและในการวัดประสิทธิภาพของระบบ ซึ่งคอมพิวเตอร์ที่ใช้นั้นมีรายละเอียดดังนี้

Brand: MSI

Model: GX640

CPU: Inter Core I5-420M

RAM: Samsung DD3 Ram 8 GB

VGA: ATI HD5850

Hard disk: Hitachi 500 GB SATA II 5200 rpm

Monitor: LED Display 15” resolution 1280 x 800

Microsoft Kinect Xbox360

3.8.2 ซอฟต์แวร์ (Software)

1. Microsoft Windows 7 Ultimate x64
2. Microsoft Kinect for Windows SDK BETA
3. Microsoft Visual Studio 2010
4. Microsoft Speech SDK 5.1
5. Adobe Photoshop CS6