

วิทยานิพนธ์นี้มีวัตถุประสงค์ในการแสดงผลภาพสำหรับข้อมูลเอกสารดิจิทัล โดยทำการแปลงข้อมูลในเอกสารจากตัวอักษรให้เป็นรูปภาพ เพื่อช่วยในการพิจารณาเปรียบเทียบความเหมือนและความแตกต่างของประเภทหรือหมวดหมู่เอกสาร ทำให้ผู้ใช้สามารถ จัดการ และจำแนกรูปแบบหรือประเภทของเอกสารได้ง่ายและรวดเร็วมากยิ่งขึ้น โดยไม่จำเป็นต้องเข้าไปพิจารณาเนื้อความในเอกสาร โดยการแสดงผลภาพมีแนวทางในการพัฒนาจากแนวคิดของทฤษฎีเคออสเกม ประยุกต์ร่วมกับการแสดงผลภาพบิตแม็บของข้อมูลอนุกรมเวลาโดยใช้วิธีการแบบแซ็ค

งานวิจัยนี้ได้ทำการวิเคราะห์รูปแบบและลักษณะต่างๆของเอกสาร โดยการปรับข้อมูลในเอกสาร และกำหนดพารามิเตอร์ที่สำคัญต่างๆ เพื่อให้การแสดงผลภาพบิตแม็บจากข้อมูลในเอกสารมีความชัดเจนและมีประสิทธิภาพ ซึ่งได้มาจากการทดลองด้วยข้อมูลจริง นอกจากนี้ยังได้ทำการทดสอบประสิทธิภาพของการแสดงผลภาพจากการพิจารณาเปรียบเทียบภาพบิตแม็บของข้อมูลเอกสาร ทั้งจากการสังเกตและใช้วิธีการจัดกลุ่มภาพบิตแม็บโดยอัตโนมัติ ซึ่งได้ผลสรุปจากการทดสอบว่า การแสดงผลภาพสำหรับข้อมูลเอกสารดิจิทัล สามารถช่วยในการพิจารณาเปรียบเทียบความเหมือนและความแตกต่างของประเภทหรือหมวดหมู่เอกสารดิจิทัลได้อย่างมีประสิทธิภาพ

The objective of this research is to visualize digital documents by converting text data in the digital documents to a bitmap image to help compare the similarities and differences of document types or categories so that the document can be easily and more conveniently clustered and managed. Users do not need to read details in the document. This visualization technique combines together the advance in Chaos Game Theory and SAX representation in Time Series bitmap visualization.

By experimenting with real data, this research analyzes the feature and format of digital documents and later adjusts document data and defines important parameters so that bitmap visualization of the document data is well-defined and effective. Moreover, this research also tests the visualization efficiency by comparing the bitmaps of the digital document through both users' observation and automatic clustering. The result shows that the bitmap visualization technique for digital document data can effectively help differentiate the documents types or categories.