รายการอ้างอิง

1. ITU-T Recommendation G.729. Coding of Speech at 8 kb/s Using Conjugate Structure Algebraic Code Excited Linear Prediction (CS-ACELP). 1996.
2. Kataoka,A., Moriya, T., and Hayashi, S. An 8 kb/s Conjugate Structure CELP (CS-CELP) Speech Code. IEEE Transaction on Signal Processing, Vol 2 (October 1998): 401-411.
3. Juan,L. , Bigin,L. , and Qiuliang F. An 8 kb/s Conjugate-Structure Algebraic CELP (CS-CELP) Speech coding. IEEE Transaction on Signal Processing, Vol 2. (October 1998) : 1729-1732.
4. S.Chompun , S.Jitapunkul , D.Tancharoen and T.Srithanason. Thai Speech Compression Using CS-ACELP Coder Based on ITU G.729 Standard. Faculty of Engineering , Chulalongkorn University , Bangkok Thailand., The Forth Symposium on Natural Language Processing 2000, SNL'2000 , Chiangmai ,Thailand.
5. พูนลาภ ลามศรีจันทร์. การพัฒนาการเข้ารหัสเสียงพูดแบบ LD-CELP ที่อัตราการเข้ารหัส 16 กิโลบิตต่อวินาที สำหรับการทำงานตามเวลาจริงโดยใช้ TMS320C50. วิทยานิพนธ์ปริญญามหาบัณฑิต สาขาวิศวกรรมไฟฟ้า จุฬาลงกรณ์มหาวิทยาลัย , 2539.
6. สุภัทรชัย ชมพันธุ์. การบีบอัดเสียงพูดภาษาไทยโดยใช้การเข้ารหัส MP-CELP ตามข้อกำหนดของ MPEG-4. วิทยานิพนธ์ปริญญามหาบัณฑิต สาขาวิศวกรรมไฟฟ้า จุฬาลงกรณ์มหาวิทยาลัย , 2543.
7. สิริ วงศ์วรชาติกาล. การรู้จำเสียงพูดไทยโดยตรงจากการเข้ารหัส G.729. วิทยานิพนธ์ปริญญามหาบัณฑิต สาขาวิศวกรรมไฟฟ้า จุฬาลงกรณ์มหาวิทยาลัย , 2543.
8. วรพจน์ พัฒนวิจิตร. การออกแบบระบบประมวลผลสัญญาณเสียงพูดดิจิตอลแบบเวลาจริงใช้ชิป TMS320C25 . วิทยานิพนธ์ปริญญามหาบัณฑิต สาขาวิศวกรรมไฟฟ้า จุฬาลงกรณ์มหาวิทยาลัย , 2539.
9. http://dspvillage.ti.com/docs/dspvillagehome.jhtml

ภาคผนวก

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

บทความทางวิชาการของผู้วิจัยที่ได้รับการตีพิมพ์

บทความทางวิชาการของผู้วิจัยที่ได้รับการตีพิมพ์ในงาน การประชุมวิชาการทางวิศวกรรมไฟฟ้าครั้งที่ 26 (EE-CON 26) ในหัวข้อ "เครื่องบันทึกเสียงพูดดิจิตอลบนพื้นฐาน TMS320C6000 โดยใช้การเข้ารหัสแบบ G.729"

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

เครื่องบันทึกเสียงพูดดิจิตอลบนพื้นฐาน TMS320C6000 โดยใช้การเข้ารหัสแบบ G.729

TMS320C6000 Based Digital Speech Recorder Using G.729 Coding

วิเชียร อูปแก้ว กฤษดา วิศวธีรานนท์ และ สุวิทย์ นาคพีระยุทธ
ภาควิชาวิศวกรรมไฟฟ้า คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ถนนพญาไท เขตปทุมวัน กรุงเทพฯ 10330
โทร 0-2218-6515 โทรสาร 0-2218-8991 Email : wichian101@chula.com

บทคัดย่อ

บทความนี้นำเสนอการสร้างเครื่องบันทึกเสียงพูดดิจิตอลที่สามารถบันทึกและเล่นเสียงพูดได้แบบเวลาจริง และใช้การเข้ารหัสและถอดรหัสแบบ CS-ACELP ตามมาตรฐาน ITU-T G.729 ที่มีอัตราการเข้ารหัสต่ำคือ 8 kbits/s (เมื่อเทียบกับการเข้ารหัสของเครื่องบันทึกเสียงพูดที่มีอยู่ในปัจจุบันที่เป็นแบบ ADPCM ซึ่งมีอัตราการเข้ารหัสที่สูงคือ 32 kbits/s) จึงทำให้ใช้หน่วยความจำในการเก็บข้อมูลน้อยกว่า โดยที่ยังคงรักษาคุณภาพของเสียงได้ใกล้เคียงกับแบบ ADPCM การเขียนโปรแกรมใช้ภาษาระดับสูงจึงทำให้สะดวกและง่ายต่อการพัฒนา และเลือกตัวประมวลผลสัญญาณดิจิตอลเป็นแบบจุดลอยที่มีความเร็วสูง จึงทำให้ผลของการคำนวณมีความถูกต้องสูงและสามารถทำงานแบบเวลาจริงได้ทัน ผลของการเข้ารหัสและถอดรหัสด้วยตัวประมวลผลสัญญาณดิจิตอลจะถูกนำไปเปรียบเทียบคุณภาพของเสียงกับผลการเข้ารหัสและถอดรหัสจากมาตรฐานของ ITU-T G.729 ที่ประมวลผลบนคอมพิวเตอร์ส่วนบุคคล จากการทดสอบบันทึกและเล่นเสียงพูดสามารถทำการบันทึกและฟังเสียงที่บันทึกได้ในระดับที่ดี

คำสำคัญ : Speech Coding , CS-ACELP , G.729 , TMS320C6000

Abstract

This article presents an implementation of digital speech recorder which can record and playback on real-time. According to ITU-T G.729 recommendation, CS-ACELP is used as an algorithm of speech coding. While ADPCM algorithm provides coding of speech at 32 kbits/s , CS-ACELP algorithm does at low bitrate, 8 kbits/s. This leads to less memory usage with nearly equivalent quality. With high level language programming, users can conveniently develop the program. High speed and floating point digital signal processor offers real-time processing with high accurate computation. The speech codec from experiments is compared in quality with standard one generated by PC using ITU-T G.729 recommendation. This digital speech recorder gives satisfied subjective result in recording and playing back.

1. คำนำ

ที่ผ่านมาการบันทึกเสียงพูดใช้เทปแม่เหล็กในการบันทึกเสียง ซึ่งทำให้เกิดข้อเสียหลายอย่าง เช่น อายุการใช้งานสั้น การเล่นหรือค้นหาข้อมูลทำได้ไม่สะดวกเท่าที่ควร มีความเสี่ยงต่อการเสียหายของข้อมูล ฯลฯ ต่อมาได้มีการพัฒนาโดยการนำอุปกรณ์อิเล็กทรอนิกส์ที่เรียกว่าชิปหน่วยความจำ มาใช้ในการเก็บข้อมูล แต่ก็ยังมีข้อเสียคือ การเข้ารหัสก่อนที่จะทำการบันทึกมีอัตราการเข้ารหัสที่สูง จึงทำให้ต้องใช้พื้นที่หน่วยความจำในการเก็บข้อมูลมากและทำให้บันทึกข้อมูลได้น้อย

บทความนี้จึงนำเสนอเครื่องบันทึกเสียงพูดดิจิตอลที่ใช้การเข้ารหัสและถอดรหัสแบบ CS-ACELP ตามมาตรฐาน ITU-T G.729 [1-5] ที่มีอัตราการเข้ารหัสต่ำคือ 8 kbits/s เมื่อเทียบกับเครื่องบันทึกเสียงพูดที่ใช้อยู่ในปัจจุบันที่มีการเข้ารหัสแบบ ADPCM มีอัตราการเข้ารหัสที่สูงคือ 32 kbits/s ทำให้ต้องใช้พื้นที่หน่วยความจำมากในการเก็บบันทึกข้อมูลเสียง ดังนั้นถ้ามีการนำเอาการเข้ารหัสแบบ CS-ACELP ตามมาตรฐานของ ITU-T G.729 มาประยุกต์ใช้จะทำให้ใช้หน่วยความจำน้อยกว่าและสามารถบันทึกข้อมูลได้มากขึ้น โดยที่ยังคงรักษาคุณภาพของเสียงไว้ได้ใกล้เคียงกับการเข้ารหัสแบบ ADPCM ในการพัฒนาเครื่องบันทึกเสียงพูดดิจิตอลจะใช้ตัวประมวลผลสัญญาณดิจิตอลตระกูล TMS320C6000 [6] ในการประมวลผลการเข้ารหัสและถอดรหัส เพราะมีความเร็วในการทำงานและยังสามารถคำนวณแบบจุดลอย(Floating point) ซึ่งให้ผลของการคำนวณถูกต้องสูงเมื่อเทียบกับการคำนวณแบบจุดตรึง(Fixed point) [7] การพัฒนาใช้โปรแกรมภาษาซีซึ่งเป็นภาษาระดับสูงจึงทำให้ง่ายและสะดวกในการพัฒนาและยังสามารถนำโค้ดไปพัฒนาใช้กับตัวประมวลผลในรุ่นอื่นหรือของบริษัทอื่นได้ง่าย เพียงแต่มีการปรับปรุงโปรแกรมในส่วนของฮาร์ดแวร์ของตัวประมวลผลรุ่นนั้นๆ ซึ่งทำให้ง่ายและสะดวกกว่าภาษาแอสเซมบลี

2. การเข้ารหัสเสียงพูดแบบ CS-ACELP ตามมาตรฐาน ITU-T G.729

การเข้ารหัสเสียงพูดโดยวิธี CS-ACELP เป็นมาตรฐานของ ITU-T ที่ถูกออกแบบมาเพื่อประมวลผลสัญญาณดิจิตอลที่ได้จากการกรองสัญญาณในช่วงแบนด์วิดท์ของระบบโทรศัพท์ และมีสัญญาณอินพุตเป็นสัญญาณแอนาลอกที่อัตราการสุ่ม 8000 Hz และเปลี่ยนเป็นรหัส PCM แบบเชิงเส้นที่มีความละเอียด 16 บิต สำหรับเป็นอินพุตให้กับตัวเข้ารหัส เพื่อคำนวณหาพารามิเตอร์ของแต่ละเฟรม ดังรายละเอียดตามตารางที่ 1

ตารางที่ 1 การจัดสรรบิตของการเข้ารหัสเสียงพูดโดยวิธี CS-ACELP

พารามิเตอร์	คำย่อ	เฟรมย่อย 1	เฟรมย่อย 2	บิต : เฟรม
Line spectrum pairs	L0, L1, L2, L3			18
Adaptive-codebook delay	P1,P2	8	5	13
Pitch-delay parity	P0	1		1
Fixed-codebook index	C1, C2	13	13	26
Fixed-codebook sign	S1, S2	4	4	8
Codebook gains (stage 1)	GA1 ,GA2	3	3	6
Codebook gains (stage 2)	GB1,GB2	4	4	8
Total				80

2.1 ตัวเข้ารหัส

หลักการของการเข้ารหัสแสดงดังรูปที่ 1 สัญญาณอินพุตจะถูกกรองแบบความถี่สูงผ่านและถูกปรับลดขนาดโดยปรีโปรเซสซิ่ง(Pre-processing) โดยสัญญาณที่ได้นี้จะเป็นสัญญาณอินพุตให้กับทุกๆวงจรที่เหลือเพื่อใช้ในการวิเคราะห์ การวิเคราะห์ LP (Linear prediction) จะแบ่งสัญญาณเสียงเป็น เฟรมๆ ละ 10 มิลลิวินาที หรือ 80 ตัวอย่างสุ่ม เพื่อใช้ในการคำนวณหาสัมประสิทธิ์ของวงจรกรองสัญญาณ LP สัมประสิทธิ์เหล่านี้จะถูกแปลงเป็นค่า LSP (Line spectrum pairs) และทำการควอนไทซ์โดยใช้วิธีการเวกเตอร์ ควอนไทซ์แบบ 2 ตอน ที่มีความละเอียด 18 บิต สัญญาณการกระตุ้น(Excitation) จะถูกคำนวณหาจากการใช้วิธีการวิเคราะห์โดยการสังเคราะห์ ในการค้นหาเพื่อให้ค่าความผิดพลาดระหว่างสัญญาณเสียงดั้งเดิมกับสัญญาณเสียงที่สังเคราะห์ขึ้นมาได้มีค่าน้อยที่สุดโดยใช้วิธีการวัดด้วย PWD (Perceptually weighted distortion) โดยใช้การกรองสัญญาณความผิดพลาดด้วยวงจรกรองสัญญาณแบบ เปอเซปทวลเวทติ้ง(perceptual weighting) ที่มีสัมประสิทธิ์ที่ได้มาจากสัมประสิทธิ์ของวงจรกรอง

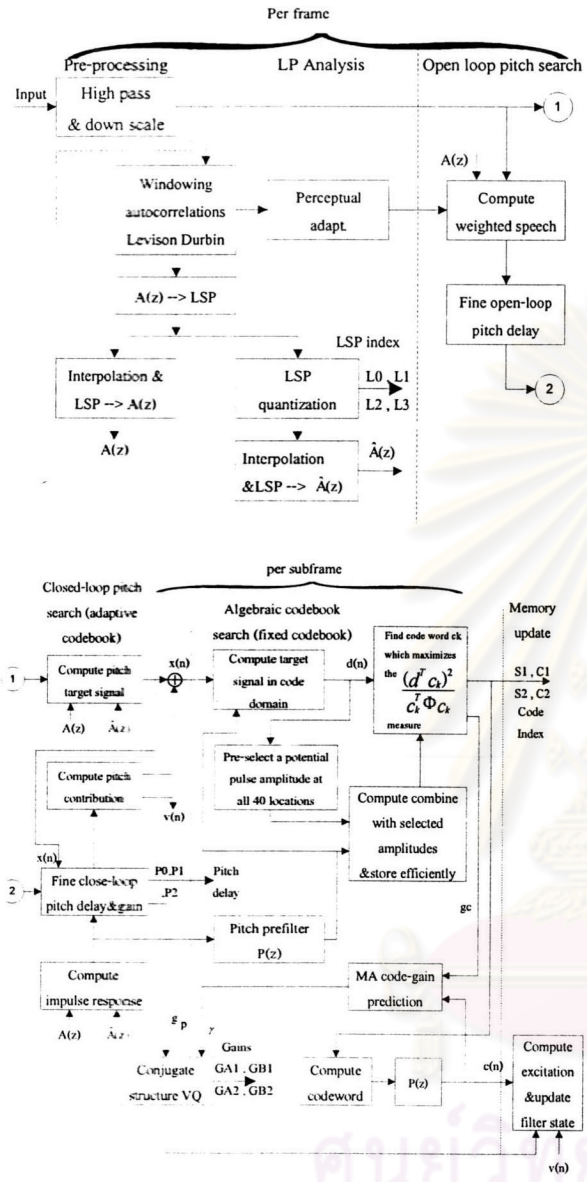
สัญญาณ LP ที่ยังไม่ควอนไทซ์ จำนวนของการทำเปอเซปทวลเวทติ้งนั้นสามารถปรับได้เพื่อให้สมรรถนะการทำงานที่ได้มีผลตอบสนองต่อสัญญาณอินพุตเท่ากันทุกๆ ความถี่ (Flat Frequency-response)

พารามิเตอร์ของการกระตุ้น(Fixed and adaptive-codebook parameter) จะถูกคำนวณหาทุกๆ เฟรมย่อย โดยแต่ละเฟรมย่อยนี้จะมีขนาด 5 มิลลิวินาที (40 ตัวอย่างสัญญาณที่อัตราการสุ่ม 8000 เฮิรตซ์) สัมประสิทธิ์ของวงจรกรองสัญญาณ LP ทั้งที่ควอนไทซ์และไม่ควอนไทซ์จะถูกใช้ในเฟรมย่อยที่ 2 ส่วนในเฟรมย่อยที่ 1 นั้นจะใช้วิธีการหาสัมประสิทธิ์ของวงจรกรองสัญญาณ LP โดยการประมาณค่าในช่วง (ทั้งที่ควอนไทซ์และไม่ ควอนไทซ์)

สำหรับค่าพิตช์แบบวงรอบเปิด(Open loop pitch) จะถูกประมาณหนึ่งครั้งในแต่ละเฟรมจากสัญญาณที่ได้จากวงจรกรองสัญญาณเปอเซปทวลเวทติ้ง สำหรับกระบวนการที่จะกล่าวต่อไปนี้จะกระทำทุกๆ เฟรมย่อย สัญญาณเป้าหมาย x(n) จะถูกคำนวณโดยการกรอง สัญญาณตกค้าง (Residual) ของ LP ผ่านวงจรกรองสัญญาณสังเคราะห์ถ่วงน้ำหนัก (Weighted synthesis) W(z)/A(z) โดยเมื่อเริ่มทำงาน วงจรกรองสัญญาณเหล่านี้จะถูกปรับให้ทันกาลจากการกรองค่าความผิดพลาดระหว่างค่าตกค้างของ LP (LP residual) กับสัญญาณกระตุ้น ซึ่งนั่นก็หมายถึงการหักล้างด้วยวงจรกรองสัญญาณแบบสังเคราะห์ถ่วงน้ำหนักที่ผลตอบสนองอินพุตศูนย์ (Zero input response) ออกจากสัมประสิทธิ์ของสัญญาณเสียง การตอบสนองอิมพัลส์ h(n) ของวงจรกรองสัญญาณสังเคราะห์ถ่วงน้ำหนัก จะถูกคำนวณหา ส่วนการวิเคราะห์หาค่าพิตช์ในวงรอบปิด(Closed-loop pitch) (เพื่อหาค่าประวิงเวลาและอัตราการขยายของชุดรหัสที่ปรับเปลี่ยนได้) จะใช้สัญญาณเป้าหมาย x(n) และการตอบสนองอิมพัลส์ h(n) โดยการค้นหารอบค่าประวิงเวลาของพิตช์แบบวงรอบเปิดทั้งค่าที่เป็นแบบจำนวนเต็มและค่าเศษส่วนด้วยความละเอียด 1/3 ค่าประวิงเวลาของพิตช์(pitch delay) จะถูกเข้ารหัสด้วยความละเอียด 8 บิตในเฟรม ย่อยแรกส่วนที่เหลือจะเข้ารหัสด้วยความละเอียด 5 บิตในเฟรมย่อยที่สอง สัญญาณของเป้าหมาย x(n) จะถูกปรับค่าโดยการนำค่าจากชุดรหัสแบบปรับเปลี่ยนได้ มาหักล้างและค่าสัญญาณของเป้าหมายที่ได้ใหม่ x'(n) นี้จะถูกใช้ในการค้นหาชุดรหัสแบบคงที่เพื่อให้ได้เอ็กไซเทชันที่ถูกต้องที่สุดต่อไป

โดยค่าโครงสร้างของชุดรหัสแบบคงที่นี้จะเป็นแบบพีชคณิต(Algebraic) ที่มีความละเอียด 17 บิต อัตราการขยายของ ชุดรหัสแบบปรับเปลี่ยนได้และชุดรหัสแบบคงที่จะถูกควอนไทซ์ที่ความละเอียด 7 บิต(โดยจะใช้ตัวทำนายสัญญาณแบบ MA (Moving average)

กับอัตราขยายของชุดรหัสคงที่) สุดท้ายค่าต่างๆ ของวงจรกรองสัญญาณจะถูกปรับค่าตามสัญญาณกระตุ้นที่ได้

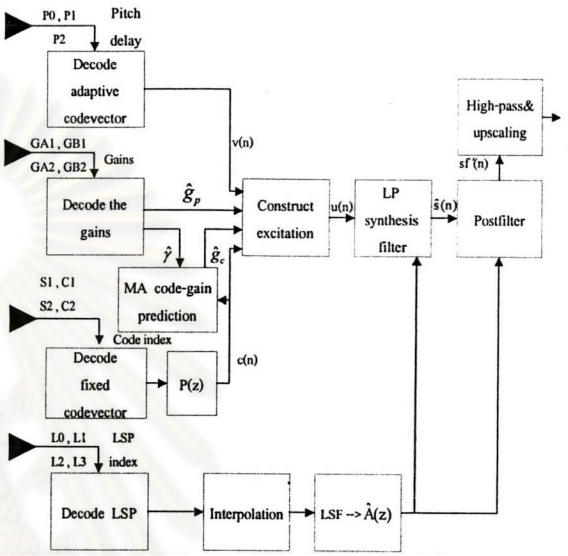


รูปที่ 1. บล็อกการทำงานของการเข้ารหัส

2.2 ตัวถอดรหัส

หลักการทำงานของตัวถอดรหัสแสดงดังรูปที่ 2 อันดับแรกพารามิเตอร์ที่รับมาได้ในรูปแบบของบิตจะถูกแยกแยะออกตามเฟรมๆ ละ 10 มิลลิวินาที และถอดรหัสออกมาเป็นค่าพารามิเตอร์ต่างๆ ได้แก่ค่าสัมประสิทธิ์ LSP ค่าประวิงเวลาของพิตช์ 2 ค่า เวกเตอร์ชุดรหัสแบบคงที่ 2 ชุด และอัตราการขยายของชุดรหัสแบบปรับเปลี่ยนได้และแบบคงที่ อย่างละ 2 ชุด สัมประสิทธิ์ LSP นั้นจะถูกดำเนินการประมาณค่าในช่วง แล้วเปลี่ยนเป็นสัมประสิทธิ์ของวงจรกรองสัญญาณ LP ของแต่ละเฟรมย่อย โดยแต่ละเฟรมย่อยที่มีขนาด 5 มิลลิวินาที (40 ตัวอย่าง) นั้นจะมีขั้นตอนการทำงานดังนี้

1. สัญญาณกระตุ้นจะถูกสร้างขึ้นโดยการรวม ชุดรหัสแบบปรับเปลี่ยนได้และชุดรหัสคงที่เข้าด้วยกันตามอัตราขยายของแต่ละตัว



รูปที่ 2. บล็อกการทำงานของตัวถอดรหัส

2. สัญญาณเสียงจะถูกสังเคราะห์ขึ้น โดยนำสัญญาณกระตุ้นที่ได้มากรองด้วยวงจรกรองสัญญาณสังเคราะห์เสียง LP
3. สัญญาณเสียงที่สังเคราะห์ได้จะนำไปผ่านวงจรโพสต์โพรเซสซิ่ง(Post-processing) ซึ่งประกอบด้วย วงจรกรองโพสฟิลเตอร์อะแดฟทีพ (Adaptive post-filter) ที่สร้างจากวงจรกรองสัญญาณแบบลองเทอม(Long-term) และแบบช็อตเทอม(Short-term) วงจรกรองสัญญาณความถี่สูงผ่าน และวงจรปรับขนาดสัญญาณ

3. เครื่องบันทึกเสียงพูดดิจิตอล

3.1 การพัฒนาโปรแกรม

การพัฒนาโปรแกรมการเข้ารหัสและถอดรหัสจะพัฒนาโดยใช้ภาษาซีเพราะสามารถเขียนและแก้ไขได้สะดวกกว่าการเขียนด้วยภาษาแอสเซมบลี การพัฒนาจะเขียนโปรแกรมบนตัวจำลองโปรแกรม(Simulator) ที่ชื่อ CCS (Code Composer Studio) ของบริษัท TI (Texas Instrument) เป็นโปรแกรมที่ใช้สำหรับตัวประมวล

ผลตระกูล TMS320CXXXX ที่ทำงานบนระบบปฏิบัติการวินโดว์เพื่อศึกษาการทำงานของตัวเข้ารหัสและถอดรหัสให้ถูกต้อง

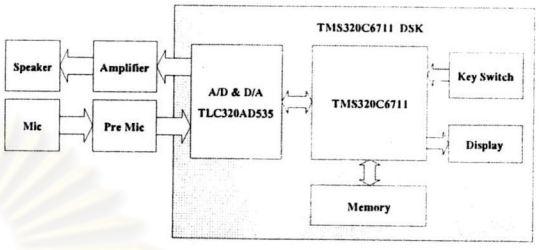
การเขียนโปรแกรมโดยใช้ตัวจำลองโปรแกรมนั้นจะแบ่งออกเป็นสองส่วนคือ ส่วนของการเข้ารหัสและการถอดรหัส เป็นการรับสัญญาณอินพุตที่อ่านจากไฟล์และสัญญาณเอาต์พุตที่ได้จะถูกเก็บไว้ในไฟล์เช่นเดียวกัน ซึ่งยังไม่เป็นแบบเวลาจริง แต่มีข้อดีคือสามารถที่จะทำการตรวจสอบผลการคำนวณและทำการแก้ไขได้สะดวกเพราะว่าการทดสอบสามารถที่จะสั่งให้ทำงานอย่างต่อเนื่องหรือสั่งให้ทำงานทีละคำสั่ง(Single step) ทำให้สามารถแก้ไขข้อผิดพลาดในโปรแกรมเป็นไปได้โดยง่ายเพราะทำให้เห็นการทำงานทุกขั้นตอน แต่โปรแกรมจำลองการทำงานมีข้อเสียคือทำงานได้ช้าถึงแม้จะใช้คอมพิวเตอร์ที่มีความเร็วสูงพอสมควรเท่าที่สามารถหาได้ในปัจจุบันก็ยังใช้เวลาในการจำลองโปรแกรมนานกว่าการทำงานแบบเวลาจริงเป็นอย่างมาก

หลังจากการพัฒนาโปรแกรมบนตัวจำลองโปรแกรมจนได้ผลที่ถูกต้องแล้วก็จะมีการปรับปรุงโปรแกรมในส่วนของการรับสัญญาณอินพุตและการส่งสัญญาณเอาต์พุต โดยใช้การอินเตอร์รัปต์รับข้อมูลและส่งข้อมูลผ่านทางพอร์ตอนุกรมเพื่อให้สามารถทำงานได้แบบเวลาจริง

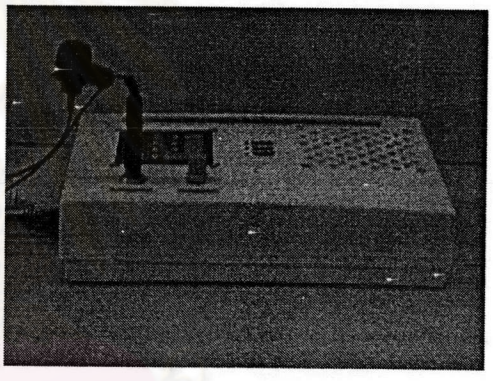
3.2 โครงสร้างทางฮาร์ดแวร์

ในการสร้างเครื่องบันทึกเสียงพูดดิจิตอลต้นแบบซึ่งมีบล็อกการทำงานดังรูปที่ 3 และรูปเครื่องบันทึกเสียงพูดดิจิตอลต้นแบบแสดงดังรูปที่ 4 การสร้างเลือกใช้ตัวประมวลผลสัญญาณดิจิตอลเบอร์ TMS320C6711 เป็นของบริษัท TI มีคุณสมบัติเป็นตัวประมวลผลสัญญาณแบบจุดลอย(Floating-point DSP) และมีโครงสร้างของการประมวลผลคำสั่งแบบ VLIW (Very long instruction word) รองรับสัญญาณนาฬิกาได้มากถึง 150 เมกกะเฮิรตซ์ สามารถประมวลผลคำสั่งแบบจุดลอยได้มากถึง 900 MFLOPS มีหน่วยประมวลผลทางคณิตศาสตร์และลอจิก (ALU) 8 หน่วย ซึ่งทำให้มีความเร็วในการประมวลผลแบบเวลาจริง หน่วยความจำขนาด 16 Mbytes ส่วนของการเชื่อมต่อกับสัญญาณแอนาลอก(Analog interface circuit : AIC) ใช้ชิปเบอร์ TLC320AD535 ทำหน้าที่แปลงสัญญาณแอนาลอกเป็นสัญญาณดิจิตอลและแปลงสัญญาณดิจิตอลเป็นสัญญาณแอนาลอก โดยติดต่อกับตัวประมวลผลผ่านทางพอร์ตอนุกรม มีอัตราการสุ่มสัญญาณ 8 KHz ประมวลผลสัญญาณขนาด 16 บิต สามารถโปรแกรมอัตราการขยายสัญญาณทางซอร์ฟแวร์ มีชุดจัดไบแอสและอัตราการขยายสำหรับไมโครโฟน พร้อมชุดขับลำโพง

ส่วนฟังก์ชันการทำงานนั้นประกอบด้วยโหมด 2 โหมดคือโหมดของการบันทึกและโหมดของการเล่นเสียง โดยควบคุมผ่านทางคีย์สวิตช์ ในส่วนของหน่วยความจำมีทั้งหมด 16 Mbytes สงวนไว้ใช้ในส่วนโปรแกรมการเข้ารหัสและถอดรหัสประมาณ 1 Mbytes ที่เหลือประมาณ 15 Mbytes สามารถใช้เก็บบันทึกเสียงได้นานประมาณ 4 ชั่วโมงโดยคิดที่อัตราการเข้ารหัสที่ 8 Kbits/s การบันทึกและเล่นแบ่งเป็น 4 ช่องคือ ช่องที่ 1 ช่องที่ 2 ช่องที่ 3 ช่องที่ 4 ช่องละ 1 ชั่วโมง ในการใช้งานการบันทึกเลือกสวิตช์โหมดบันทึกจากนั้นเลือกช่องที่จะทำการบันทึก และกดสวิตช์เริ่มต้น(Start) ถ้าต้องการหยุดการบันทึกให้กดปุ่มหยุด(Stop) และในการใช้งานในโหมดเล่นเสียงก็เช่นกันเริ่มจากการเลือกโหมดการเล่นเสียงจากนั้นเลือกช่องในการเล่น และกดสวิตช์เริ่มต้น ถ้าต้องการหยุดการบันทึกให้กดปุ่มหยุด



รูปที่ 3 บล็อกการทำงานของเครื่องบันทึกเสียงพูดดิจิตอล

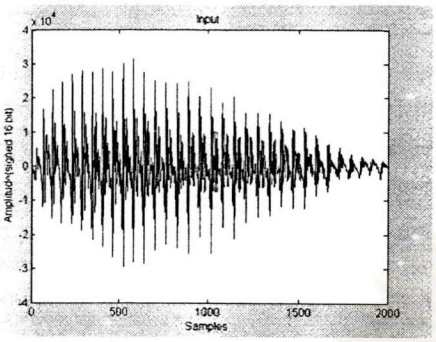


รูปที่ 4 เครื่องบันทึกเสียงพูดดิจิตอล

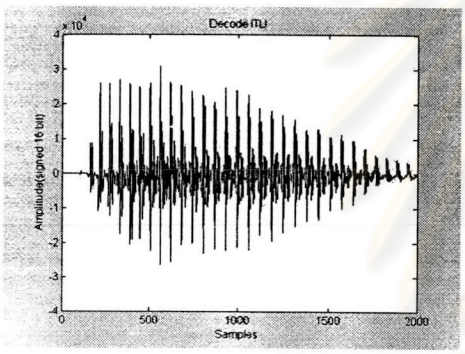
4. การทดสอบและผลการทดสอบ

ในขั้นตอนแรกจะทดสอบคุณภาพของการเข้ารหัสด้วยดีเอสพีที่คำนวณแบบจุดลอย (Floating-point) โดยใช้ค่าอัตราส่วนกำลังของสัญญาณต่อกำลังของสัญญาณรบกวน (Power Signal-to-Noise Ratio:PSNR) เพื่อเปรียบเทียบผลกับค่าที่ได้จากมาตรฐาน ITU ที่คำนวณแบบจุดตรึง(Fixed-point) ด้วยคอมพิวเตอร์ส่วนบุคคล ในการทดสอบใช้สัญญาณเสียงพูดเสียงเดียวกัน โดยใช้เสียงพูด "หนึ่ง" ถึง "สิบ" เป็นสัญญาณอินพุตที่ใช้กับการเข้ารหัสและถอดรหัสของมาตรฐาน ITU และเป็นอินพุตให้กับการเข้ารหัสและถอดรหัสในการทดลองโดยใช้ดีเอสพี ตัวอย่างของสัญญาณอินพุต สัญญาณที่ผ่านการถอดรหัสของมาตรฐาน ITU และสัญญาณที่ผ่านการถอดรหัสโดยใช้ดีเอสพีของคำว่า "สอง" แสดงดังรูปที่ 5 และคำนวณหาค่า PSNR ของ

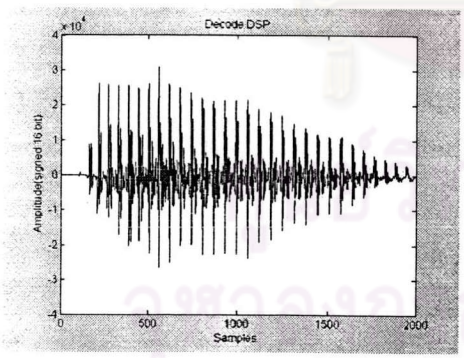
แต่ละคำพูด ผลของการคำนวณแสดงดังตารางที่ 2 ผลจากการทดลองจะเห็นว่าค่า PSNR ที่ได้จากการทดลองมีค่าที่ใกล้เคียงกับค่าที่คำนวณจากมาตรฐานของ ITU และบางค่ามีค่าที่มากกว่าค่า PSNR ที่ได้จากการเข้ารหัสจากมาตรฐาน ITU



(ก) สัญญาณเสียงคำว่า " สอง "



(ข) สัญญาณการถอดรหัสจากมาตรฐาน ITU



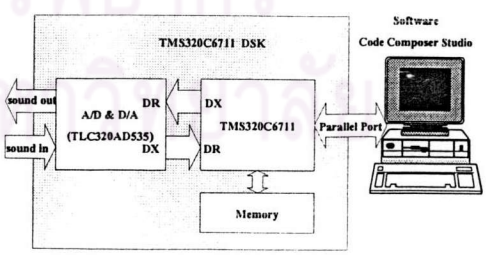
(ค) สัญญาณการถอดรหัสจากดีเอสพี

รูปที่ 5 สัญญาณเสียงคำว่า " สอง"

ขั้นตอนที่สองจะทดสอบการทำงานแบบเวลาจริงโดยใช้การควบคุมการทำงานด้วยโปรแกรม CCS แสดงบล็อกการทำงานในรูปที่ 6 การทดสอบแบบเวลาจริงนั้นจะทำการรับสัญญาณเสียงอินพุตมาทำการเข้ารหัสและส่งสัญญาณเสียงที่ได้จากการถอดรหัสผ่านทางวงจรการเชื่อมต่อสัญญาณแอนะลอกโดยใช้การอินเตอร์รัปต์ การประมวลผลจะประมวลผลทีละ 80 ตัวอย่างสุ่ม การทำงานจะทำการประมวลผลทั้งการเข้ารหัสและการถอดรหัสพร้อมกันๆ โดยกระบวนการประมวลผลทั้งการเข้ารหัสและการถอดรหัสจะต้องเสร็จสิ้นภายในเวลา 10 มิลลิวินาทีหรือ 80 ตัวอย่างสุ่ม โดยก่อนที่สัญญาณอินพุตของเฟรมถัดไปจะเข้ามา จากการทดสอบโดยการป้อนสัญญาณอินพุตเสียงพูดและให้โปรแกรมทำการนับจำนวนครั้งของการอินเตอร์รัปต์ที่เกิดขึ้นภายในการคำนวณเสร็จสิ้น 1 เฟรมปรากฏว่าค่าจำนวนนับที่เกิดขึ้นคือ 68 ครั้ง แสดงให้เห็นว่าสามารถประมวลผลการเข้ารหัสและถอดรหัสได้แบบเวลาจริง จากนั้นทำการป้อนสัญญาณเสียงพูดและฟังเสียงที่ได้จากการเข้ารหัสและถอดรหัสนั้นสามารถรับฟังเสียงได้ถูกต้อง

ตารางที่ 2 ผลการทดลองคำนวณค่า PSNR ของการเข้ารหัสเสียง

ตัวอย่างที่	เสียงต้นฉบับ	ค่า PSNR ของ ITU (dB)	ค่า PSNR ที่เข้ารหัสด้วยดีเอสพี (dB)
1	หนึ่ง	1.574	1.708
2	สอง	6.181	6.008
3	สาม	5.188	5.241
4	สี่	2.070	1.822
5	ห้า	4.693	4.467
6	หก	5.079	4.764
7	เจ็ด	5.391	4.897
8	แปด	4.897	4.569
9	เก้า	6.408	6.107
10	สิบ	2.068	2.183



รูปที่ 6. บล็อกการทำงานที่ควบคุมโดยโปรแกรม CCS

ขั้นตอนที่สามจะเป็นการทดสอบการทำงานในลักษณะของเครื่องบันทึกเสียงมีบล็อกการทำงานดังรูปที่ 3 โดยควบคุมการทำงานผ่านทางคีย์สวิตช์โดยทำการบันทึกทีละช่องให้ครบทั้ง 4 ช่องจากนั้นทำการเล่นเสียงที่ได้ทำการบันทึกไว้ของแต่ละช่องปรากฏว่าสามารถบันทึกและเล่นเสียงได้ตามเวลาที่กำหนดไว้ทุกช่องโดยคุณภาพของเสียงอยู่ในเกณฑ์ที่ดีดังแสดงรายละเอียดในตารางที่ 3

ตารางที่ 3 ผลการทดลองของการบันทึกเสียง

ช่อง/โหมด	ช่องที่ 1	ช่องที่ 2	ช่องที่ 3	ช่องที่ 4
บันทึก	ได้ครบ 1 ชม.	ได้ครบ 1ชม.	ได้ครบ 1ชม.	ได้ครบ 1ชม.
เล่นเสียง	ได้ชัดเจน	ได้ชัดเจน	ได้ชัดเจน	ได้ชัดเจน

5. สรุปผล

บทความนี้ได้นำเสนอการสร้างเครื่องบันทึกเสียงพูดดิจิตอลที่ใช้การเข้ารหัสตามมาตรฐาน ITU-T G.729 ที่อัตราการเข้ารหัสต่ำคือ 8 kbits/s ทำให้ใช้หน่วยความจำในการเก็บบันทึกน้อยกว่าเครื่องบันทึกเสียงในปัจจุบันที่มีการเข้ารหัสแบบ ADPCM ถึง 4 เท่า โดยที่ยังคงรักษาคุณภาพของเสียงไว้ได้ในระดับที่ใกล้เคียงกัน และจากการทดลองบันทึกและเล่นเสียงสามารถบันทึกและเล่นได้ตามเวลาที่กำหนดและให้คุณภาพเสียงในระดับที่ดี

6. กิตติกรรมประกาศ

ขอขอบพระคุณท่านอาจารย์ที่ปรึกษา อาจารย์ที่ปรึกษาร่วมที่ได้ให้คำแนะนำและสนับสนุนเครื่องมือและอุปกรณ์ในการทดลอง คุณสุภัทรชัย ชมพันธ์ นิสิตปริญญาเอก ที่ได้ให้คำแนะนำ

เอกสารอ้างอิง

[1] ITU-T Recommendation G.729, "Coding of Speech at 8 kb/s Using Conjugate Structure Algebraic Code Excited Linear Prediction (CS-ACELP)," March 1996

[2] Kataoka,A., Moriya, T., and Hayashi, S. "An 8 kb/s Conjugate Structure CELP (CS-CELP) Speech Code". IEEE Transaction on Signal Processing Vol 2 (October 1998): 401-411.

[3] Juan,L. , Bigin,L. , and Qiuliang F. "An 8 kb/s Conjugate-Structure Algebraic CELP (CS-CELP) Speech coding". IEEE Transaction on Signal Processing Vol 2. (October 1998) : 1729-1732.

[4] S.Chompun , S.Jitapunkul , D.Tancharoen and T.Srithanason. "Thai Speech Compression Using CS-ACELP Coder Based on ITU G.729", Standard. Faculty of Engineering , Chulalongkorn University Bangkok Thailand., The Forth Symposium on Natural Language Processing 2000, SNL'2000 , Chiangmai , Thailand.

[5] สุภัทรชัย ชมพันธุ์. "การบีบอัดเสียงพูดภาษาไทยโดยใช้การเข้ารหัส MP-CELP ตามข้อกำหนดของ MPEG-4". วิทยานิพนธ์ปริญญามหาบัณฑิต สาขาวิศวกรรมไฟฟ้า จุฬาลงกรณ์มหาวิทยาลัย , 2543.

[6] http://dspvillage.ti.com/docs/dspvillagehome.jhtml

[7] พูนลาภ ลามศรีจันทร์. " การพัฒนาการเข้ารหัสเสียงพูดแบบ LD-CELP ที่อัตราการเข้ารหัส 16 กิโลบิตต่อวินาที สำหรับการทำงานตามเวลาจริงโดยใช้ TMS320C50". วิทยานิพนธ์ปริญญามหาบัณฑิต สาขาวิศวกรรมไฟฟ้า จุฬาลงกรณ์มหาวิทยาลัย , 2539.

วิเชียร อุปแก้ว สำเร็จการศึกษาปริญญาตรีวิศวกรรมศาสตรบัณฑิต สาขาวิศวกรรมไฟฟ้า จากศูนย์กลางสถาบันเทคโนโลยีราชมงคล ปี พ.ศ. 2542 จากนั้นเข้ารับราชการในตำแหน่งอาจารย์ ประจำภาควิชาวิศวกรรมไฟฟ้า ศูนย์กลางสถาบันเทคโนโลยีราชมงคล ปัจจุบันกำลังศึกษาระดับปริญญาโท สาขาวิศวกรรมไฟฟ้า ที่จุฬาลงกรณ์มหาวิทยาลัย

กฤษดา วิศวธีรานนท์ สำเร็จการศึกษาระดับปริญญาตรี และ โท สาขาวิศวกรรมไฟฟ้าจากมหาวิทยาลัยเกียวโต ประเทศญี่ปุ่น ปัจจุบันดำรงตำแหน่งรองศาสตราจารย์ ประจำภาควิชาวิศวกรรมไฟฟ้า คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

สุวิทย์ นาคพีระยุทธ สำเร็จการศึกษาระดับปริญญาตรี และโท สาขาวิศวกรรมไฟฟ้า จากจุฬาลงกรณ์มหาวิทยาลัย ปัจจุบันเป็นอาจารย์ประจำภาควิชาวิศวกรรมไฟฟ้า คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ประวัติผู้เขียนวิทยานิพนธ์

นายวิเชียร อูปแก้ว เกิดวันที่ 21 มีนาคม 2518 ที่จังหวัดร้อยเอ็ด สำเร็จการศึกษาระดับปริญญาตรีวิศวกรรรมศาสตร์บัณฑิต สาขาวิศวกรรมไฟฟ้า จากศูนย์กลางสถาบันเทคโนโลยีราชมงคล เมื่อปี พ.ศ. 2542 และเข้ารับราชการในตำแหน่งอาจารย์ ประจำภาควิชาวิศวกรรมไฟฟ้า คณะวิศวกรรมศาสตร์ ศูนย์กลางสถาบันเทคโนโลยีราชมงคลในปีเดียวกัน จนถึงปี พ.ศ. 2544 ลาศึกษาต่อในหลักสูตรวิศวกรรมศาสตร์มหาบัณฑิต สาขาวิศวกรรมไฟฟ้า จุฬาลงกรณ์มหาวิทยาลัย

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย