

บทคัดย่อ

รูปภาพแต่ละรูปจะมีคุณลักษณะของรูปแตกต่างกัน กัน เช่น สี (Color) รูปทรง (Shape) ของอ็อบเจกต์ (Object) ต่างๆ แต่รูปภาพเหล่านี้ อาจจะมี ความหมายที่เหมือนกันในระดับที่มนุษย์เข้าใจ (High level conceptualization) แต่ข้อจำกัดของระบบจำแนกรูปภาพในปัจจุบันคือจะจำแนกรูปภาพตามคุณลักษณะของรูปภาพเป็นหลักซึ่งทำให้การจำแนกรูปภาพมีประสิทธิภาพต่ำเนื่องจากจะได้รูปที่มีคุณลักษณะเหมือนกันอยู่ในกลุ่มเดียวกัน แต่มีความหมายหรือเนื้อหาของรูปภาพที่ต่างกันโดยสิ้นเชิง ดังนั้นงานวิจัยนี้จึงทำการพัฒนากรอบวิธีวิจัยเพื่อขจัดปัญหาดังกล่าวโดยใช้วิธีที่เรียกว่า “ถุงของคำ” (bag-of-visual words, BVW) มาช่วยในการแทนคุณลักษณะของรูปภาพ ซึ่งวิธีการนี้จะมีการทำงานอยู่บนพื้นฐานของโมเดลเชิงพื้นที่เวกเตอร์ (vector space model) ซึ่งเป็นการคำนวณทางสถิติ งานวิจัยนี้ได้มีแนวคิดใหม่ (novelty) ดังนี้คือ 1) แนวคิดในการแปลงโครงสร้างของคุณลักษณะระดับต่ำของรูปภาพจากรูปแบบเวกเตอร์ (Vector space model) ให้เป็นโครงสร้างแบบลำดับชั้น (Hierarchical model) 2) แนวคิดในการรวมข้อมูลต่างชนิดกันไว้ในออนโทโลยีเดียวกัน เพื่อเพิ่มประสิทธิภาพในการค้นคืนรูปภาพ ซึ่งในการวิจัยนี้ทำการรวมข้อมูลประเภทตัวอักษรและคุณลักษณะระดับต่ำของรูปภาพไว้ด้วยกัน และ 3) แนวคิดที่ลดข้อจำกัดความไม่สมบูรณ์ของออนโทโลยี โดยใช้ LSI มาช่วยในการค้นหาข้อมูล ซึ่งช่วยเพิ่มประสิทธิภาพการจำแนกรูปภาพตีความหมายรูปภาพ และค้นหารูปภาพได้ดียิ่งขึ้น

Abstract

Images that have a different visual appearance may be semantically related using a higher level conceptualization. However, image classification and retrieval systems tend to rely only on the low-level visual structure within images. This research presents a framework to deal with this limitation by exploiting the well-known bag-of-visual words (BVW) vector space model, to represent visual content. The novelty of this research is threefold. *First*, a novel method is used to restructure the vector space model of visual words with respect to a structural ontology model in order to resolve visual synonym and polysemy problems. *Second*, the idea to fuse different types of information into a unified ontology. The *third* contribution in this research is that a technique to effectively handle an incomplete ontology using LSI. It is expected that the proposed method can significantly improve image classification, interpretation and retrieval performance.