บทที่ 1

บทนำ

ในการเข้ารหัสเสียงพูดที่มีคุณภาพของเสียงที่สูงนั้นเดิมจำเป็นที่จะต้องใช้อัตราการส่งสูง เช่น G.711 64 kbps log PCM หรือ G.721 32 kbps ADPCM อันเป็นมาตรฐานนานาชาติที่ตั้งขึ้นโดย International Telecommunication Union หรือ ITU-T (เดิมคือ Consultative Committee of International Telegraph and Telephone หรือ CCITT) ซึ่งเป็นที่ยอมรับและมีการใช้งานกันอย่างกว้างขวางในปัจจุบัน ความต้องการการเข้ารหัสเสียงที่ให้คุณภาพสูงโดยมีอัตราข้อมูลต่ำกว่านี้ทำให้มีการพัฒนาวิธีการใหม่ๆออกมามากมาย เช่น Code-Excited Linear Prediction (CELP), Multipulse Linear Predictive Coding (MPLPC), Adaptive Predictive Coding (APC), Adaptive Transform Coding (ATC) เป็นต้นสามารถให้คุณภาพเสียงที่สูงในอัตรา 16 กิโลบิตต่อวินาที แต่เทคนิกต่าง ๆเหล่านี้ทำให้เกิดช่วงเวลาประวิงค่อนข้างสูงประมาณ 40-60 มิลลิวินาที เนื่องมาจากการเก็บข้อมูลเสียงจำนวนมากไว้เพื่อใช้ในการคำนวณความซ้ำซ้อน(redundancy) ในปี 1988 ITU-Tต้องการกำหนดมาตรฐานของการเข้ารหัสเสียงตัวใหม่ที่อัตรา 16 กิโลบิตต่อวินาที โดยที่มีช่วงเวลาประวิงต่ำกว่า 2 มิลลิวินาที ซึ่งในที่สุดได้เลือก Low-Delay CELP (LD-CELP) ที่นำเสนอโดย Juin-Hwey Chen [1] ในปี 1992 เป็นมาตรฐานของการเข้ารหัสเสียง ITU-T G.728 16kbps LD-CELP [2] มีช่วงเวลาประวิงทางเดียวเพียง 0.625 มิลลิวินาทีและได้รับการยอมรับว่าคุณภาพของเสียงที่ได้ไม่ต่ำกว่าของ ADPCM ถึงแม้ว่าจะมีการเข้ารหัสแบบไม่เข้าจังหวะอย่างต่อเนื่องถึง 3 ครั้งก็ตาม อันเป็นผลที่ได้มาจากการนำโพสต์ฟิลเตอร์(Postfilter)[3]มาช่วยเพิ่มคุณภาพของเสียงที่ด้านถอดรหัส

LD-CELP มีปริมาณการคำนวณที่สูงมาก การสร้างเครื่องเข้ารหัสจริงโดยใช้ตัวประมวลผลสัญญาณดิจิตอลแบบจุดลอย (floating-point) รุ่น DSP32C ของ AT&T ใช้การทำงานถึง 19 ล้านคำสั่งใน 1 วินาทีสำหรับทั้งการเข้ารหัสและถอดรหัส ในการกำหนดมาตรฐานระบุไว้ว่าสามารถสร้างเครื่องเข้ารหัสจริงได้โดยใช้ตัวประมวลผลสัญญาณดิจิตอลแบบจุดตรึง (fixed-point) อันจะเป็นการประหยัดค่าใช้จ่ายได้มากกว่า [4] การสร้างเครื่องเข้ารหัสในวิทยานิพนธ์นี้ได้เลือกใช้ TMS320C50 ซึ่งเป็นตัวประมวลผลสัญญาณดิจิตอลแบบจุดตรึง 16 บิตของบริษัทเทกซัสอินสตรูเมนต์ (Texas Instrument หรือ TI)ที่มีการใช้งานกันอย่างกว้างขวาง โดยประกอบมาเป็นชุดพัฒนาโปรแกรมสำเร็จรูปที่เรียกว่า Digital signal processing Starter Kit (DSK) รุ่นที่ใช้ในงานวิจัยนี้สามารถทำงานได้ 20 ล้านคำสั่งใน 1 วินาที (โดยที่รุ่นสูงสุดในปัจจุบันสามารถทำงานได้ถึง 40 ล้านคำสั่งใน 1 วินาที) และมีหน่วยความจำภายในเป็นแรมขนาด 9 กิโลเวอร์ดสำหรับโปรแกรมและ 1 กิโลเวอร์ดสำหรับข้อมูล เหตุผลที่เลือกใช้ตัวประมวลผลตัวนี้เพราะมีความเหมาะสมทั้งทางด้านความสามารถในการคำนวณและราคาประหยัดกว่าแบบจุดลอยมาก ทั้งยังมีการนำไปใช้งานกันอย่างกว้างขวางเป็นที่ยอมรับกันโดยทั่วไป งานวิจัยนี้จะเน้นด้านการพัฒนาโปรแกรมที่ถูกต้องและรวดเร็วตามมาตรฐาน โดยจะยังไม่จัดการในเรื่องการรับส่งรหัสที่ได้กับระบบภายนอก ซึ่งเป็นเรื่องของระบบสื่อสารหรือวงจรเชื่อมต่อที่ไม่เกี่ยวข้องกับอัลกอริธึมที่ใช้งาน ระบบจะเข้ารหัสสัญญาณเสียงแล้วถอดรหัสกลับมาเป็นสัญญาณเสียงออกมาทันทีโดยไม่มีการเก็บรหัสเสียงไว้หรือส่งออกมา

