

การเปลี่ยนข้อมูลเสียงเป็นการเคลื่อนไหวของรูปปากตามลักษณะเสียงสระภาษาไทย



นาย พิชัย เอี่ยมธนาอนันต์

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

สาขาวิชาวิทยาการคณนา ภาควิชาคณิตศาสตร์

คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2548

ISBN 974-14-3774-9

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

AUDIO CONVERSION TO LIP SHAPE MOVEMENT FOR THAI VOWELS

Pichai lamtana-anan

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Science Program in Computational Science

Department of Mathematics

Faculty of Science

Chulalongkorn University

Academic Year 2005

ISBN 974-14-3774-9

พิชัย เตียมธนาอนันต์ : การเปลี่ยนข้อมูลเสียงเป็นการเคลื่อนไหวของรูปปากตามลักษณะเสียง
 สระภาษาไทย. (AUDIO CONVERSION TO LIP SHAPE MOVEMENT FOR THAI
 VOWELS), อ. ที่ปรึกษา : ศ. ดร.ชิตชนก เหลือสินทรัพย์, อ.ที่ปรึกษาร่วม :
 รศ. ดร. สมชาย จิตะพันธ์กุล , 68หน้า. ISBN 974-14-3774-9.

ในกระบวนการสร้างภาพเคลื่อนไหวเพื่อเลียนแบบการเคลื่อนไหวของสิ่งมีชีวิตนั้น จำเป็นต้อง
 มีการศึกษาถึงรูปแบบการเคลื่อนไหวของสิ่งมีชีวิตที่เราต้องการเลียนแบบการเคลื่อนไหวนั้น การพูดเอง
 ก็มีรูปแบบของการเคลื่อนไหวที่มีความสัมพันธ์กับการออกเสียงพูด ซึ่งสามารถสร้างแบบจำลองเพื่อ
 การเลียนแบบได้เช่นกัน งานวิจัยชิ้นนี้ได้ทำการศึกษาถึงความสัมพันธ์ของการออกเสียงสระที่มีต่อการ
 เคลื่อนไหวของรูปปาก โดยพิจารณาองค์ประกอบของรูปปากในลักษณะของจุดพิกัด และทำการเก็บ
 ข้อมูลเพื่อสร้างแบบจำลองด้วย Backpropagation Neural Network ขอบเขตของการศึกษากำหนดไว้
 ที่การเคลื่อนไหวของรูปปากที่เกิดจากเสียงสระเดี่ยวจำนวน 16 ตัว ที่ทำการพูดในลักษณะไม่ต่อเนื่อง
 และพิจารณาในลักษณะ 2 มิติที่ไม่คำนึงถึงการเปลี่ยนแปลงของกล้ามเนื้อที่มีผลจากการพูดออก
 เสียงนอกจากนี้ ยังได้ทำการเปรียบเทียบประสิทธิภาพของแบบจำลองที่สร้างขึ้นด้วยข้อมูลของผู้พูด
 เพียงคนเดียวกับแบบจำลองที่สร้างขึ้นจากข้อมูลของผู้พูดหลายคน ด้วยชุดข้อมูลเดียวกัน ผลจากการ
 วิจัยแสดงให้เห็นว่า เราสามารถทำการสร้างภาพเคลื่อนไหวของรูปปากตามเสียงสระได้ในระดับหนึ่ง
 แม้ว่าจะมีเสียงสระบางตัวที่แบบจำลองไม่สามารถสร้างขึ้นมาก็ได้ตาม

ภาควิชา..... คณิตศาสตร์.....ลายมือชื่อนิลิต.....
 สาขาวิชา.....วิทยาการคนนา.....ลายมือชื่ออาจารย์ที่ปรึกษา.....
 ปีการศึกษา 2548.....ลายมือชื่ออาจารย์ที่ปรึกษาร่วม.....

