รู้จักและสร้างเครื่องถอดรหัส MP3 ด้วยมือของคุณ ตอน 1 มารู้จักกับ MP3

ลภน สุภาพ

ปัจจุบันนี้คอมพิวเตอร์ได้เข้ามามีส่วนร่วมในการดำเนินชีวิตประจำวันอย่างแทบจะหลีกเลี่ยงไม่ได้ และคงมีน้อยคนนักที่ไม่เคยได้ยินคำว่า “MP3” แต่จะมีซักกี่คนที่รู้ว่าเจ้าคำๆ นี้ว่ามันมีต้นกำเนิดอย่างไรและทำไมมันจึงเป็นที่นิยมมากขนาดนี้ บทความนี้เป็นตอนแรกของมินิซีรีย์ “รู้จักและสร้างเครื่องถอดรหัส MP3 ด้วยมือของคุณ” ซึ่งจะเริ่มตั้งแต่ความรู้เบื้องต้นเกี่ยวกับ MP3 ว่ามีที่มาอย่างไร ต่อด้วยการสร้างบอร์ดทดลองถอดรหัส MP3 การนำบอร์ดทดลองไปใช้งานร่วมกับคอมพิวเตอร์ จนถึงสุดท้ายจะกลายเป็นเครื่องถอดรหัส MP3 ที่สามารถทำงานได้โดยไม่ต้องอิงกับคอมพิวเตอร์ (Standalone) โดยอ่านข้อมูลออกมาจากแผ่นซีดีรอม และแสดงผลเป็นภาษาไทย

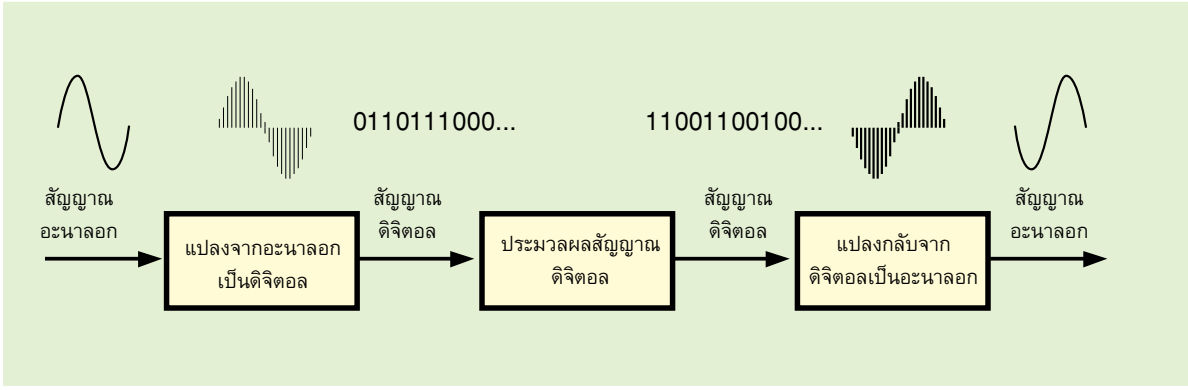
โครงงานชุดใหม่ที่ผสมผสานบทความเพื่อให้รู้จักกับเทคโนโลยี MP3 ที่ใช้กันอย่างแพร่หลายในปัจจุบัน ไล่เรื่อยไปจนสามารถสร้างตัวถอดรหัส MP3 ที่ไม่ต้องอาศัยคอมพิวเตอร์ขึ้นได้ด้วยตัวคุณเอง

ก่อนจะมาเป็น MP3

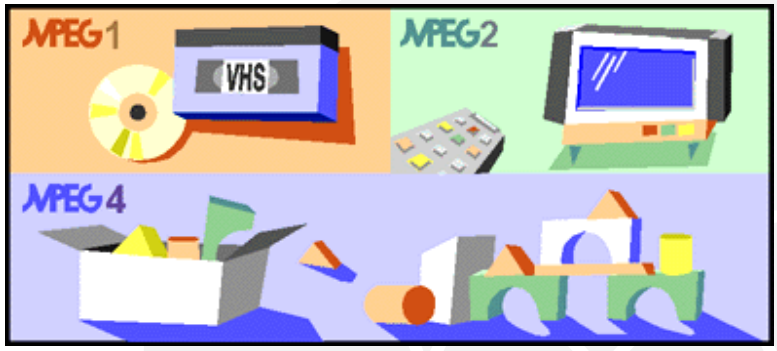
ก่อนที่จะพูดถึง MP3 ก็คงจะต้องขอกล่าวถึงการเก็บข้อมูลเสียง (Audio) ในปัจจุบันเสียก่อน แรกเริ่มเดิมทีนั้นข้อมูลประเภทเสียงจะถูกเก็บอยู่ในรูปแบบของสัญญาณอะนาลอก ซึ่งเราก็คงคุ้นเคยกันดีในรูปของเทปคาสเซ็ตต์หรือสำหรับคนรุ่นเก่าหน่อยก็คงไม่มีใครไม่รู้จักแผ่นเสียง กล่าวคือข้อมูลที่เก็บอยู่สามารถอ่านออกมาได้ในรูปของสัญญาณไฟฟ้าที่มีแรงดันต่างๆ กันต่อเนื่องกันไปเรื่อยๆ เมื่อนำสัญญาณเหล่านี้มาผ่านวงจรขยายและส่งต่อไปยังลำโพงก็จะทำให้เราได้ยินเสียงต้นฉบับที่ได้บันทึกมา

แต่เนื่องจากการเก็บข้อมูลแบบอะนาลอกนั้นไม่สามารถเก็บข้อมูลของเสียงได้ครบทั้งหมด นอกจากนั้นเมื่อเก็บไว้เป็นระยะเวลานานมากๆ คุณภาพของ

สงวนลิขสิทธิ์โดย บริษัท ซีเอ็ดยูเคชั่น จำกัด (มหาชน)



รูปที่ 1 การประมวลผลข้อมูลเสียงในระบบดิจิตอล



รูปที่ 2 มาตรฐาน MPEG และการใช้งาน

สื่อที่ใช้เก็บข้อมูลก็จะเสื่อมลงทำให้ข้อมูลเสียงที่อ่านออกมาได้แย่ลงไปอีก จึงมีการคิดค้นการเก็บข้อมูลเสียงแบบใหม่ในรูปของสัญญาณดิจิตอล หรือที่เราคุ้นเคยกันเป็นอย่างดีในปัจจุบันก็คือซีดีรอมนั่นเอง

การเก็บข้อมูลเสียงในแผ่นซีดีรอมนั้นเป็นข้อมูลในแบบดิจิตอล แต่ข้อมูลเสียงนั้นเป็นข้อมูลแบบอะนาลอกดังนั้นก่อนที่จะเก็บข้อมูลได้นั้นจำเป็นต้องมีการแปลงจากอะนาลอกเป็นดิจิตอลเสียก่อน การแปลงข้อมูลจากอะนาลอกเป็นดิจิตอลนั้นใช้หลักการ “สุ่มตัวอย่าง” เป็นพื้นฐานในการแปลง กล่าวคือขณะที่ข้อมูลเสียงซึ่งเป็นสัญญาณอะนาลอกถูกป้อนเข้ามายังตัวแปลง ตัวแปลงจะเก็บค่าความแรง (แอมพลิจูด) ของสัญญาณ ณ ขณะนั้นไว้ แล้วแปลงเป็นเลขฐาน 2 เมื่อแปลงเสร็จก็จะไปเก็บค่าความแรงจากสัญญาณอะนาลอกต่อไปเรื่อย ๆ อุปกรณ์ที่แปลงสัญญาณนี้เราเรียกว่า วงจรแปลงสัญญาณอะนาลอกเป็นดิจิตอล (Analog to Digital Convertor : ADC)

จากหลักการที่กล่าวมาเห็นได้ว่ายิ่งมีการเก็บตัวอย่างจำนวนมากเท่าไหร่ข้อมูลที่แปลงออกมาก็จะเหมือนข้อมูลต้นฉบับมากขึ้นเท่านั้น แต่ปัญหาที่ตามมาก็คือ มีข้อมูลที่ต้องเก็บเป็นจำนวนมหาศาลเลยทีเดียว ในปัจจุบันข้อมูลที่เก็บในแผ่นซีดีรอมเพลงนั้นใช้อัตราการสุ่ม (Sampling Frequency) 44.1 กิโลเฮิรตซ์ สาเหตุที่ต้องใช้ความถี่นี้ก็เนื่องมากจากทฤษฎีของไนควิซ (Nyquist Theorem) ซึ่งกล่าวไว้ว่า “ถ้าต้องการแปลงข้อมูลจากอะนาลอกเป็นดิจิตอลโดยไม่ให้ข้อมูลผิดเพี้ยนจะต้องมีความถี่ในการสุ่มเก็บตัวอย่าง อย่างน้อย 2 เท่าของความถี่ข้อมูลที่ต้องการเก็บ” และเนื่องจากหูมนุษย์รับฟังเสียงความถี่สูงสุดได้ไม่เกิน 20 กิโลเฮิรตซ์ ดังนั้นอัตราการสุ่มของแผ่นซีดีรอมเพลงจึงเป็น 44.1 กิโลเฮิรตซ์ โดยในการสุ่มหนึ่งครั้งจะแปลงเป็นข้อมูลดิจิตอลขนาด 16 บิต เก็บทั้งแชนเนลซ้ายและขวา ดังนั้นปริมาณข้อมูลทั้งหมดที่จะเกิดขึ้นในเวลา 1 วินาทีจะเท่ากับ 44,100×16×2 = 1,411,200 บิตเลยทีเดียว และถ้าเราลองคำนวณดูว่าเพลง 1 เพลงที่มีความยาวประมาณ 4 นาที จะต้องเก็บข้อมูลปริมาณเท่าไร

จากปัญหาที่กล่าวมาข้างต้นนั้นจึงทำให้เกิดความพยายามที่จะแก้ปัญญานี้ซึ่งพระเอกที่จะมาแก้ปัญหานี้ก็คือ MPEG นั่นเอง

แล้ว MPEG มันดียังไง

คำว่า MPEG นั้นย่อมาจาก Moving Picture Experts Group ซึ่งเป็นชื่อของกลุ่มบุคคลที่ร่วมมือกันสร้างมาตรฐานสากล (International Standard) เพื่อใช้ในการเข้ารหัสข้อมูลภาพและเสียงที่อยู่ในรูปของสัญญาณดิจิตอลก่อตั้งขึ้นตั้งแต่ปี ค.ศ. 1988 โดยมาตรฐานที่สร้างขึ้นนั้นถูกบรรจุเป็น

ตารางที่ 1 อัตราการบีบอัดข้อมูลและความเร็วในการส่งข้อมูลจากเครื่องอ่านของข้อมูลที่ถูกบีบอัดตามมาตรฐาน MPEG-1

มาตรฐานการบีบอัด	อัตราส่วน	ความเร็วในการส่งข้อมูล
MPEG-1 Layer1	1:4	384kb/s
MPEG-1 Layer2	1:6 ถึง 1:8	256-192kb/s
MPEG-1 Layer3	1:10 ถึง 1:12	128-115kb/s

มาตรฐานสากล ISO/IEC หมายความว่าบุคคลทั่วไปทุกๆ คนสามารถนำมาตรฐานนี้มาใช้เข้ารหัสหรือถอดรหัสข้อมูลของตัวเองได้โดยไม่ต้องขออนุญาตหรือจ่ายค่าลิขสิทธิ์ให้กับผู้ใด

มาตรฐาน MPEG นั้นแบ่งย่อยออกเป็นกลุ่มๆ ตามชนิดของข้อมูลที่ถูกเข้ารหัสและการนำไปใช้งาน ซึ่งในปัจจุบันแบ่งออกได้เป็น 5 กลุ่ม ดังนี้

MPEG-1 เข้ารหัสข้อมูลภาพและเสียง ใช้ในระบบวิดีโอซีดี และเสียงเพลง

MPEG-2 เข้ารหัสข้อมูลภาพและเสียง ใช้ในระบบโทรทัศน์ดิจิตอล และดีวีดี

MPEG-4 เข้ารหัสข้อมูลมัลติมีเดียที่ใช้กันอยู่ในเว็บเพจ

MPEG-7 เป็นมาตรฐานในการใส่คำอธิบายชนิดของข้อมูลมัลติมีเดียลงในตัวของมันเอง ใช้ประโยชน์ในด้านการหาข้อมูลบนระบบอินเทอร์เน็ต

MPEG-21 เป็นมาตรฐานใหม่ล่าสุดที่เพิ่งกำหนดขึ้นเมื่อเดือนมิถุนายนปีนี้ ว่าด้วยเรื่องเกี่ยวกับ Multimedia Framework

แล้ว MP3 อยู่ตรงไหนล่ะ?

