

หัวข้อวิทยานิพนธ์

การจัดหมวดสถานที่โดยใช้การรู้จำตัวอักษรในฉากแบบ
องค์รวม

ผู้เขียน

นายภวิศ ธรรมสอน

ปริญญา

วิศวกรรมศาสตรมหาบัณฑิต (วิศวกรรมคอมพิวเตอร์)

อาจารย์ที่ปรึกษา

อาจารย์ ดร. กานต์ ปทานุคม

บทคัดย่อ

การจัดหมวดสถานที่เป็นขั้นตอนวิธีที่กำหนดขึ้นมาเพื่อให้คอมพิวเตอร์สามารถจำแนกภาพสถานที่ต่างๆ ให้อยู่ในหมวดหมู่สถานที่ต่างๆ ให้มีผลลัพธ์ใกล้เคียงกับที่มนุษย์ทำ ซึ่งสามารถนำไปประยุกต์ใช้ในระบบการสำรวจอัตโนมัติหรือระบบปรับการทำงานตามสภาพสถานที่ต่างๆ ได้ แต่ในปัจจุบันระบบการรู้จำดังกล่าวยังมีปัญหาเรื่องการแยกหมวดสถานที่ที่มีลักษณะปรากฏต่างๆ คล้ายกันซึ่งทำให้ได้ผลความแม่นยำในการจัดหมวดที่ค่อนข้างต่ำ แต่เมื่อพิจารณาวิธีการที่มนุษย์โดยทั่วไปใช้แยกแยะสถานที่ พบว่าคนสามารถแยกแยะสถานที่ได้จากข้อมูลคำในภาพฉากที่สามารถบรรยายบริบทต่างในสถานที่นั้นๆ ได้ งานวิจัยครั้งนี้มีวัตถุประสงค์ที่จะพัฒนาขีดความสามารถของระบบการจัดหมวดสถานที่ต่อจากโมเดล D-TAP [1] ที่สามารถใช้ข้อมูลจากการรู้จำคำในภาพฉากมาคำนวณคะแนนความเป็นไปได้ในแต่ละหมวดสถานที่ที่กำหนดไว้ โดยจะปรับปรุงในส่วนขั้นตอนการสกัดลักษณะเด่นจากพื้นที่คำให้สามารถจับรายละเอียดจากพื้นที่คำได้มากขึ้น เพิ่มประยุกต์ใช้แผนภาพความเด่นเข้ามาช่วยกำหนดคะแนนระดับความมั่นใจของคำต่างๆ และปรับปรุงวิธีการคำนวณคะแนนของสถานที่ให้สามารถเพิ่มความเป็นไปได้ที่มีคำบางคำปรากฏอยู่หลายครั้ง ผลการทดสอบพบว่า การปรับปรุงการทำงานของโมเดลการแยกหมวดสถานที่ที่เสนอในงานวิจัยนี้ได้ความแม่นยำดีขึ้นกว่าวิธีฐานและวิธีอื่นๆ ที่นำมาเปรียบเทียบในชุดข้อมูลทั้งภาษาไทยและภาษาอังกฤษ

Thesis Title Place Categorization Using Holistic Scene Text Recognition

Author Mr. Phawis Thammasorn

Degree Master of Engineering (Computer Engineering)

Advisor Dr. Karn Patanukhom

ABSTRACT

Place categorization is a process that allow computer to classify places into predefined category similar to what humans do. Good performance from the process can lead to automatic intelligent surveying system or better location-based service from system of computer. However, place categorization process is from scene feature still have limited performance when categorizing scene that have similar appearance which usually achieve low categorizing accuracies. Generally, there are texts on the scene that indicates the contextual clue that human can use to distinguish the categories of places. This research aims to develop better performance of the categorization system that uses scene text information to classify scenes. This research aim to improve the existing model of D-TAP [4] that can extract features from scene texts and use them for scene classification. The improvement are modifications of feature extraction method that able to capture better detail of scene text, weighting scheme of different text areas on word score calculation, and changes in categories likelihood score calculation to suit varying number of text word occurrences in the scenes. The results show that the modifications improve performance and achieve better accuracy from baseline and other compared method in datasets of Thai and English languages.