4572417523 : MAJOR COMPUTATIONAL SCIENCE

KEY WORD: NEURAL NETWORK / AUDIO TO VISUAL CONVERSION / LIP SYNCHRONIZATION / MULTIMEDIA COMMUNICATION / LIP ANIMATION

PICHA IAMTANA-ANAN: AUDIO CONVERSATION TO LIP SHAPE MOVEMENT FOR THAI VOWELS. THESIS ADVISOR: PROF. CHIDCHANOK LURSINSAP , THESIS COADVISOR : ASSOC. PROF. SOMCHAI JITAPUNKUL , 68 pp. ISBN 974-14-3774 -9.

In the animation process for simulating the movement of living creatures, it is necessary to study the forms of movement of the simulated object. Speaking has the forms of movement related to pronouncing which can be simulated in the same manner. This research studies the relation between the pronunciation of vowels effecting to the movement of lip shapes. This is done by considering the shape in forms of coordinates and collecting the data for modeling by using backpropagation network. The scope of this study is limited to the lip movements of 16 mono-vowels in non-continuous and 2-dimensional situation regardless of the effect from muscular pronunciation. In addition, the performances of single-speaker model and multi-speaker model are compared under the same data set. The result confirms that it is moderately possible to animate the lip movement according to the vowel sound even though there are some vowel sounds cannot be simulated.

Department Mathematics.....Student's signature.....
Field of study Computational Science.....Advisor's signature.....
Academic year 2005.....Co-advisor's signature.....

กิตติกรรมประกาศ

ขอขอบพระคุณศาสตราจารย์ ดร. ชิดชนก เหลือสินทรัพย์ อาจารย์ที่ปรึกษา และรองศาสตราจารย์ ดร. สมชาย จิตะพันธ์กุล อาจารย์ที่ปรึกษาร่วมสำหรับการให้คำแนะนำ คำปรึกษาต่าง ๆ รวมถึงความกรุณาในการตรวจทานแก้ไขวิทยานิพนธ์ฉบับนี้สามารถสำเร็จ ลุล่วงได้ด้วยดี

ขอขอบคุณ ดร. สิริพันธ์ สงวนสินธุกุล ประธานกรรมการสอบวิทยานิพนธ์และผู้ช่วยศาสตราจารย์ ดร. บุญเสริม กิจสิริกุล กรรมการสอบวิทยานิพนธ์ ที่ช่วยสละเวลาในการตรวจสอบ และให้คำแนะนำแก้ไขวิทยานิพนธ์ฉบับนี้จนสำเร็จเรียบร้อยเป็นอย่างดี

ขอขอบคุณคุณกุลวดี สุพรรณธะริดา, คุณจุไรรัตน์ สำเร็จผลดี, คุณกมล ศุภรักษ์จินดา และคุณวิสันชนา ศิวะโสภา สำหรับความช่วยเหลือในการเก็บข้อมูลในการทำวิทยานิพนธ์

ขอขอบคุณรองศาสตราจารย์ สุชาดา ศิริพันธ์ คุณเมธีและสมาชิกใน Avic ทุกคนสำหรับความห่วงใยในการทำวิทยานิพนธ์ ช่วยให้คำแนะนำ และคำปรึกษาในการแก้ปัญหาต่างๆ

ขอขอบคุณบัณฑิตวิทยาลัยสำหรับทุนอุดหนุนการทำวิทยานิพนธ์

ท้ายที่สุดขอขอบพระคุณครอบครัวที่ให้การสนับสนุนทั้งกำลังใจและกำลังใจที่ดีมาโดยตลอด และที่ขาดไม่ได้คือ คุณมธุรส แสงไพโรจน์ ที่ช่วยเป็นกำลังใจและสนับสนุนในทุกๆ อย่าง และช่วยผลักดันให้สามารถผ่านอุปสรรคและปัญหาในระหว่างการทำวิทยานิพนธ์มาได้

สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	ง
บทคัดย่อภาษาอังกฤษ.....	จ
กิตติกรรมประกาศ.....	ฉ
สารบัญ.....	ช
สารบัญตาราง.....	ฅ
สารบัญภาพ.....	ญ
บทที่.....	
1 บทนำ.....	1
1.1 ความเป็นมาและความสำคัญของปัญหา.....	1
1.2 วัตถุประสงค์ของการวิจัย.....	2
1.3 ขอบเขตของการวิจัย.....	2
1.4 ข้อจำกัดของการวิจัย.....	2
1.5 ประโยชน์ที่คาดว่าจะได้รับ.....	3
1.6 วิธีดำเนินการวิจัย.....	3
1.7 ลำดับขั้นตอนในการเสนอผลการวิจัย.....	3
1.8 เอกสารและงานวิจัยที่เกี่ยวข้อง.....	4
2 ความรู้เบื้องต้นและทฤษฎีพื้นฐานที่ใช้ในงานวิจัย.....	7
2.1 สัทศาสตร์ : กลไกการเปล่งเสียง.....	7
2.2 Neural Network หรือระบบเครือข่ายประสาทเทียม.....	12
2.3 Audio-to-Visual Conversion.....	15

บทที่	หน้า
3 การสร้างแบบจำลองการเคลื่อนไหวของรูปปาก.....	17
3.1 การเก็บข้อมูล.....	17
3.1.1 การกำหนดจุดพิกัดรอบปากเพื่อทำการเก็บข้อมูล.....	18
3.1.2 การถ่ายภาพวิดีโอเพื่อบันทึกการเคลื่อนไหวของรูปปาก.....	19
3.1.3 การคัดเลือกภาพที่จะนำมาเป็นตัวแทนการเคลื่อนไหวของรูปปาก.....	20
3.1.4 การระบุจุดพิกัดรอบปากด้วย Algorithm.....	21
3.2 การสร้างแบบจำลองด้วย Backpropagation Neural Network.....	23
3.3 วิธีการสร้างภาพเคลื่อนไหวจากแบบจำลองและการประเมิน.....	27
ประสิทธิภาพของแบบจำลอง	
4 ผลการสร้างแบบจำลอง.....	31
5 สรุปผลการวิจัย และข้อเสนอแนะ.....	42
5.1 สรุปผลการวิจัย.....	42
5.2 ข้อเสนอแนะ.....	43
รายการอ้างอิง.....	44
ภาคผนวก.....	46
ภาคผนวก ก.....	47
ภาคผนวก ข.....	49
ภาคผนวก ค.....	57
ประวัติผู้เขียนวิทยานิพนธ์.....	58

สารบัญตาราง

ตาราง	หน้า
1.1 รายการงานวิจัยที่เกี่ยวกับการสร้างรูปปากด้วยวิธีต่างๆ.....	6
3.1 ค่า Index ของเสียงสระเดี่ยวที่ใช้ในการวิจัยทั้ง 16 เสียง.....	19
3.2 ค่า index ของลำดับการเคลื่อนไหวและลักษณะของภาพ.....	20
3.3 จำนวน Hidden Nodes ที่เหมาะสมสำหรับแบบ.....	25
จำลองที่สร้างจากตัวอย่างบุคคลเดี่ยวชุดที่ 1 และ 2	
3.3 จำนวน Hidden Nodes ที่เหมาะสมสำหรับแบบ.....	26
จำลองที่สร้างจากตัวอย่างหลายบุคคล	
4.1 ค่า SSE โดยรวมของเสียงสระ 1- 16 ในการทดสอบทั้ง 3 แบบ.....	39
4.2 ค่าเปอร์เซ็นต์ของความสอดคล้องในทิศทางการเคลื่อนที่.....	40
ของจุดพิกัดของแบบจำลองกับจุดพิกัดจริง	