ในการสร้างจริงหรือการการอิมพลีเมนต์เครื่องเข้ารหัสแบบ LD-CELP บน TMS320C50 นี้มีหน่วยความจำจำกัดจึงต้องเขียนโปรแกรมให้สั้นที่สุดเพื่อให้สามารถบรรจุลงในหน่วยความจำภายในได้ทั้งหมด โปรแกรมต้องมีความเร็วพอที่จะทำการคำนวณค่าต่างๆของสัญญาณตัวปัจจุบันทั้งหมดเสร็จสิ้นก่อนที่สัญญาณตัวใหม่จะเข้ามาจึงจะได้การทำงานตามเวลาจริง และที่สำคัญการอิมพลีเมนต์แบบจุดตรึง 16 บิต เมื่อเปรียบเทียบกับแบบจุดลอย 32 บิตหรือจำนวนจริง(real number)จะ

มีข้อเสียเปรียบในเรื่องความละเอียดของข้อมูลที่น้อยกว่า(ปัญหาเรื่อง finite word length)จะต้องมีการปรับค่าความละเอียดของข้อมูลอยู่เสมอเพื่อป้องกันข้อมูลขนาดเกินพิกัด(overflow)ข้อมูลขนาดเล็กเกินไป(underflow)และเพื่อให้ความละเอียดของข้อมูลใกล้เคียงกับแบบจุดลอยมากที่สุด ในการอิมพลีเมนต์แบบจุดตรึงของโปรแกรมตัวเดียวกันการอิมพลีเมนต์ที่มีการปรับความละเอียดของข้อมูลที่ใกล้เคียงกับจุดลอยมากกว่าจะให้คุณภาพของการอิมพลีเมนต์ที่ดีกว่า แต่อย่างไรก็ตามเนื่องจากข้อจำกัดด้านขนาดของข้อมูลของจุดตรึงผลที่ได้ในที่สุดก็คงมีความละเอียดน้อยกว่าจุดลอยอยู่เสมอ

วัตถุประสงค์

เพื่อศึกษาหลักการทำงานและพัฒนาเครื่องเข้ารหัสและถอดรหัสเสียงแบบ LD-CELP สำหรับการทำงานตามเวลาจริงที่อัตรา 16 กิโลบิตต่อวินาทีโดยอิงมาตรฐาน ITU-T G.728 บน TMS320C50 ซึ่งเป็นตัวประมวลผลสัญญาณดิจิตอลแบบจุดตรึงที่มีหน่วยความจำภายในสำหรับโปรแกรม 9 กิโลเวอร์ดและสำหรับข้อมูล 1 กิโลเวอร์ดสามารถคำนวณได้ 20 ล้านคำสั่งใน 1 วินาที

ขอบเขตของวิทยานิพนธ์

1. จำลองการทำงานของ LD-CELP บนคอมพิวเตอร์ส่วนบุคคล โดยใช้โปรแกรมช่วยในการคำนวณ MATLAB เพื่อทำความเข้าใจในหลักการทำงานของการเข้ารหัสและเป็นโปรแกรมอ้างอิงที่ใช้ช่วยในการทดสอบโปรแกรมเข้ารหัสตามเวลาจริง

2. เขียนโปรแกรมด้วยภาษาแอสเซมบลีของ TMS320C50 เพื่อเข้ารหัสเสียงพูดแถบความถี่ไม่เกิน 4 กิโลเฮิรตซ์พร้อมกับถอดรหัสเสียงออกมาทันที ในขั้นนี้เครื่องจะสามารถใช้เข้ารหัสและถอดรหัสเสียงอย่างต่อเนื่องตามเวลาจริงโดยไม่มีการเก็บรหัสเสียงไว้หรือส่งออก

3. ทำการประเมินคุณภาพของการการอิมพลีเมนต์แบบจุดตรึงนี้ โดยเปรียบเทียบขนาดของสัญญาณเสียงที่ได้ว่าใกล้เคียงเพียงใดกับเสียงที่ได้จากโปรแกรม MATLAB ซึ่งเป็นแบบจุดลอย

ประโยชน์ที่ได้รับ

1. ได้โปรแกรมเข้ารหัสและถอดรหัสเสียงพูดแบบ LD-CELP ที่สามารถทำงานได้ตามเวลาจริงบน TMS320C50

2. เมื่อนำโปรแกรมไปใช้บนฮาร์ดแวร์ของ TMS320C50 ที่เหมาะสมที่สามารถรับส่งข้อมูลกับภายนอกได้เร็วพอ เช่น Evaluation Module (EVM) จะสามารถเก็บรหัสเสียงหรือเสียงที่ถูกถอดรหัสแล้วไว้ในหน่วยความจำสำรองได้ เช่น ฮาร์ดดิสค์ หรือส่งผ่านระบบสื่อสารไปยังปลายทางได้ตามเวลาจริง

3. การเข้ารหัสเสียงสามารถนำไปพัฒนางานด้านอื่น ๆได้ เช่นเครื่องป้องกันการดักฟังโทรศัพท์โดยใช้ร่วมกับโมเด็ม (Modem) การรู้จำเสียงพูดจากรหัสเสียงโดยตรง เป็นต้น

4. สามารถทดสอบคุณภาพของเสียงสัญญาณโทน (tone) ต่าง ๆเช่นโมเด็ม, facsimile (FAX), Dual Tone Multi Frequency (DTMF) ว่าเมื่อผ่านการเข้ารหัสและถอดรหัสออกมาแล้วนำไปใช้ได้เหมือนเดิมหรือไม่

5. สามารถทดสอบคุณภาพของเสียงเมื่อมีการเข้ารหัสแบบไม่เข้าจังหวะอย่างต่อเนื่องหลายครั้งได้ตามเวลาจริง

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย