รายงานการวิจัย

การศึกษาและวิจัยการทำงานของระบบค้นคืนสารสนเทศโดยใช้อัลกอริทึม VIPS

A Study of Information Retrieval System Using Vision based Pages Segmentation (VIPS) Algorithms

นางสาวสุธีรา พันธุ์ธีรานุรักษ์

ได้รับทุนสนับสนุนงานวิจัยจากเงินรายได้ ประจำปีงบประมาณ 2554

คณะวิศวกรรมศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

รายงานการวิจัย

การศึกษาและวิจัยการทำงานของระบบค้นคืนสารสนเทศโดยใช้อัลกอริทึม VIPS
A Study of Information Retrieval System Using Vision based Pages Segmentation (VIPS) Algorithms

นางสาวสุธีรา พันธุ์ธีรานุรักษ์

ได้รับทุนสนับสนุนงานวิจัยจากเงินรายได้ ประจำปีงบประมาณ 2554
คณะวิศวกรรมศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

กิตติกรรมประกาศ

โครงการวิจัยนี้สำเร็จลุล่วงได้เป็นอย่างดีโดยได้รับความช่วยเหลือจาก คุณเกศินี ทองตันไตรย์ คุณชญานี จารุพัฒนะสิริกุล และ คุณชนิกตา ธรรมสรางกูร อดีตนักศึกษาของภาควิชาวิศวกรรมสารสนเทศ คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง ที่ช่วยทำการทดลองต่าง ๆ

นางสาวสุธีรา พันธุ์ธีรานุรักษ์

ชื่อโครงการ (ภาษาไทย) การศึกษาและวิจัยการทำงานของระบบค้นคืนสารสนเทศโดยใช้อัลกอริทึม VIPS

ชื่อโครงการ (ภาษาอังกฤษ) A Study of Information Retrieval System Using Vision based Pages Segmentation (VIPS) Algorithms

แหล่งเงิน เงินรายได้คณะวิศวกรรมศาสตร์

ประจำปีงบประมาณ 2554 จำนวนเงินที่ได้รับการสนับสนุน 64,900 บาท

ระยะเวลาทำการวิจัย 1 ปี ตั้งแต่ 1 ตุลาคม 2553 ถึง 30 กันยายน 2554

ชื่อ-สกุล หัวหน้าโครงการ และผู้ร่วมโครงการวิจัย พร้อมระบุ หน่วยงานต้นสังกัดและ อีเมล์

นางสาวสุธีรา พันธุ์ธีรานุรักษ์

สาขาวิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง kpsuthee@kmitl.ac.th

คำสำคัญ (Keywords) Information Retrieval System, Web-based Retrieval Systems, Relevant Feedback Algorithms

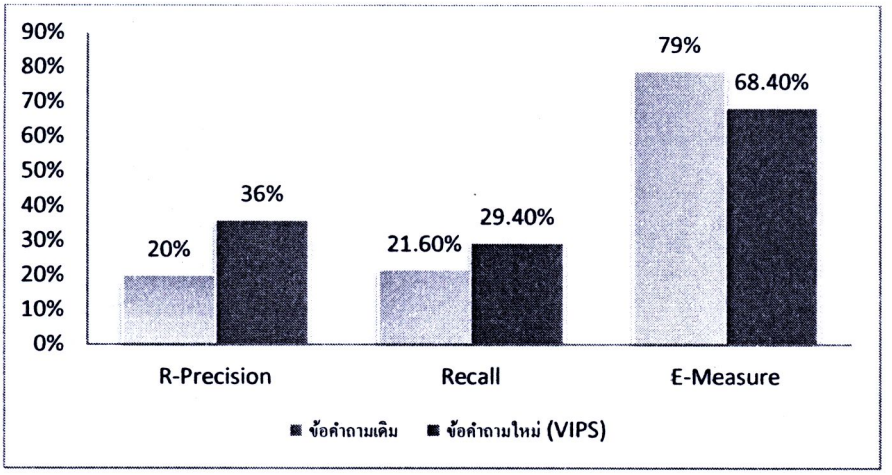
บทคัดย่อ

โครงงานวิจัยนี้เป็นการศึกษาการทำงานของระบบค้นคืนสารสนเทศบนเว็บ โดยมีการเก็บการค้นคืนย้อนกลับจากผู้ใช้และการใช้อัลกอริทึม Vision Base Page Segmentation (VIPS) เพื่อทำการแบ่งเว็บเพจที่ได้จากการค้นคืนออกเป็นบล็อก และหาคำที่จะนำมาเพิ่มในข้อคำถามเดิมเพื่อสร้างเป็นข้อคำถามใหม่ โดยได้ทำการจำลองระบบการค้นคืนสารสนเทศบนเว็บเพื่อทำการเก็บการค้นคืนย้อนกลับจากผู้ใช้ แล้วนำเว็บเพจที่ผู้ใช้เลือกมาหาข้อคำถามใหม่เปรียบเทียบกับการค้นคืนย้อนกลับจากผู้ใช้แล้วนำเว็บที่ผู้ใช้เลือกไปแบ่งเป็นบล็อกโดยใช้อัลกอริทึมวีไอพีเอส จากนั้นให้ผู้ใช้เลือกบล็อกที่แบ่งได้ซึ่งผู้ใช้เห็นว่าเกี่ยวข้องกับความต้องการอีกครั้งหนึ่ง จากผลการทดลองพบว่าข้อคำถามใหม่ที่เกิดจากการเพิ่มคำ โดยนำอัลกอริทึมวีไอพีเอสมาใช้นั้นทำให้ได้ผลการค้นคืนที่เกี่ยวข้องกับสิ่งที่ผู้ใช้ต้องการมากขึ้น

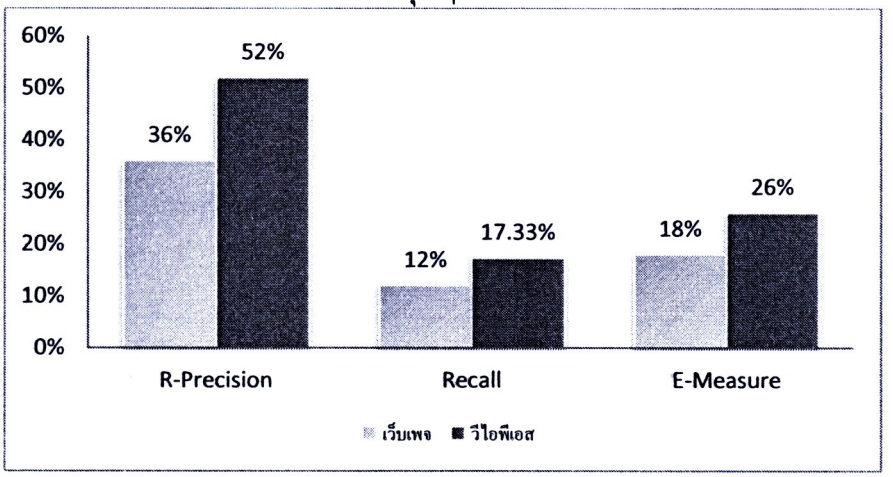
ABSTRACT

This research project aims to study web search system using relevant feedback and Vision-based Page Segmentation (VIPS) algorithm. Web information retrieval using relevant feedback was simulated to compare results of two experiments. One is result from web search system using relevant feedback. Another is web search system using relevant feedback and VIPS algorithm. In our experiments, we can show that the expansion terms from using VIPS algorithm is meet user requirements more than web search using relevant feedback.

รูปผลงานวิจัย



รูปที่ 1 กราฟแสดงการเปรียบเทียบประสิทธิภาพผลการค้นคืนย้อนกลับจากข้อคำถามเดิมและการใช้อัลกอริทึมวีไอพีเอสของทุก ๆ ข้อคำถาม



รูปที่ 2 กราฟแสดงการเปรียบเทียบประสิทธิภาพผลการค้นคืนย้อนกลับจากการเพิ่มข้อคำถามที่มาจากทั้งหน้าเว็บเพจและการใช้อัลกอริทึมวีไอพีเอสของทุกๆข้อคำถาม

จากรูปกราฟที่ 1 และรูปกราฟที่ 2 จะเห็นว่าผลการเปรียบเทียบการประเมินประสิทธิภาพของการค้นคืนย้อนกลับจากการเพิ่มข้อคำถามที่มาจากหน้าเว็บเพจนั้นมีค่าน้อยกว่าการใช้อัลกอริทึมวีไอพีเอสในทุกๆค่า ดังนั้นอัลกอริทึมวีไอพีเอสมีส่วนช่วยเพิ่มประสิทธิภาพในการค้นคืนให้ดีขึ้นและได้เว็บเพจที่เกี่ยวข้องกับสิ่งที่ผู้ใช้ต้องการมากยิ่งขึ้น

สารบัญ

หน้า

สารบัญ (ต่อ)

หน้า

สารบัญ (ต่อ)

หน้า

สารบัญตาราง

ตารางที่ หน้า

สารบัญรูป

รูปที่ หน้า

สารบัญรูป (ต่อ)

รูปที่ หน้า