บรรณาธิการปริทัศน์/Editorial Corner

การสื่อสารระหว่างมนุษย์กับคอมพิวเตอร์ จากพิมพ์ดีด สัมผัสจอ และสั่งงานด้วยเสียง ไปสู่กริยาท่าทาง

Human & Computer Communication from Typing, Screen Touch, and Speech to Gesture

โกสินทร์ จำนงไทย*
ภาควิชาวิศวกรรมอิเล็กทรอนิกส์และโทรคมนาคม คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี

Kosin Chamnongthai*
Electronic and Telecommunication Engineering Department, Faculty of Engineering, King Mongkut's University of Technology Thonburiu, Bangkok, Thailand

*Corresponding Author, E–mail: kosin.cha@kmutt.ac.th
DOI: 10.14416/j.kmutnb.2023.03.005
© 2023 King Mongkut's University of Technology North Bangkok. All Rights Reserved.

มนุษย์สามารถออกเสียงที่ซับซ้อนพูดคุยด้วยภาษาต่างๆ ทำให้มนุษย์สามารถสื่อสารระหว่างมนุษย์ด้วยกันได้ ซึ่งต่างจากสัตว์ที่ไม่สามารถพูดคุยและแสดงความรู้สึกอย่างละเอียดอ่อนและซับซ้อนได้ เสียงมนุษย์ซึ่งอยู่ในรูปของสัญญาณที่ประกอบด้วยสารสนเทศสามารถส่งจากคนหนึ่งไปยังคนอื่นในระยะใกล้ได้ สัญญาณเหล่านี้อยู่ในย่านความถี่ต่ำ มีพลังงานน้อย ไม่สามารถส่งผ่านตัวกลางที่เป็นบรรยากาศของโลกแล้วสะท้อนไปที่ไกลๆ ได้ ต่อมามนุษย์มีความจำเป็นต้องติดต่อสื่อสารกันในระยะที่ไกลขึ้น จึงเกิดความต้องการเทคโนโลยีเพื่อการสื่อสารทางไกล การสื่อสารทางไกลที่เรียกว่าโทรคมนาคมได้เริ่มขึ้นราวปลายศตวรรษที่ 19 โดยการนำสัญญาณที่อยู่ในย่านความถี่สูงที่สามารถส่งผ่านบรรยากาศไปในระยะไกลมาเป็นสัญญาณพาห์ (Carrier) เพื่อเป็นพาหะส่งสารสนเทศไปที่ไกลๆ ในเวลานั้นมนุษย์คิดค้นเทคนิคการสอดแทรกสารสนเทศลงในสัญญาณพาห์ที่เรียกกันว่า "มอดดูเลชัน (Modulation)" หรือ "การกล้ำสัญญาณ" ในรูปแบบต่างๆ เช่น AM, FM, PM เป็นต้น ทำให้เทคโนโลยีการส่งสัญญาณไปสู่แดนไกลเกิดเป็นจริงขึ้น เมื่อมนุษย์สามารถส่งสารสนเทศโดยการส่งสัญญาณดังกล่าวไปยังมนุษย์ในที่ต่างๆ ที่อยู่ไกลในโลกนี้ได้ ธุรกิจที่เกี่ยวกับการสื่อสารจึงเกิดขึ้นจากเทคโนโลยีนี้ ได้แก่ กิจการวิทยุ กิจการโทรทัศน์ เป็นต้น

โทรคมนาคมในระยะต้นยังอยู่ในรูปการประมวลผลสัญญาณแบบแอนะล็อก ในราวต้นคริสต์ศตวรรษที่ 20 มนุษย์ได้ค้นพบทฤษฎีสุ่ม (Sampling Theory) ทำให้ความเป็นไปได้ในการแปลงสัญญาณระหว่างแอนะล็อกกับดิจิทัลเกิดขึ้น เป็นที่ทราบกันดีโดยทั่วไปว่า สัญญาณในโลกนี้ส่วนใหญ่อยู่ในรูปของสัญญาณแอนะล็อกซึ่งประมวลผลได้เที่ยงตรงยากเมื่อเทียบกับสัญญาณดิจิทัล มนุษย์ต้องการประมวลผลสัญญาณในรูปของดิจิทัลมากกว่าแอนะล็อกเพราะง่ายและเที่ยงตรงกว่า ในช่วงแรกการแปลงสัญญาณแอนะล็อกให้เป็นสัญญาณดิจิทัลและการแปลงกลับกันยังไม่สามารถทำได้อย่างสมบูรณ์ ทำให้การประมวลผลสัญญาณดิจิทัลไม่แพร่หลาย ต่อมาเมื่อเครื่องแปลงสัญญาณระหว่างแอนะล็อกกับดิจิทัลสมบูรณ์ การแปลงสัญญาณแอนะล็อกที่มีอยู่ในธรรมชาติให้เป็นสัญญาณดิจิทัลเพื่อการประมวลผล แล้วแปลงกลับมาเป็นสัญญาณแอนะล็อกกลายเป็นเรื่องที่ทำกันอย่างแพร่หลาย

การอ้างอิงบทความ: โกสินทร์ จำนงไทย, "การสื่อสารระหว่างมนุษย์กับคอมพิวเตอร์ จากพิมพ์ดีด สัมผัสจอ และสั่งงานด้วยเสียง ไปสู่กริยาท่าทาง," *วารสารวิชาการพระจอมเกล้าพระนครเหนือ*, ปีที่ 33, ฉบับที่ 4, หน้า 1–3, เลขที่บทความ 234-006840, ต.ค.–ธ.ค. 2566.

จนถึงปัจจุบัน ปัญหาในงานวิจัยเกี่ยวกับการสื่อสารและการประมวลผลสัญญาณทั้งฮาร์ดแวร์และซอฟต์แวร์จึงมีแนวโน้มไปทางดิจิทัลมากกว่าแอนะล็อก

มนุษย์มีเครื่องมือสื่อสารระหว่างมนุษย์ด้วยกันที่อยู่ไกลกัน ทำให้เกิดความเป็นไปได้ในการสื่อสารระหว่างบุคคลและการกระจายเสียง เทคโนโลยีสื่อสารทางไกลระหว่างมนุษย์ด้วยกันได้รับการพัฒนาจนเป็นที่ยอมรับได้อย่างแพร่หลายจนถึงปัจจุบัน ในขณะเดียวกัน ความต้องการระบบอัตโนมัติในสังคมมนุษย์เกิดขึ้น มนุษย์มีความจำเป็นต้องสั่งงานเครื่องจักรอัตโนมัติหรือคอมพิวเตอร์ จึงเกิดความจำเป็นในการสั่งงานโดยการพูดคุยกันหรือการสื่อสารระหว่างมนุษย์กับคอมพิวเตอร์ เรียกกันว่า อินเตอร์เฟสระหว่างมนุษย์กับคอมพิวเตอร์ (Human-machine Interface) ในอดีตอินเตอร์เฟสนี้เริ่มเกิดขึ้นในรูปแบบที่มนุษย์ต้องพิมพ์อักษรต่างๆ ที่แสดงเจตนาผ่านคีย์บอร์ดเพื่อสั่งงานคอมพิวเตอร์ ทำให้ผู้ใช้หลายคนจำเป็นต้องฝึกฝนการพิมพ์คีย์บอร์ดของคอมพิวเตอร์ซึ่งเหมือนแป้นพิมพ์ดีดให้เกิดความคล่องตัว แต่การพิมพ์ดีดไม่ใช่สิ่งปรกติที่มนุษย์กับมนุษย์คุยกัน มนุษย์ต้องการสิ่งที่สะดวกสบายที่จะคุยกับคอมพิวเตอร์โดยไม่ต้องฝึกฝนทักษะและเป็นธรรมชาติ นักวิจัยและพัฒนาจึงคิดค้นจอสัมผัส (Touch Screen) ที่ทำให้มนุษย์สามารถพูดคุยกับคอมพิวเตอร์โดยผ่านการสัมผัสภาพบนจอ ทำให้ผู้ใช้ไม่จำเป็นต้องฝึกฝนและจำตำแหน่งตัวอักษรต่างๆ ในแป้นพิมพ์ เทคโนโลยีจอสัมผัสทำให้เกิดความยืดหยุ่นในการสื่อสารระหว่างมนุษย์กับคอมพิวเตอร์มากยิ่งขึ้นในปัจจุบันแต่ก็ยังต้องอาศัยทักษะที่ต้องฝึกฝนบ้างและยังไม่ใช่สิ่งที่มนุษย์รู้สึกสะดวก

ตามปรกติมนุษย์พูดคุยกับมนุษย์ด้วยกันด้วยเสียง ภาษา กริยาท่าทาง เป็นต้น นักวิจัยและนักพัฒนางานเกี่ยวกับอินเตอร์เฟสระหว่างมนุษย์กับคอมพิวเตอร์จึงเริ่มนำเทคโนโลยีการประมวลผลสัญญาณเสียง การรู้จำเสียงมนุษย์ การประมวลผลภาษาธรรมชาติ การเข้าใจกริยาท่าทาง ฯลฯ มาประยุกต์กับระบบของอินเตอร์เฟสระหว่างมนุษย์กับคอมพิวเตอร์ การวิจัยและพัฒนาระบบรู้จำเสียงมีการพัฒนาเป็นเวลานาน มีการแก้ปัญหาในเชิงแอลกอริธึม และในปัจจุบันมีผลงานที่เริ่มนำมาประยุกต์ในงานจริงบ้างแล้ว ประกอบกับคอมพิวเตอร์มีการพัฒนาให้สามารถประมวลผลข้อมูลได้เร็วขึ้นมาก ทำให้การรู้จำเสียงและการสั่งงานด้วยเสียงกลายเป็นอินเตอร์เฟสระหว่างมนุษย์กับคอมพิวเตอร์ที่ใช้งานได้จริงและกำลังเป็นที่นิยมมากยิ่งขึ้น คาดกันว่า ในทางปฏิบัติการสั่งงานด้วยเสียงจะเข้ามาแทนการสั่งงานด้วยจอสัมผัสในชีวิตประจำวันของมนุษย์ในเร็วๆ นี้ เนื่องจากเป็นวิธีการสื่อสารที่สะดวกกับมนุษย์

อย่างไรก็ตาม ในเวลาพูดคุยและเวลาสั่งงาน มนุษย์ไม่ได้ใช้เฉพาะเสียงเท่านั้น มนุษย์ยังมีกริยาท่าทางจากอวัยวะต่างๆ โดยธรรมชาติ เช่น การชี้ด้วยนิ้ว [1] การใช้สายตาจ้อง การขยิบตา การส่ายศีรษะ เป็นต้น ซึ่งนับเป็นสารสนเทศที่สำคัญในการสื่อสาร นักวิจัยและพัฒนาที่เกี่ยวกับอินเตอร์เฟสระหว่างมนุษย์กับคอมพิวเตอร์จึงทำการวิจัยเพื่อค้นหาวิธีการสื่อสารด้วยกริยาท่าทางต่างๆ กริยาท่าทางเหล่านี้เป็นการแสดงเจตนาที่สำคัญอย่างหนึ่งของมนุษย์ วิธีการที่ใช้กริยาท่าทางเหล่านี้จะเสริมให้คอมพิวเตอร์และหุ่นยนต์สามารถสื่อสารกับมนุษย์ได้เยี่ยงมนุษย์กับมนุษย์โดยสมบูรณ์ ซึ่งนอกจากมนุษย์จะสามารถสื่อสารกับคอมพิวเตอร์และหุ่นยนต์ได้สะดวกยิ่งขึ้นแล้ว ยังทำให้คอมพิวเตอร์และหุ่นยนต์เข้าใจเจตนาของมนุษย์มากยิ่งขึ้น และทำให้หลีกเลี่ยงอุบัติเหตุที่เกิดจากการเข้าใจผิดในระหว่างการทำงานร่วมกันระหว่างมนุษย์กับหุ่นยนต์และคอมพิวเตอร์ [2] นอกจากนี้คนทุพพลภาพบางประเภท เช่น ทุพพลภาพทางสมองที่ไม่สามารถพูดหรือสัมผัสได้ ฯลฯ ก็สามารถใช้ประโยชน์จากฟังก์ชันนี้ในการสื่อสารกับคนปรกติได้เช่นกัน

งานวิจัยและพัฒนาความรู้พื้นฐานเกี่ยวกับการสื่อสารระหว่างมนุษย์กับมนุษย์เริ่มเข้าสู่สภาวะที่ใช้งานได้จริงและเป็นที่ยอมรับ ไม่ค่อยมีการเปลี่ยนแปลงในเชิงพื้นฐานนัก ในขณะที่การสื่อสารระหว่างมนุษย์กับคอมพิวเตอร์ยังอยู่ในสภาพที่ต้องอาศัยการวิจัยและค้นคว้าอีกมากเพื่อพัฒนาให้สามารถนำมาประยุกต์ในชีวิตประจำวันของมนุษย์ได้ นับเป็นโอกาสดีสำหรับนักวิจัยและเป็นความหวังของผู้ใช้ในการสร้างผลงานวิจัยที่เกี่ยวกับการเข้าใจกริยาท่าทางและเจตนามนุษย์เพื่อให้ระบบสื่อสารระหว่างมนุษย์กับคอมพิวเตอร์ในอนาคตมีความสมบูรณ์ขึ้น

เอกสารอ้างอิง

[1] S. Yeamkuan and K. Chamnongthai, "3D point-of-intention determination using a multimodal fusion of hand pointing and eye gaze for a 3D display," *SENSORS*, vol. 21, pp. 1–31, 2021.

[2] S. Yeamkuan, K. Chamnongthai, and W. Pichitwong, "A 3D point-of-intention estimation method using multimodal fusion of hand pointing, eye gaze and depth sensing for collaborative robots," *IEEE Sensors Journal*, vol. 22, no. 3, pp. 2700–2710, 2022.

ศาสตราจารย์ ดร.โกสินทร์ จำนงไทย
กองบรรณาธิการ