

หัวข้อวิทยานิพนธ์	การรู้จำเสียงพูดคำไทยแบบไม่ขึ้นกับผู้พูดด้วยนิวรัลเน็ตเวิร์ค
นักศึกษา	นายเสรี ปานซาง
อาจารย์ผู้ควบคุมวิทยานิพนธ์	รศ.ดร. ชม กิมปาน
ระดับการศึกษา	วิศวกรรมศาสตรมหาบัณฑิต สาขาวิชาวิศวกรรมไฟฟ้า
	สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
พ.ศ.	2540

บทคัดย่อ

วิทยานิพนธ์ฉบับนี้มีจุดมุ่งหมายเพื่อศึกษาพฤติกรรมและหาอัตราการรู้จำเสียงพูดคำไทยโดยๆ จำนวน 20 เสียงทดสอบ แบบไม่ขึ้นกับผู้พูด เสียงทดสอบที่ใช้ทดลองคือ "ศูนย์", "หนึ่ง", "สอง", "สาม", "สี่", "ห้า", "หก", "เจ็ด", "แปด", "เก้า", "สิบ", "เปิด", "ปิด", "หมุน", "ยก", "วาง", "ช่อง", "ซ้าย", "ขวา", "ไฟ" ในขบวนการ ปริ-โปร เซลซึ่ง สัญญาณในแกนเวลาของเสียงที่ใช้ทดลองจะถูกแปลงให้เป็น สัญญาณดิจิตอล แล้วใช้การแปลงฟูเรียร์อย่างรวดเร็ว ร่วมกับการแปลงข้อมูลในแกนความถี่ ให้อยู่ในแกน ฮาร์โมนิคของความถี่ [3] เพื่อลดขนาดของข้อมูลและขยายรายละเอียดในช่วงความถี่ที่มีการเปลี่ยนแปลงมาก ได้เป็นสเปกโตรแกรมของเสียงซึ่งจะนำไปใช้ในการหาคุณลักษณะจำเพาะ โดยการหาค่าสัมประสิทธิ์สหสัมพันธ์ ระหว่างสเปกโตรแกรมของเสียงที่ต้องการ กับ กลุ่มสเปกโตรแกรมเสียงหลัก แล้วนำไปจัดเรียงแบบมากไปน้อย แล้วเลือกออกมา 10 เปอร์เซนต์แรก เพื่อนับคะแนนความถี่ของเสียงแต่ละเสียง ผลที่ได้จะแสดงคุณลักษณะจำเพาะของเสียงทดสอบนั้น เมื่อนำคุณลักษณะจำเพาะของเสียงทดสอบกลุ่มหนึ่ง มาฝึกสอนให้กับนิวรัลเน็ตเวิร์คแบบแบคพรอบพาเกชั่น โดยกำหนดเป้าหมายที่สอดคล้องกับเสียงที่นำมาฝึกสอน จนกระทั่งค่าผิดพลาดลดลงถึงค่าที่ยอมรับได้ แล้วนำข้อมูลคุณลักษณะจำเพาะของเสียงอีกกลุ่มหนึ่งมาทดสอบ นิวรัลเน็ตเวิร์คจะสามารถแยกแยะระดับความสัมพันธ์ของเสียงที่นำมาทดสอบนั้นสอดคล้องกับเป้าหมายที่เคยเรียนรู้มา

จากผลการดำเนินงาน อัตราการรู้จำ คุณลักษณะจำเพาะของเสียงทดสอบโดยไม่ขึ้นกับผู้พูด 20 เสียง พูดยเสียงทดสอบละ 3 ครั้ง จากกลุ่มเสียงที่ใช้ฝึกสอน 10 คน โดยใช้กลุ่มเสียงทดสอบ 12 คน เป็น 73.89 % เมื่อลดจำนวนเสียงทดสอบลงมา เหลือ 11 เสียง คือ "ศูนย์", "หนึ่ง", "สอง", "สาม", "สี่", "ห้า", "หก", "เจ็ด", "แปด", "เก้า", "สิบ" อัตราการรู้จำเป็น 80.05 % อัตราการรู้จำคุณลักษณะจำเพาะของเสียงทดสอบโดยขึ้นกับกลุ่มผู้พูด 20 เสียง ภายในกลุ่ม 22 คน เป็น 84.32 % และอัตราการรู้จำเสียงทดสอบ 11 เสียง ภายในกลุ่มผู้พูด 12 คน เป็น 92.42 % จากผลการทดลองพบว่าอัตราการรู้จำจะแปรผกผันตามจำนวนผู้พูดและเสียงทดสอบ การวิเคราะห์แพตเทิร์นของเสียงแบบเทียบกับแบบอ้างอิงที่ใช้อัลกอริทึมแบบเดียวกัน ไม่สามารถนำไปใช้รู้จำเสียงทดสอบจำนวนมาก แบบไม่ขึ้นกับผู้พูดได้