

การจัดทำบริการคำบรรยายแทนเสียง ในช่วงเวลาเดียวกันหรือใกล้เคียง กับเวลาออกอากาศ

RESEARCH AND DEVELOPMENT OF A NEAR REAL-TIME CAPTIONING SERVICE

โยธิน สิริบดีกุล^{1*}, ณัฐนันท์ ทัดพิทักษ์กุล², อนันต์ลดา โชติมงคล³
สุชาติ อรุณนธิกุล⁴, กมลพรรณ พันพึ้ง⁵, สุภาพันท์ เกตุคำ⁶

Yothin Sitthibodeekul^{1*}, Nattanun Thatphithakkul², Ananlada Chotimongkol³
Suchart Arunnitikul⁴, Kamolpun Punpuing⁵, Suphaphan Katekum⁶

องค์การกระจายเสียงและแพร่ภาพสาธารณะแห่งประเทศไทย^{1 และ 4}
ศูนย์เทคโนโลยีอิเล็กทรอนิกส์ และคอมพิวเตอร์แห่งชาติ^{2, 3, 5 และ 6}

Thai Public Broadcasting Service: Thai PBS, Bangkok 10210 Thailand^{1&4}
National Electronics and Computer Technology Center : NECTE National Science
and Technology Development Agency : NSTDA, Pathum Thani Province 12120 Thailand^{2, 3, 5&6}

Corresponding author E-mail: yothins@thaipbs.or.th^{1*}

บทคัดย่อ

การจัดทำบริการคำบรรยายแทนเสียงในช่วงเวลาเดียวกันหรือใกล้เคียงกับเวลาออกอากาศ เป็นการพัฒนาด้อยอดจากระบบออกอากาศโทรทัศน์ภาคพื้นดินระบบดิจิทัล (DVB-T2) ของสถานีโทรทัศน์ไทยพีบีเอส เพื่อให้รองรับข้อมูลนำเข้าในรูปแบบการส่งข้อความอย่างต่อเนื่องโดยไม่มีการจัดบรรทัดไว้ล่วงหน้า และไม่มีช่องว่างระหว่างคำของคำบรรยายแทนเสียงในช่วงเวลาเดียวกันๆ สำหรับภาษาไทยได้ การปรับปรุงระบบฯ มีทั้งส่วนฮาร์ดแวร์และซอฟต์แวร์ โดยทดสอบเชื่อมต่ออุปกรณ์ 2 แนวทาง คือ 1) อุปกรณ์ Screen U8000 ร่วมกับซอฟต์แวร์ Q-Live และ 2) อุปกรณ์ VT3 รุ่น Peach ร่วมกับซอฟต์แวร์ PIMmala Subtitle Gateway พบว่าอุปกรณ์ VT3 เหมาะสมกับการแสดงผลภาษาไทยมากกว่า เนื่องจากสามารถแสดงผลข้อความแบบทยอยโดยไม่ต้องอาศัยช่องว่างระหว่างคำได้ แต่เนื่องจากแบนด์วิดท์ (Bandwidth) ที่จัดสรรไว้ให้กับบริการนี้ ถูกจำกัดอยู่ที่ 100 กิโลบิตต่อวินาที ทำให้ไม่สามารถแสดงข้อความใหม่ที่ละตัวอักษรได้ ต้องรอแสดงข้อความใหม่ทุก 500 มิลลิวินาที ส่งผลให้ค่าเวลาหน่วงของการแสดงผลคำบรรยายแทนเสียงเมื่อเทียบกับเสียงพูดมีค่าสูงขึ้น โดยมีค่าเฉลี่ยอยู่ที่ 6.1 วินาที จากการทดสอบออกอากาศเสมือนจริงกับโครงข่ายสำรอง และทดสอบการยอมรับจากคนพิการทางการได้ยินด้วยคลิปวิดีโอรายการข่าวที่มีค่าความถูกต้องเฉลี่ยร้อยละ 80-90 พบว่าระบบที่พัฒนาขึ้นสามารถออกอากาศได้จริงในเชิงเทคนิค และผู้ใช้อยอมรับรูปแบบการแสดงผลและความถูกต้องที่ระดับนี้ได้

คำสำคัญ : คำบรรยายแทนเสียงแบบปิด เวลาจริง ระบบออกอากาศ โทรทัศน์ระบบดิจิทัล อัตราบิต

Abstract

A near real-time captioning service is developed based on the DVB-T2 broadcasting system of Thai PBS television channel. To produce a real-time closed-captioning service for a live TV program in Thailand, a live stream text receiving module must be developed and integrated into the existing broadcasting system to handle a real-time text stream that has no sentence or word boundary in Thai. From a preliminary experiment, the solution that comprises of the PIMmala subtitle gateway and the VT3 PEACH DVB-sub generator is chosen over the Screen's solution (The Q-Live software and the U8000 DVB-sub generator) as it can handle live text stream which has no word boundary in Thai more appropriately. From five technical experiments (including a simulated broadcasting experiment using a spare MUX) and two user experiments, it has been shown that the systems developed under this study can produce and broadcast real-time closed captions that are acceptable by people with hearing disability which is the targeted group of this study. The caption accuracies of live news programs are around 80-90%. In terms of delay, a caption appears on a television screen 6.1 seconds after its corresponding speech on average. Due to bandwidth limitation, which the regulator allocates only 100 Kbps for the closed-caption service, a caption has to be displayed in a group of characters every 500 milliseconds instead of character-by-character which increases the delay.

Keywords : Closed caption, Real-time, Broadcasting system, Digital television, Bit rate

ความเป็นมา

เนื่องจากสำนักงานคณะกรรมการกิจการกระจายเสียง กิจการโทรทัศน์ และกิจการโทรคมนาคมแห่งชาติ (สำนักงาน กสทช.) มีนโยบายมุ่งเน้นในการเข้าถึงหรือรับรู้ข้อมูลข่าวสารอย่างทั่วถึงและเสมอภาค และกระตุ้นให้เกิดการจัดบริการตามประกาศ กสทช. เรื่อง การส่งเสริมและคุ้มครองสิทธิของคนพิการให้เข้าถึงหรือรับรู้ และใช้ประโยชน์จากรายการของกิจการโทรทัศน์ พ.ศ. 2559 ข้อ 5 กำหนดหลักเกณฑ์ให้ผู้รับใบอนุญาตมีการจัดบริการล่ามภาษามือ (Sign language) คำบรรยายแทนเสียง (Captioning) เสียงบรรยายภาพ (Audio description) หรืออื่นใด ระบุให้จัดบริการให้เหมาะสมกับรูปแบบของแต่ละรายการ โดยต้องมีสัดส่วนของบริการคำบรรยายแทนเสียงอย่างน้อยร้อยละ 40 และจำเป็นต้องให้บริการเพิ่มขึ้นถึงร้อยละ 60 ของสัดส่วนรายการที่เป็นข่าวสารหรือสาระที่เป็นประโยชน์ต่อสาธารณะภายในระยะเวลาที่กำหนด รวมเวลาของรายการที่มีบริการดังกล่าวเฉลี่ยต้องไม่น้อยกว่า 180 นาทีต่อวัน (สำนักงานคณะกรรมการกิจการกระจายเสียง กิจการโทรทัศน์ และกิจการโทรคมนาคมแห่งชาติ, 2560, น. 8-9)

ในปัจจุบันยังไม่มีการจัดบริการคำบรรยายแทนเสียงในช่วงเวลาเดียวกัน หรือใกล้เคียงกับเวลาออกอากาศ (Near real-time captioning) กับรายการโทรทัศน์ที่ออกอากาศสดในประเทศไทย เช่น รายการเหตุการณ์สด

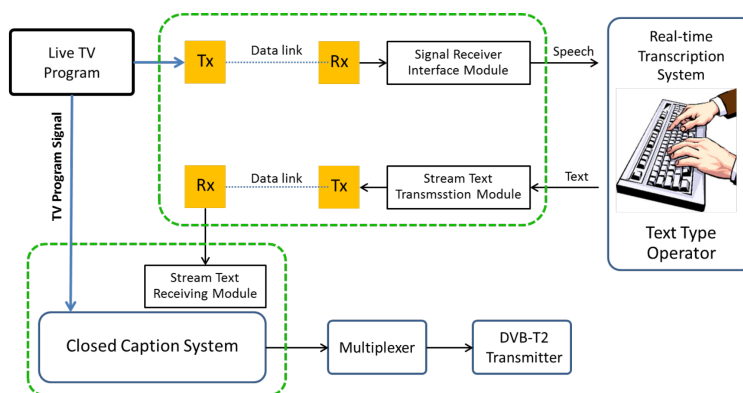
รายการข่าว รายการกีฬา ฯลฯ จากการทบทวนความรู้และประเด็นที่เกี่ยวข้องในการจัดทำคำบรรยายแทนเสียงในช่วงเวลาเดียวกันๆ ในกิจการโทรทัศน์ของต่างประเทศที่ได้ดำเนินการมาก่อน ได้แก่ สหรัฐอเมริกา สหราชอาณาจักร และญี่ปุ่น พบว่าการจัดทำคำบรรยายแทนเสียงในช่วงเวลาเดียวกันๆ มีความแตกต่างและมีความซับซ้อนทางเทคนิคมากกว่าการจัดทำคำบรรยายแทนเสียงสำหรับรายการที่บันทึกไว้ล่วงหน้า (Pre-recorded captions) เนื่องจากคำบรรยายแทนเสียงในช่วงเวลาเดียวกันๆ อยู่ในรูปแบบการส่งข้อความอย่างต่อเนื่อง (Text stream) เพื่อให้ทันต่อเวลาจริงมากที่สุด แตกต่างจากคำบรรยายแทนเสียงในรายการที่บันทึกไว้ล่วงหน้า ซึ่งอยู่ในรูปแบบข้อมูลที่ระบุตำแหน่งเวลา (Time coded) ของวิดีโอ หรือประโยค สามารถกำกับเวลาในการแสดงผลไว้ล่วงหน้าก่อนออกอากาศ

ดังนั้นเพื่อให้ได้แนวทางและกระบวนการในการจัดทำบริการคำบรรยายแทนเสียงในช่วงเวลาเดียวกันหรือใกล้เคียงกับเวลาออกอากาศ ทั้งในเชิงเทคนิคและการบริหารจัดการซึ่งสามารถนำไปใช้ปฏิบัติการได้จริง จึงจำเป็นต้องศึกษาและปรับปรุงระบบออกอากาศทั้งในส่วนของฮาร์ดแวร์และซอฟต์แวร์ เพื่อให้สามารถรองรับข้อมูลนำเข้าในรูปแบบการส่งข้อความอย่างต่อเนื่อง และแสดงผลคำบรรยายแทนเสียงบนจอโทรทัศน์ตามมาตรฐานโทรทัศน์ภาคพื้นดินระบบดิจิทัล (DVB-T2) ได้ นอกจากนี้ ระบบที่พัฒนาขึ้นสำหรับทำหน้าที่แสดงผลข้อความต้องสามารถรองรับการแสดงผลภาษาไทย โดยเฉพาะการแสดงผลข้อความแบบต่อเนื่องโดยไม่มีการจัดบรรทัดไว้ล่วงหน้าได้ เนื่องจากภาษาไทยไม่มีช่องว่างระหว่างคำเหมือนในภาษาอังกฤษ ซึ่งสามารถนำไปใช้ปฏิบัติการได้จริง *"สำหรับการออกอากาศรายการโทรทัศน์ในประเทศไทย"*

แนวทางการศึกษาวิจัย

การวิจัยเพื่อพัฒนาระบบบริการจัดทำคำบรรยายแทนเสียงในช่วงเวลาเดียวกันๆ แบ่งออกเป็น 2 ส่วน คือ 1) ส่วนระบบออกอากาศของสถานีโทรทัศน์ และ 2) ส่วนระบบถอดความเสียงแบบใกล้เคียงเวลาจริง ในส่วนระบบออกอากาศของสถานีโทรทัศน์ การศึกษาของโครงการนี้เป็นการพัฒนาต่อยอดจากระบบฮาร์ดแวร์และเครือข่ายที่มีอยู่เดิมของสถานีโทรทัศน์ไทยพีบีเอส โดยจะปรับปรุงหรือแก้ไขระบบอุปกรณ์เฉพาะในส่วนการออกอากาศระบบโทรทัศน์ดิจิทัล ให้สามารถบริการคำบรรยายแทนเสียงใกล้เคียงกับเวลาจริงได้ โดยส่วนที่ต้องปรับปรุงคือ โมดูล (Module) ที่ใช้จัดทำคำบรรยายแทนเสียงแบบปิดที่มีอยู่เดิม ซึ่งใช้กับรายการที่บันทึกไว้ล่วงหน้าและมีคำบรรยายแทนเสียงที่ได้จัดเตรียมไว้ล่วงหน้าในฐานข้อมูลแล้วเช่นกัน การให้บริการคำบรรยายแทนเสียงในช่วงเวลาเดียวกันๆ ต้องพัฒนามอดูลนี้ให้สามารถรับข้อความที่ได้จากการถอดคำบรรยายแทนเสียงในรูปแบบการส่งข้อความอย่างต่อเนื่อง โดยการปรับปรุงระบบส่งสัญญาณคำบรรยายแทนเสียงแบบปิด (Closed captions transmission system) ซึ่งมีหน้าที่สร้างคำบรรยายแทนเสียงตามมาตรฐาน DVB-Subtitle โดยเพิ่มชุดอุปกรณ์ต่อพ่วงในการรับข้อความต่อเนื่องแบบเวลาจริง (Live stream text receiving module) ที่รองรับการแสดงผลข้อความภาษาไทยแบบต่อเนื่องโดยไม่มีการจัดบรรทัดไว้ล่วงหน้าได้ นอกจากนี้ยังต้องพัฒนาส่วนต่อเชื่อม (Interface module) ระหว่างระบบถอดความเสียงพบกับระบบออกอากาศ (Signal receiver interface module) เพื่อรับสัญญาณเสียงจากระบบออกอากาศของสถานีโทรทัศน์มายังระบบถอดความเสียง

พูดแบบใกล้เคียงเวลาจริง และส่งข้อความที่ได้จากการถอดความกลับไปยังระบบออกอากาศของสถานีโทรทัศน์ในรูปแบบของการถ่ายทอดสด ดังแสดงในภาพที่ 1 โดยส่วนที่จะต้องมีการพัฒนาเพิ่มขึ้นแสดงในกรอบสีเขียว



ภาพที่ 1 โครงสร้างของระบบให้บริการคำบรรยายแทนเสียงสำหรับรายการโทรทัศน์ที่ออกอากาศสด

ที่มา : โยธิน สิทธิพิศกุล และคณะ, 2561

หมายเหตุ : ปรับปรุงจากรายงานฉบับสมบูรณ์โครงการศึกษาวิจัยและพัฒนาระบบการให้บริการคำบรรยายแทนเสียงในเวลาเดียวกัน หรือใกล้เคียงกับเวลาออกอากาศ

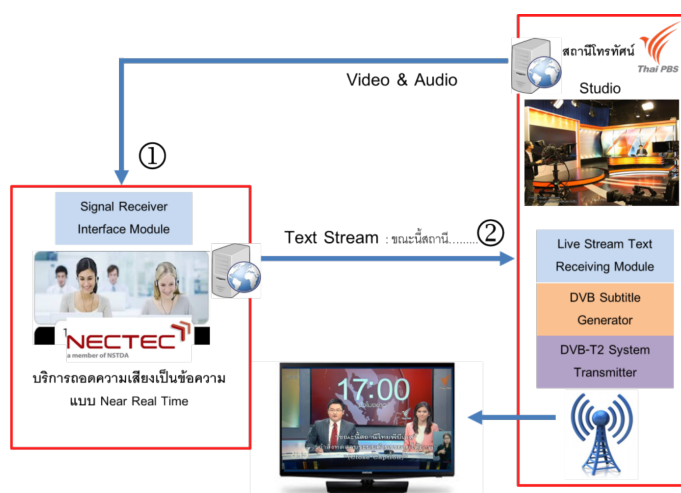
ส่วนระบบถอดความเสียงแบบใกล้เคียงเวลาจริง เนื่องจากความเร็วในการพูดของบุคคลโดยทั่วไปอยู่ที่ 120-160 คำต่อวินาที แต่ความเร็วในการพิมพ์ของเจ้าหน้าที่พิมพ์ดีดที่มีความชำนาญอยู่ที่เพียง 50-60 คำต่อวินาที ซึ่งไม่เพียงพอต่อการจัดทำคำบรรยายเสียงในเวลาเดียวกันๆ ในต่างประเทศได้มีการนำเทคโนโลยีรู้จำเสียงพูดแบบอัตโนมัติ (Automatic Speech Recognition--ASR) มาประยุกต์ใช้ สำหรับภาษาที่ระบบรู้จำเสียงพูดมีความถูกต้องสูง เช่น ญี่ปุ่น อังกฤษ (Imai, 2012, pp. 1-9; Saraclar, Riley, Bocchieri, and Goffin, 2002, pp. 1741-1744) แต่สำหรับภาษาไทย ระบบรู้จำเสียงพูดยังมีข้อจำกัดในเรื่องของความถูกต้อง โดยมีความถูกต้องอยู่ในระดับร้อยละ 80 (Chunwijitra, Chotimongkol, and Wutiwiwatchai, 2015, pp. 2385-2389) และหากมีเสียงรบกวน ผู้พูดพูดเร็วหรือพูดไม่ชัดเจน ความถูกต้องจะยิ่งลดต่ำลง จำเป็นต้องมีการแก้ไขข้อความให้ถูกต้องมากกว่าร้อยละ 90 เสียก่อน เพื่อให้คนพิการทางการได้ยินสามารถเข้าใจเนื้อหาข้อความได้ (สำนักงาน กสทช., 2560, น. 11) โดยกระบวนการทั้งหมดต้องทำภายในเวลาจำกัด เพื่อให้ได้คำบรรยายแทนเสียงในเวลาใกล้เคียงเวลาจริงมากที่สุด

จากข้อจำกัดข้างต้น ทีมวิจัยจึงเลือกใช้เทคนิคการถอดความเสียงแบบช่วยกันพิมพ์ของเนคเทค ซึ่งใช้ผู้ถอดความหลายคนช่วยกันถอดความด้วยแป้นพิมพ์คีย์บอร์ดแบบปกติ โดยเจ้าหน้าที่แต่ละคนจะรับผิดชอบถอดความเสียงเฉพาะช่วงสั้นๆ เสียงช่วงถัดไปจะถูกถอดความโดยเจ้าหน้าที่คนอื่น เมื่อนำข้อความที่เจ้าหน้าที่

แต่ละคนถอดความได้มาเรียงต่อกัน จะได้ข้อความที่สมบูรณ์ภายในเวลาจริง ระบบถอดความเสียงพูดแบบช่วยกันพิมพ์ใช้โปรโตคอล WebRTC (Web Real-Time Communication) ในการส่งข้อความ เนื่องจาก WebRTC เป็นโปรโตคอลแบบทันต่อเวลาจริง สามารถส่งข้อมูลออกมาทุกครั้งเมื่อมีการกดปุ่มตัวอักษรบนแป้นพิมพ์ โดยไม่ต้องกดปุ่ม Enter หรือ Send เมื่อพิมพ์ถอดความจบ ทำให้สามารถส่งข้อความจากผู้ถอดความมายังระบบออกอากาศได้อย่างต่อเนื่องเพื่อให้ทันต่อเวลาจริงมากที่สุด (Mozilla, 2018, online)

การทดสอบการเชื่อมต่อกับอุปกรณ์

การทดสอบการเชื่อมต่ออุปกรณ์เพื่อออกแบบระบบบริการคำบรรยายแทนเสียงในช่วงเวลาเดียวกัน มีการทดสอบใน 2 ส่วน คือ 1) Signal receiver interface module และ 2) Live stream text receiving module ดังแสดงในภาพที่ 2



ภาพที่ 2 แผนผังการทดสอบการเชื่อมต่อกับอุปกรณ์

ที่มา : โยธิน ลิทธิบัติกุล และคณะ, 2561

หมายเหตุ : ปรับปรุงจากรายงานฉบับสมบูรณ์โครงการศึกษาวิจัยและพัฒนาการจัดการจัดทำบริการคำบรรยายแทนเสียงในช่วงเวลาเดียวกัน หรือใกล้เคียงกับเวลาออกอากาศ

การทดสอบการเชื่อมต่ออุปกรณ์ส่วนที่ 1) Signal receiver interface module โดยการส่งสัญญาณเสียงที่ดึงจากห้องส่งของสถานีโทรทัศน์ไทยพีบีเอสผ่านอินเทอร์เน็ตไปยังระบบถอดความเสียงพูดของเนคเทค และส่งสัญญาณเสียงกลับมายังสถานีโทรทัศน์ไทยพีบีเอสอีกครั้ง พบว่าสัญญาณเสียงที่ส่งกลับมามีลักษณะเป็นเสียงสะท้อน (Echo) ของเสียงที่ส่งไป แสดงให้เห็นว่าระยะเวลาหน่วง (Delay) ที่เกิดจากการส่งสัญญาณเสียงไป-กลับ ผ่านเครือข่ายอินเทอร์เน็ตน้อยกว่า 100 มิลลิวินาที ซึ่งเป็นระยะเวลาที่น้อยมาก

การทดสอบในส่วนที่ 2) ทดสอบการรับข้อความที่ได้จากการถอดความเสียงในรูปแบบข้อความต่อเนื่องของ Live stream text receiving module รวมทั้งการแปลงข้อความที่ได้รับให้เป็นคำบรรยายแทนเสียงแบบปิดตามมาตรฐาน DVB-Subtitle (DVB-sub generator) โดยได้ทดสอบใน 2 แนวทาง ดังนี้

แนวทางที่ 1 ใช้เครื่อง Screen U8000 เป็น DVB-sub generator และใช้ซอฟต์แวร์ WinCAPS Quantum Q-Live (Q-Live) เป็น Live stream text receiving module

แนวทางที่ 2 ใช้เครื่อง VT3 รุ่น Peach ทำหน้าที่เป็น DVB-sub generator และใช้ซอฟต์แวร์ PIMmala Subtitle Gateway ที่พัฒนาขึ้นใหม่โดย หจก. นีว่า เทคโนโลยี (ห้างหุ้นส่วนจำกัด นีว่า เทคโนโลยี, 2561, น. 1-2) ทำหน้าที่เป็น Live stream text receiving module

