



เสียงกับงาน มัลติมีเดีย

ประดิษฐ์ แสงไกร

โดย ส่วนใหญ่แล้วผู้ที่ใช้คอมพิวเตอร์ มักจะรับรู้ข้อมูลหลังจากการประมวลผลเมื่อใช้คอมพิวเตอร์ผ่านทางจอแสดงผล ไม่ว่าจะเป็นการพิมพ์ตัวอักษรหรือการคลิกเมาส์ เราจะได้เห็นว่า การรับรู้ข้อมูลข่าวสารดังกล่าวมักจะเป็นการรับรู้ผ่านการมองเห็นแทบทั้งสิ้น แต่เมื่อเปรียบเทียบกับการใช้ชีวิตประจำวันยังมีอีกส่วนหนึ่งที่เรารับรู้ควบคู่ไปกับการมองเห็น นั่นก็คือการได้ยิน ดังนั้นในการออกแบบสื่อมัลติมีเดียต่างๆ เราจะต้องให้ความสำคัญกับเสียงในน้ำหนักที่เท่ากันกับภาพ โดยเฉพาะงานด้านวิดีโอเกม เพราะผู้ออกแบบวิดีโอเกมเล็งเห็นถึงความสำคัญในการสร้างความสมจริงให้กับผู้เล่น เราจึงเห็นว่า วิดีโอเกมในปัจจุบันได้มีการกำหนดข้อจำกัดในส่วนของฮาร์ดแวร์ต่างๆ รวมถึงการตัดเสียงในคอมพิวเตอร์ เพื่อรองรับความต้องการนั้นๆ ด้วยเช่นกัน ไม่เพียงเท่านั้น สถาปนิก มัณฑนากร นักออกแบบเวที นักสังคมวิทยา นักดนตรี และศิลปินที่ทำงานเกี่ยวกับเสียง จึงต้องรู้และเข้าใจบทบาทของเสียงที่มีต่อสัมผัสการรับรู้ของมนุษย์และผลของปัจจัยภายนอกที่มีต่อเสียงดังกล่าว

เมื่อเรานำทั้งเสียงกับภาพเข้ามาใช้ในงานของเราแล้ว ทั้งสองสิ่งจะเกื้อหนุนและเป็นแรงผลักดันให้ฝ่ายใดฝ่ายหนึ่งมีน้ำหนักในการนำเสนอมากขึ้น เนื่องจากเสียงจะเป็นปัจจัยบ่งบอกถึงขอบเขตของสถานที่ที่เกิดเสียงไปด้วย ซึ่งปัจจัยที่ทำให้เรารับรู้ขอบเขตที่เกิดเสียงนั้นคือสภาพเสียงก้อง (Reverberation) เมื่อสถานที่แตกต่างกันออกไป ลักษณะเสียงก้องที่เกิดขึ้นย่อมแตกต่างกันออกไปด้วย แต่การรับรู้ที่ต่างจากการรับรู้ทางตา เพราะการรับรู้เสียงที่ต่างนี้เกิดจากนิยามที่สมองสร้างขึ้นจากประสบการณ์เดิมว่า สภาพแวดล้อมของเสียงนั้นๆ เป็นอย่างไร

ลักษณะทางกายภาพของเสียงเกิดจากการเปลี่ยนแปลงของอากาศรอบๆ ตัวเรา ส่งผลให้เกิดคลื่นเสียง หลังจากนั้นสมองจะแปรสัญญาณที่รับรู้ผ่านหู เกิดเป็นคลื่นไฟฟ้าที่ความถี่เดียวกัน และรับรู้เสียงดังกล่าว โดยทั่วไปเรามักใช้หน่วยวัดความถี่ของเสียงเป็นเฮิรตซ์ (Hertz) หรือรอบต่อวินาที ทั้งนี้ หลังจากเสียงใดๆ เกิดขึ้นก็ตาม สมองของมนุษย์จะกำหนดรับรู้ตามสภาพของเสียงนั้นๆ โดยทั่วไปเวลาที่เรารับรู้เสียงที่มีความถี่ต่างๆ สมองจะสั่งการว่าเสียงสูงเกิดขึ้นในสถานที่ซึ่งแตกต่างจากเสียงต่ำ ดังจะเห็นได้ว่าเวลาที่เรารับรู้เสียงสูง เราจะมีสติว่าเสียงลอยขึ้น ส่วนเวลาที่เรารับรู้เสียงต่ำ เสียงจะก่อดลต่ำ หูของมนุษย์สามารถรับรู้เสียงได้ 20 ถึง 20000 เฮิรตซ์ และจะตอบสนองย่านความถี่ของเสียงพูดมนุษย์ได้ดีที่สุด แต่มีข้อสังเกตอยู่บางประการ นั่นคือ มนุษย์จะไม่สามารถแยกแยะการเปลี่ยนแปลงความถี่ได้ทุกครั้ง เพราะปัจจัยสำคัญที่เกี่ยวข้องกับการรับรู้ก็อย่างหนึ่งคือความเข้มของเสียง (Intensity)



เสียงสร้างปรากฏการณ์การรับรู้ของมนุษย์ได้หลายประการ อาทิ

เสียงทำให้เรารับรู้ต้นกำเนิดของเสียง

ในบางครั้ง เมื่อเราฟังเสียงโดยปราศจากภาพ ก็สามารถทำให้เรารับรู้ได้ว่าเสียงนั้นๆ มีที่มาอย่างไร และมีผลต่อการรับรู้โดยตรง ยกตัวอย่างเช่น เราสามารถระบุวัตถุได้ทันทีเมื่อได้ยินเสียงแก้วแตก เสียงเดินบนสนามหญ้า เป็นต้น ปรากฏการณ์ของเสียงในลักษณะนี้เกิดขึ้นในชีวิตประจำวันของทุกคน ไม่ว่าจะเป็นในสำนักงาน บนรถประจำทาง หรือในรถยนต์ส่วนบุคคล ทั้งนี้ถือว่าการยากที่มนุษย์จะรับรู้เสียงลักษณะต่างๆ ดังกล่าวในเวลาเดียวกัน เนื่องจากมีความขัดแย้งกันอยู่ ผู้ออกแบบเสียงจึงหลีกเลี่ยงในการใช้เสียงในกรณีนี้พร้อมๆ กัน ในการวิจัยบางครั้งระบุได้ว่าปัจจัยในด้านเสียงเป็นตัวแปรสำคัญอย่างหนึ่งที่ส่งผลต่อผลการวิจัย ดังจะเห็นได้จากการทดลองในเรื่องเงื่อนไขแบบคลาสสิก (Classical Conditioning)¹

เสียงทำให้เรารับรู้ถึงข้อความแฝง

ในชีวิตประจำวันของเรานั้น เราได้ยินเสียงเพลงประกอบโฆษณามากมาย เพลงเหล่านั้นทำหน้าที่เปรียบเสมือนภาพสัญลักษณ์แทนสินค้าชิ้นต่างๆ ไม่ใช่แค่เพลงประกอบโฆษณาเท่านั้น ที่สามารถสร้างความหมายแฝงให้กับเสียง เราสามารถพบกรณีศึกษาดังกล่าวจากภาพยนตร์ ละครโทรทัศน์ รายการวิทยุ หรือแม้กระทั่งระบบปฏิบัติการบนเครื่องคอมพิวเตอร์

เสียงเป็นตัวจำลองสถานการณ์

ในการใช้งานแอปพลิเคชันต่างๆ เสียงนับเป็นส่วนสำคัญในการป้อนข้อมูลกลับสู่ผู้ใช้ เพราะในความเป็นจริงแล้ว เสียงกับภาพเป็นสิ่งที่อยู่คู่กันอย่างหลีกเลี่ยงมิได้ อย่างเช่นในวิดีโอเกม เราคาดหวังว่าจะได้ยินเสียงปืนของคู่ต่อสู้ดังขึ้นเมื่ออยู่ใกล้เรามากขึ้นเรื่อยๆ ในทางกลับกัน เสียงดังกล่าวก็น่าจะเบาลงเมื่ออยู่ไกลเราออกไป ในที่นี้คู่ต่อสู้คือการรับรู้ด้านภาพของเรานั้นเอง เช่นเดียวกัน เมื่อเราเห็นภาพคนเดินอยู่ในโถงทางเดินเราย่อมคาดหวังว่าจะได้ยินเสียงก้องและเสียงสะท้อนที่เกิดจากการตอบสนองต่อเสียงนั้นๆ ด้วยเช่นกัน