สารบัญญภาพ

ภาพประกอบ	หน้า
2.1 ตำแหน่งเกิดเสียงในกระบวนการออกเสียง.....	8
2.2 แบบจำลองการทำงานที่ไม่เป็นเชิงเส้นของโนด.....	13
3.1 แผนผังขั้นตอนการดำเนินการวิจัย.....	17
3.2 จุดพิกัดที่ถูกนำมาใช้ในการสร้างภาพเคลื่อนไหวในงานวิจัยชิ้นอื่น.....	18
3.3 จุดพิกัดของรูปปากทั้ง 11 จุดที่นำมาใช้ในการวิจัย.....	18
3.4 ภาพในช่วงการเคลื่อนไหวที่เลือกเก็บมาเป็นข้อมูล.....	20
3.5 ภาพจุดพิกัดรอบริมฝีปากที่ทำการแยกเอาเฉพาะกลุ่มของจุดพิกัดออกมาแล้ว.....	21
3.6 ภาพที่ทำการบันทึกจุดพิกัดลงบนภาพโดยตรงเพื่อใช้เป็นชุดข้อมูลอ้างอิง.....	22
3.7 ข้อมูลจุดพิกัดที่ได้จากเก็บข้อมูลด้วย Algorithm.....	23
3.8 แผนผังการทำงานของแบบจำลองแบบ Multi-Model ที่สร้างด้วย Backpropagation Neural Network	28
3.9 ตัวอย่างทิศทางการเคลื่อนที่ของจุดพิกัดที่สร้างจากแบบจำลองและจุดพิกัด ที่บันทึกจากตัวอย่างผู้พูด	29
4.1 เปรียบเทียบจุดพิกัดที่แบบจำลองได้กับข้อมูลจุดพิกัดจริงในช่วงที่ออกเสียง “อะ”.....	32
โดยสีน้ำเงินแทนจุดพิกัดที่จำลองได้ และสีแดงแทนจุดพิกัดจากข้อมูลจริง	
4.2 เปรียบเทียบจุดพิกัดที่แบบจำลองได้กับข้อมูลจุดพิกัดจริงในช่วงที่ออกเสียง “อุ”.....	32
4.3 รูปของจุดพิกัดที่แบบจำลองได้ทำการสร้างเปรียบเทียบกับข้อมูลจุดพิกัดจริง.....	33
ในช่วงเสียง “อะ”	
4.4 รูปของจุดพิกัดที่แบบจำลองได้ทำการสร้างเปรียบเทียบกับข้อมูลจุดพิกัดจริงใน.....	34
แบบจำลองที่สร้างจากข้อมูลรวม และทำการทดสอบในเสียงสระ “อา”	
4.5 รูปของจุดพิกัดที่แบบจำลองได้ทำการสร้างเปรียบเทียบกับข้อมูลจุดพิกัดจริง.....	34
ในแบบจำลองที่สร้างจากข้อมูลรวม และทำการทดสอบในเสียงสระ “อิ”	
4.6 แผนภูมิเปรียบเทียบค่า SSE ของการทดสอบแบบจำลองที่สร้างจากตัวอย่าง.....	35
บุคคลเดียว ชุดที่ 1 กับแบบจำลองที่สร้างจากตัวอย่างบุคคลเดียว ชุดที่ 2	
ทดสอบชุดข้อมูล parameters ของตัวอย่างบุคคลที่ 1	
4.7 แผนภูมิเปรียบเทียบค่า SSE ของการทดสอบแบบจำลองที่สร้างจากตัวอย่าง.....	36
บุคคลเดียว ชุดที่ 1 กับแบบจำลองที่สร้างจากตัวอย่างบุคคลเดียว ชุดที่ 2	
ทดสอบชุดข้อมูล parameters ของตัวอย่างบุคคลที่ 2	

ภาพประกอบ	หน้า
4.8 แผนภูมิเปรียบเทียบค่า SSE ของการทดสอบแบบจำลองที่สร้างจากตัวอย่าง.....	37
หลายบุคคลกับแบบจำลองที่สร้างจากตัวอย่างบุคคลเดียว ชุดที่ 1 ทดสอบ	
ชุดข้อมูล parameters ของตัวอย่างบุคคลที่ 1	
4.9 แผนภูมิเปรียบเทียบค่า SSE ของการทดสอบแบบจำลองที่สร้างจากตัวอย่าง.....	38
หลายบุคคลกับแบบจำลองที่สร้างจากตัวอย่างบุคคลเดียว ชุดที่ 2 ทดสอบ	
ชุดข้อมูล parameters ของตัวอย่างบุคคลที่ 2	