

ห้องสมุดงานวิจัย สำนักงานคณะกรรมการวิจัยแห่งชาติ



247146

การออกแบบและประเมินส่วนรับเข้าควบคุมควันเล็ง

นายสุกเคช ฉันทร์สวัญ

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต

สาขาวิชาวิศวกรรมคอมพิวเตอร์ ภาควิชาวิศวกรรมคอมพิวเตอร์

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2553

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

60095301



247146

การออกแบบและประเมินส่วนรับเข้าควบคุมด้วยเสียง



นายศุภเดช ฉันทจรัสวิชัย

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต
สาขาวิชาวิศวกรรมคอมพิวเตอร์ ภาควิชาวิศวกรรมคอมพิวเตอร์
คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2553
ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย



5 1 7 0 4 8 0 5 2 1

A DESIGN AND EVALUATION OF VOICE-CONTROLLED INPUT

Mr. Supadaech Chanjaradwichai

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 2010

Copyright of Chulalongkorn University

หัวข้อวิทยานิพนธ์

การออกแบบและประเมินส่วนรับเข้าควบคุมด้วยเสียง

โดย

นายสุกเดช ฉันทจรัสวิชัย

สาขาวิชา

วิศวกรรมคอมพิวเตอร์

อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก

ผู้ช่วยศาสตราจารย์ ดร. โปรตปราน บุญพุกกณะ

อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม

ผู้ช่วยศาสตราจารย์ ดร.อดิวงค์ สุชาโต

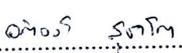
คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้รับวิทยานิพนธ์ฉบับนี้เป็น
ส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาโท

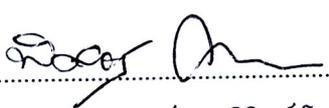

..... คณบดีคณะวิศวกรรมศาสตร์
(รองศาสตราจารย์ ดร.บุญสม เลิศหิรัญวงศ์)

คณะกรรมการสอบวิทยานิพนธ์


..... ประธานกรรมการ
(ศาสตราจารย์ ดร.ประภาส จงสถิตย์วัฒนา)


..... อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก
(ผู้ช่วยศาสตราจารย์ ดร. โปรตปราน บุญพุกกณะ)


..... อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม
(ผู้ช่วยศาสตราจารย์ ดร.อดิวงค์ สุชาโต)


..... กรรมการภายนอกมหาวิทยาลัย
(ผู้ช่วยศาสตราจารย์ ดร. พิศิษฎ์ โภคารัตน์กุล)

ศุภเดช ฉันทจรัสวิชัย : การออกแบบและประเมินส่วนรับเข้าควบคุมด้วยเสียง (A DESIGN AND EVALUATION OF VOICE-CONTROLLED INPUT). อ.ที่ปรึกษาวิทยานิพนธ์หลัก : ผศ. ดร.โปรดปราน บุญยพุกกณะ, อ.ที่ปรึกษาวิทยานิพนธ์ร่วม: ผศ. ดร.อดิวงค์ สุชาโต, 92 หน้า

247146

งานวิจัยนี้ได้เสนอถึงวิธีการใช้เสียงเพื่อเป็นส่วนรับเข้าของคอมพิวเตอร์ ซึ่งเน้นให้ครอบคลุมงานที่จำเป็นในการใช้งานคอมพิวเตอร์ แต่ไม่รวมถึงงานที่เกี่ยวกับการพิมพ์ข้อความ และใช้คำสั่งในการควบคุมให้น้อยที่สุด วิธีการควบคุมแบ่งเป็นหลายสถานะ ซึ่งประกอบด้วยสถานะเลือกรูปแบบการควบคุม สถานะเลือกเป้าหมายที่ใช้ในการเปลี่ยนตำแหน่งของตัวชี้ตำแหน่งหรือวัตถุที่ถูกเลือก และสถานะเพิ่มเติมที่ใช้ในการทำการดำเนินการเพิ่มเติมไปยังเป้าหมาย ประกอบด้วย การคลิก การคลิกสองครั้ง การคลิกขวา การเลื่อน การลาก การปล่อย และการเรียกใช้วัตถุที่ถูกเลือก ทั้งหมดนี้ควบคุมโดยใช้เสียงฮัมที่แบ่งเป็นสี่ระดับเสียงที่แตกต่างกัน และเสียงเสียดแทรก เสียงฮัมแทนการสั่งให้เคลื่อนที่ และเสียงเสียดแทรกใช้ในการหยุดการเคลื่อนที่ หรือการเลือกวัตถุ จากผลการทดลองพบว่าวิธีการที่นำเสนอมีความแม่นยำมากกว่าไวคอลลอยสติก และยังมี ความเหนือในการใช้งานที่น้อยกว่า แต่ใช้เวลาในการทำงานที่ได้รับมอบหมายมากกว่า

ภาควิชาวิศวกรรมคอมพิวเตอร์..... ลายมือชื่อนิสิต..... *ศุภเดช ฉันทจรัสวิชัย*.....
 สาขาวิชาวิศวกรรมคอมพิวเตอร์..... ลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์หลัก..... *[Signature]*.....
 ปีการศึกษา ..2553..... ลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์ร่วม..... *อดิวงค์ สุชาโต*.....

5170480521 : MAJOR COMPUTER ENGINEERING

KEYWORDS : NON-VERBAL / VOICE-CONTROLLED / VOICE INPUT / VOICE CURSOR / HUMMING / FRICATIVE

247146

SUPADAECH CHANJARADWICHAI : A DESIGN AND EVALUATION OF VOICE-CONTROLLED INPUT. ADVISOR : ASST. PROF. PROADPRAN PUNYABUKKANA, Ph.D., CO-ADVISOR : ASST. PROF. ATIWONG SUCHATO, Ph.D., 92 pp.

This thesis proposes a method for controlling a computer via speech input that covers all key computer operations except for typing-based input with a small set of voice command. The method consists of multiple states, including: ‘control method selection state’, ‘target selection state’ for changing the position of the cursor or selected item, and ‘option state’ for applying additional operations to the target. The operations are clicking, double-clicking, right-clicking, scrolling, dragging, dropping and activating selected item. All are controlled by hum sounds with four different pitch types and a fricative sound. Hum sounds are dedicated to movement while the fricative sound is for stopping or activation. We have found that the proposed method showed better precision than Vocal Joystick while it was less tiring to use despite that longer intervals were usually needed to complete given tasks.

Department:Computer Engineering... Student’s Signature
 Field of Study:Computer Engineering... Advisor’s Signature
 Academic Year: ..2010..... Co-advisor’s Signature

กิตติกรรมประกาศ

วิทยานิพนธ์นี้สำเร็จลุล่วงได้ด้วยดีก็ด้วยความกรุณาจากท่านอาจารย์ที่ปรึกษา ผศ. ดร. อติ-
วงศ์ สุชาโต, ผศ. ดร. โปรดปราน บุญยพุกกณะ ตลอดจนเพื่อน ๆ และพี่ ๆ ที่ได้สละเวลาในการให้
คำปรึกษา ชี้แนะแนวทางในงานวิจัยต่าง ๆ ซึ่งผู้เขียนวิทยานิพนธ์ไม่ได้มีพื้นฐานทางด้าน
เทคโนโลยีเสียงพูด และการออกแบบวิธีการควบคุมคอมพิวเตอร์มาก่อนมาก่อน ถ้าไม่ได้ความ
ช่วยเหลือของท่านเหล่านี้ วิทยานิพนธ์นี้ก็มิอาจสำเร็จลุล่วงไปได้

นอกจากนี้ ผมต้องขอขอบพระคุณคณะกรรมการสอบวิทยานิพนธ์ ซึ่งได้แก่ ศ. ดร.
ประภาส จงสถิตย์วัฒนา และ ผศ.ดร. พิศิษฎ์ โภคารัตน์กุล ที่ได้สละเวลาเป็นกรรมการสอบ
วิทยานิพนธ์ และให้คำแนะนำอันเป็นประโยชน์ ซึ่งช่วยเพิ่มคุณค่าให้แก่งานวิจัยนี้

ขอขอบคุณ ผศ. ดร. อติวงศ์ สุชาโต ภาควิชาวิศวกรรมคอมพิวเตอร์ และบัณฑิตวิทยาลัย ที่
ได้ให้เงินทุนสนับสนุนงานวิจัย และค่าใช้จ่ายต่าง ๆ ซึ่งช่วยให้งานวิจัยนี้สำเร็จลุล่วงได้ดี

ขอขอบคุณ ผศ. ดร. อติวงศ์ สุชาโต และ ผศ. ดร. โปรดปราน บุญยพุกกณะ ที่ได้ให้โอกาส
ผมในการเข้าศึกษาในระดับปริญญาโท

สุดท้ายนี้ขอขอบคุณพี่ ๆ และน้อง ๆ ในห้องปฏิบัติการ SLS และ ATL และน้อง ๆ ใน
ภาควิชาคอมพิวเตอร์ และที่อื่น ๆ ที่มาร่วมช่วยทดสอบ โปรแกรม และเป็นตัวอย่างเสียงในการ
พัฒนาโปรแกรม ซึ่งถ้าไม่มีบุคคลนั้น งานวิจัยนี้คงไม่สามารถสำเร็จลุล่วงได้ดี ขอขอบคุณไอซ์ที่
ช่วยเหลือในด้านต่าง ๆ และช่วงหาแนวทางเพื่อช่วยให้สามารถผ่านพ้นช่วงเวลาที่กำลังลำบากมาได้

สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	ง
บทคัดย่อภาษาอังกฤษ.....	จ
กิตติกรรมประกาศ.....	ฉ
สารบัญ.....	ช
สารบัญตาราง.....	ฅ
สารบัญภาพ.....	ฉ
บทที่ 1 บทนำ.....	1
1.1 ความเป็นมาและความสำคัญของปัญหา.....	1
1.2 วัตถุประสงค์ของการวิจัย.....	2
1.3 ขอบเขตของการวิจัย.....	2
1.4 ขั้นตอนการดำเนินงานวิจัย.....	2
1.5 ประโยชน์ที่คาดว่าจะได้รับจากการวิจัย.....	3
1.6 เนื้อหาในวิทยานิพนธ์.....	3
1.7 งานตีพิมพ์.....	4
บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง.....	5
2.1 ทฤษฎีที่เกี่ยวข้อง.....	5
2.1.1 การแปลงฟูรีเยร์แบบวิฤต.....	5
2.1.2 สเปกโตรแกรม.....	6
2.1.3 อัตรสหสัมพันธ์.....	8
2.1.4 เสียงชนิดต่างๆ.....	9
2.2. ทบทวนวรรณกรรม.....	13
บทที่ 3 การออกแบบส่วนรับเข้าส่งการด้วยเสียง.....	18
3.1 การศึกษารูปแบบการควบคุมคอมพิวเตอร์.....	18
3.2 การออกแบบวิธีการใช้ตัวชี้ตำแหน่ง.....	19
3.2.1 สถานะเลือกรูปแบบการควบคุม.....	21
3.2.2 สถานะเลือกเป้าหมาย.....	21
3.2.3 สถานะเพิ่มเติม.....	23
3.3 ส่วนของการเลือกเสียงที่เหมาะสม.....	24

	หน้า
3.4 ส่วนของการรู้จำเสียง.....	26
3.4.1 การวิเคราะห์ระดับกรอบสัญญาณ.....	26
3.4.2 การวิเคราะห์ระดับส่วนของกรอบสัญญาณ.....	27
3.5 การเชื่อมโยงคำสั่งกับเสียงในการควบคุม.....	31
3.6 ตัวอย่างการควบคุม.....	31
3.7 วิธีการพัฒนาโปรแกรม.....	35
บทที่ 4 วิธีการทดลอง.....	36
4.1 จุดประสงค์ของการทดลอง.....	36
4.2 ส่วนรับเข้าด้วยเสียงที่นำมาเปรียบเทียบ.....	36
4.3 ระบบที่ใช้ในการทดลอง.....	37
4.4 ผู้ทดลอง.....	38
4.5 วิธีการทดลอง.....	38
4.5.1 การทดลองที่ 1.....	38
4.5.2 การทดลองที่ 2.....	42
บทที่ 5 ผลการทดลอง และการอภิปรายผลการทดลอง.....	47
5.1 ผลการทดลองที่ 1.....	47
5.2.1 การคำนวณความเร็วในการใช้งานส่วนรับเข้าด้วยเสียงที่นำเสนอ.....	49
5.2 ผลการทดลองที่ 2.....	50
5.2.1 ปัญหาในการทดลองใช้ส่วนรับเข้าด้วยเสียง.....	50
5.2.2 ความถูกต้องในการรู้จำเสียง.....	50
5.2.3 รูปแบบการเคลื่อนที่ที่พัฒนาขึ้นมาใหม่.....	56
5.2.4 ความเร็วในการเข้าถึงเป้าหมาย.....	60
5.2.5 ความแม่นยำในการเข้าถึงเป้าหมาย.....	64
5.2.6 ความเหนื่อยในการใช้งาน.....	67
5.2.7 ความนิยมในการเลือกใช้.....	68
บทที่ 6 สรุปผลการวิจัย และข้อเสนอแนะ.....	69
รายการอ้างอิง.....	72
ภาคผนวก.....	75
ประวัติผู้เขียนวิทยานิพนธ์.....	92

สารบัญตาราง

	หน้า
ตารางที่ 3.1 แสดงสิ่งเพิ่มเติมที่ปรากฏในรูปแบบการเคลื่อนที่ต่าง ๆ.....	24
ตารางที่ 3.2 คุณสมบัติที่ใช้ในการแบ่งแยกเสียงชนิดต่าง ๆ.....	27
ตารางที่ 5.1 เวลาในการตอบสนองของตัวชี้ตำแหน่ง VJ และ U3I.....	47
ตารางที่ 5.2 เวลาในการตอบสนองของส่วนรับเข้าด้วยเสียงที่นำเสนอ.....	47
ตารางที่ 5.3 ผลการทดลองในส่วนของคุณถูกต้องในการรู้จำเสียงของส่วนรับเข้าด้วยเสียงที่ นำเสนอ.....	52
ตารางที่ 5.4 ผลการทดลองในส่วนของคุณถูกต้องในการรู้จำเสียงของตัวชี้ตำแหน่ง VJ.....	54
ตารางที่ 5.5 รูปแบบการเคลื่อนที่ที่ผู้ทดลองเลือกใช้ในการควบคุมด้วยส่วนรับเข้าด้วยเสียงที่ นำเสนอ.....	57
ตารางที่ 5.6 จำนวนคำสั่งที่ใช้ในการทำงานย่อยที่ 2 และงานย่อยที่ 4.....	58
ตารางที่ 5.7 เวลาที่ใช้ในการทำงานย่อยด้วยส่วนรับเข้าด้วยเสียงที่นำเสนอ.....	62
ตารางที่ 5.8 เวลาที่ใช้ในการทำงานย่อยด้วยตัวชี้ตำแหน่ง VJ.....	63
ตารางที่ 5.9 จำนวนครั้งที่ตัวชี้ตำแหน่งเคลื่อนที่เลยเป้าหมาย.....	65
ตารางที่ 5.10 ร้อยละของการที่ตัวชี้ตำแหน่งเคลื่อนที่เลยเป้าหมาย.....	66
ตารางที่ 5.11 ผลของแบบสอบถามด้านความเหนื่อย.....	67
ตารางที่ 5.12 ผลของแบบสอบถามด้านความนิยมในการเลือกใช้.....	68
ตารางที่ ก.1 ผลการทดลองด้วยส่วนรับเข้าด้วยเสียงที่นำเสนอของผู้ทดลอง M1.....	76
ตารางที่ ก.2 ผลการทดลองด้วยตัวชี้ตำแหน่ง VJ ของผู้ทดลอง M1.....	77
ตารางที่ ก.3 ผลการทดลองด้วยส่วนรับเข้าด้วยเสียงที่นำเสนอของผู้ทดลอง M2.....	78
ตารางที่ ก.4 ผลการทดลองด้วยตัวชี้ตำแหน่ง VJ ของผู้ทดลอง M2.....	79
ตารางที่ ก.5 ผลการทดลองด้วยส่วนรับเข้าด้วยเสียงที่นำเสนอของผู้ทดลอง M3.....	80
ตารางที่ ก.6 ผลการทดลองด้วยตัวชี้ตำแหน่ง VJ ของผู้ทดลอง M3.....	81
ตารางที่ ก.7 ผลการทดลองด้วยส่วนรับเข้าด้วยเสียงที่นำเสนอของผู้ทดลอง M4.....	82
ตารางที่ ก.8 ผลการทดลองด้วยตัวชี้ตำแหน่ง VJ ของผู้ทดลอง M4.....	83
ตารางที่ ก.9 ผลการทดลองด้วยส่วนรับเข้าด้วยเสียงที่นำเสนอของผู้ทดลอง F1.....	84
ตารางที่ ก.10 ผลการทดลองด้วยตัวชี้ตำแหน่ง VJ ของผู้ทดลอง F1.....	85
ตารางที่ ก.11 ผลการทดลองด้วยส่วนรับเข้าด้วยเสียงที่นำเสนอของผู้ทดลอง F2.....	86
ตารางที่ ก.12 ผลการทดลองด้วยตัวชี้ตำแหน่ง VJ ของผู้ทดลอง F2.....	87
ตารางที่ ก.13 ผลการทดลองด้วยส่วนรับเข้าด้วยเสียงที่นำเสนอของผู้ทดลอง F3.....	88

	หน้า
ตารางที่ ก.14 ผลการทดลองด้วยตัวชี้ตำแหน่ง VJ ของผู้ทดลอง F3.....	89
ตารางที่ ก.15 ผลการทดลองด้วยส่วนรับเข้าด้วยเสียงที่นำเสนอของผู้ทดลอง F4.....	90
ตารางที่ ก.16 ผลการทดลองด้วยตัวชี้ตำแหน่ง VJ ของผู้ทดลอง F4.....	91

สารบัญภาพ

	หน้า
รูปที่ 2.1 สัญญาณในโดเมนเวลา.....	5
รูปที่ 2.2 สัญญาณในโดเมนของความถี่.....	6
รูปที่ 2.3 ความสัมพันธ์ระหว่างสัญญาณทางเวลา และสเปกโตรแกรม.....	7
รูปที่ 2.4 ผลลัพธ์จากการคำนวณหาค่าอัตราสัมพันธ์ของฟังก์ชันรายคาบ.....	8
รูปที่ 2.5 ผลลัพธ์จากการคำนวณหาค่าอัตราสัมพันธ์ของฟังก์ชันที่ไม่เป็นรายคาบ.....	9
รูปที่ 2.6 ความแตกต่างของเสียงฮัม.....	10
รูปที่ 2.7 ความแตกต่างของเสียงสระ.....	11
รูปที่ 2.8 แสดงความแตกต่างของเสียงสระ และเสียงเสียดแทรก.....	12
รูปที่ 2.9 ความแตกต่างของเสียงเสียดแทรก ที่มีฐานกรณ์บริเวณริมฝีปาก.....	13
รูปที่ 2.10 ส่วนต่อประสานกับผู้ใช้ของตัวชี้ตำแหน่ง VJ.....	16
รูปที่ 2.11 การฮัมในลักษณะต่างๆ เพื่อใช้ควบคุมตัวชี้ตำแหน่ง U3I.....	17
รูปที่ 3.1 แผนภาพแสดงสถานะของ โปรแกรม.....	20
รูปที่ 3.2 ตัวอย่างโปรแกรมในส่วนของรายการเลือกรูปแบบการเคลื่อนที่.....	21
รูปที่ 3.3 แผนภาพแสดงกระบวนการในการแบ่งแยกเสียงชนิดต่าง ๆ.....	28
รูปที่ 3.4 แผนภาพแสดงเหตุการณ์ในการวิเคราะห์ระดับส่วนของกรอบสัญญาณ.....	29
รูปที่ 3.5 แผนภาพแสดงกระบวนการในระดับส่วนของกรอบสัญญาณเสียงที่เป็นเสียงฮัม.....	30
รูปที่ 3.6 หน้าต่างเริ่มต้นของตัวชี้ตำแหน่งที่พัฒนาขึ้นมา.....	31
รูปที่ 3.7 หน้าต่างเมื่อเข้าสู่การเคลื่อนที่แบบที่ใช้เลือกวัตถุในรายการเริ่มต้นของวินโดวส์.....	32
รูปที่ 3.8 ตัวอย่างการเคลื่อนที่เพื่อเข้าถึงในรายการ.....	33
รูปที่ 3.9 รายการของสถานะเพิ่มเติมของรูปแบบที่ใช้เลือกวัตถุในรายการเริ่มต้นของวินโดวส์.....	33
รูปที่ 3.10 ขั้นตอนการเลือกวัตถุในรายการ.....	34
รูปที่ 3.11 ตัวอย่างการใช้งานในรูปแบบการเคลื่อนที่แบบต่อเนื่อง.....	34
รูปที่ 3.12 รายการของสถานะเพิ่มเติมของรูปแบบการเคลื่อนที่ของตัวชี้ตำแหน่งแบบต่อเนื่อง.....	34
รูปที่ 3.13 ภาพรวมการทำงานของโปรแกรม.....	35
รูปที่ 4.1 รายละเอียดของระบบประมวลผลของเครื่องคอมพิวเตอร์ที่ใช้ในการทดสอบ.....	37
รูปที่ 4.2 ช่วงเวลาในช่วงเริ่มการเคลื่อนที่.....	39
รูปที่ 4.3 ช่วงเวลาในช่วงหยุดการเคลื่อนที่.....	40
รูปที่ 4.4 ช่วงเวลาในการตอบสนองต่อเสียงฮัม.....	41
รูปที่ 4.5 ช่วงเวลาในการตอบสนองต่อเสียงเสียดแทรก.....	42

	หน้า
รูปที่ 4.6 ลำดับการเข้าถึงของงานย่อยที่ 1.....	43
รูปที่ 4.7 ลำดับการเข้าถึงของงานย่อยที่ 2.....	44
รูปที่ 4.8 ลำดับการเข้าถึงของงานย่อยที่ 3.....	45
รูปที่ 4.9 หน้าต่างของโปรแกรมไมโครซอฟท์เพนที.....	45
รูปที่ 5.1 การใช้ทรัพยากรของหน่วยประมวลผล.....	48