มนุษย์ไวต่อการจับทิศทางเสียงเป็นอย่างมาก เนื่องจากเป็นผลของสัญชาตญาณในการเอาตัวรอดของมนุษย์ การได้ยินถือเป็นการเฝ้าระวังภัยที่อาจจะมาถึงตัวเรา ดังนั้น การออกแบบวิดีโอเกมจึงจำเป็นต้องใช้ลำโพงมากกว่า 2 ตัวขึ้นไป เพื่อให้อรรถรสในการจับทิศทางเสียงเป็นไปอย่างสมจริง อีกทั้งยังต้องสร้างปรากฏการณ์ที่เกิดจากทิศทางของเสียง เช่น ดอปเปลอร์เอฟเฟกต์ (Doppler effect)² และเสียงสะท้อนอีกด้วย

¹ เงื่อนไขแบบคลาสสิก เป็นทฤษฎีทางจิตวิทยาในการเสริมแรงรูปแบบหนึ่ง ซึ่งคิดค้นโดยอีวาน พาฟลอฟ (Ivan Pavlov) นักจิตวิทยาชาวรัสเซีย ซึ่งทฤษฎีดังกล่าวสามารถเห็นได้ชัดจากการทดลองของพาฟลอฟครั้งหนึ่ง ซึ่งพาฟลอฟ ได้วางเงื่อนไขกับสุนัขของเขาเอง โดยการให้สัญญาณกระดิ่งก่อนการให้อาหาร แล้วสังเกตปฏิกิริยาตอบสนองของสุนัข แล้วค่อยๆ ลดจำนวนการให้สัญญาณกระดิ่งลงตามลำดับ แต่ปรากฏว่าสุนัขก็ยังคงมีปฏิกิริยาตอบสนองเหมือนเดิม

² ดอปเปลอร์เอฟเฟกต์ หรือ ดอปเปลอร์ชิฟต์ (Doppler Effect, Doppler Shift) เป็นปรากฏการณ์ทางเสียง เมื่อแหล่งกำเนิดเสียงหรือผู้สังเกตการณ์เคลื่อนที่ ส่งผลให้ความยาวช่วงคลื่นเปลี่ยนแปลงและมีผลกับความถี่ของเสียงของผู้สังเกตการณ์



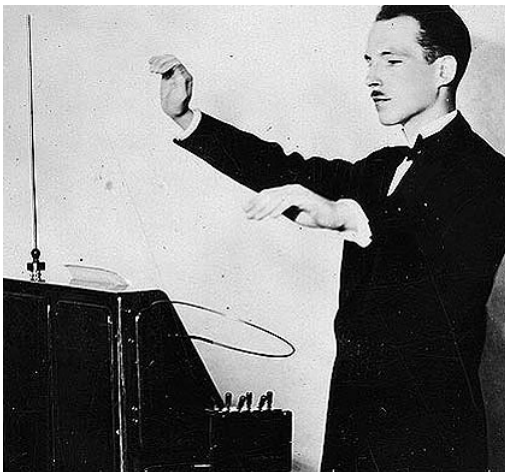
เสียงเป็นสื่อสร้างอารมณ์

เสียงสามารถเป็นตัวสร้างอารมณ์ต่างๆได้ เนื่องด้วยเสียงมีผลต่ออารมณ์ในส่วนของการจิตใต้สำนึก ไม่ว่าจะเป็นเสียงผู้คนโห่ร้องดีใจ หรือเสียงเพลงปลุกใจให้รักชาติต่างๆ เราจึงสังเกตได้ว่าในปัจจุบันสื่อมัลติมีเดียต่างๆ มักจะสอดแทรกเสียงที่สร้างอารมณ์เพิ่มเข้าไปมากขึ้น อีกทั้งยังสร้างอารมณ์และความรู้สึกถึงสิ่งที่ยื่นนอกจอแสดงผลอีกด้วย

และจากปรากฏการณ์ข้างต้น เราสามารถแยกแยะวิธีการสร้างปฏิสัมพันธ์ระหว่างเสียงกับผู้ใช้งาน แอปพลิเคชันได้เป็น 4 กรณี นั่นคือ

1. ผู้ใช้งานเป็นผู้สร้างเสียงที่ตัวผู้ใช้เองหรือผู้อื่นสามารถได้ยินได้

ในกรณีแรกนี้ จุดประสงค์หลักของผู้ใช้งานคือต้องการดัดแปลง, ควบคุม หรือสร้างเสียงขึ้น ซึ่งในกรณีนี้ อุปกรณ์สร้างเสียงไม่จำเป็นจะต้องเป็นสิ่งที่สามารถสร้างเสียงได้โดยตรงเท่านั้น อย่างเช่นเครื่องดนตรีต่างๆ หรืออุปกรณ์ในห้องบันทึกเสียง การใช้งานในกรณีนี้ อาจจะเป็นการสื่อสารเบื้องต้นให้ผู้อื่นควบคุมการสร้างเสียง หรือให้สัญญาณกับบุคคลอื่นก็ได้ ยกตัวอย่างเช่นเครื่องเทเรมิน (Theremin) และเครื่องเล่นแผ่นเสียง ทั้งสองสิ่งนี้จะผลิตเสียงก็ต่อเมื่อมีการปฏิสัมพันธ์ในส่วนของการเคลื่อนไหวร่างกายของผู้ใช้ หรือแม้กระทั่งการใช้ฮาร์ดแวร์อื่นๆเพื่อควบคุมองค์ประกอบของดนตรีอย่างเช่นการประยุกต์ใช้วีโมท (Wii mote) เพื่อควบคุมเสียงต่างๆ และการใช้องค์ประกอบจากภาพวาดเพื่อนำมาสร้างเสียงต่างๆ ก็จัดอยู่ในกรณีนี้ด้วยเช่นกัน



รูปที่ 1 เครื่องเทเรมิน เล่นโดยผู้คิดค้น นายลีอง เทเรมิน



รูปที่ 2 อุปกรณ์วีโมท

2. ผู้ใช้งานสร้างสัญญาณเสียงโดยรับค่าจากอุปกรณ์อย่างใดอย่างหนึ่ง

ในปัจจุบัน ภาษามือได้เข้ามามีบทบาทสำคัญในการผลิตสัญญาณเสียงเป็นอย่างมาก เราสามารถใช้ อุปกรณ์อย่างเช่นมิดีคีย์บอร์ดเพื่อการควบคุมค่าบางอย่างหรือใช้แทนคำสั่งต่างๆได้



3. ผู้ใช้งานสร้างเสียงเพื่อใช้เป็นข้อมูลขาเข้า

ในกรณีนี้ ผู้ใช้งานจะต้องเป็นผู้สร้างเสียงใดๆ ก็ตามโดยปราศจากเครื่องมือช่วยใดๆ หลังจากนั้นเราจะใช้ไมโครโฟนเพื่อนเป็นตัวแปลงสัญญาณเสียงให้เป็นสัญญาณไฟฟ้า เพื่อส่งต่อไปยังการประมวลผลส่วนอื่นๆ แล้วใช้องค์ประกอบของเสียงเพื่อควบคุมปัจจัยต่างๆ อาทิ ใช้ความดังของเสียงเพื่อควบคุมความสว่างของภาพ เป็นต้น นอกจากนี้ยังมีปัจจัยอื่นๆ เช่นระดับเสียง หรือสีล้นของเสียงที่เราสามารถนำมาใช้ควบคุมปัจจัยภายนอกอื่นๆ ได้อีกด้วย

4. ผู้ใช้พูดหรือเปล่งเสียงเพื่อใช้เป็นข้อมูลขาเข้า

ในกรณีนี้แตกต่างจากในข้อ 3 ตรงที่ในข้อ 3 จะเป็นเสียงอะไรก็ได้ที่ผู้ใช้งานสร้างขึ้นโดยปราศจากตัวช่วย ส่วนในข้อ 4 นี้ จำเป็นจะต้องเป็นเสียงพูดเท่านั้น ซึ่งเราจะเรียกวิธีการนี้ว่า Speech Recognition ซึ่งหมายถึงการเปล่งเสียงเพื่อใช้แทนคำสั่งต่างๆ

ซอฟต์แวร์สำหรับสร้างสื่อมัลติมีเดียต่างๆ ในปัจจุบันได้พัฒนาวิธีการรับค่าจากผู้ใช้ในกรณีต่างๆ ดังกล่าว โดยเน้นที่จุดมุ่งหมายแตกต่างกันออกไป ฟรีแวร์ที่ได้รับความนิยมในการสร้างงานมัลติมีเดีย อย่างเช่น Processing หรือชุดคำสั่ง openFrameworks สำหรับภาษา C++ ต่างก็มีชุดคำสั่งเพื่อรองรับการทำงานด้านเสียงเบื้องต้นเพื่อการประยุกต์ใช้ในแอปพลิเคชันต่างๆ อย่างเช่น ซอฟต์แวร์ Processing นั้น จะมีไลบรารี Minim ซึ่งทำหน้าที่วิเคราะห์หรือสังเคราะห์เสียง รวมทั้งประมวลผลสัญญาณด้วย ส่วนชุดไลบรารี open Frameworks ก็มีไลบรารีเอฟมอดเอ็กซ์ (FMOD Ex) ซึ่งเป็นชุดคำสั่งที่ใช้ในการพัฒนาวิดีโอเกมในทุกๆ แพลตฟอร์ม ไม่ว่าจะเป็นเอกซ์บ็อกซ์, เพลย์สเตชัน 3 หรือไอโฟน ด้วยเหตุผลที่ว่าพัฒนาวิดีโอเกมจำเป็นต้องใช้ภาษา C++ ซึ่งเป็นภาษาที่มีจุดเด่นอยู่การสร้างภาพ 3 มิติ จึงสะดวกที่จะใช้ไลบรารีดังกล่าวมากกว่า ซึ่งแตกต่างจากจุดเด่นของ Processing ซึ่งเป็นซอฟต์แวร์ที่สามารถส่งข้อมูลไปยังบอร์ดอาร์ดูโน (Arduino) ซึ่งเป็นบอร์ดประมวลผลที่เป็นที่นิยมใช้กันแพร่หลายในหมู่ผู้ที่ทำงานมัลติมีเดียในขณะนี้ แต่ถ้าหากจุดประสงค์ของผู้ใช้คือการสังเคราะห์เสียง โดยใช้เทคนิคที่ซับซ้อน การใช้ซอฟต์แวร์เพื่อการสังเคราะห์เสียงโดยเฉพาะอย่างเช่น Puredata, Max/MSP, Csound และ Supercollider จะทำหน้าที่ได้ดีกว่า แต่ถึงแม้ว่า Puredata และ Max/MSP จะสามารถทำงานได้ ทั้งเรื่องภาพและเสียง แต่จากประสบการณ์ของผู้เขียนพบว่าปัญหาในการทำงานร่วมกันระหว่างภาพและเสียง โดยเฉพาะเมื่อมีการเรนเดอร์ภาพสามมิติที่ต้องใช้ทรัพยากรภายในเครื่องคอมพิวเตอร์เป็นจำนวนมาก จะส่งผลให้อัตราเฟรมภาพ (Frame Rate) ผิดเพี้ยนได้ง่าย ดังนั้นจึงอยู่ที่จุดประสงค์ของผู้เขียนแอปพลิเคชัน ว่าต้องการใช้งานเน้นหนักในด้านใดเป็นสำคัญ



บรรณานุกรม

- Noble J. 2009. Programming Interactivity : **A Designer's Guide to Processing, Arduino and openFrameworks.**
Sebastopol, CA. : O'Reilly Media Inc.