

เทคโนโลยี

ฉบับที่ 24,779 วันพุธที่ 16 สิงหาคม พ.ศ. 2560 หน้า 25

รอบรู้ไอที รอบโลกเทคโนโลยี

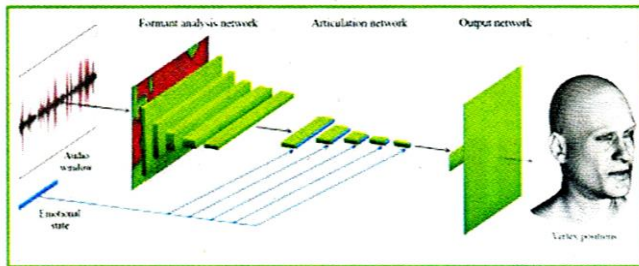
สร้างโมเดลหน้าคนสามมิติจากการเรียนรู้เชิงลึก

ถ้าคุณผู้อ่านประจำคอลัมน์วันพุธของผมพอจะจำได้ว่าผมเคยเขียนถึงเรื่อง Deep Learning หรือ การเรียนรู้เชิงลึก ที่ในปัจจุบันเริ่มมีการนำมาประยุกต์ใช้ในหลากหลายวงการมากขึ้น

ซึ่งถ้าคุณผู้อ่านจำได้ผมเคยเขียนถึง การประกวดความงามของใบหน้ามนุษย์ที่นำเทคนิคการเรียนรู้เชิงลึกมาตัดสินในงานประกวด Beauty.AI ซึ่งจัดขึ้นครั้งแรกไปเมื่อต้นปีที่แล้ว (ค.ศ. 2016) โดยการแข่งขันของบริษัทไมโครซอฟท์และบริษัทเอ็นวีเดีย (NVIDIA) ซึ่งก็ถือว่าเป็นการประกวดที่นำเทคนิคการเรียนรู้เชิงลึกมาตัดสินความงามของมนุษย์เป็นครั้งแรกครั้งหนึ่งของโลกเลยทีเดียว

แต่มาวันนี้ในปี ค.ศ. 2017 บริษัทเอ็นวีเดียยังไม่หยุดเล่นกับเทคนิคการเรียนรู้เชิงลึกอยู่แค่นั้นครับ เพราะที่งานประชุมวิชาการระดับนานาชาติ SIGGRAPH งานประชุมวิชาการด้านคอมพิวเตอร์กราฟิกและการได้คอบที่มีชื่อเสียงของโลกที่จัดขึ้นที่เมืองลอสแอนเจลิส รัฐแคลิฟอร์เนีย สหรัฐอเมริกา ในระหว่างวันที่ 30 กรกฎาคม ถึง 3 สิงหาคม ค.ศ. 2017 ที่ผ่านมานี้

มีงานวิจัยจากบริษัทเอ็นวีเดียนำทีมโดย Tero Karras ที่ได้ใช้เทคนิคการเรียนรู้

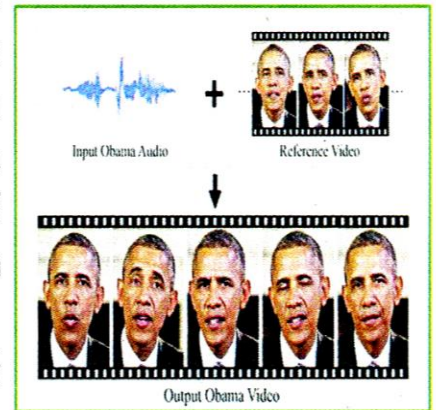


นำเทคนิคการเรียนรู้เชิงลึกมาช่วยในการแปลงวิดีโอคนจริงมาเป็นโมเดลหน้าสามมิติ (ภาพ www.2017.siggraph.org)

เชิงลึกเข้ามาช่วยในการแก้ปัญหาในการแปลงวิดีโอมนุษย์คนจริงมาเป็นโมเดลหน้าสามมิติที่ทำได้ค่อนข้างน่าสนใจมาก โดยข้อมูลที่ใช้เทรนสอนระบบในการเรียนรู้เชิงลึกนั้น เป็นวิดีโอยาวเพียงไม่กี่นาทีเท่านั้นครับ แต่ผลลัพธ์โมเดลสามมิติที่ได้นั้นมีความถูกต้องดีพอสมควรเลย

เพราะในอดีตมาจนถึงปัจจุบันนั้น วิธีที่นิยมใช้ในการทำใบหน้าขณะพูดของตัวละครหลักหรือตัวละครรองในภาพยนตร์และเกมคอมพิวเตอร์กราฟิกให้ดูคล้ายกับมนุษย์จริงนั้น จะทำโดยการใช้ต้นแบบเป็นวิดีโอของนักแสดงที่เป็นมนุษย์จริง ๆ มาผ่านซอฟต์แวร์ในการแปลงเป็นใบหน้าโมเดลสามมิติ แล้วก็ค่อย ๆ ให้ทางทีมงานมาทำการปรับปรุงแก้ไขโมเดลใบหน้าสามมิตินี้ให้เหมือนกับหน้านักแสดงที่เป็นมนุษย์จริงให้มากที่สุด

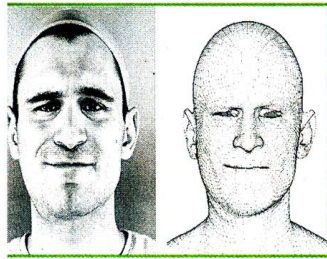
ซึ่งต้องบอกว่าขั้นตอนพวกนี้จะใช้



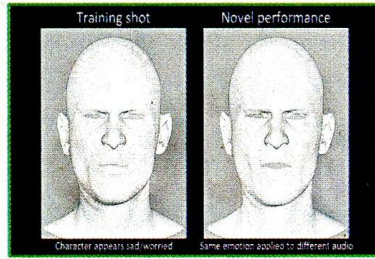
อีกงานวิจัยหนึ่งในงาน SIGGRAPH 2017 ที่ใช้การวิเคราะห์เสียงไว้ถ้ามีโอกาสจะมาเล่าขยายความให้ในครั้งต่อไป (ภาพ www.s2017.siggraph.org)

ระยะเวลาและทรัพยากรค่อนข้างสูง ซึ่งงานวิจัยของบริษัทเอ็นวีเดียนี้ก็ทำการแก้ปัญหาโดยการนำเทคนิคการเรียนรู้เชิงลึกมาใช้เทรนสอนระบบ ทำให้ใช้ระยะเวลาในการเทรนเหลือเพียงแค่ประมาณ 3-5 นาทีเท่านั้น

แถมนอกจากการสร้างโมเดลใบหน้าสามมิติโดยข้อมูลที่ใช้เทรนสอนระบบนั้น



โมเดลหน้าสามมิติที่สร้างขึ้น
(ภาพ www.s2017.siggraph.org)



ข้อมูลที่ใช้เทรนระบบนี้เป็นวิดีโอชาวเอเชีย
ไม่กี่นาทีเท่านั้น (ภาพ www.s2017.siggraph.org)

ซึ่งงานวิจัย
แนวนี้เชื่อกันว่าจะ
สามารถทำให้ต้นทุน
ในการสร้างโมเดล
ใบหน้าของเกม
คอมพิวเตอร์กราฟิกมี
ราคาที่ถูกลงมาก รวม

เป็นวิดีโออย่างที่กล่าวไปแล้วนั้น ทีม
วิจัยของบริษัทเอ็นวีเดียยังนำเสนอถึง
การใช้เสียงพูดของคนเข้ามาช่วยใน
การเทรนระบบด้วย ซึ่งทำโดย
การนำกราฟคลื่นเสียงมาเชื่อมโยงกับ
จุดตำแหน่งต่าง ๆ ที่มีการขยับของ
โมเดลใบหน้า เพื่อให้ใบหน้านั้น
สามารถขยับตามเสียงพูดได้อย่างทันที
ทันใด

โดยเขาได้ทดลองกับคนทั้ง
เพศชายและหญิง รวมไปถึงการออก
เสียงในสำเนียงและภาษาต่าง ๆ ที่แตกต่างกัน
ตั้งแต่ภาษาอังกฤษ และ ภาษาฝรั่งเศส
ไปจนถึงภาษาเยอรมัน ภาษาอิตาลี และ
ภาษาสเปน เป็นต้น เพราะฉะนั้นการสร้าง
ใบหน้าของตัวละครที่พูดในภาษาต่าง ๆ ใน
ภาพยนตร์หรือเกมคอมพิวเตอร์กราฟิกที่ต้อง
มีเสียงพากย์หลายภาษาก็จะสามารถทำได้
ง่ายขึ้น

ไปถึงไม่จำเป็นต้องจ้างนักแสดงมาเล่นเป็น
ตัวละครแต่ละตัวละครให้ครบทุกตัว แต่ใช้
ปัญญาประดิษฐ์เข้ามาช่วยวาดใบหน้าของตัว
ละครที่อาจมีจำนวนหลายตัวละครได้
เห็นไหมครับว่ารอบโลกเทคโนโลยี
สารสนเทศในศตวรรษที่ 21 ยังมีนวัตกรรม
สมัยใหม่ให้นำศึกษา นำค้นหา และนำมา
มาพัฒนาต่อยอดอยู่อย่างเป็นพลวัตจริง ๆ
พวกเราคนไทยเองก็อย่าตกขบวนรถไฟ
แห่งการวิจัยและพัฒนาพวกนี้นะครับ.

พศ.ดร.ชุตินันต์ เกดวิบูลย์เวช
สถาบันนวัตกรรมและบริการศาสตร์ (นิด้า)
chutisant.ker@nida.ac.th