ชื่อเรื่องวิทยานิพนธ์ การใช้งานซิฟต์ร่วมกับวิธีการจัดกลุ่มในระบบการแปลภาษามือไทยด้วยวิธีการสะกดนิ้วมือ

ผู้เขียน นายวัฒนพงศ์ สุทธภักดิ์

ปริญญา วิศวกรรมศาสตรมหาบัณฑิต (วิศวกรรมคอมพิวเตอร์)

อาจารย์ที่ปรึกษาวิทยานิพนธ์ ผศ.ดร.ศันสนีย์ เอื้อพันธ์วิริยะกุล

บทคัดย่อ

ในการสื่อสารระหว่างบุคคลปกติและผู้ที่มีความบกพร่องทางการได้ยิน นั้นบุคคลปกติจำเป็นต้องศึกษาเรียนรู้ลักษณะท่าทางของมือที่ได้มีการกำหนดไว้เป็นมาตรฐานภาษามือ ซึ่งจำเป็นต้องใช้ระยะเวลาในการศึกษาเรียนรู้ จึงจะสามารถใช้ภาษามือในการสื่อสารได้ ทำให้เกิดความยุ่งยากซับซ้อนในการที่จะศึกษาเรียนรู้ภาษามือ

ภาษามือที่ใช้กันโดยทั่วไปมีอยู่ด้วยกัน 2 แบบ คือภาษามือที่เป็นการสะกดนิ้วมือตามตัวอักษร และภาษามือที่เป็นการเคลื่อนไหวท่าทางตามลักษณะคำ ซึ่งในงานวิจัยนี้เราสนใจในการสร้างระบบแปลภาษามือภาษาไทยที่เป็นการสะกดนิ้วมือตามตัวอักษร ถึงแม้ว่าผู้ที่มีความบกพร่องทางการได้ยินที่สามารถใช้การสะกดนิ้วมือในการสื่อสารสามารถเขียนหนังสือในการสื่อสารได้ด้วย แต่การที่มีระบบนี้ก็เพื่อช่วยให้การสื่อสารระหว่างผู้มีความบกพร่องทางการได้ยินและบุคคลทั่วไป เป็นไปอย่างรวดเร็วมากขึ้นและสะดวกมากกว่า การเขียนหนังสือในการสื่อสาร และระบบนี้ยังสามารถเป็นพื้นฐานในการสร้างระบบแปลภาษามือภาษาไทยที่เป็นการเคลื่อนไหวท่าทางได้อีกด้วย

วิทยานิพนธ์ฉบับนี้นำเสนอวิธีการแปลภาษามือ แบบสะกดนิ้วมือภาษาไทย โดยจะเลือกตัวอักษรและสระภาษาไทยรวม 19 ตัวอักษร เพื่อรู้จำท่าทางของมือโดยประยุกต์ใช้ Scale Invariant Feature Transform (SIFT) ร่วมกับวิธีการจัดกลุ่มแบบ Fuzzy C-Means (FCM) และ Hard C-Means (HCM) โดยทำการทดสอบระบบกับท่าทางของมือที่แสดงถึงตัวอักษร และคำที่เกิดจากการประสม

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

ตัวอักษรและสระจำนวน 42 คำ โดยใช้ผู้ทำการทดลองทั้งหมด 21 คน แบ่งเป็น 2 กลุ่มคือ กลุ่มที่ 1 (ผู้ทดลอง 10 คน) เก็บรูปเพื่อใช้ทดสอบและเก็บรูปเพื่อใช้เป็นฐานข้อมูลที่ใช้ในการเปรียบเทียบ (signature library) และอีกกลุ่ม (ผู้ทดลอง 11 คน) ใช้สำหรับทดสอบเพื่อเปรียบเทียบผลความถูกต้องเท่านั้น โดยขั้นตอนในการเก็บรูปที่ใช้เป็นฐานข้อมูลในการเปรียบเทียบ จะทำการบันทึกวีดิทัศน์ของผู้ทำการทดลองที่แสดงท่าทางมือของแต่ละตัวอักษร แล้วทำการแยกเฟรมวีดิทัศน์ออกเป็นรูปภาพ โดยจะทำการคัดเลือกรูปที่ชัดที่สุดของท่าสะกดนิ้วมือตัวอักษรละ 10 รูปในแต่ละผู้ทำการทดลอง แล้วทำการตัดเฉพาะบริเวณมือ โดยไม่มีการตัดแต่งรูป สำหรับรูปที่ใช้ในการทดสอบเกิดจากการบันทึกวีดิทัศน์ของแต่ละผู้ทำการทดลองแล้วทำการแยกเฟรมวีดิทัศน์ด้วยอัตรา 30 เฟรมต่อวินาที ซึ่งเท่ากับอัตราการบันทึกภาพเคลื่อนไหวของกล้องที่ใช้ในการทดลอง ในการเลือกตัวอักษรที่ใช้ในการทดลองจะเป็นตัวอักษรที่เกิดจากท่ามือที่ใช้แขนข้างเดียว และไม่มีการเคลื่อนไหว โดยผู้ทำการทดลองจะต้องสวมใส่เสื้อแขนยาวสีทึบ และมีฉากหลังเป็นสีทึบ จำนวนของรูปที่ใช้เป็นฐานข้อมูลในการเปรียบเทียบมี 1,900 รูป จำนวนรูปที่ใช้ในการทดสอบการสะกดนิ้วมือแบบเปรียบเทียบเฉพาะตัวอักษรมี 3,990 รูป และจำนวนรูปที่ใช้ในการทดสอบคำที่เกิดจากการประสมตัวอักษรมี 570,106 รูป ในการทดลองเปรียบเทียบการประสมคำ จะมีวิธีการเลือกรูปสะกดนิ้วมือของตัวอักษรในแต่ละคำ แบ่งเป็น 2 วิธีคือการเลือกด้วยมือโดยจะเลือกรูปสะกดนิ้วมือที่ชัดที่สุดและการเลือกรูปสะกดนิ้วมืออัตโนมัติ

ผลการทดลองเปรียบเทียบเฉพาะตัวอักษรวิธีซิฟต์ให้ความถูกต้องเฉลี่ยร้อยละ 61.74 ซึ่งน้อยกว่าวิธีประยุกต์ซิฟต์ร่วมกับการจัดกลุ่มด้วย HCM และ FCM คือร้อยละ 62.26 และ 64.16 ตามลำดับ ผลการทดลองจากการเปรียบเทียบคำที่ได้จากตัวอักษร โดยวิธีการเลือกเฟรมด้วยมือ ให้ความถูกต้องเฉลี่ยร้อยละ 11.26%, 8.14%, 8.98% โดยวิธีซิฟต์, ซิฟต์ร่วมกับ HCM และ ซิฟต์ร่วมกับ FCM ตามลำดับ วิธีการเลือกเฟรมอัตโนมัติให้ความถูกต้องเฉลี่ยร้อย 1.02% 0.81% และ 0.81% โดยวิธี ซิฟต์, ซิฟต์ร่วมกับHCM และ ซิฟต์ร่วมกับ FCM ตามลำดับ เหตุผลที่ความถูกต้องของการเปรียบเทียบคำให้ผลที่ต่ำเกิดจากความผิดพลาดสะสมจากการประสมตัวอักษรที่มีความผิดพลาดสูงเช่น “ต” “น” และ“พ” เป็นต้น รวมไปถึงความผิดพลาดของรูปในฐานข้อมูลการเปรียบเทียบเพราะมีการตัดรูปบริเวณมือผิดพลาด แต่อย่างไรก็ตามวิธีการที่ได้นำเสนอให้ความถูกต้องที่สูงในการทดลองเปรียบเทียบตัวอักษรโดยที่การประยุกต์ใช้ซิฟต์ร่วมกับการจัดกลุ่มให้ผลที่ใกล้เคียงกับวิธีซิฟต์ โดยที่ความเร็วในการคำนวณสูงกว่า

ผลการทดลองของซิฟต์ร่วมกับ HCM และ FCM ถือได้ว่า เทียบเท่ากับผลการทดลองที่ได้จากซิฟต์อย่างเดียว แต่เวลาในการคำนวณของซิฟต์ร่วมกับ HCM และ FCM ใช้เวลาการประมวลผลเร็วกว่าวิธีซิฟต์อย่างเดียวประมาณ 5 เท่า

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

Thesis Title Incorporating SIFT with Clustering Algorithm in Thai Finger Spelling Translation System

Author Mr.Wattanapong Suttapak

Degree Master of Engineering (Computer Engineering)

Thesis Advisor Asst. Prof. Dr. Sansanee Auephanwiriyakul

ABSTRACT

To communicate with a hearing impaired person, we need to learn the standard sign language. It takes time to study the language because it is very difficult and very complex. Hence, communication with a hearing impaired person is not easy.

Normally, there are 2 types of sign language, i.e., finger spelling and gesture sign language. We are interested in developing the translation system for Thai finger spelling sign language in this research. Although a hearing impaired person who can use finger spelling sign language can communicate through writing, this system can help fastening and providing an easy support for the communication. This system can be improved to translate Thai gesture sign language in the future as well.

In this research, we implement a translation system for Thai finger spelling sign language using Scale-invariant Feature Transform (SIFT) with Fuzzy C-Means (FCM) clustering and that with Hard C-means (HCM) clustering. In particular, there are 19 alphabets and vowels altogether, and 42 words from these alphabets and vowels utilized in the system. We collect testing video data set from 21 subjects 10 out of 20 subjects are also collected for signature library. All subjects are asked to wear a black shirt with long sleeves and stand in front of the dark background. Each video is sampling with the rate of 30 frames per second. To create the signature library,

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

the image frame with the most clearly position of the finger will be chosen. Then we cut the area around the hand without adjusting the image. For the testing data set, we process the whole image. There are 1900 images in the signature library and 3990 testing images. The number of testing images for words is 570,106. To translate each word, we manually select the good frames and translate each frame separately. We compare the result with that from the automatically frame selection as well.

The best result of alphabets and vowels finger spelling sign language from SIFT, SIFT with HCM and SIFT with FCM are 61.74%, 62.26% and 64.16%, respectively. The best result of the manually frame selection in word translation system are 11.26%, 8.14% and 8.98% using SIFT, SIFT with HCM and SIFT with FCM, respectively. While the best result of the automatically frame selection in word translation system are 1.02%, 0.81% and 0.81% using SIFT, SIFT with HCM and SIFT with FCM, respectively. The reason of very low correct word translation in both systems may be because there are some alphabets (i.e., “(t)ต” “(n)น” “(p)พ” and etc.) that are very similar to each other. The words that we used are comprised of those alphabets. However the result of the proposed algorithms comparable with that from SIFT but the time process is much lower than SIFT.

The results of SIFT with HCM and that with FCM are comparable with that from SIFT. The processing time of SIFT with HCM and that with FCM is approximately 5 times faster than that of SIFT.

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved