

ระบบระบุผู้พูดภาษาไทยด้วยวิธีไดนามิกส์ไทม์วอร์ปิง

วารินทร์ อัจฉริยะกุลพร¹, ชัย วุฒิวิวัฒน์ชัย², จุฬารัตน์ ต้นประเสริฐ³ หน่วยปฏิบัติการวิจัยและพัฒนาวิศวกรรมภาษาและซอฟต์แวร์ ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ สำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติ ชั้น 22 อาคารมหานครยิบซัม ถ.ศรีอยุธยา เขตราชเทวี กรุงเทพฯ 10400 อีเมล: ¹varin@nectec.or.th, ²chai@nectec.or.th, ³chulak@nectec.or.th

ABSTRACT -- This paper proposes a closed set, text-dependent speaker identification system for Thai language. Speaking text used in this system is Thai digit 0-9. From our preliminary experiments, a pattern matching technique namely Dynamic Time Warping (DTW) has shown a very good performance, especially for text-dependent system. This work consequently focuses on DTW based system with several comparative studies. Changes of number of sound references and some significant DTW parameters, e.g. time-alignment window, are evaluated. Three algorithms for practical implementation of DTW are proposed and compared. Various kinds of popular speech features are conducted in experiment. Furthermore, experiment on longer speaking-text using concatenation of isolated digits is performed with a new efficient decision technique, proposed especially for concatenated speech system.

Keywords -- Thai language speaker identification, Dynamic time warping

บทคัดย่อ -- บทความฉบับนี้นำเสนอระบบระบุผู้พูดสำหรับภาษาไทยแบบระบบปิด และกำหนดคำพูดตายตัวคือเสียงตัวเลขภาษาไทย 0-9 จากการทดลองเบื้องต้นพบว่าการเทียบเคียงแบบไดนามิกส์ไทม์วอร์ปิง (Dynamic Time Warping : DTW) เป็นวิธีที่มีประสิทธิภาพสูง และเหมาะสำหรับระบบระบุผู้พูดชนิดกำหนดคำพูดตายตัว ในบทความนี้จึงเป็นการวิจัยต่อเนื่องเน้นหนักเฉพาะการใช้เทคนิค DTW โดยมีการทดลองเปรียบเทียบผลการระบุผู้พูด เมื่อเปลี่ยนแปลงจำนวนเสียงอ้างอิง เปลี่ยนแปลงค่าตัวแปรสำคัญในกระบวนการ DTW เช่นค่ากรอบของการเทียบจุด (Time-alignment Window) เปลี่ยนแปลงกฎการตัดสินใจ พร้อมทั้งนำเสนอและเปรียบเทียบเทคนิคการพัฒนา DTW ในทางปฏิบัติ 3 แบบ นอกจากนี้ยังได้ทำการทดลองเปรียบเทียบผล เมื่อใช้ค่าลักษณะสำคัญของเสียงแบบต่างๆ และการทดลองสำหรับเสียงพูดที่ยาวขึ้น โดยใช้เสียงของตัวเลขโดดต่อกัน สำหรับการทดลองกับเสียงตัวเลขต่อเนื่องนี้ ได้นำเสนอวิธีการตัดสินใจแบบใหม่ซึ่งให้ผลดีกว่าเดิมอีกด้วย

คำสำคัญ -- การระบุผู้พูดสำหรับภาษาไทย, ไดนามิกส์ไทม์วอร์ปิง