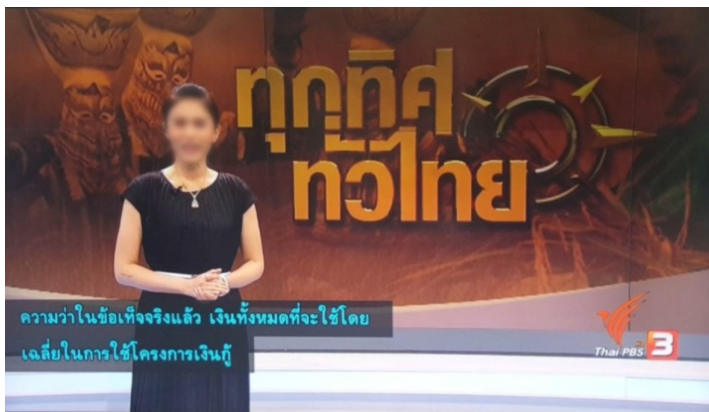
จากการทดสอบในแนวทางที่ 1 พบว่า Q-Live ออกแบบมาสำหรับการแสดงผลภาษาอังกฤษหรือภาษาที่มีช่องว่างคั่นระหว่างคำเป็นหลัก เมื่อนำมาใช้กับภาษาที่ไม่มีช่องว่างระหว่างคำแต่มีช่องว่างระหว่างวลี เช่น ภาษาไทย การแสดงผลข้อความจึงไม่สมบูรณ์ เกิดปัญหาในการจัดบรรทัดเมื่อพบวลีที่มีความยาวมาก ดังแสดงในภาพที่ 3 เนื่องจาก Q-Live ไม่สามารถทำการยกคำไปขึ้นบรรทัดใหม่ (Line wrapping) ได้ วลีที่ยาวจะถูกเลื่อนลงมาอยู่บรรทัดถัดไป และเกิดช่องว่างขนาดใหญ่ในบรรทัดบน เมื่อพยายามปรับตั้งค่าให้สามารถแสดงผลรายตัวอักษรได้ พบว่าข้อความทั้งข้อความเกิดการกะพริบ เนื่องจาก Q-Live ทำการประมวลผลข้อความแล้วนำไปเรนเดอร์ (Render) ภาพใหม่ทุกครั้ง เมื่อสอบถามไปยังบริษัท Screen พบว่าการใช้งานกับภาษาที่ไม่มีช่องว่างระหว่างคำดังภาษาไทย จะใช้การแสดงผลทีละบรรทัด ซึ่งจะเพิ่มเวลาหน่วงในการแสดงผล และเบื้องต้นไม่สามารถปรับซอฟต์แวร์ให้รองรับการแสดงผลในภาษาไทยได้



ภาพที่ 3 ตัวอย่างการแสดงผลคำบรรยายแทนเสียงบนหน้าจอโทรทัศน์ด้วยซอฟต์แวร์ Q-Live

ที่มา : โยธิน ลิทธิติกุล และคณะ (2561)

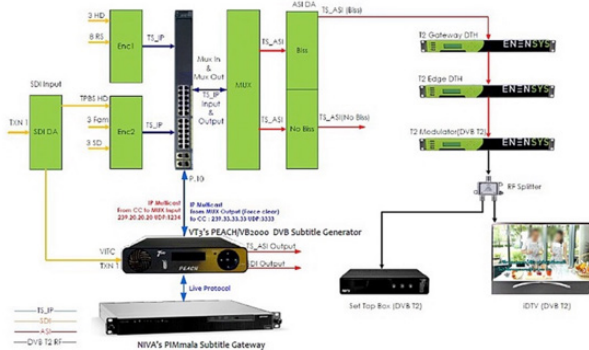
จากการทดสอบในแนวทางที่ 2 โดยใช้ซอฟต์แวร์ PIMmala Subtitle Gateway ควบคุมการแสดงผล ปรับขนาดตัวอักษร สีตัวอักษร ตำแหน่งสูง-ต่ำ การจัดวางกึ่งกลาง ชิดซ้าย-ชิดขวา และรูปแบบการแสดงผล ตัวอักษร พบว่าสามารถแสดงผลแบบทยอยออกได้ โดยแบ่งการแสดงผลออกเป็นส่วนๆ แล้ว จึงแสดงข้อความในแต่ละส่วนตามลำดับ และกำหนดจังหวะเวลาที่เหมาะสมต่อระบบการส่งออกอากาศ ดังแสดงในภาพที่ 4 ส่งผลให้ข้อความปรากฏขึ้นบนจอภาพค่อนข้างเร็ว และสามารถอ่านได้ง่ายกว่าการแสดงผลในแนวทางที่ 1 ซึ่งเหมาะกับการแสดงผลภาษาไทยที่ไม่มีช่องว่างระหว่างคำในประโยคได้



ภาพที่ 4 ตัวอย่างการแสดงผลคำบรรยายแทนเสียงบนหน้าจอโทรทัศน์ด้วยซอฟต์แวร์ PIMma

ที่มา : โยธิน สิทธิพิบัติกุล และคณะ (2561)

จากผลการทดสอบเปรียบเทียบกับระบบอุปกรณ์และซอฟต์แวร์ในแนวทางที่ 1 และแนวทางที่ 2 ได้ผลเป็นที่ประจักษ์สามารถสรุปได้ว่า ระบบอุปกรณ์และซอฟต์แวร์ในแนวทางที่ 2 (PIMmla Subtitle Gateway) เหมาะสมกับการแสดงผลภาษาไทยมากกว่าแนวทางที่ 1 (Q-Live) และยังคงมีการปรับปรุงพัฒนาซอฟต์แวร์ให้รองรับการแสดงผลภาษาไทยได้อย่างต่อเนื่อง ดังนั้น จึงสามารถออกแบบโครงสร้างของระบบออกอากาศเพื่อให้บริการคำบรรยายแทนเสียงในช่วงเวลาเดียวกันฯ ดังแสดงในภาพที่ 5



ภาพที่ 5 โครงสร้างของระบบออกอากาศเพื่อให้บริการคำบรรยายแทนเสียงในช่วงเวลาเดียวกันๆ ที่เหมาะสม

ที่มา : โยธิน สิทธิพิบัติกุล และคณะ (2561)

ผลการศึกษา

ผลการทดสอบเชิงเทคนิค

คณะผู้วิจัยดำเนินการทดสอบระบบบริการจัดทำคำบรรยายแทนเสียงในช่วงเวลาเดียวกันๆ แบบปิดเชิงเทคนิค จำนวน 5 ครั้ง เพื่อทดลองการทำงานของส่วนที่เกี่ยวข้องให้สามารถจัดทำผลงานบริการดังกล่าวได้อย่างมีประสิทธิภาพที่สุด โดยใช้รูปแบบการต่อเชื่อมมอดูลและอุปกรณ์ต่างๆ ในการผลิตคำบรรยายแทนเสียงแบบปิดตามมาตรฐาน DVB-Subtitle เพื่อออกอากาศในระบบออกอากาศโทรทัศน์ภาคพื้นดินระบบดิจิทัลของสถานีโทรทัศน์ ดังแสดงในภาพที่ 5

หลังจากทดสอบการต่อเชื่อมองค์ประกอบต่างๆ ของระบบให้บริการคำบรรยายแทนเสียงในช่วงเวลาเดียวกันๆ แบบปิด พบว่าผู้ใช้สามารถเลือกเปิด-ปิดคำบรรยายแทนเสียงภาษาไทยผ่านเมนูของเครื่องรับสัญญาณโทรทัศน์ภาคพื้นดินฯ ได้ ด้านการแสดงผลคำบรรยายบนหน้าจอโทรทัศน์ พบว่าสามารถปรับรูปแบบตัวอักษร สีของตัวอักษร สีของพื้นหลัง และการจัดวางตำแหน่งของตัวอักษรบนหน้าจอโทรทัศน์ได้เช่นกัน

ผลการทดสอบความเสถียรของระบบออกอากาศเมื่อมีการให้บริการคำบรรยายแทนเสียงในช่วงเวลาเดียวกันๆ แบบปิดร่วมด้วย โดยการวัดค่าอัตราบิต (Bit rate) ที่ใช้ในการออกอากาศ พบว่าการแสดงผลคำบรรยายแทนเสียงทีละตัวอักษร (Character) เพื่อให้ทันต่อเวลาจริง ระบบต้องจัดรูปแบบและแสดงผลข้อความใหม่ค่อนข้างถี่ อาจทำให้เกิดปัญหาการใช้อัตราบิตสูงกว่าที่ได้รับการจัดสรร (Bandwidth overshoot) ซึ่งจะกระทบต่อการทำงานของระบบโครงข่าย (MUX) และอาจทำให้โครงข่ายล้นได้ ทั้งนี้ ปัจจุบันสำนักงาน กสทช. จัดสรรแบนด์วิดท์สำหรับบริการคำบรรยายแทนเสียงไว้ที่ 100 กิโลบิตต่อวินาที (สำนักงานคณะกรรมการกิจการกระจายเสียง กิจการโทรทัศน์ และกิจการโทรคมนาคม, 2560, น. 15 และ 36) โดยสถานีโทรทัศน์ไทยพีบีเอสได้กำหนดค่าอัตราบิตไว้ที่ 80 กิโลบิตต่อวินาที เพื่อป้องกันปัญหาไว้อีกชั้นหนึ่ง และปรับรูปแบบการแสดงผลคำบรรยายแทนเสียงจากเดิมแบบทีละตัวอักษร เป็นรูปแบบใหม่ตามค่าเวลาที่ตั้งไว้แทน (Refresh rate) โดยได้ทดสอบหาค่า Refresh rate ที่มีค่าความถี่สูงสุด เพื่อให้ได้การแสดงผลทันต่อเวลา แต่ยังคงมีอัตราบิตอยู่ในกรอบแบนด์วิดท์ที่ได้จัดสรรไว้สำหรับบริการคำบรรยายแทนเสียง

จากการทดสอบพบว่า ในการแสดงผลคำบรรยายแทนเสียงแบบ HD 720 ความถี่ในการส่งข้อมูลสูงสุดที่ไม่เกิดปัญหาใช้อัตราบิตสูงกว่าที่ได้รับการจัดสรร คือ Refresh rate ที่ 500 มิลลิวินาที ส่วนการแสดงผลคำบรรยายแทนเสียงแบบ SD 625 ความถี่ในการส่งข้อมูลสูงสุดที่ไม่เกิดปัญหาใช้อัตราบิตสูงกว่าที่ได้รับการจัดสรร คือ Refresh rate ที่มากกว่าหรือเท่ากับ 250 มิลลิวินาที ซึ่งสามารถแสดงผลคำบรรยายแทนเสียงได้ทันต่อเวลาจริงและมีความต่อเนื่องมากกว่าการแสดงผลแบบ HD 720

หลังจากทดสอบและปรับปรุงระบบให้บริการคำบรรยายแทนเสียงในช่วงเวลาเดียวกันๆ แบบปิดในประเด็นต่างๆ ข้างต้นแล้ว ได้มีการทดสอบออกอากาศบริการคำบรรยายแทนเสียงเสมือนจริงกับโครงข่ายสำรอง ในวันที่ 6 มกราคม พ.ศ. 2561 โดยทดสอบรับชมจากเครื่องรับสัญญาณโทรทัศน์ภาคพื้นดินประเภท STB (Set-top box) และเครื่องรับสัญญาณโทรทัศน์ภาคพื้นดินประเภท iDTV (Integrated digital television) ซึ่งสามารถเลือกเปิด-ปิดคำบรรยายแทนเสียงได้จากรีโมทคอนโทรล ดังแสดงในภาพที่ 6



ภาพที่ 6 การเปิดบริการคำบรรยายแทนเสียงบนเครื่องรับสัญญาณโทรทัศน์ ประเภท iDTV

ที่มา : โยธิน สิทธิพิศกุล และคณะ (2561)

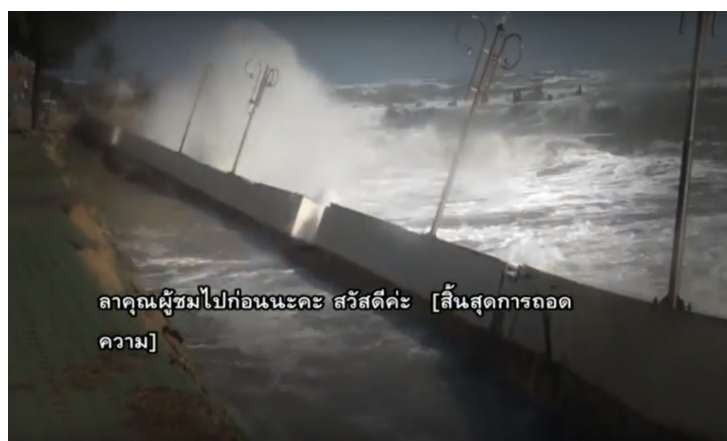
สำหรับค่าความถูกต้องในการถอดความเสียงรายการข่าวอยู่ที่ร้อยละ 80-90 โดยมีค่าเฉลี่ยอยู่ที่ร้อยละ 86.0 เมื่อใช้เจ้าหน้าที่ถอดความ 5 คน ดังแสดงในตารางที่ 1 หากใช้ผู้ถอดความน้อยกว่า 5 คน จะส่งผลให้ค่าความถูกต้องลดลงค่อนข้างมาก เนื่องจากรายการข่าวมีอัตราการพูดค่อนข้างเร็ว สำหรับระยะเวลาหน่วยวัดจากระยะห่างระหว่างเวลาที่ผู้ชมได้ยินเสียงพูดกับเวลาที่ผู้ชมเห็นคำบรรยายแทนเสียงของประโยคนั้น ปรากฏบนจอโทรทัศน์ จากการวิเคราะห์ค่าเวลาหน่วง พบว่ามีค่าเฉลี่ยอยู่ที่ 6.1 วินาที

ตารางที่ 1 ตารางแสดงผลการทดสอบความถูกต้องในการถอดความ และค่าเวลาหน่วง (Delay) เฉลี่ย

ชื่อรายการโทรทัศน์	ค่าความถูกต้องในการถอดความเฉลี่ย (ร้อยละ)	ค่าเวลาหน่วง (Delay) เฉลี่ย (วินาที)
กันข่าว 16.00 น.	88.7	7.2
กันข่าว 17.00 น.	85.8	5.8
เดินหน้าประเทศไทย	85.1	4.5
ข่าวค่ำ มิติใหม่ทั่วไทย	84.4	6.8
ผลเฉลี่ย	86.0	6.1

ที่มา : โยธิน สิทธิพิศกุล และคณะ (2561)

เนื่องจากคำบรรยายแทนเสียงจะปรากฏช้ากว่าเสียงพูด และจำเป็นต้องเผื่อเวลาให้ผู้รับชมสามารถอ่านข้อความได้ครบถ้วน เมื่อสิ้นสุดการถอดความเสียง ระบบให้บริการคำบรรยายแทนเสียงในช่วงเวลาเดียวกันๆ จะส่งข้อความ “[สิ้นสุดการถอดความ]” เพื่อให้ผู้รับชมทราบ ดังแสดงในภาพที่ 7 และส่งอักขระพิเศษ “๕” (Unicode U+22F3) เพื่อให้โปรแกรม PIMMala subtitle gateway หยุดแสดงผลคำบรรยายแทนเสียง โดยจะคงคำบรรยายแทนเสียงหลังจบรายการระยะหนึ่ง ก่อนหยุดการแสดงผล



ภาพที่ 7 ข้อความ “[สิ้นสุดการถอดความ]” ปรากฏขึ้นเมื่อจบรายการที่มีบริการคำบรรยายแทนเสียง
ที่มา : <https://youtu.be/PV4uJF7Q-Vk> (นาทีที่ 1.57)

ผลการทดสอบกับผู้ใช้

คณะผู้วิจัยฯ ดำเนินการทดสอบระบบบริการจัดทำคำบรรยายแทนเสียงในช่วงเวลาเดียวกันๆ แบบปิดกับผู้ใช้กลุ่มคนพิการทางการได้ยิน จำนวน 2 ครั้ง รวมจำนวนผู้เข้าร่วมทดสอบทั้งสิ้น 233 คน เพื่อรับทราบความคิดเห็นและความต้องการเกี่ยวกับการจัดทำระบบบริการดังกล่าว ผ่านการตอบแบบสอบถาม โดยให้รับชมคลิปวิดีโอรายการโทรทัศน์ที่มีคำบรรยายแทนเสียงรูปแบบต่างๆ จากเครื่องรับโทรทัศน์ โดยแบ่งผู้ใช้เป็นกลุ่มย่อยจำนวน 10-15 คนต่อเครื่องรับโทรทัศน์ 1 เครื่อง ดังแสดงในภาพที่ 8 เพื่อให้ใกล้เคียงกับการรับชมโทรทัศน์โดยทั่วไป การทดสอบโดยใช้คลิปวิดีโอช่วยให้เปรียบเทียบคำบรรยายแทนเสียงของเนื้อหารายการเดียวกันในรูปแบบที่ต่างกัน และสามารถทดสอบกับผู้ใช้จำนวนมากโดยแบ่งเป็นรอบการทดสอบได้



ภาพที่ 8 การทดสอบรับชมคลิปวิดีโอรายการโทรทัศน์ที่มีคำบรรยายแทนเสียงในช่วงเวลาเดียวกันฯ
แบบปิดจากเครื่องรับโทรทัศน์ ของกลุ่มคนพิการทางการได้ยิน

ที่มา : โยธิน สิทธิบัติกุล และคณะ (2561)

ผลการทดสอบกับผู้ใช้งานซึ่งประเมินผ่านแบบสอบถาม สามารถสรุปประเด็นตามปัจจัยที่มีผลต่อการใช้งานบริการคำบรรยายแทนเสียง 3 ปัจจัย ได้ดังนี้

1. ความเร็วในการพูดของผู้พูด พบว่าคำบรรยายแทนเสียงที่ผู้พูดมีอัตราความเร็วในการพูดแตกต่างกัน ผู้ใช้มีความเห็นว่าสามารถอ่านคำบรรยายแทนเสียงได้ทันไม่แตกต่างกัน
2. รูปแบบการแสดงผลคำบรรยายแทนเสียง พบว่าลักษณะการแสดงผลการไหลของตัวอักษร (Text flow) เช่น จังหวะการเลื่อนตัวอักษรเพื่อขึ้นบรรทัดใหม่ ความสม่ำเสมอของตัวอักษรที่ปรากฏขึ้นใหม่บนหน้าจอ มีผลต่อความยากง่ายในการอ่านคำบรรยายแทนเสียง
3. ความถูกต้องในการถอดความ พบว่าผู้ใช้ร้อยละ 47.1 ยอมรับข้อผิดพลาดในการถอดความที่สังเกตเห็นได้ เนื่องจากมีความเห็นว่าข้อผิดพลาดไม่ร้ายแรง ไม่กระทบกับเนื้อหาของรายการ ขณะที่ผู้ใช้ร้อยละ 51.9 ไม่ยอมรับข้อผิดพลาด เนื่องจากมีความเห็นว่าข้อผิดพลาดที่พบส่งผลให้อ่านคำบรรยายแทนเสียงได้ยาก และมีผลต่อการเรียนรู้ภาษาเขียนของคนหูหนวก ทั้งนี้ ความถูกต้องในการถอดความเสียงแบบทันต่อเวลาสำหรับรายการข่าว ที่นำมาใช้ในการทดสอบอยู่ที่ร้อยละ 80-90

ผลการทดสอบอุปกรณ์ภาคออกอากาศ

คณะผู้วิจัย ได้นำเครื่องรับสัญญาณโทรทัศน์ภาคพื้นดินฯ ประเภท STB (DVB-T2 Set-top box) ที่มีจำหน่ายในท้องตลาด มาทดสอบการรับสัญญาณเสมือนการแพร่ภาพออกอากาศรายการโทรทัศน์จริง และตรวจสอบการแสดงผลของข้อความคำบรรยายแทนเสียงในช่วงเวลาเดียวกันฯ บนหน้าจอภาพ ในวันที่ 6 มกราคม 2561 เวลา 10.00-20.15 น. เพื่อทดสอบใน 2 ประเด็น คือ 1) เครื่องรับสัญญาณโทรทัศน์ภาคพื้นดินฯ ประเภท STB ยี่ห้อต่างๆ สามารถรับสัญญาณคำบรรยายแทนเสียงในช่วงเวลาเดียวกันฯ แบบปิดได้หรือไม่ และมีความแตกต่างกันหรือไม่อย่างไร 2) หากเปลี่ยนแปลงค่าพารามิเตอร์ที่ใช้ในการส่งคำบรรยายแทนเสียงในช่วงเวลาเดียวกันฯ แบบปิด จะเกิดผลอย่างไรต่อเครื่องรับสัญญาณโทรทัศน์ภาคพื้นดินฯ ที่นำมาทดสอบเพื่อนำมาใช้ในการกำหนดค่าพารามิเตอร์ของระบบ Subtitle generator ต่อไป

ทั้งนี้ คณะผู้วิจัยออกแบบการทดสอบโดยใช้เครื่องรับสัญญาณโทรทัศน์ภาคพื้นดินฯ ประเภท STB จำนวน 8 เครื่องที่มียี่ห้อแตกต่างกัน ส่งสัญญาณคลื่นโทรทัศน์ระบบดิจิทัลโดยการต่อสายสัญญาณภาพและเสียงชนิด HDMI จากเครื่องรับสัญญาณโทรทัศน์ภาคพื้นดินฯ แบบ STB ไปที่เครื่องรับโทรทัศน์ และบันทึกรายการโทรทัศน์ที่ทดสอบให้บริการระบบคำบรรยายแทนเสียงในช่วงเวลาเดียวกันฯ ลงในหน่วยความจำ (Flash memory) ด้วยเครื่องรับสัญญาณฯ ที่มีคุณสมบัติสามารถบันทึกรายการโทรทัศน์ได้ในระหว่างดำเนินการทดลอง ผลที่ได้จากการทดสอบ แสดงในตารางที่ 2

ตารางที่ 2 สรุปผลการทดสอบเครื่องรับสัญญาณโทรทัศน์ภาคพื้นดินฯ ประเภท STB (Set-top box)

เครื่องรับสัญญาณโทรทัศน์ภาคพื้นดินฯ ประเภท STB	พารามิเตอร์ของระบบ		การแสดงผลคำบรรยายแทนเสียง ในช่วงเวลาเดียวกันฯ
	Refresh Rate (Milliseconds)	ระบบภาพ	
จำนวน 4 เครื่อง (4 ยี่ห้อ)	500	SD 625	แสดงผลสมบูรณ์เป็นปกติ
	500	HD 720	แสดงผลสมบูรณ์เป็นปกติ
	250	SD 625	แสดงผลสมบูรณ์เป็นปกติ
	250	HD 720	แสดงผลสมบูรณ์เป็นปกติ
จำนวน 4 เครื่อง (4 ยี่ห้อ)	500	SD 625	แสดงผลสมบูรณ์เป็นปกติ
	500	HD 720	ข้อความคำบรรยายแทนเสียงกะพริบ
	250	SD 625	ข้อความคำบรรยายแทนเสียงกะพริบ
	250	HD 720	ข้อความคำบรรยายแทนเสียงกะพริบ

ที่มา : โยธิน สิทธิพิศกุล และคณะ (2561)

การอภิปรายผลการศึกษา

การทดสอบเชิงเทคนิค

การทดสอบออกอากาศบริการคำบรรยายแทนเสียงแบบเสมือนจริง ใช้ความละเอียดการแสดงผลคำบรรยายแทนเสียงประเภท SD โดยมี Refresh rate ที่ 250 มิลลิวินาที ตามผลการทดสอบเชิงเทคนิคครั้งที่ 2 อย่างไรก็ดี เมื่อทดสอบออกอากาศผ่านโครงข่าย พบว่าการแสดงผลที่ปรากฏบนจอโทรทัศน์ผ่านเครื่องรับสัญญาณโทรทัศน์ภาคพื้นดินฯ ประเภท STB และประเภท iDTV มีความถี่น้อยกว่าค่าที่ DVB Subtitle generator ตั้งไว้และไม่สม่ำเสมอ อาจเนื่องจากความเร็วในการแปลงข้อมูล (Encode) เพื่อส่งสัญญาณภาพไปออกอากาศ จึงปรับใช้การแสดงผลประเภท SD 625 โดยมี Refresh rate ที่ 500 มิลลิวินาที

การทดสอบกับผู้ใช้

จากการวิเคราะห์ผลการทดสอบกับผู้ใช้ เพื่อประเมินปัจจัยที่มีผลต่อการใช้งานบริการคำบรรยายแทนเสียงพบว่า แม้คำบรรยายแทนเสียงที่ผู้พุดมีอัตราความเร็วในการพูดแตกต่างกัน แต่ผู้ใช้มีความเห็นว่าสามารถอ่านคำบรรยายแทนเสียงได้ทันไม่แตกต่างกัน ทั้งนี้ จากข้อกำหนดของ กสทช. ที่กำหนดให้การถอดความคำบรรยายแทนเสียงต้องมีข้อความครบถ้วนตามเสียงพูดยังไม่อนุญาตให้สรุปความ จึงยังไม่สามารถแสดงผลข้อความให้ช้าลงหรือมีปริมาณข้อความน้อยลงได้ เพื่อแสดงผลข้อความให้ทันต่อเวลา อย่างไรก็ดี จากการสอบถามเพิ่มเติมพบว่า ลักษณะการแสดงผลการไหลของตัวอักษร เช่น จังหวะการเลื่อนตัวอักษรเพื่อขึ้นบรรทัดใหม่ ความสม่ำเสมอของตัวอักษรที่ปรากฏขึ้นใหม่บนหน้าจอ มีผลต่อความยากง่ายในการอ่านคำบรรยายแทนเสียง จึงปรับปรุงให้ตัวอักษรที่ปรากฏขึ้นใหม่มีความสม่ำเสมอมากขึ้น และเพิ่มการรอแสดงผลเมื่อมีการเลื่อนบรรทัดเพื่อให้ผู้ใช้สามารถอ่านทัน จากนั้นจึงทดสอบกับผู้ใช้อีกครั้ง

โดยกลุ่มผู้ใช้ที่เข้าร่วมการทดสอบทั้ง 2 ครั้ง จำนวน 30 คน ทั้งหมดให้ความเห็นว่า การปรับปรุงการแสดงผลการไหลของข้อความดังกล่าว ช่วยให้อ่านคำบรรยายแทนเสียงได้ง่ายขึ้น

ด้านความเข้าใจเนื้อหาของรายการโทรทัศน์จากการอ่านบริการคำบรรยายแทนเสียงโดยการตอบคำถามพบว่า ความถูกต้องในการตอบคำถามเกี่ยวกับเนื้อหารายการมีความสอดคล้องกับความเห็น

ด้านการอ่านคำบรรยายแทนเสียงได้ทัน โดยกลุ่มผู้ใช้ที่ตอบว่าอ่านคำบรรยายแทนเสียงได้ทันจะมีคะแนนความถูกต้องโดยเฉลี่ยสูงกว่ากลุ่มผู้ใช้ที่ตอบว่าอ่านไม่ทัน แสดงให้เห็นว่าการอ่านคำบรรยายแทนเสียงได้ทันมีความสัมพันธ์กับความเข้าใจเนื้อหาของรายการโทรทัศน์อย่างมีนัยสำคัญ

การทดสอบอุปกรณ์ภาคออกอากาศ

เมื่อกำหนดค่าพารามิเตอร์ของ Subtitle generator เป็นระบบ HD 720 และโครงข่ายสัญญาณที่ใช้ในการทดสอบ จำกัดแบนด์วิดท์สูงสุดไว้ที่ไม่เกิน 80 กิโลบิตต่อวินาที และเพื่อค่าของความสามารถในการรับสัญญาณหรือบันทึกได้สูงสุด (Head room) กรณีที่อาจจะเกิดการใช้อัตราบิตสูงกว่าที่ได้รับการจัดสรรไว้ที่ 100 กิโลบิตต่อวินาที ส่งผลให้เครื่องรับสัญญาณโทรทัศน์ภาคพื้นดินฯ ประเภท STB จำนวน 4 ยี่ห้อ พบปัญหา คือ การแสดงผลของข้อความครบถ้วน แต่ภาพสั่นและกะพริบเป็นจังหวะทุกค่า Refresh rate ส่วนเครื่องรับสัญญาณฯ อีก 4 ยี่ห้อ สามารถแสดงผลได้สมบูรณ์ แสดงให้เห็นว่าเครื่องรับสัญญาณฯ ที่ไม่ได้ออกแบบมาเพื่อรองรับปริมาณข้อมูลมากเพียงพอ จะไม่สามารถแสดงผลข้อความคำบรรยายแทนเสียงในช่วงเวลาเดียวกันฯ เป็นปกติได้ และเมื่อนำไฟล์วิดีโอที่บันทึกได้จากเครื่องรับสัญญาณฯ ที่แสดงผลข้อความคำบรรยายแทนเสียงกะพริบกลับมาย้อนดูอีกครั้งด้วยโปรแกรม VLC ที่สามารถถอดรหัสคำบรรยายแทนเสียงแบบปิดได้ ปรากฏว่าข้อความคำบรรยายแทนเสียงแสดงผลเป็นปกติ ดังนั้นจึงสามารถสรุปได้ว่าเครื่องรับสัญญาณโทรทัศน์ภาคพื้นดินฯ ประเภท STB ที่มีขายอยู่ในตลาดช่วงเวลา 4 ปีที่ผ่านมาถึงปัจจุบัน คาดว่าสามารถรองรับอัตราข้อมูลของระบบคำบรรยายแทนเสียงในช่วงเวลาเดียวกันฯ ได้สูงสุดระหว่าง 80-100 กิโลบิตต่อวินาที หากอัตราข้อมูลมากกว่านี้ จะส่งผลให้เครื่องแปลงสัญญาณโทรทัศน์บางยี่ห้อหรือบางรุ่นไม่สามารถแสดงผลคำบรรยายแทนเสียงในช่วงเวลาเดียวกันฯ ตามโครงการนี้ได้

บทสรุปและข้อเสนอแนะ

บทสรุปและข้อเสนอแนะเชิงเทคนิค

จากผลการศึกษาพบว่า สถานีโทรทัศน์ไทยพีบีเอสสามารถดำเนินการจัดทำบริการคำบรรยายแทนเสียงในช่วงเวลาเดียวกันฯ ได้จริงในเชิงเทคนิค และกลุ่มผู้ใช้งานคนพิการทางการได้ยินซึ่งเป็นกลุ่มเป้าหมายหลักยอมรับผลการทดสอบว่าอยู่ในระดับการใช้งานที่เข้าใจได้ อย่างไรก็ตาม จากผลการทดสอบเชิงเทคนิค ยังพบข้อจำกัดที่ยังไม่สามารถแก้ไขได้ในปัจจุบัน จำเป็นต้องพัฒนาเพิ่มเติมเพื่อเพิ่มคุณภาพบริการคำบรรยายแทนเสียงในช่วงเวลาเดียวกันฯ ดังนี้

ด้านความถูกต้องของการถอดความด้วยเทคนิคการแบ่งพยางค์สำหรับรายการข่าว มีค่าความถูกต้องเฉลี่ยอยู่ที่ร้อยละ 80-90 จากการทดสอบกับผู้ใช้ พบว่าผู้ใช้งานส่วนน้อยยอมรับความถูกต้องที่ระดับนี้ได้ ในขณะที่ผู้ใช้งานส่วนน้อยต้องการความถูกต้องที่สูงกว่านี้ ปัจจัยที่มีผลต่อความถูกต้องในการถอดความ ได้แก่ คุณภาพของเสียงพูด คำศัพท์ และชื่อเฉพาะ ทั้งนี้ ปัจจัยด้านคุณภาพของเสียงอาจปรับปรุงได้ยาก เนื่องจากรายการข่าวอาจมีการสัมภาษณ์นอกสถานที่หรือผ่านทางโทรศัพท์ จึงมีคุณภาพของเสียงจำกัดตามรูปแบบช่องทางการสื่อสารนั้นๆ ส่วนปัจจัยด้านคำศัพท์สามารถปรับปรุงได้ โดยให้เจ้าหน้าที่ถอดความศึกษาบทสกริปต์ที่ใช้ในการ

อ่านช้าก่อน ซึ่งควรมีการพัฒนาส่วนต่อเชื่อมสำหรับดึงข้อมูลสคริปต์ข่าวเพื่อส่งต่อไปยังระบบถอดความเสียงพูดแบบทันที

ด้านความทันต่อเวลา การจัดสรรอัตราบิตไว้ให้กับบริการคำบรรยายแทนเสียงของประเทศไทยอยู่ที่ 100 กิโลบิตต่อวินาที ส่งผลให้ไม่สามารถแสดงผลคำบรรยายแทนเสียงทีละตัวอักษรที่จำเป็นต้องใช้แบนด์วิดท์เกินกว่าที่กำหนด คือ 500 มิลลิวินาที และใช้ได้สำหรับการแสดงผลแบบ SD เท่านั้น ส่วนการแสดงผลแบบ HD จำเป็นต้องใช้ค่าความถี่ที่น้อยกว่านี้ การศึกษาจึงต้องปรับใช้เป็นการแสดงข้อความปรากฏบนจอโทรทัศน์ทีละกลุ่มตัวอักษร ส่งผลให้ค่าเวลาหน่วงของการแสดงผลคำบรรยายแทนเสียงเมื่อเทียบกับเสียงพูดเพิ่มขึ้นจากการแสดงผลคำบรรยายแทนเสียงแบบทีละตัวอักษร โดยเฉลี่ยอยู่ที่ 6.1 วินาที ดังนั้น จึงควรเพิ่มอัตราบิตให้กับบริการคำบรรยายแทนเสียงประเภท SD 625 เป็น 200 กิโลบิตต่อวินาที

ด้านรูปแบบการแสดงผลคำบรรยายแทนเสียง เนื่องจากข้อจำกัดด้านแบนด์วิดท์ข้างต้น จึงต้องแสดงผลคำบรรยายแทนเสียงบนจอโทรทัศน์ทีละกลุ่มตัวอักษร ซึ่งอาจแสดงผลไม่ครบเต็มคำ ส่งผลให้ขาดความต่อเนื่องในการอ่านข้อความ การแก้ปัญหาจึงอาศัยโปรแกรมตัดคำเข้ามาช่วย อย่างไรก็ตาม โปรแกรมตัดคำภาษาไทยที่มีอยู่ในปัจจุบันยังไม่สามารถรองรับข้อความที่ไม่สมบูรณ์จากการพิมพ์ข้อความแบบต่อเนื่อง และอาจมีการแก้ไขข้อความหลังจากที่ส่งข้อมูลไปยังโปรแกรมตัดคำในตอนต้น จึงต้องพัฒนาโปรแกรมตัดคำภาษาไทยที่สามารถรองรับข้อความจากระบบถอดความเสียงพูดแบบทันทีขึ้นมาใหม่เช่นกัน

บทสรุปและข้อเสนอแนะเชิงนโยบาย

1. สำนักงาน กสทช. ควรพิจารณาปรับนโยบายแนวทางการบริหารจัดการอัตราบิตที่เหลืออยู่ให้สามารถนำมาใช้ประโยชน์ได้อย่างคุ้มค่ามากขึ้นกับบริการที่เกิดขึ้นในทีวีดิจิทัล โดยจัดสรรอัตราบิตให้ถึง 200 กิโลบิตต่อวินาที ทั้งนี้ ปัจจุบันการให้บริการโครงข่ายโทรทัศน์ภาคพื้นดินในระบบดิจิทัลของบริษัท อสมท จำกัด (มหาชน) กรมประชาสัมพันธ์ และ องค์การกระจายเสียงและแพร่ภาพสาธารณะแห่งประเทศไทย (ส.ส.ท.) ยังคงมีแบนด์วิดท์เหลืออยู่ ซึ่งสามารถนำอัตราบิตที่เหลือจากการให้บริการโครงข่ายโทรทัศน์ฯ มาใช้ประโยชน์ได้
2. ควรมีการศึกษาวิจัยเพื่อทดลองดำเนินการจริงของสถานีโทรทัศน์ โดยการส่งสัญญาณแพร่ภาพพร้อมกันหลายสถานี และการทำงานของหน่วยบริการโครงข่ายเพิ่มเติม
3. ควรมีศูนย์บริการการจัดทำคำบรรยายแทนเสียงที่เป็นหน่วยงานกลาง ทำหน้าที่สนับสนุนการจัดบริการคำบรรยายแทนเสียงในช่วงเวลาเดียวกันฯ ให้กับสถานีโทรทัศน์ที่มีความต้องการ เพื่อเป็นการบูรณาการใช้ทรัพยากรร่วมกันหลายหน่วยงาน โดยที่แต่ละหน่วยงานไม่ต้องแยกจัดทำบริการเอง

4. ควรร่วมกับสถาบันการศึกษา สมาคม หรือกลุ่มของคนพิการทางการได้ยินซึ่งมีสมาชิกเป็นคุณหนูหรือหญิงสาว เผยแพร่ความรู้ความเข้าใจเรื่องการใช้งานบริการคำบรรยายแทนเสียงในระยะเวลาเดียวกันๆ แบบปิดในระบบโทรทัศน์ดิจิทัล เพื่อส่งเสริมการใช้งานระบบบริการคำบรรยายแทนเสียงของกลุ่มคนดังกล่าว

กิตติกรรมประกาศ

ขอขอบคุณ คุณวีระ ชิตชอบ สำหรับข้อมูลทางเทคนิคของอุปกรณ์ VT3 รุ่น Peach และซอฟต์แวร์ PIMmala Subtitle Gateway และคำแนะนำทางเทคนิคในการทดสอบต่อเชื่อมอุปกรณ์ทั้งสองในการทดสอบในงานวิจัยนี้ และขอขอบคุณ คุณเดชาวัต จิงจุฑพรชัย สำหรับคำแนะนำวิธีลดความเสี่ยงพูดของเจ้าหน้าที่บริการลดความเสี่ยงพูดในระยะเวลาเดียวกันๆ

108