เมื่ออ่านหัวข้อที่แล้วจบหลายคนอาจสงสัยว่า MP3 ที่เรากำลังพูดถึงมันไปอยู่ซะตรงไหนล่ะเนี่ย ดังที่กล่าวไปแล้วว่า มาตรฐาน MPEG-1 ใช้ในการเข้ารหัสข้อมูลภาพและเสียง แต่ตัว MPEG-1 เองนี้ยังแบ่งออกเป็น 3 เลเยอร์ (Layer) ตามความสามารถและความซับซ้อนในการเข้ารหัสข้อมูล โดยเลเยอร์ 1 มีความซับซ้อนในการเข้ารหัสน้อยทำให้บีบอัดข้อมูลได้น้อย และในทางกลับกันเลเยอร์ 3 มีความซับซ้อนในการเข้ารหัสมากที่สุดทำให้บีบอัดข้อมูลได้มากที่สุด แต่ทั้งสามเลเยอร์จะมีพื้นฐานการบีบอัดเหมือนกันทั้งหมด

MP3 ที่เราจะพูดถึงในบทความนี้ จริงๆ แล้วชื่อเต็มๆ ของมันคือ **MPEG-1 Layer3** ไม่ใช่ MPEG 3 อย่างที่หลายคนเรียกกัน และเพื่อป้องกันความสับสน ในส่วนต่อไปของบทความนี้ถ้ามีการกล่าวถึงคำว่า MPEG จะขอหมายความถึง MPEG-1และคำว่า MP3 หมายถึง MPEG-1 Layer3

จะเห็นได้ว่า MP3 มีกระบวนการในการเข้ารหัสที่ซับซ้อนที่สุดทำให้บีบอัดข้อมูลได้มาก หรือทำให้เหลือขนาดของข้อมูลเล็กมาก แต่ขนาดข้อมูลที่เล็กลงนี้ใช่ว่าจะไม่เสียอะไรไปเลย เนื่องจากการบีบอัดข้อมูลแบบ MPEG นั้นเป็นการบีบอัดข้อมูลแบบมีการสูญเสีย (Lossy Compression) ดังนั้นขนาดของข้อมูลที่เล็กลงนี้จึงมีการสูญเสียข้อมูลบางส่วนไปแต่ส่วนที่สูญเสียไปนั้นไม่มีผลต่อเสียงที่มนุษย์ได้ยินเมื่อมีการแปลงกลับ

การบีบอัดข้อมูลตามมาตรฐาน MPEG-1 นั้นได้ผลลัพธ์เป็นข้อมูลที่เรียงต่อกัน (Data Stream) ซึ่งจะรองรับทั้งในแบบ 1 ช่องสัญญาณและ 2 ช่องสัญญาณ แยกออกเป็นระบบเสียงต่างๆ ได้ถึง 4 ระบบคือ

1. ระบบโมโน คือได้ผลลัพธ์เป็นข้อมูลเพียง 1 ช่องสัญญาณ ซึ่งอาจจะเป็นเสียงจากช่องซ้ายหรือขวาก็ได้

2. ระบบดูอัลโมโน (Dual-Mono) ได้ข้อมูลผลลัพธ์ออกมา 2 ช่องสัญญาณ โดยช่องหนึ่งเป็นเสียงจากลำโพงฝั่งซ้าย อีกช่องหนึ่งเป็นเสียงจากลำโพงฝั่งขวา

3. ระบบสเตอริโอ ข้อมูลที่ได้ประกอบด้วย 2 ช่องสัญญาณเช่นกัน แต่ช่องหนึ่งจะเป็นผลรวมของเสียงลำโพงซ้ายกับลำโพงขวา อีกช่องหนึ่งเป็นผลต่างของลำโพงซ้ายกับลำโพงขวา

4. ระบบจอยท์-สเตอริโอ (Joint-Stereo) มีลักษณะคล้ายกับระบบสเตอริโอ แต่จะมีการรวมสัญญาณที่ความถี่ต่ำๆ ไว้ในช่องสัญญาณเดียว และแยกเสียงที่ความถี่สูงขึ้นมาเหมือนกับระบบสเตอริโอ ปรกติทั่วไปเนื่องจากธรรมชาติของหูมนุษย์จะแยกแยะตำแหน่งของแหล่งกำเนิดเสียงได้ยากถ้าเสียงนั้นมีความถี่ต่ำๆ

จากตอนต้นของบทความนี้เห็นได้ว่าข้อมูลที่อ่านได้จากซีดีรอมเพลงธรรมดานั้นมีจำนวนถึง 1,411,200 บิตต่อความยาว 1 วินาที ดังนั้นความเร็วในการส่งข้อมูลจากเครื่องอ่านไปยังตัวแปลงข้อมูลต้องสูงตามไปด้วยก็คือประมาณ 1.4 เมกะบิตต่อวินาที จากตารางที่ 1 แสดงให้เห็นถึงอัตราการบีบอัดข้อมูลและความเร็วในการส่งข้อมูลจากเครื่องอ่านของข้อมูลที่ถูกบีบอัดตามมาตรฐาน MPEG-1 เลเยอร์ต่างๆ โดยอ้างอิงเสียงที่ได้จากการแปลงกลับให้อยู่ในระดับคุณภาพเสียง

สงวนลิขสิทธิ์โดย บริษัท ซีเอ็ดยูเคชั่น จำกัด (มหาชน)

ตารางที่ 2 ความสัมพันธ์ระหว่างคุณภาพเสียงที่ต้องการกับขนาดของข้อมูลที่ถูกบีบอัด

คุณภาพเสียง	แบนด์วิดธ์	ระบบเสียง	อัตราการส่งข้อมูล	อัตราส่วนการบีบอัดข้อมูล
โทรศัพท์	2.5kHz	โมโน	8kb/s	96:1
ดีกว่าคลื่น shortwave	4.5kHz	โมโน	16kb/s	48:1
ดีกว่าคลื่นวิทยุ AM	7.5kHz	โมโน	32kb/s	24:1
ใกล้เคียงวิทยุ FM	11kHz	สเตอริโอ	56 ถึง 64kb/s	26 ถึง 24:1
ใกล้เคียงวิทยุ CD	15kHz	สเตอริโอ	96kb/s	16:1
CD	>15kHz	สเตอริโอ	112 ถึง128kb/s	14 ถึง 12:1

ตารางที่ 3 แสดงให้เห็นถึงเวลาที่เสียไปในการแปลงข้อมูล

เลเยอร์	อัตราส่วนการบีบอัด	เวลาที่ใช้ในการถอดรหัส
1	4:1	19ns
2	6:1	35ns
3	12:1	59ns

ซีดี ซึ่งเห็นได้ว่าข้อมูลที่บีบอัดตามมาตรฐาน MP3 นั้นจะเหลืออัตราความเร็วในการส่งข้อมูลเพียง 128 กิโลบิตต่อวินาที นั่นหมายความว่าข้อมูลเสียงเพลง 1 นาทีจากเดิมที่มีขนาดประมาณ 10 เมกะไบต์จะเหลือเพียงประมาณ 1 เมกะไบต์เท่านั้น นั่นหมายความแผ่นซีดีรอมจากเดิมที่เก็บเพลงได้ 14 - 18 เพลงรวมความยาวประมาณ 65 นาที ถ้านำมาเก็บข้อมูลที่ถูกเข้ารหัส MP3 จะสามารถเก็บเพลงได้มากกว่า 200 เพลง รวมความยาวมากกว่า 600 นาที

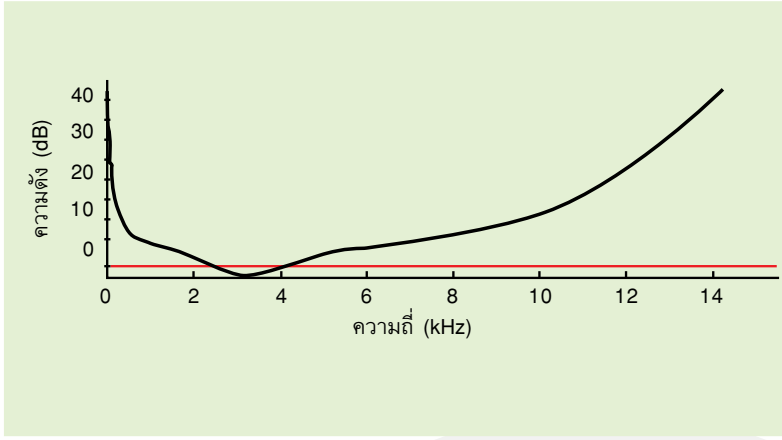
คุณภาพเสียงที่ได้

การบีบอัดข้อมูลตามมาตรฐาน MPEG นั้น สามารถกำหนดได้ว่าต้องการให้ข้อมูลที่ได้มีคุณภาพเสียงในระดับใด ถ้าต้องการให้มีคุณภาพเสียงที่ดีก็จะมีขนาดของข้อมูลใหญ่กว่าข้อมูลที่ต้องการคุณภาพเสียงที่ด้อยลงมา ขนาดของข้อมูลที่ถูกบีบอัดสามารถแทนได้ด้วยอัตราเร็วในการส่งข้อมูล เนื่องจากถ้าอัตราเร็วในการส่งข้อมูลสูง หมายความว่าใน 1 วินาทีต้องใช้ข้อมูลในจำนวนที่สูงตามไปด้วย

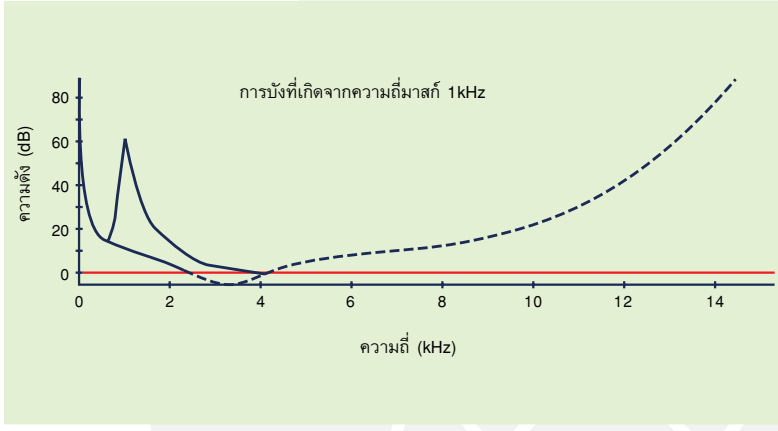
ต่อไปเรามาดูกันถึงความสัมพันธ์ระหว่างคุณภาพเสียงที่ต้องการกับขนาดของข้อมูลที่ถูกบีบอัด จากตารางที่ 2 แสดงให้เห็นถึงความสามารถของการบีบอัดแบบ MP3 ได้อย่างชัดเจน ในกรณีที่เราต้องการคุณภาพเสียงในระดับของโทรศัพท์ จะเหลืออัตราการส่งข้อมูลเพียง 8 กิโลบิตต่อวินาทีเท่านั้น (ปัจจุบันระบบโทรศัพท์ดิจิตอลส่งข้อมูลเสียงด้วยความเร็ว 64 กิโลบิตต่อวินาทีตามมาตรฐาน ISDN) หรือถ้าเราต้องการเสียงที่มีคุณภาพในระดับวิทยุ FM โดยใช้การบีบอัดแบบ MP3 ก็จะใช้อัตราการส่งข้อมูลเพียง 64 กิโลบิตต่อวินาทีเท่านั้นด้วย ความเร็วเท่านี้ทำให้เราสามารถสร้างระบบวิทยุดิจิตอลซึ่งส่งข้อมูลดิจิตอลออกอากาศให้เป็นความจริงได้ และในขณะนี้มีนักศึกษา ของไทยเราทำปริญญานิพนธ์ในด้านนี้อยู่เช่นกัน

แต่ความสามารถในการบีบอัดที่สูงก็ใช่ว่าจะได้มาโดยไม่เสียอะไรเลย เนื่องจากกระบวนการในการเข้ารหัสที่ซับซ้อนทำให้การแปลงข้อมูลต้องใช้เวลาช่วงหนึ่งในการทำงาน จากตารางที่ 3 แสดงให้เห็นถึงเวลาที่เสียไปในการแปลงข้อมูล

จากบทความในช่วงแรกที่ผ่านมาคงพอทำให้รู้จักกับการบีบอัดข้อมูลตามมาตรฐาน MP3 กันไปพอสมควรแล้ว ในส่วนถัดไปเรามาดูกันว่าการบีบอัดข้อมูลของมาตรฐาน MPEG นั้นเค้าทำกันอย่างไร ทำไมจึงทำให้ข้อมูลมีขนาดลดลงได้โดยเสียงที่ได้ยังอยู่ในระดับคุณภาพเท่ากับซีดี



รูปที่ 3 ความไวต่อเสียงความถี่ต่าง ๆ ของหูมนุษย์



รูปที่ 4 การบังทางด้านความถี่ (Frequency Masking) ที่ความถี่มาสก์ 1 กิโลเฮิรตซ์

ลดขนาดข้อมูลกันอย่างไร

การลดขนาดข้อมูลตามมาตรฐาน MPEG นั้นใช้พฤติกรรมในการได้ยินเสียงของมนุษย์มาเป็นเครื่องมือในการลดขนาดข้อมูล ก่อนที่เราจะไปดูว่านำมาใช้เป็นเครื่องมือได้อย่างไร เรามาคุยกันเรื่องพฤติกรรมการได้ยินเสียงของมนุษย์กันก่อนดีกว่า

อย่างที่รู้กันโดยทั่วไปแล้วว่า หูของมนุษย์นั้นมีขีดจำกัดในด้านการรับฟัง โดยขีดจำกัดอยู่ที่ด้านความถี่ 20 - 20,000 เฮิรตซ์ นั่นหมายความว่ามนุษย์จะได้ยินเสียงในช่วงความถี่ที่สูงกว่า 20 เฮิรตซ์แต่ต่ำกว่า 20 กิโลเฮิรตซ์เท่านั้น ถ้าอยู่เกินช่วงนี้ไปหูเราจะไม่ได้ยินเสียงนั้น ๆ แต่ใช่ว่าหูเราจะมีความไวต่อทุก ๆ ความถี่ที่เราได้ยินเท่า ๆ กัน จากการทดลองของนักวิทยาศาสตร์โดยสร้างห้องปิดที่ไม่มีเสียงรบกวนจากภายนอกขึ้น ให้ผู้ทดสอบเข้าไปข้างในห้อง จากนั้นก็เริ่มสร้างเสียงที่ความถี่ค่าหนึ่ง (ที่หูมนุษย์ได้ยิน) โดยค่อย ๆ เพิ่มความดังของเสียงขึ้นเรื่อย ๆ จนหูของผู้ทดสอบได้ยินแล้วบันทึกค่าไว้ แล้วก็เปลี่ยนความถี่ที่ใช้ทดลองไปเรื่อย ๆ จนครบตลอดย่านความถี่ที่มนุษย์ได้ยิน นำข้อมูลที่ได้มาวาดกราฟจะได้ดังรูปที่ 3

จากรูปที่ 3 จะเห็นได้อย่างชัดเจนว่าหูของมนุษย์เรามีความไวต่อเสียงแตกต่างกันที่ความถี่ต่าง ๆ กัน โดยจะมีความไวมากที่สุดในช่วงความถี่ 2 - 4 กิโลเฮิรตซ์ และถ้าเป็นความถี่ที่ต่ำมาก ๆ หรือสูงมาก ๆ ก็ยิ่งต้องใช้ความดังมาก ๆ เพื่อให้หูเราได้ยิน การทดลองในช่วงที่ผ่านมาเป็นการทดลองกับความถี่เพียงความถี่เดียว ต่อไปเราจะมาดูว่าถ้ามีเสียง 2 ความถี่ที่ใกล้เคียงกันกำเนิดพร้อม ๆ กันหูเราจะได้ยินอย่างไร

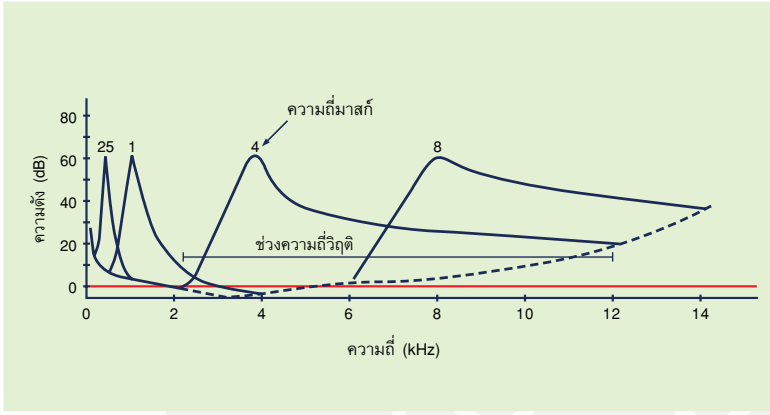
ในห้องปิดห้องเดิมสร้างเสียงความถี่ 1 กิโลเฮิรตซ์ที่ระดับความดัง 60 ดีบี จากนั้นก็สร้างเสียงที่มีความถี่ใกล้เคียงกันขึ้นมา เช่น 0.9 กิโลเฮิรตซ์ แล้วเพิ่มความดังขึ้นเรื่อย ๆ จนกว่าหูจะได้ยิน แล้วบันทึกค่าไว้ แล้วปรับความถี่ที่สร้างขึ้นเป็นความถี่ที่ 2 ทดลองไปเรื่อย ๆ แล้วบันทึกค่า นำมาวาดกราฟจะได้ดังรูปที่ 4

จากรูปที่ 4 เห็นได้ว่าถ้ามีความถี่หนึ่งดังขึ้นมา เสียงความถี่ที่ใกล้เคียงที่จะทำให้เราได้ยินนั้นจำเป็นต้องมีความดังมาก ๆ หากมีความดังน้อยกว่าเส้นกราฟหมายความว่าเราจะไม่ได้ยินเสียงนั้น ๆ ไปเลย คุณสมบัติของหูมนุษย์ในจุดนี้ทำให้เกิดย่านความถี่วิกฤต (Critical Band) ถ้ามองจากรูปก็คือช่วงความถี่ที่อยู่ในส่วนฐานของสามเหลี่ยม และเราเรียกความถี่ที่เป็นยอดของสามเหลี่ยมนี้ว่าความถี่มาสก์ (Masking Tone) สรุปแล้วย่านความถี่วิกฤตก็คือย่านความถี่ในช่วงที่ได้ยินหรือแยกแยะเสียงได้ยากถ้ามีเสียงความถี่มาสก์ของแบนด์นั้น ๆ ยังอยู่ เราเรียกคุณสมบัติข้อนี้ของหูมนุษย์ว่าการบังทางความถี่ (Frequency Masking)

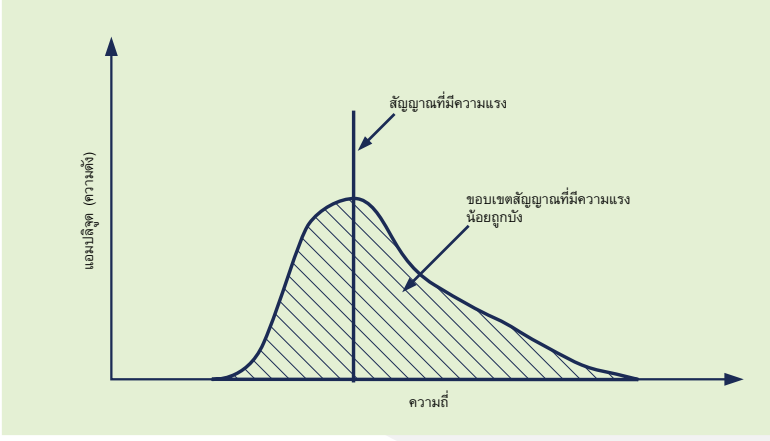
นอกจากการไม่ได้ยินเสียงที่อยู่ในย่านความถี่วิกฤตแล้ว ยังมีคุณสมบัติอีกข้อหนึ่งที่ทำให้หูเราไม่ได้ยินเสียงไป

สงวนลิขสิทธิ์โดย บริษัท ซีเอ็ดยูเคชั่น จำกัด (มหาชน)

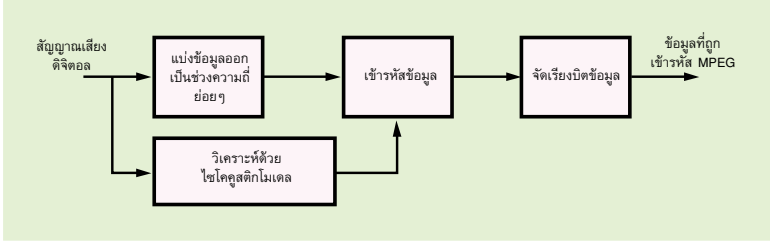
ชั่วคราว เราเรียกคุณสมบัติข้อนี้ว่า การบังชั่วคราว (Temporal Masking) คือเมื่อเราได้ยินเสียงที่เป็นความถี่มาสก์ดังขึ้นมาในระดับหนึ่ง หลังจากเสียงนั้นหยุดลงเราจะต้องเสียเวลาช่วงหนึ่งก่อนที่เราจะได้ยินเสียงที่มีความถี่ใกล้เคียงกับความถี่มาสก์นั้นๆ รูปที่ 5 เป็นกราฟที่ได้จากการทดลองสร้างเสียงความถี่มาสก์ 1 กิโลเฮิรตซ์ ที่มีความดัง 60 ดีบี หลังจากนั้นปิดเสียงที่เป็นความถี่มาสก์ ทดลองสร้างความถี่ใกล้เคียง (ในที่นี้คือ 1.1 กิโลเฮิรตซ์) ที่ความดังค่าหนึ่งขึ้นมาจับเวลาที่เสียไปก่อนจะได้ยินเสียงนั้น แล้วทดลองซ้ำโดยเปลี่ยนระดับความดังของเสียงที่เกิดขึ้นทีหลังไปเรื่อยๆ นำข้อมูลทั้งหมดมาวาดกราฟ ซึ่งจะเห็นได้ว่ายิ่งเสียงที่เกิดขึ้นทีหลังมีความดังน้อยเท่าไหร่ ก็ยิ่งมีระยะที่เราไม่ได้ยินเสียงยาวขึ้นเท่านั้น



รูปที่ 5 ย่านความถี่วิกฤต (Critical Band) ที่ความถี่มาสก์ต่าง ๆ



รูปที่ 6 ขอบเขตของช่วงความถี่ที่ถูกบังในการใช้ไซโคอคูสติก (Psychoacoustic) วิเคราะห์



รูปที่ 7 กระบวนการเข้ารหัส MPEG

คุณสมบัติด้านการฟังทั้งหมดที่กล่าวมานั้นรวมเรียกว่าไซโคอคูสติกโมเดล (Psychoacoustic Model) ถือเป็นเครื่องมือสำคัญในการลดขนาดข้อมูลของการบีบอัดตามมาตรฐาน MPEG ซึ่งมีลำดับกระบวนการในการบีบอัดดังนี้

1. นำข้อมูลเสียงดิจิตอลป้อนเข้าฟิลเตอร์เพื่อแยกเสียงออกเป็นช่วงความถี่ย่อยๆ (Subbands) ซึ่งมีความกว้างเท่ากับย่านความถี่วิกฤต จำนวน 32 ช่วงความถี่ เรียกขั้นตอนนี้ว่า Sub-band Filtering

2. ใช้ไซโคอคูสติกโมเดลเป็นเครื่องมือในการวิเคราะห์ข้อมูลส่วนที่ไม่มีผลต่อการได้ยินของมนุษย์ออกไป โดยพิจารณาระหว่างช่วงความถี่ 2 ช่วงที่ติดกัน และพิจารณาย่อยลงไปในแต่ละช่วงความถี่ด้วย

3. ถ้าวิเคราะห์แล้วพบว่าเสียงช่วงใดไม่มีผลต่อการได้ยิน ให้ตัดข้อมูลส่วนนั้นออกไป ไม่นำไปเข้ารหัสในส่วนถัดไป

4. นำข้อมูลที่เหลือมาเข้ารหัส ซึ่งจะมีวิธีที่แตกต่างกันขึ้นอยู่กับแต่ละเลเยอร์ แต่จะไม่ขอกล่าวในที่นี้เพราะเกรงว่าจะทำให้เนื้อหาหนักจนเกินไป

โครงสร้างข้อมูลของไฟล์ MP3

ข้อมูลที่ถูกบีบอัดตามมาตรฐาน MP3 นั้นจะอยู่ในลักษณะของเฟรมข้อมูล โดยในแต่ละเฟรมข้อมูลจะมีส่วนประกอบภายในอยู่ 4 ส่วนคือ

● หัวข้อมูล (Header) เป็นข้อมูลขนาด 32 บิต แสดงลักษณะทั่วไปของไฟล์นั้น ๆ

● ส่วนตรวจสอบความผิดพลาด (CRC) เป็นข้อมูลขนาด 16 บิต ใช้ตรวจสอบข้อมูลภายในเฟรมว่าถูกต้องหรือไม่จะมีหรือไม่มีก็ได้

● ข้อมูลข้างเคียง (Side Information) มีขนาด 17 หรือ 32 ไบต์ (17 ไบต์สำหรับระบบโมโน 32 ไบต์สำหรับระบบอื่น ๆ) เป็นส่วนที่เก็บองค์ประกอบที่ใช้ในการถอดรหัส

● ข้อมูลหลัก (Main Data) มีความยาวขึ้นอยู่กับอัตราการส่งข้อมูล (Baud rate) และอัตราการสุ่มข้อมูลในการแปลงกลับเป็นสัญญาณอะนาลอก (Sampling Frequency)

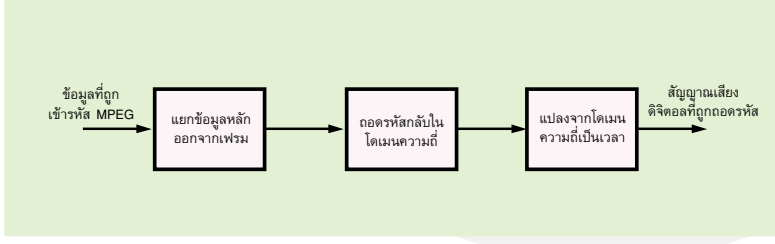
นำ MP3 มาใช้งาน

เมื่อรู้ที่มาที่ไปของ MP3 กันไปเรียบร้อยแล้วก็ถึงเวลาที่เราจะนำมาใช้ประโยชน์กันเสียที ความสามารถที่เด่นชัดของ MP3 นั่นก็คือการลดขนาดข้อมูลลงในปริมาณมากโดยยังคงคุณภาพของเสียงไว้ในระดับซีดี ดังนั้นการใช้งาน MP3 เกือบทั้งหมดในปัจจุบันจึงเกี่ยวข้องกับเสียงเพลงซึ่งให้ความบันเทิงเป็นหลัก สามารถแยกออกได้เป็นสองส่วนใหญ่ ๆ ส่วนแรกคือการนำข้อมูลเสียงมาเข้ารหัสบีบอัดให้เป็น MP3 และอีกส่วนหนึ่งคือการนำไฟล์ MP3 มาถอดรหัสแปลงกลับเป็นข้อมูลเสียงเหมือนเดิม โดยการใช้งานทั้งสองส่วนนั้นในระยะแรก (ประมาณสองปีที่แล้ว) ต้องใช้คอมพิวเตอร์ที่มีความสามารถในการทำงานสูงเท่านั้นจึงจะสามารถทำได้ แต่ในปัจจุบันด้วยความก้าวหน้าทางเทคโนโลยีทำให้ความสามารถในการประมวลผลของซีพียูมีความเร็วที่สูงขึ้นมาก การใช้งาน MP3 จึงแพร่หลายดังที่เป็นอยู่ในปัจจุบัน

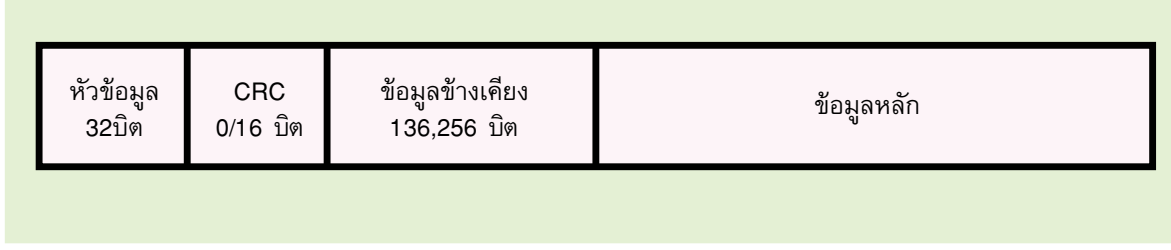
สำหรับการใช้งานในส่วนแรกคือการเข้ารหัสข้อมูลเสียงเป็นไฟล์ MP3 นั้นจำเป็นต้องใช้การประมวลผลข้อมูลปริมาณมหาศาล เนื่องจากต้องมีการวิเคราะห์ข้อมูลเสียงทั้งหมดด้วยไซโคอคูสติกโมเดลแล้วนำไปเข้ารหัสต่อ จึงจำเป็นต้องใช้ความสามารถในการประมวลผลของตัวประมวลผลที่สูงตามไปด้วย ถ้าเป็นเมื่อก่อนหากนำเพลงที่มีความยาว 4 - 5 นาที มาแปลงเป็นไฟล์ MP3 นั้นอาจจะต้องใช้เวลาถึง 20 นาที แต่ด้วยความสามารถของซีพียูในปัจจุบันซึ่งทำงานที่ความเร็วมากกว่า 500 เมกะเฮิรตซ์ทำให้เราสามารถเข้ารหัสข้อมูลเสียงเป็น MP3 ได้ในระดับเรียลไทม์ (Realtime) คือได้ข้อมูลผลลัพธ์ออกมาแทบจะในทันทีที่ป้อนเสียงเข้าไป

ส่วนการใช้งานอีกส่วนหนึ่งคือการถอดรหัสข้อมูล MP3 กลับมาเป็นเสียงนั้นจะอาศัยการประมวลผลที่น้อยกว่าเนื่องจากไม่ต้องวิเคราะห์ข้อมูลใด ๆ เลย เพียงแต่นำข้อมูล MP3 มาจัดเรียงให้ถูกลำดับและถอดรหัสเท่านั้นก็จะได้ข้อมูลเสียงกลับมา ระบบที่ใช้จึงไม่จำเป็นต้องมีความสามารถที่สูงมากเหมือนกับการเข้ารหัสเสียงไปเป็น MP3 (แต่ก็ยังต้องใช้การประมวลผลที่สูงถ้าเทียบกับการทำงานของโปรแกรมทั่วไป)

โปรแกรมที่ใช้สำหรับเข้ารหัสข้อมูลเสียงเป็น MP3 หรือถอดรหัส MP3 กลับเป็นข้อมูลเสียงที่ใช้งานบนคอมพิวเตอร์ในปัจจุบันมีอยู่เป็นจำนวนมาก มีทั้งในแบบใช้ได้ฟรี (Freeware, Shareware) และต้องเสียเงินซื้อ หากใครอยากได้โปรแกรมเหล่านี้ก็สามารถเข้าไปดาวน์โหลดได้จากอินเทอร์เน็ตมาใช้งานได้ทันทีโดยอาจจะเข้าไปที่ http://www.download.com หรือที่ http://www.mp3.com ก็ได้ ตัวอย่างของโปรแกรมถอดรหัส MP3 ที่คุ้นหูกันอย่างดีก็คือ WinAmp นั่นเอง

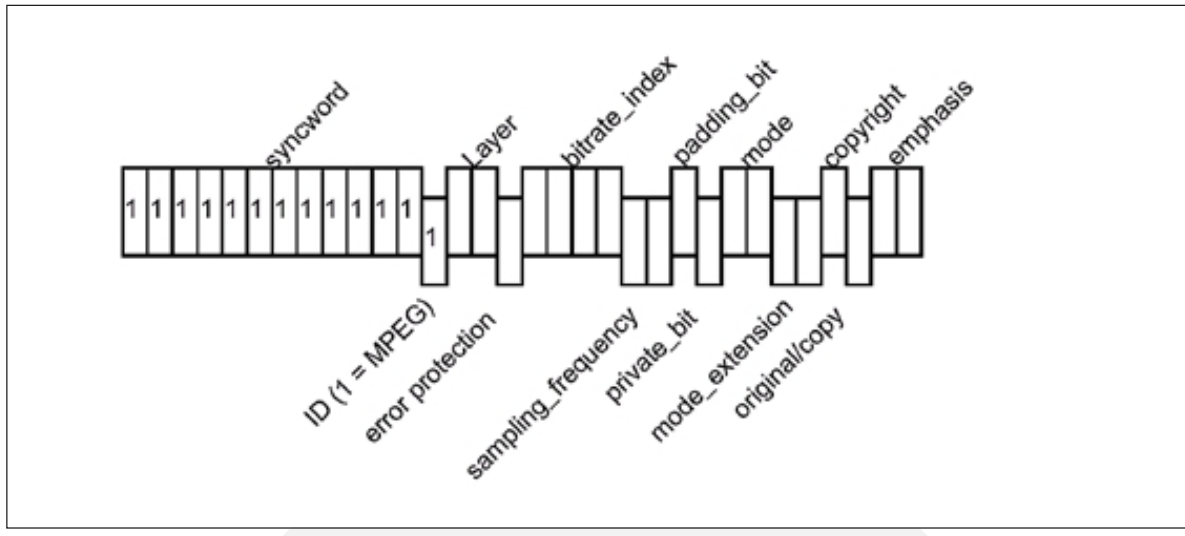


รูปที่ 8 กระบวนการถอดรหัส MPEG



รูปที่ 9 โครงสร้างของไฟล์ข้อมูล MP3 ใน 1 เฟรม

สงวนลิขสิทธิ์โดย บริษัท ซีเอ็ดยูเคชั่น จำกัด (มหาชน)



รูปที่ 10 ส่วนประกอบภายในหัวข้อมูล (Header)

MP3 ต้องคู่กับคอมพิวเตอร์เท่านั้นหรือ?

หลายคนอาจจะเกิดความสงสัยว่า ในเมื่อการใช้งานทั้งในส่วนของการเข้ารหัสและถอดรหัสล้วนแล้วแต่ใช้โปรแกรมบนคอมพิวเตอร์ทั้งนั้นเลย นั่นก็หมายความว่าถ้าไม่มีคอมพิวเตอร์แล้วไม่สามารถใช้งานMP3 ได้เลยหรือ?

ถ้าเป็นเมื่อ 2 ปีที่แล้วก็คงจะเป็นคำตอบว่าใช่อย่างหลีกเลี่ยงไม่ได้ แต่จากข้อดีในด้านความเล็กของมันจึงมีความพยายามที่จะนำ MP3 ไปแทนที่เครื่องเล่นคอมแพคดิสก์ที่ใช้ตามบ้านและในรถยนต์ ซึ่งระยะแรกจำเป็นที่จะต้องนำเครื่องคอมพิวเตอร์ทั้งเครื่องมาตัดส่วนที่ไม่ใช้งานออกและปรับลดขนาดให้เล็กลงเพื่อใช้ถอดรหัส MP3 ซึ่งวิธีนี้จะทำให้ราคาของเครื่องถอดรหัสสูงมากเนื่องจากเปรียบเสมือนซื้อเครื่องคอมพิวเตอร์ทั้งเครื่องมาทำงานเพียงอย่างเดียวคือถอดรหัส MP3 ราคาของเครื่องจะอยู่ในหลักสองหมื่นบาทขึ้นไป

แต่ในช่วงต่อมาก็มีบริษัทผลิตชิปเดี่ยว (Single Chip) ที่สามารถถอดรหัส MP3 ได้โดยไม่ต้องอาศัยอุปกรณ์เพิ่มเติมมากมายเหมือนเครื่องคอมพิวเตอร์ ทำให้เครื่องถอดรหัส MP3 ที่ไม่ต้องอิงกับคอมพิวเตอร์เป็นจริงขึ้นมาเกิดความตื่นตัวเป็นอย่างมากในวงการอิเล็กทรอนิกส์ในต่างประเทศ หลังจากนั้นก็เกิดเครื่องถอดรหัส MP3 ตามมาอย่างมากมายมีทั้งที่ใช้งานภายในบ้าน หรือติดรถยนต์ ไปจนถึงเครื่องเล่นที่พกพาได้ โดยสื่อในการบันทึกข้อมูล MP3 ก็มีต่างๆ กันไปตั้งแต่ฮาร์ดดิสก์ ซีดีรอม แฟลชเมโมรี ไปจนถึงการฟังผ่านระบบเน็ตเวอร์ก

เครื่องถอดรหัสที่เป็นเป้าหมายของมินิซีรีย์ชุดนี้สามารถใช้งานได้ภายในบ้าน แต่ถ้ามีการดัดแปลงอีกเล็กน้อยก็จะสามารถนำไปใช้งานภายในรถยนต์ได้ เนื่องจากได้ออกแบบให้มีระบบบัฟเฟอร์ข้อมูลป้องกันการขาดช่วงของข้อมูลจากการสั่นสะเทือน (Anti Shock) ได้นาน 7 วินาที หรือถ้าใครมีความคิดในการดัดแปลงก็สามารถนำไปสร้างเป็นเครื่องเล่นพกพาได้โดยเปลี่ยนสื่อบันทึกข้อมูลเป็นชนิดอื่นเช่นแฟลชเมโมรี่ เนื่องจากเครื่องต้นแบบนี้สามารถทำงานได้ที่ไฟเลี้ยงต่ำสุด 0.9 โวลต์ นั่นก็หมายความว่าถ่านไฟฉายเพียงก้อนเดียวก็สามารถทำให้มันทำงานได้ และนี่ก็เป็นบทพิสูจน์ว่า "คนไทยก็ทำได้"

สำหรับในตอนถัดไปจะเจาะลงไปถึงชิปที่ใช้ถอดรหัส MP3 ไปจนถึงการสร้างบอร์ดทดลองถอดรหัส MP3 อย่าลืมติดตามกันให้ได้นะครับ สำหรับบุคคลใดต้องการรายละเอียดเพิ่มเติมเกี่ยวกับ MP3 สามารถหาได้จาก http://www.mpeg.org, http://www.mp3.com, http://www.layer3.org

www.se-ed.com

แหล่งแลกเปลี่ยนความรู้ที่ไม่มีวันสิ้นสุด

เชิญแวะเยี่ยมเวบบอร์ดของเราเพื่อสอบถามปัญหาและแลกเปลี่ยนความรู้ด้านอิเล็กทรอนิกส์,อุตสาหกรรม,วิทยาศาสตร์,คอมพิวเตอร์และอินเทอร์เน็ต,ความรู้ทั่วไป,การบริหารงานและหนังสือ