

ชื่อ : นายณัฐวุฒิ เอ็มโอชะ
ชื่อวิทยานิพนธ์ : การตรวจจับตัวอักษรประดิษฐ์ในวิดีโอข่าว
สาขาวิชา : เทคโนโลยีคอมพิวเตอร์
สถาบันเทคโนโลยีพระจอมเกล้าพระนครเหนือ
ที่ปรึกษาวิทยานิพนธ์ : ผู้ช่วยศาสตราจารย์ ดร.สุพจน์ นิตย์สุวัฒน์
ปีการศึกษา : 2547

บทคัดย่อ

T168005

วิทยานิพนธ์นี้กล่าวถึงกระบวนการตรวจจับตัวอักษรประดิษฐ์ในวิดีโอข่าว โดยทำการแยกภาพในแต่ละเฟรมของวิดีโอออกมาเป็นภาพนิ่ง แล้วทำการวิเคราะห์หาบริเวณที่เป็นตัวอักษรประดิษฐ์ทั้งภาษาไทยและภาษาอังกฤษในทุกๆ 5 เฟรม โดยใช้การประเมินค่าผลรวมกราเดียน จากนั้นทำการพิจารณาพื้นที่ตัวอักษรที่ตรวจจับได้ร่วมกันจากหลายเฟรมเพื่อให้ได้ข้อมูลที่สมบูรณ์ที่สุด แล้วทำการแปลงข้อมูลเป็นภาพขาว-ดำ (Binarization) เพื่อให้ได้ภาพอักษรขาว-ดำ ก่อนส่งข้อมูลนี้ไปยังโปรแกรมรู้จำตัวอักษร (OCR) เพื่อทำการแปลงภาพขอบเขตตัวอักษรที่ได้เป็นรหัสอักษรแอสกีที่คอมพิวเตอร์สามารถประมวลผลได้ต่อไป

จากการวิจัยพบว่าสามารถตรวจจับตัวอักษรประดิษฐ์จากภาพนิ่งตัวอย่างได้อย่างถูกต้อง 85.57 % และจากวิดีโอตัวอย่างได้อย่างถูกต้อง 90.79% และกระบวนการพิจารณาความสัมพันธ์ระหว่างเคอร์เนล (Kernel) กับพื้นหลังของภาพ ซึ่งเป็นการย้อนกลับค่าโทนสีเทาของเคอร์เนล เมื่อพบว่าลักษณะพื้นหลังของเคอร์เนลนั้นไม่สัมพันธ์กับลักษณะพื้นหลังของภาพร่วมกับกระบวนการแปลงข้อมูลเป็นภาพขาว-ดำแบบสัมประสิทธิ์ชนิดปรับตัวได้ (Locally Adaptive Threshold) สามารถให้ผลเป็นภาพขาว-ดำที่น่าพอใจเมื่อสังเกตด้วยสายตา สามารถสังเกตได้ว่าตัวอักษรมีความสมบูรณ์กว่า เมื่อเปรียบเทียบกับแปลงข้อมูลเป็นเลขฐานสองแบบสัมประสิทธิ์รวม (Global Image Threshold) และพื้นหลังมีการรบกวน (Noise) น้อยกว่าการแปลงข้อมูลเป็นภาพขาว-ดำแบบสัมประสิทธิ์ชนิดปรับตัวได้ ที่ไม่ใช้กระบวนการพิจารณาเคอร์เนลร่วมด้วย

(วิทยานิพนธ์มีจำนวนทั้งสิ้น 99 หน้า)



ประธานกรรมการที่ปรึกษาวิทยานิพนธ์

Name : Mr. Nuttawut M'ocha
Thesis Title : Extraction of Artificial Text in News Video
Major Field : Computer Technology
King Mongkut's Institute of Technology North Bangkok
Thesis Advisor : Assistant Professor Dr. Supot Nitsuwat
Academic Year : 2004

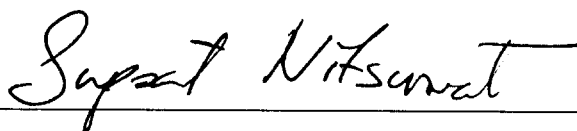
Abstract

T168005

This thesis concerns about the extraction of artificial text in news video, focused on both Thai and English characters. The first stage of the system is text detection. The purpose is to detect the text area from the still images that are captured from every 5 frames of video sequences. The second stage is to track all detected data and improve the quality of detected gray scale images by a multi-frame integration stage. The final stage is binarization. The purpose is to binarize the detected gray scale images by using the locally adaptive threshold method with an additional kernel consideration process before sending the binarized data to OCR application.

The efficiency of the system for still images detection is 85.57% and 90.79% for the video sequences.

(Total 99 pages)



Chairperson