

โครงร่างวิทยานิพนธ์

ชื่อเรื่อง (ภาษาไทย)	เฟรมเวิร์กของส่วนต่อประสานกับผู้ใช้ผ่านศีรษะมนุษย์เสมือนโดยใช้โมโนสโคปิกฟิชแทงก์วีอาร์
ชื่อเรื่อง (ภาษาอังกฤษ)	USER INTERFACE FRAMEWORK USING VIRTUAL HUMAN HEAD WITH MONOSCOPIC FISH-TANK VR
เสนอโดย	นาย อรรถกร พูนศิลป์
รหัสประจำตัว	457 06407 21
สาขาวิชา / ภาควิชา	วิศวกรรมคอมพิวเตอร์ / วิศวกรรมคอมพิวเตอร์
คณะ	วิศวกรรมศาสตร์
สถานที่ติดต่อ	ห้อง 17-05 ชั้น 17 ตึก 4 คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
โทรศัพท์	09-677-6966
E-Mail	g45aps@cp.eng.chula.ac.th
อาจารย์ที่ปรึกษาวิทยานิพนธ์	อ.ดร.อรรถวิทย์ สุขแสง attawith@cp.eng.chula.ac.th

บทคัดย่อ

งานวิจัยนี้มีจุดประสงค์เพื่อสร้างระบบต้นแบบสำหรับติดต่อกับผู้ใช้ (User Interface) ในลักษณะโมเดลสามมิติของศีรษะมนุษย์ที่เน้นให้มีลักษณะเหมือนจริงมากที่สุด ซึ่งโมเดลนี้จะได้รับการปรับเปลี่ยนมุมมองให้สัมพันธ์กับตำแหน่งดวงตาของผู้ใช้ได้แบบเวลาจริง ดวงตาของโมเดลสามารถมองมายังผู้ใช้งาน รวมถึงการขยับปากเพื่อให้สอดคล้องกับคำพูดที่จะพูดกับผู้ใช้ โดยให้ผู้ใช้งานเกิดความรู้สึกเสมือนว่ามีมนุษย์ที่สามารถสื่อสารได้อยู่ภายในจอภาพคอมพิวเตอร์จริงๆ

คำสำคัญ (ภาษาไทย)	การแสดงผลภาพนามธรรมสามมิติ, ส่วนต่อประสานกับผู้ใช้, การประมวลผลภาพ, การมองเห็นด้วยคอมพิวเตอร์, โอเพนจีแอล, ความจริงเสมือน
คำสำคัญ (ภาษาอังกฤษ)	3-D Visualization, User Interface, Image Processing, Computer Vision, OpenGL, Virtual Reality

1. ความเป็นมาและความสำคัญของปัญหา

ระบบคอมพิวเตอร์ตั้งแต่อดีตมาจนถึงปัจจุบันได้มีพัฒนาการอย่างมากมาย ทั้งความเร็วในการประมวลผล หน่วยความจำ รวมถึงระบบเพิ่มขยายต่างๆ เพื่อเพิ่มความสามารถให้มากยิ่งขึ้น แต่สำหรับการปฏิสัมพันธ์ (User Interface) กับผู้ใช้นั้น ระบบส่วนใหญ่ยังคงกำหนดให้ใช้แป้นพิมพ์ และเมาส์ในการป้อนข้อมูลเป็นหลัก ทำให้ในบางกรณีผู้ใช้ไม่สามารถใช้งานได้อย่างสะดวกและเป็นธรรมชาติเท่าที่ควร เช่น กรณีของผู้ที่ไม่คุ้นเคยกับคีย์บอร์ดและเมาส์ หรือคนพิการ นอกจากนี้ เมื่อผ่านการประมวลผลแล้วระบบคอมพิวเตอร์ก็จะแสดงผลลัพธ์ออกมาในรูปแบบตัวอักษร ภาพสองมิติ หรือภาพสามมิติที่ส่วนใหญ่ไม่มีปฏิสัมพันธ์ใดๆ กับผู้ใช้ จึงทำให้การนำเสนอขาดความน่าสนใจ

ในปี 1968 Ivan Sutherland [12] ได้พัฒนางานวิจัยต้นแบบของ Head-mounted display (จอภาพแบบสวมศีรษะ) ซึ่งมีชื่อเสียงมาก เพราะสามารถแสดงผลโลกเสมือนจริงสามมิติที่มีปฏิสัมพันธ์กับผู้ใช้ได้อย่างเป็นธรรมชาติ ทำให้ศาสตร์ทางด้านความจริงเสมือน (Virtual Reality) มีความเป็นไปได้ในทางปฏิบัติมากขึ้น แต่อย่างไรก็ตาม อุปกรณ์นี้ยังมีอุปสรรคและข้อจำกัดในด้านการใช้งานค่อนข้างมาก เนื่องมาจากความละเอียดต่ำ อุปกรณ์มีราคาแพง และหาซื้อได้ยาก นอกจากนี้ Sutherland ได้รับการยอมรับว่าเป็นคนแรกที่นำศัพท์คำว่า Virtual Reality (ความจริงเสมือน) มาใช้ ซึ่งมีความหมายว่าเป็นระบบการแสดงผลภาพสามมิติที่มีคุณภาพเพียงพอที่จะกระตุ้นการรับรู้ของผู้ใช้ให้มีความรู้สึกเสมือนว่าอยู่ในสิ่งแวดล้อมนั้นจริงๆ



รูปที่ 1 อุปกรณ์ Head-mounted display [15]

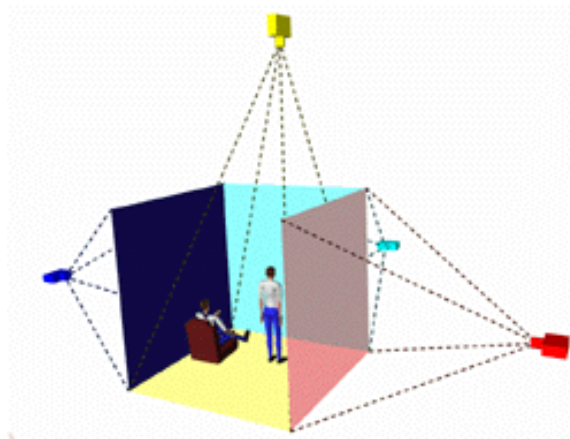
ต่อมาเมื่อต้นปี 1993 Colin Ware [14] ได้นำเสนอระบบความจริงเสมือนรูปแบบใหม่ที่แสดงผลผ่านจอมอนิเตอร์ธรรมดาที่หาซื้อได้ทั่วไป โดยใช้ระบบการแสดงผลแบบ Stereoscopic (แสดงภาพสองภาพพร้อมๆ กัน สำหรับตาสองข้าง ใช้ร่วมกับแว่นสามมิติ) และใช้ระบบการติดตามตำแหน่งศีรษะของผู้ใช้ (Head Tracking) มาร่วมด้วย เพื่อปรับเพอซเพกทีฟ (Perspective) สำหรับการแสดงผลภาพสามมิติให้สัมพันธ์กับตำแหน่งดวงตาของผู้ใช้ในขณะนั้น ระบบนี้ผู้พัฒนาได้ตั้งชื่อว่า Fish Tank VR เนื่องจากมีลักษณะคล้ายกับการมองเข้าไปในตู้ปลาซึ่งเปรียบได้กับโลกความจริงเสมือนนั่นเอง

ด้วยการแสดงผลจากจอมอนิเตอร์ของคอมพิวเตอร์ Desktop PC ธรรมดา ทำให้ระบบนี้มีข้อดีตรงที่มีความละเอียดและความคมชัดในการแสดงผลอย่างมาก แสดงภาพเคลื่อนไหวได้ดี ในขณะที่มีค่าใช้จ่ายน้อยมากหากเทียบกับระบบแบบเก่า อุปกรณ์ต่างๆ สามารถหาได้ทั่วไป และมีค่าใช้จ่ายในการบำรุงรักษาที่ต่ำ



รูปที่ 2 ตัวอย่างระบบ Fish Tank VR [16][17]

ต่อมาในเวลาไล่เลี่ยกัน เมื่อปลายปี 1993 มีนักวิจัยอีกกลุ่มได้เสนอระบบความจริงเสมือนในอีกรูปแบบหนึ่ง ตั้งชื่อว่า CAVE (CAVE Automatic Virtual Environment) [3] เป็นระบบที่ใช้ Projector ฉายภาพเวดล้อมสามมิติ แบบ Stereo ลงบนฉากรับภาพ ที่ตั้งฉากกันเป็นรูปกล่องลูกบาศก์เปรียบเสมือนผนังของห้อง มีขนาดเท่ากับห้องๆ หนึ่ง สามารถรองรับได้หลายคน โดยแต่ละคนต้องสวมแว่นตาสามมิติ และสามารถเดินไปเดินมาได้โดยอิสระภายใน ห้อง ซึ่งมีระบบตรวจจับตำแหน่งของศีรษะเพื่อปรับภาพสามมิติให้สอดคล้องกับมุมมองของผู้นำกลุ่มอีกด้วย แบบจำลองระบบสามารถแสดงได้ดังรูปที่ 3



รูปที่ 3 แบบจำลองระบบ CAVE [18]

ระบบ CAVE มีข้อดีเพิ่มขึ้นมาหลายอย่าง เช่น มีมุมมองที่กว้างขึ้น มีความละเอียดสูงกว่าแบบ Head-mounted display (ความละเอียดเทียบเคียงได้กับ HDTV หรือสองเท่าของ HDTV) รองรับผู้ใช้ได้หลายคน และเนื่องจากว่าฉากรับภาพแบนราบ รวมถึงมีระบบ Geometric Correction จึงทำให้ภาพที่ได้มีการบิดเบี้ยว (Distortion) ที่น้อยมากถ้าเทียบกับระบบอื่นๆ

จากลักษณะต่างๆ ของระบบ Virtual Reality ที่กล่าวมา Demiralp และ Laidlaw [5] ได้ทำการทดลองเปรียบเทียบข้อดีข้อเสีย ของระบบ CAVE เทียบกับ Fish Tank VR ในแง่ของประโยชน์ในการใช้งานจริง โดยมีผู้เข้า

ทำการทดสอบที่มาจากหลากหลายสาขาวิชา ทดลองใช้โปรแกรมสามมิติในสาขานั้นๆ แล้ววัดความพึงพอใจ ผลที่ได้คือผู้ใช้ถึงสามในสี่พึงพอใจกับระบบ Fish Tank VR มากกว่า ด้วยเหตุผลดังนี้

- Fish Tank VR ให้ภาพที่คมชัดและสว่างมากกว่าระบบ CAVE
- Fish Tank VR มีพื้นที่การแสดงผลเล็กกว่า จึงมองเห็นองค์รวมและความสัมพันธ์ระหว่างวัตถุต่างๆ ได้ง่ายกว่า
- Fish Tank VR ให้ความสะดวกสบายในการใช้งานมากกว่า ไม่จำเป็นต้องยืนและเดินไปเดินมาเหมือนระบบ CAVE

จากสิ่งที่ได้กล่าวมาทั้งหมด ทำให้ระบบ Virtual Reality บนพีซีธรรมดาที่สามารถสร้างสภาพแวดล้อมที่น่าประทับใจและเป็นธรรมชาติแก่ผู้ใช้นั้นมีความเป็นไปได้ ผู้วิจัยจึงเกิดแนวคิดว่าจะมีระบบติดต่อกับผู้ใช้ (User Interface) แนวใหม่ที่แสดงผลอยู่บนพื้นฐานของ Fish Tank VR เพราะจะทำให้ได้เปรียบระบบติดต่อกับผู้ใช้แบบเก่าซึ่งเป็นแบบ 2 มิติ หรือ 3 มิติทั่วไปอยู่ค่อนข้างมาก

เมื่อพิจารณาถึงระบบติดต่อกับผู้ใช้แบบดั้งเดิมที่ประกอบไปด้วยเมนูคำสั่ง ทูลบาร์ต่างๆ ผู้ใช้ส่วนใหญ่ที่ยังไม่คุ้นเคยกับคอมพิวเตอร์หรือไม่เคยใช้คอมพิวเตอร์มาก่อนมักจะประสบปัญหาในการใช้โปรแกรม ผู้ใช้งานคนอาจจะไม่เข้าใจว่าสิ่งต่างๆ ศัพท์เทคนิคต่างๆ ที่ปรากฏบนหน้าจอคืออะไร และมีหน้าที่อะไร บางคนก็อาจจะไม่คุ้นเคยกับการใช้คีย์บอร์ด ทำให้ล่าช้าในการป้อนข้อมูล สิ่งต่างๆ เหล่านี้เป็นอุปสรรคที่ส่งผลทำให้ต้องใช้เวลาพอสมควรในการเรียนรู้ว่าจะใช้งานได้

กรณีหนึ่งที่เห็นได้ชัดคือโครงการต่างๆ ที่จัดซื้อคอมพิวเตอร์ให้กับโรงเรียนในชนบท มีวัตถุประสงค์ให้เด็กนักเรียนในชนบทได้มีโอกาสเรียนรู้คอมพิวเตอร์และสามารถเข้าถึงข้อมูลในอินเทอร์เน็ต แต่เมื่อระบบต่างๆ ติดตั้งพร้อมแล้ว ก็เกิดปัญหาครูอาจารย์ใช้งานไม่เป็น หรือใช้งานได้เพียงแค่อ้างอิง เด็กนักเรียนก็ไม่สามารถใช้งานได้เนื่องจากอ่านยังไม่คล่อง หรือเจอสัญลักษณ์แปลกๆ ที่ไม่ทราบความหมาย ไม่ทราบว่าต้องเอาเมาส์ไปกดตรงไหนจึงจะให้คอมพิวเตอร์ทำงานได้ตามต้องการ บางคนก็อาจจะเลิกเล่นความตั้งใจและท้อแท้ และเบื่อกับที่จะใช้คอมพิวเตอร์ไปเลย

จะดีสักเพียงใด หากเปิดคอมพิวเตอร์แล้วมีใบหน้าผู้หญิงปรากฏขึ้นมาบนจอ (รูปที่ 4) พร้อมกับทักทายว่า “สวัสดีจ๊ะน้อง..... วันนี้ดูสิหน้าน้อง..... ไม่ดีเลย มีอะไรไม่สบายใจหรือเปล่า เลี้ยวที่เปิดเพลงให้ฟังนะ” แล้วโปรแกรมฟังเพลงก็ถูกเรียกขึ้นมาทำงานอัตโนมัติ

หรือ

“เมื่อวานพิมพ์งานค้างไว้เนี่ย วันนี้จะทำต่อหรือเปล่านั้น จะได้เรียกโปรแกรมพิมพ์งานให้”

ผู้ใช้อาจสั่งงานคอมพิวเตอร์ได้ เช่น

“คอมพิวเตอร์ ช่วยช่วยดูรูปถ่ายไปเนี่ย แล้วบันทึกไว้หน่อย” จากนั้นใบหน้าบนจอก็เหลือบมองไปยังรูปถ่ายและบันทึกภาพนั้นเก็บไว้โดยผ่านกล้องที่ติดอยู่บนจอภาพ

จากสิ่งต่างๆ ที่กล่าวมา ผู้วิจัยจึงได้เกิดแนวคิดที่จะออกแบบและวางรากฐานส่วนติดต่อกับผู้ใช้บนคอมพิวเตอร์พีซีเสียใหม่ให้มีความเสมือนจริง เป็นธรรมชาติ และมีความน่าสนใจ ผู้ใช้ไม่จำเป็นต้องใช้เวลาเรียนรู้การใช้งาน เพราะผู้ใช้สามารถสื่อสารกับคอมพิวเตอร์ได้โดยไม่ต้องทำอะไรกับการสื่อสารกับมนุษย์ด้วยตนเอง

แต่การที่จะทำให้ระบบรับรู้ เข้าใจในภาษาพูด สีหน้าและท่าทางของผู้ใช้นั้นยังมีอุปสรรคเนื่องด้วยปัจจัยต่างๆ เช่นความเร็วในการประมวลผล เทคโนโลยีปัญญาประดิษฐ์ การประมวลผลภาษาธรรมชาติ และการประมวลผลภาพดิจิทัลที่ในปัจจุบันยังมีข้อจำกัดอยู่มาก แต่ในอนาคตเมื่อเทคโนโลยีเหล่านี้มีความก้าวหน้าเพียงพอ ก็สามารถนำมารวมกับงานวิจัยชิ้นนี้เพื่อสร้างเป็นระบบการติดต่อกับผู้ใช้ที่สมบูรณ์แบบ

เพราะฉะนั้น ในงานวิจัยนี้จึงเน้นเฉพาะส่วนของการพัฒนาโมเดลสามมิติของศีรษะมนุษย์ให้สามารถมีปฏิกิริยาต่างๆ ต่อผู้ใช้ได้ เช่น รับรู้สภาพแวดล้อมรอบๆ ผู้ใช้ มองไปยังผู้ใช้ ขยับปากสนทนากับผู้ใช้ตามสคริปต์ที่กำหนดไว้ก่อน รวมถึงการเปลี่ยนมุมมองของโมเดลให้สัมพันธ์กับตำแหน่งดวงตาของผู้ใช้ในแบบเวลาจริงโดยนำ Fish Tank VR มาประยุกต์ใช้ ซึ่งจะทำให้ผู้ใช้เกิดความรู้สึกเสมือนว่ามีศีรษะมนุษย์อยู่ภายในจอมอนิเตอร์จริงๆ

ด้วยส่วนติดต่อกับผู้ใช้แบบนี้ จะช่วยเพิ่มความน่าสนใจในการใช้คอมพิวเตอร์ได้อีกมาก การใช้โมเดลศีรษะมนุษย์จะมีผลต่อจิตใจทำให้ผู้ใช้รู้สึกว่าคอมพิวเตอร์เป็นมิตรมากขึ้นและสามารถเป็นเพื่อนกับเราได้ (โดยเฉพาะเด็กๆ) ทำให้ชอบที่จะใช้งาน และไม่รู้สึกเบื่อ



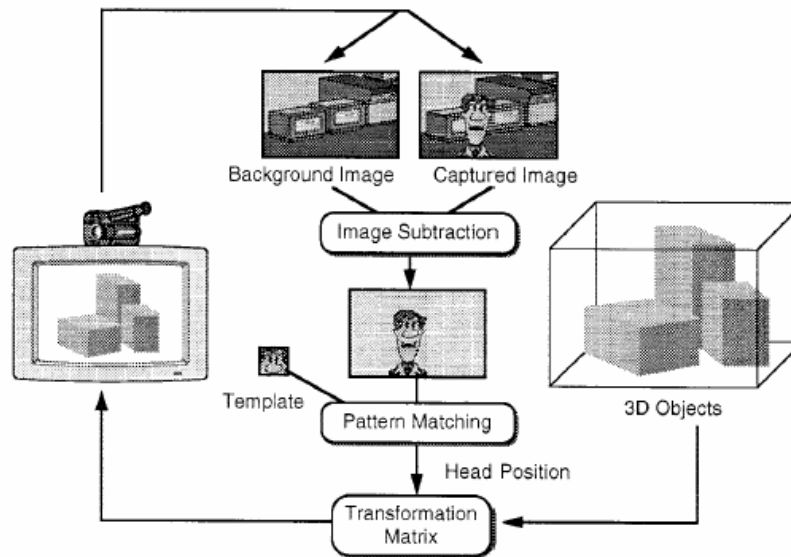
รูปที่ 4 แนวคิดในการนำโมเดลศีรษะมนุษย์มาใช้ในการติดต่อกับผู้ใช้

2. งานวิจัยที่เกี่ยวข้อง

Jun Rekimoto [10] ได้สร้างระบบ Fish Tank Virtual Reality ขึ้นมาในปี 1995 โดยใช้กล้องจับภาพผู้ใช้เพื่อติดตามตำแหน่งของศีรษะในแบบเวลาจริง แล้ววัดประสิทธิภาพในการทำงานและการแสดงผลแบบสามมิติบนจอภาพสำหรับการประมวลผลเพื่อติดตามตำแหน่งของศีรษะ (Head Tracking) นั้น แบ่งออกเป็นสองขั้นตอนคือ

- 1.) ค้นหาบริเวณที่เป็นใบหน้าผู้ใช้ โดยจะเก็บภาพฉากหลัง (background) ไว้ก่อน เมื่อเริ่มประมวลผลจึงลบภาพฉากหลังออกจากภาพผู้ใช้ที่ได้รับมาจากกล้องฟิสิกส์ต่อฟิสิกส์ เพื่อให้มีความแม่นยำที่ดีขึ้น ได้เลือกใช้ระบบสีแบบ YUV ในการประมวลผล ทำให้ได้ภาพบริเวณผู้ใช้ที่ชัดเจนถึงแม้จะมีฉากหลังที่ซับซ้อน
- 2.) ระบบจะค้นหาตำแหน่งกึ่งกลางใบหน้าโดยใช้การจับคู่กับแม่แบบ (Template Matching) บริเวณบางส่วนของใบหน้าผู้ใช้จะถูกเก็บไว้ก่อนในรูปของแม่แบบ เมื่อได้รับภาพบริเวณใบหน้าผู้ใช้เข้ามา ระบบจะคำนวณ Correlation Coefficients ระหว่างแม่แบบกับทุกๆ บริเวณบนใบหน้า แล้วเลือกจุดที่ได้ค่าสูงที่สุดเป็นตำแหน่งกึ่งกลางใบหน้า

เมื่อได้ตำแหน่งกึ่งกลางของใบหน้าแล้ว จึงนำไปคำนวณหาตำแหน่งในสามมิติ (x, y, z) ของผู้ใช้ที่เทียบกับกล้อง โดยมีข้อกำหนดว่าผู้ใช้ต้องมีระยะห่างกับกล้องที่คงที่ หลังจากนั้นระบบจะนำพิกัดที่ได้ไปปรับ Perspective และตำแหน่งกล้องเพื่อฉาย โมเดลสามมิติให้สัมพันธ์กับตำแหน่งของผู้ใช้ การทำงานของโปรแกรมทั้งหมดแสดงได้ดังรูปที่ 5

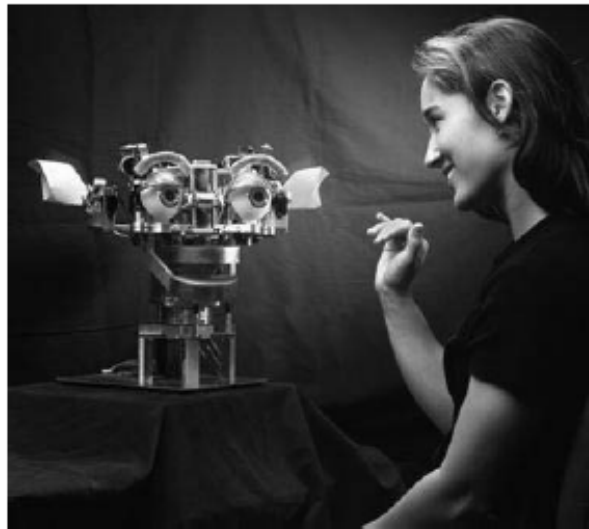


รูปที่ 5 แสดงการทำงานของโปรแกรมใน [10]

ภาพโมเดลสามมิติที่ใช้ทดสอบความลึกนั้น เป็นภาพโมเดลต้นไม้ที่มีกิ่งก้านจำนวนสามต้นที่วางไว้ที่ความลึกแตกต่างกัน โดยให้ผู้ใช้ทดสอบทำการแยกแยะความลึกของต้นไม้ในส่วนต่างๆ ผลการทดลองปรากฏออกมาว่า เมื่อใช้ระบบการติดตามตำแหน่งของศีรษะเข้ามาช่วยด้วย สามารถเพิ่มความถูกต้องในการทำแบบทดสอบ โดยมีความผิดพลาดเพียงแค่ 5% เมื่อเทียบกับการไม่ใช้ตำแหน่งของศีรษะซึ่งมีความผิดพลาดสูงถึง 21.3%

ทางผู้วิจัยที่ทำงานวิจัยชิ้นนี้มีความเชื่อว่ากล้องที่ติดอยู่บนจอภาพจะกลายเป็นอุปกรณ์รับข้อมูลเข้ามาตรฐานแบบที่สาม (นอกเหนือจากแป้นพิมพ์และเมาส์) ในอนาคตอันใกล้ เนื่องจากมีประโยชน์อย่างมหาศาลในด้านต่างๆ โดยเฉพาะศาสตร์ทางด้านการปฏิสัมพันธ์ระหว่างคอมพิวเตอร์และผู้ใช้ (Human-computer interaction)

นอกจากนี้ ยังมีงานวิจัยอีกชิ้นหนึ่งที่คล้ายคลึงกับงานวิจัยที่จะนำเสนอ งานวิจัยชิ้นนี้เป็นการสร้างหุ่นยนต์ (เฉพาะส่วนของศีรษะและใบหน้า) สามารถแสดงออกทางอารมณ์บนใบหน้า มีปฏิสัมพันธ์กับผู้ใช้ได้ในแบบต่างๆ กัน จุดมุ่งหมายหลักคือต้องการให้เป็นหุ่นยนต์ที่สามารถสื่อสารกับมนุษย์ได้อย่างเป็นธรรมชาติมากที่สุด และสามารถเป็นเพื่อนเล่นกับมนุษย์ได้ พัฒนาโดย Cynthia Breazeal [1] [2] ซึ่งหุ่นยนต์มีชื่อว่า “Kismet” มีลักษณะดังรูปที่ 6



รูปที่ 6 หุ่นยนต์ Kismet สามารถมีปฏิสัมพันธ์กับผู้ใช้ผ่านการแสดงออกทางใบหน้า [1][2]

มนุษย์สามารถสื่อสารกับ Kismet ได้โดยใช้เสียงพูดและท่าทางประกอบ เพราะฉะนั้นผู้ใช้งานไม่จำเป็นต้องเรียนรู้วิธีการใช้ใดๆ ก็สามารถสื่อสารกับ Kismet ได้อย่างมีประสิทธิภาพ

สำหรับการออกแบบ หุ่น Kismet มีกล้องและไมโครโฟนอย่างละสองตัว (Stereo) , มีระบบการสังเคราะห์เสียงพูด (Speech Synthesizer) , และอวัยวะบนใบหน้าเพื่อแสดงอารมณ์ในแบบต่างๆ กัน ซึ่งจะมีลักษณะการแสดงออกทางอารมณ์และความสามารถเทียบได้กับเด็กทารกอายุหกเดือน

3.) วัตถุประสงค์

1. เพื่อสร้างกรอบการทำงาน (Framework) สำหรับระบบติดต่อกับผู้ใช้ในลักษณะโมเดลสามมิติของศีรษะมนุษย์ที่เน้นให้ผู้ใช้งานสามารถใช้งานได้อย่างเป็นธรรมชาติ
2. เพื่อให้คอมพิวเตอร์สามารถสื่อสารกับมนุษย์ได้อย่างมีประสิทธิภาพมากยิ่งขึ้น โดยคอมพิวเตอร์สามารถแสดงออกโดยใช้ภาษาท่าทาง เช่น การใช้สายตา การยิ้ม เป็นต้น

3. เพื่ออำนวยความสะดวกให้กับผู้ที่ต้องการสร้างระบบติดต่อกับผู้ใช้ที่ผู้ใช้สามารถใช้งานได้ง่ายและไม่จำเป็นต้องใช้เวลาในการศึกษาเรียนรู้

4.) แนวคิดและวิธีการวิจัย

ภาพจำลองของงานวิจัยนี้แสดงโดยรูปด้านล่าง



รูปที่ 7 แสดงภาพจำลองของการนำ Fish Tank VR มาประยุกต์ใช้ในงานวิจัย

จากรูปจะเห็นว่าภาพสามมิติถูกแสดงผลโดยขึ้นอยู่กับมุมมองหรือตำแหน่งของผู้ใช้ ยกตัวอย่างเช่นถ้าหากผู้ใช้เคลื่อนที่ไปทางขวา ระบบก็ควรจะปรับเปลี่ยนมุมในการแสดงให้เห็นด้านซ้ายของโมเดล โดยที่มุมในการแสดงผลโมเดลต้องสอดคล้องกับมุมที่ผู้ใช้กระทำกับจอภาพ (เปรียบเสมือนจอภาพเป็นหน้าต่าง ที่มีศีรษะมนุษย์อยู่ภายใน)

เนื่องจากระบบจะต้องรับรู้ตำแหน่งศีรษะของผู้ใช้เทียบกับจอภาพ ในงานวิจัยจึงได้นำกล้องวงจรปิดติดตั้งไว้บนจอมอนิเตอร์เพื่อรับภาพผู้ใช้เข้าไปประมวลผล เมื่อได้ตำแหน่งศีรษะแล้วจึงปรับมุมมองและเพอสเปกทิฟของโมเดลสามมิติให้สัมพันธ์กัน เพราะฉะนั้นจะอธิบายแนวคิดแยกเป็นสามส่วน ส่วนแรกคือการตรวจหาตำแหน่งศีรษะของผู้ใช้ในภาพที่ได้จากกล้อง (x, y) เพื่อคำนวณหาตำแหน่งของผู้ใช้ในรูปแบบพิกัดสามมิติที่เทียบกับจอภาพ (x, y, z) ส่วนที่สองจะกล่าวถึงการนำตำแหน่งพิกัดสามมิติที่ได้ตั้งเป็นตำแหน่งศูนย์กลางของการโปรเจกชัน (Center of Projection) และการปรับเพอสเปกทิฟใน OpenGL ให้สัมพันธ์กับผู้ใช้งาน สำหรับส่วนที่สามจะกล่าวถึงการนำระบบนี้ไปสร้างการปฏิสัมพันธ์ระหว่างคอมพิวเตอร์และผู้ใช้งาน (Human-Computer Interaction) ในรูปแบบต่างๆ

4.1) การตรวจหาตำแหน่งผู้ใช้เทียบกับจอภาพ

การตรวจหาตำแหน่งในระบบพิกัดสามมิติของผู้ใช้จากภาพดิจิทัลสองมิติสามารถทำได้โดยใช้เทคนิค Camera Calibration [8] ซึ่งเกี่ยวข้องกับพารามิเตอร์ของกล้องแบบ Intrinsic และ Extrinsic ดังนี้

Intrinsic camera parameters คือค่าที่บ่งบอกถึงลักษณะเฉพาะของกล้องนั้นๆ ซึ่งได้แก่

- ความยาวโฟกัส (Focal length) ซึ่งก็คือระยะระหว่างเลนส์ของกล้องและฉากรับภาพ
- ตำแหน่งของจุดกึ่งกลางภาพในระบบพิกัดจุด (Principal point)
- ขนาดของแต่ละจุดภาพ (Effective pixel size)
- สัมประสิทธิ์การบิดเบี้ยวในแนวรัศมีของเลนส์ (Radial distortion coefficient)

Extrinsic camera parameters คือค่าที่บ่งบอกถึงความสัมพันธ์เชิงระยะทาง (Spatial Relationship) ระหว่างกล้อง (Camera) กับโลกภายนอก (World) ได้แก่

- Rotational Matrix คือ เมตริกซ์การหมุน
- Translation Vector คือ เวกเตอร์การเคลื่อนที่

ถ้าหากใช้กล้องแบบรูเข็มต่างๆ ไปแล้ว จุดในพิกัดสามมิติ M และจุดที่ปรากฏบนภาพ m เป็นไปตามความสัมพันธ์ดังนี้

$$m = A[Rt]M$$

โดยที่ A คือเมตริกซ์ของ Intrinsic parameters:

$$A = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}, \text{ เมื่อ}$$

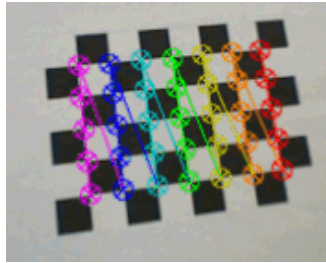
(c_x, c_y) คือพิกัดของจุดกึ่งกลางภาพ หรือ Principal point

(f_x, f_y) คือระยะโฟกัสบนแกน x และ y

(R, t) คือ Extrinsic parameters: เมตริกซ์การหมุน R และเวกเตอร์การเคลื่อนที่ t

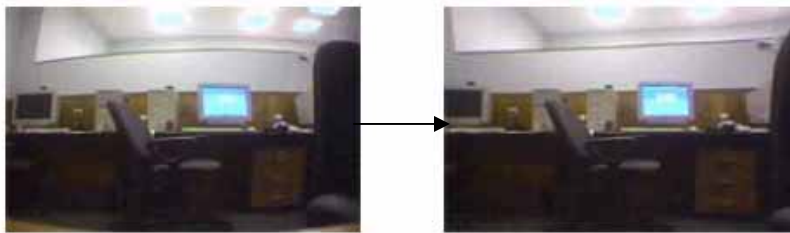
เมื่อทราบถึงความสัมพันธ์ระหว่างจุดในโลกของผู้ใช้กับจุดที่ปรากฏบนภาพก็สามารถแปลงตำแหน่งของจุดสองมิติบนภาพไปเป็นจุดบนพิกัดสามมิติในโลกของผู้ใช้ได้ แต่ในการแปลงจุดในพิกัดสองมิติให้อยู่ในพิกัดสามมิตินั้นใช้จุดเดียวไม่เพียงพอ (เพราะการแปลงโดยใช้จุดสองมิติเพียงจุดเดียวจะได้ผลลัพธ์เพียงแค่แนวโปรเจกชันของจุดในระนาบสามมิติ) จำเป็นต้องใช้หลายๆ จุดที่มีรูปแบบการเรียงตัวที่แน่นอน (Pattern) จึงจะเพียงพอสำหรับการคำนวณ

สำหรับในงานวิจัยนี้ กำหนดให้คอมพิวเตอร์รับภาพผู้ใช้โดยกล้องวงจรปิดที่ต่อกับการ์ดจับภาพวีดีโอ (Video capture card) กล้องจะถูกติดตั้งอยู่บนจอมอนิเตอร์ (เป็นตำแหน่งที่กล้องสามารถจับภาพสี่ระยะได้ดีที่สุด) ก่อนที่จะนำกล้องไปใช้งาน จะต้องผ่านการ Calibration โดยการนำเซตของจุด (x, y, z) บน Reference frame (เช่น Pattern จุดมุมตารางหมากรุกดังรูปที่ 8 และเซตของจุด (x, y) บนภาพที่ได้จากกล้องมาคำนวณหา Intrinsic parameters



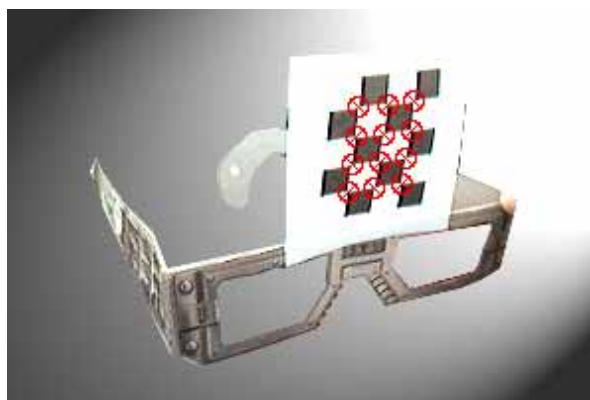
รูปที่ 8 ลายตารางหมากรุกที่ใช้ในการ Calibrate เพื่อหา Intrinsic parameters [8]

เมื่อได้ Intrinsic parameters แล้วจึงเริ่มรับภาพผู้ใช้เข้ามาจากกล้อง แต่ภาพที่ได้จากกล้องส่วนใหญ่จะมีการบิดเบี้ยวตามแนวรัศมี หากนำไปใช้คำนวณหาตำแหน่งของศีรษะภายหลังอาจให้ผลที่ผิดพลาดได้ เพราะฉะนั้นจึงมีความจำเป็นที่จะต้องปรับแก้ภาพให้มีความเที่ยงตรงเสียก่อน โดยการนำ Radial distortion coefficient ที่ได้จาก Intrinsic parameters มาร่วมคำนวณเพื่อทำ Undistortion ซึ่งได้ผลเป็นดังรูปด้านล่าง



รูปที่ 9 รูปจากกล้องที่มีการบิดเบี้ยวตามแนวรัศมี (ซ้าย) และภาพที่ปรับแก้การบิดเบี้ยวแล้ว (ขวา) [8]

จากนั้นจึงเข้าสู่กระบวนการตรวจหาตำแหน่งของศีรษะ โดยระหว่างการใช้งานผู้ใช้จำเป็นต้องสวมแว่นติดแพทเทิร์นที่มีลักษณะเป็นลายตารางหมากรุกขนาด 5 x 4 ช่อง เพื่อตรวจจับจุดมุมทั้ง 12 จุดใช้สำหรับเป็น Pattern ในการคำนวณหา Extrinsic parameters โดยแต่ละจุดมุมที่ตรวจจับได้แสดงโดยวงกลม ดังรูป



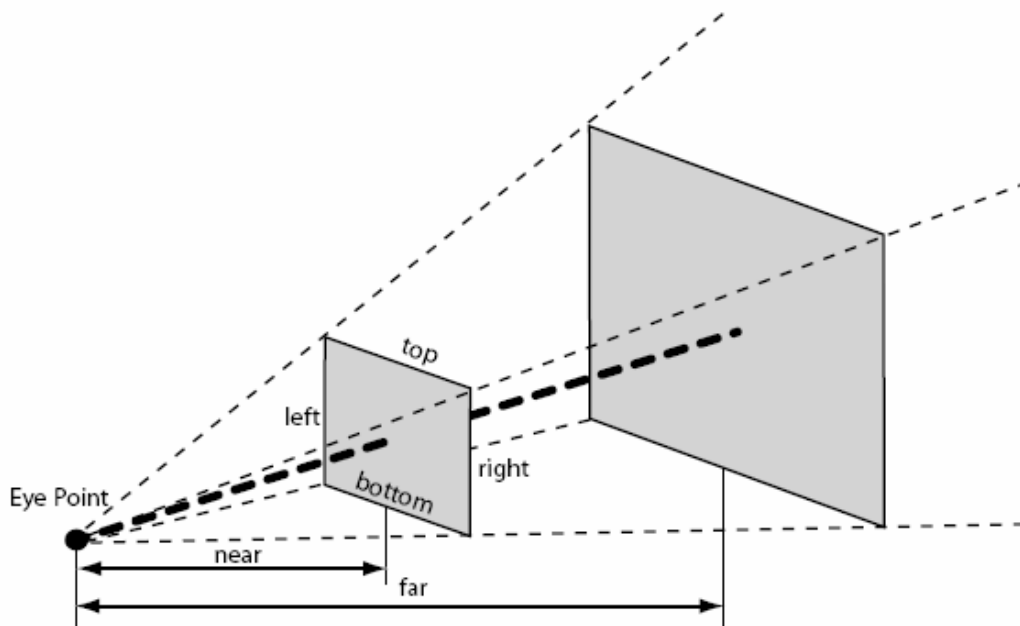
รูปที่ 10 อุปกรณ์เพื่อช่วยในการตรวจจับตำแหน่งของศีรษะ

ข้อดีของการใช้การตรวจจุดมุมของตารางหมากรุกแทนการตรวจอวัยวะบนใบหน้าคือ วิธีการนี้มีความรวดเร็วในการประมวลผล และมีความแม่นยำสูงกว่า และไม่ขึ้นกับลักษณะหน้าตาของผู้ใช้ ไม่ว่าผู้ใช้คนใด เพียงแค่สวมอุปกรณ์สวมศีรษะก็สามารถใช้งานในระบบได้ทันที

เมื่อได้จุดมุมทั้ง 12 จุดแล้ว จึงนำไปคำนวณหา Extrinsic parameters จะได้ Translation Vector (ตำแหน่งของแพทเทิร์นเทียบกับกล้อง) และ Rotational Matrix (มุมที่แพทเทิร์นกระทำกับกล้อง) เมื่อรู้ตำแหน่งและการหมุนของแพทเทิร์น ก็สามารถประมาณตำแหน่งของตาและมุมมองที่ผู้ใช้กระทำกับกล้องได้ (การใช้แว่นทำให้ แพทเทิร์นมีตำแหน่งคงที่เทียบกับตา) จากนั้นระบบจะต้องคำนวณการย้ายแกนเพื่อให้จุดอ้างอิงเปลี่ยนจากกล้องไปอยู่ที่ตำแหน่งกึ่งกลางจอภาพ (เนื่องจากผู้ใช้มองไปยังจอภาพ) จึงจะได้ตำแหน่งของผู้ใช้เทียบกับจอภาพตามต้องการ

4.2) การปรับเพออสเปกทีฟใน OpenGL

OpenGL [11] คือไลบรารีที่ช่วยในการจัดการและแสดงกราฟฟิคสามมิติ ซึ่งการแสดงผลกราฟฟิคและโมเดลสามมิติต่างๆ ออกทางหน้าจอจะกระทำผ่านกล้อง (Camera) เป็นหลัก (ภาพที่แสดงออกทางหน้าจอคือภาพที่กล้องจับอยู่) ซึ่งกล้องใน OpenGL มีคุณสมบัติหลายอย่างเหมือนกับกล้องที่เราใช้กันจริงๆ อาทิเช่น กล้องมีออชัพมักจะถอดเปลี่ยนเลนส์ เช่น เปลี่ยนเป็นเลนส์มุมกว้าง (Wide) หรือเลนส์เทเล (Tele) ซึ่งจะมีมุมรับภาพที่แตกต่างกันออกไป กล้องใน OpenGL ก็สามารถทำแบบนี้ได้เช่นกัน โดยการปรับเพออสเปกทีฟ (Perspective) เมื่อนำมาใช้ในระบบ Fish Tank VR จะเป็นดังรูปที่ 11

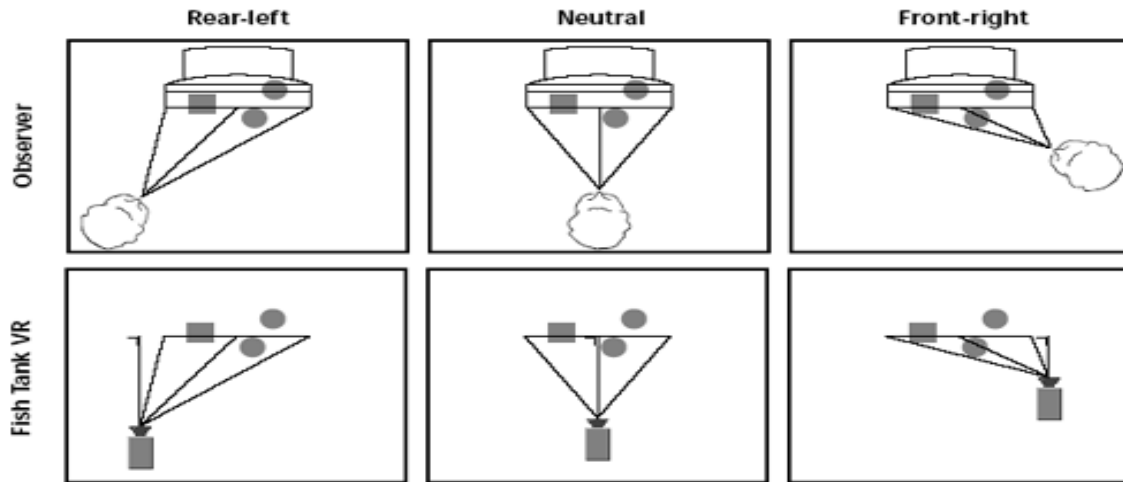


รูปที่ 11 รูปแบบของเพออสเปกทีฟใน OpenGL เมื่อใช้ใน Fish Tank VR [11]

ตำแหน่งศีรษะของผู้ใช้เทียบกับจอภาพที่ได้ในหัวข้อที่แล้ว จะถูกนำมาเซตเป็นตำแหน่งให้จุด Eye Point

เทียบกับ Near plane ซึ่งวัตถุสามมิติที่อยู่ระหว่าง Near plane และ Far plane จะถูกฉาย (หรือโปรเจ็ก) ไปยัง Near plane โดยมีศูนย์กลางของการโปรเจ็กอยู่ที่จุด Eye point ก่อให้เกิดเป็นภาพสองมิติเพื่อแสดงออกทางจอภาพ

รูปที่ 11 แสดงการตั้งเฟออสเพกทีฟเมื่อศีรษะผู้ใช้อยู่ในแนวเดียวกับจอภาพและมองตั้งฉากกับจอภาพ ถ้าหากผู้ใช้เคลื่อนศีรษะไปทางซ้ายหรือขวาและมองทำมุมเอียงกับจอภาพ จะต้องเคลื่อนตำแหน่งกล้อง (Eye Point) ใน OpenGL ให้สัมพันธ์กัน (แต่ฉากรับภาพจะยังคงอยู่ที่เดิม ไม่เปลี่ยนแปลง) และกล้องต้องมีทิศตั้งฉากกับฉากรับภาพ Near plane เสมอ เพราะฉะนั้นการปรับเฟออสเพกทีฟจะเป็นดังรูปที่ 12



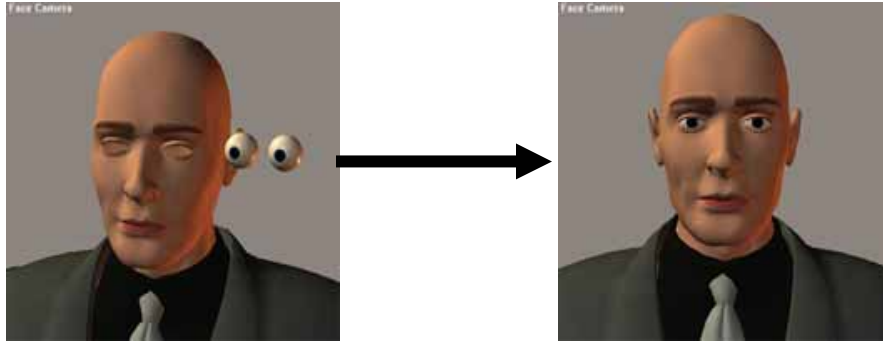
รูปที่ 12 การปรับเฟออสเพกทีฟเมื่อศีรษะผู้ใช้ทำมุมต่างๆ กันกับจอภาพ

จากรูปจะเห็นว่าฉากรับภาพ Near plane ไม่จำเป็นต้องสมมาตรกับกล้อง ซึ่ง OpenGL ได้เตรียมคำสั่งสำหรับการตั้งเฟออสเพกทีฟเช่นนี้ไว้แล้ว คือ glFrustum()

4.3) การสร้างปฏิสัมพันธ์กับผู้ใช้งาน

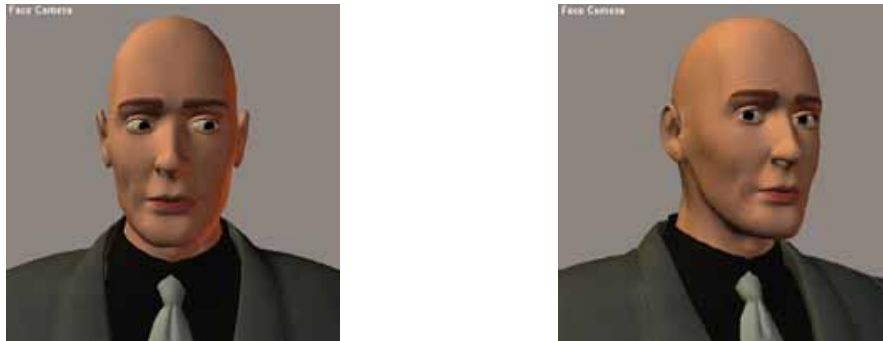
เมื่อสามารถปรับมุมมองของโมเดลสามมิติให้สัมพันธ์กับตำแหน่งศีรษะของผู้ใช้ได้แล้ว กระบวนการสร้างระบบ Fish Tank VR จึงเสร็จสมบูรณ์ ขั้นตอนต่อไปที่จะพิจารณาคือการนำระบบ Fish Tank VR ที่มีอยู่มาสร้างปฏิสัมพันธ์กับผู้ใช้งานในรูปแบบต่างๆกัน โดยผ่านโมเดลสามมิติศีรษะมนุษย์เสมือนจริง ซึ่งการปฏิสัมพันธ์กับผู้ใช้ประกอบไปด้วยการทำให้โมเดลศีรษะสามารถมองไปยังผู้ใช้ หรือบริเวณต่างๆ ในโลกของผู้ใช้ และสามารถขยับปากสนทนากับผู้ใช้ตามสคริปต์ที่กำหนดไว้ได้

เพื่อให้สามารถบรรลุวัตถุประสงค์ตามที่กล่าวมา จึงได้ออกแบบโมเดลศีรษะมนุษย์ [6][9] โดยใช้โมเดลสามชิ้นหลักๆ คือ ส่วนพื้นผิวศีรษะ ดวงตาซ้าย และดวงตาขวา สำหรับส่วนพื้นผิวศีรษะจะเป็นพื้นผิวที่มีลักษณะกลวงภายในและมีช่องเปิดที่บริเวณดวงตา เมื่อใส่โมเดลดวงตาทั้งสองเข้าไปแล้วก็จะมิลักษณะคล้ายตามนุษย์ดังรูปที่



รูปที่ 13 โมเดลทั้งสามชิ้น (ซ้าย) และเมื่อนำมาประกอบกัน (ขวา)

หากต้องการให้โมเดลเปลี่ยนทิศทางการมอง ก็สามารถทำได้โดยการหมุนดวงตาทั้งสองข้างรอบแกน y หากจะให้มองตรงมายังผู้ใช้ ก็เพียงแต่หมุนให้ดวงตามีทิศไปยังตำแหน่งของกล้อง ตัวอย่างแสดงดังรูปที่ 14



รูปที่ 14 โมเดลมองไปยังด้านขวาของผู้ใช้ (ซ้าย) และโมเดลมองไปยังผู้ใช้เมื่อผู้ใช้เคลื่อนตัวไปทางซ้าย (ขวา)



รูปที่ 15 โมเดลขยับปากเพื่อสื่อสารกับผู้ใช้

ในรูปที่ 15 แสดงการขยับปากของโมเดลเพื่อสื่อสารกับผู้ใช้ การขยับปากทำได้โดยการยืด / หด พื้นผิว บริเวณปาก ซึ่งภายในช่องปากก็จะเป็นพื้นผิวของลิ้นและฟัน ทำให้ผู้ใช้รู้สึกเสมือนว่าโมเดลสามารถพูดได้ เสียงพูดจะ ถูกสังเคราะห์โดยใช้หลักการของ Text to Speech ซึ่งจะพูดตามสคริปต์ที่ได้กำหนดไว้ก่อนแล้ว

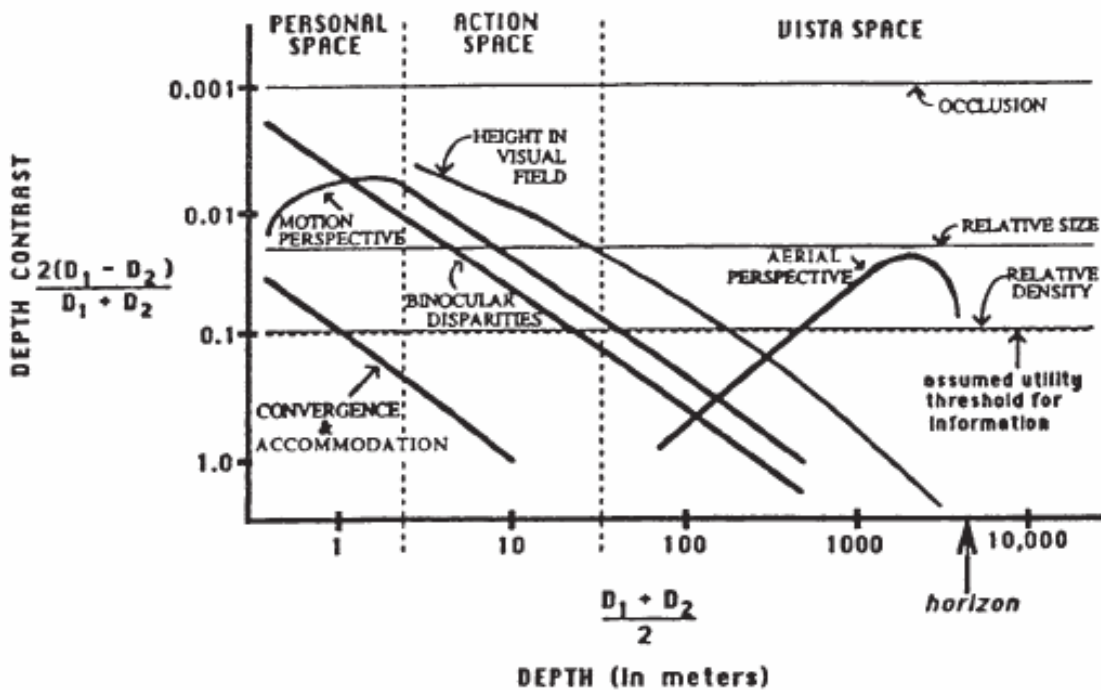
อย่างไรก็ตาม การที่ระบบ Fish Tank VR ใช้จอมอนิเตอร์ธรรมดาในการแสดงผล ทำให้มีข้อจำกัดด้านความเสมือนจริงอยู่พอสมควร และไม่สามารถที่จะทำให้ผู้ใช้เกิดความรู้สึกเสมือนจริงร้อยเปอร์เซ็นต์ [13] แต่เราก็สามารถ สร้างความสมจริงในระดับหนึ่งที่ยอมรับได้ โดยใช้เทคนิคต่างๆ กัน ซึ่งก็มีหลายๆ เทคนิคที่ใช้กันอย่างแพร่หลายในงานศิลปะ ภาพวาด และวงการภาพยนตร์ โดยมีนักวิจัยหลายๆ ท่าน เช่น James E. Cutting [4] ได้นำเสนอและเปรียบเทียบถึงปัจจัยต่างๆ ที่มีผลต่อการรับรู้ความสมจริงและความลึกของมนุษย์ซึ่งสามารถสรุปได้ดังนี้

- 1.) **Occlusion / Interposition** คือการบดบังกันของวัตถุ โดยพื้นที่บางส่วนของวัตถุหนึ่งถูกบังโดยวัตถุอื่น วัตถุที่มีบางส่วนถูกบัง มนุษย์จะรับรู้ว่าอยู่ไกลออกไปมากกว่าวัตถุที่มาบัง เทคนิคนี้เป็นเทคนิคที่มนุษย์ สามารถรับรู้ได้เพียงว่าวัตถุอยู่ด้านหน้าหรือด้านหลังเมื่อเทียบกับอีกวัตถุหนึ่ง แต่ไม่สามารถรู้ได้ว่าอยู่ห่างกันเพียงใด
- 2.) **Height in the visual field** คือความสัมพันธ์ระหว่างความสูงของฐานวัตถุในภาพและระยะห่างจาก ผู้ใช้ หากวัตถุมีระยะไกลออกไปมากขึ้น ผู้ใช้จะเห็นส่วนฐานจะอยู่สูงกว่าฐานวัตถุที่อยู่ใกล้ (ซึ่งจริงๆ แล้ววัตถุต่างวางอยู่บนพื้นในระดับเดียวกัน)
- 3.) **Relative size** คือ ขนาดของวัตถุที่ผู้ใช้รับรู้ เมื่อวัตถุอยู่ในระยะต่างๆ กัน ถ้าหากมีวัตถุสองอันรูปร่าง เหมือนกัน อันหนึ่งมีขนาดเล็ก และอีกอันมีขนาดใหญ่ ผู้ใช้จะรับรู้ว่ามีขนาดใหญ่อยู่ใกล้กว่าวัตถุ ที่มีขนาดเล็ก
- 4.) **Relative density** คือ ความหนาแน่นของพื้นผิว (Textures) ที่ผู้ใช้รับรู้ต่อพื้นที่การมองเห็นๆ ถ้าหาก วัตถุหรือพื้นผิวอยู่ห่างออกไป ผู้ใช้จะมองเห็นถึงความหนาแน่นที่มากขึ้น (เกี่ยวข้องกับ Relative Size ที่ มีขนาดเล็กลง)
- 5.) **Aerial perspective** คือผลกระทบจากสภาพอากาศที่มีต่อวัตถุในระยะต่างๆ กัน เช่น มีหมอก หรือ ฝุ่น คว้น ทำให้วัตถุที่อยู่ไกลออกไปมีค่าความเปรียบต่าง (Contrast) น้อยกว่า, มัวกว่าวัตถุที่อยู่ใกล้ เนื่องจาก แสงต้องเดินทางผ่านอากาศเป็นระยะทางที่ไกลกว่า
- 6.) **Binocular disparity** คือความต่างของภาพที่ตกลงบนเรตินาในตาแต่ละข้าง ตาของคนเรานั้นมีสองข้าง จะรับภาพที่มีความแตกต่างกันสองภาพ ทำให้สมองรับรู้ได้ถึงความลึก หากวัตถุอยู่ใกล้ตามาก ในสอง ภาพนั้นวัตถุก็จะมีตำแหน่งที่แตกต่างกันมาก แต่หากวัตถุอยู่ไกลตา วัตถุจะมีตำแหน่งที่แตกต่างกันน้อย หรือไม่แตกต่างเลย
- 7.) **Accommodation** คือการปรับรูปร่างของเลนส์ตา (ปรับโฟกัส) เพื่อให้บริเวณที่เราสนใจตกลงบนเรติ นามีความชัดเจนมากที่สุด บริเวณอื่นที่อยู่นอกโฟกัสจะเบลอ (เปรียบเสมือนระยะชัดลึก หรือ Depth of Field ที่ใช้กันในวงการถ่ายภาพ)
- 8.) **Convergence** คือมุมระหว่างทิศทางการมองของตาทั้งสองข้างเมื่อจ้องมองไปยังวัตถุหนึ่งๆ (angle between foveal axes) หากวัตถุที่อยู่ใกล้ตามาก มุมจะมีขนาดใหญ่เนื่องจากตาทั้งสองข้างจะเอียงเข้า

หากันเพื่อโฟกัสบริเวณด้านหน้าจมูก แต่หากวัตถุอยู่ไกล มุมจะเข้าใกล้ศูนย์ เพราะทิศทางกรมองของตาทั้งสองข้างแทบจะขนานกัน

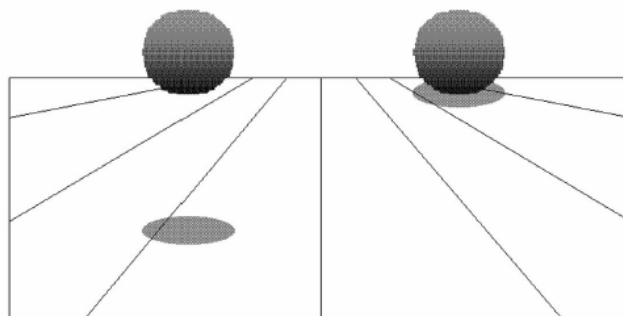
- 9.) **Motion Parallax / Motion Perspective** เกิดขึ้นเมื่อผู้สังเกตเคลื่อนที่ เช่น ผู้สังเกตจะเห็นว่าวัตถุที่อยู่ใกล้เคลื่อนที่ผ่านไปจากมุมมองเร็วกว่าวัตถุที่อยู่ไกล (ในกรณีที่มีมุมมองของผู้สังเกตเคลื่อนที่ ซ้าย / ขวา) หรือจะสังเกตเห็นว่าวัตถุที่อยู่ใกล้มีการเปลี่ยนแปลงขนาด (ใหญ่ขึ้น หรือเล็กลง) ที่เร็วกว่าวัตถุที่อยู่ไกล (จากการเคลื่อนที่เข้าไปใกล้ขึ้น / ห่างออก)

รูปด้านล่างแสดงความสัมพันธ์ระหว่างความแตกต่างของความลึกที่ผู้สังเกตรับรู้และสามารถแยกแยะได้เมื่อเทียบกับระยะห่างระหว่างผู้สังเกตและวัตถุ โดยใช้เทคนิคต่างๆ ที่ได้กล่าวมา



รูปที่ 16 เปรียบเทียบระยะที่ได้ผลดีของเทคนิคต่างๆ [4]

นอกจากที่กล่าวมาแล้ว ยังมีเทคนิคอื่นๆ เพิ่มเติมอีก เช่นงานวิจัยของ Geoffrey [7] ที่ได้นำการใส่เงา (Shadow) มาใช้เพื่อเพิ่มความแตกต่างในด้านความลึกให้กับวัตถุ ด้วยเทคนิคนี้ช่วยให้ผู้ใช้สามารถแยกความแตกต่างความลึกของวัตถุได้ดียิ่งขึ้น ดังรูปที่ 17



รูปที่ 17 การใช้เงาเพื่อแสดงความลึก [7]

เพื่อเพิ่มความสมจริงในด้านความลึก ให้ผู้ใช้รับรู้ความลึกของโมเดลได้ดีขึ้น จึงนำเทคนิคตามที่ได้กล่าวไว้มาประยุกต์ใช้ เช่น ใส่วัตถุเพิ่มเข้าไปในโลกสามมิติ ดังรูปที่ 18



รูปที่ 18 เพิ่มวัตถุเพื่ออ้างอิงความลึก ในมุมมองตรง (ซ้าย) และมุมมองในมุมข้าง (ขวา)

จากรูปที่ 18 หากเปรียบเทียบกับรูปก่อนหน้า จะเห็นว่าโมเดลต่างๆ มีมิติมากขึ้น ประกอบกับการเปลี่ยนมุมมองไปตามตำแหน่งผู้ใช้ ทำให้ผู้ใช้สามารถรับรู้ความลึกและความเสมือนจริงได้ดีมากขึ้น

5.) ขอบเขตการวิจัย

1. ระบบที่พัฒนาขึ้นสามารถแสดงโมเดลสามมิติที่ระยะมนุษย์ให้สัมพันธ์กับมุมมองของผู้ใช้
2. โมเดลที่ระยะสามารถมองไปยังจุดต่างๆ รอบตัวผู้ใช้ได้
3. โมเดลที่ระยะสามารถขยับปากเพื่อแสดงอารมณ์ ได้แก่ ยิ้ม บึ้ง และปกติ
4. โมเดลที่ระยะสามารถขยับปากเพื่อสนทนากับผู้ใช้ตามคำพูดที่ได้กำหนดไว้ล่วงหน้า
5. การควบคุมโมเดลที่ระยะสามารถกระทำโดยผ่านสคริปต์
6. ผู้ใช้ต้องเคลื่อนไหวอย่างจำกัดให้อยู่ภายในมุมรับภาพของกล้อง
7. ผู้ใช้ต้องสวมแว่นตาที่ติดแพทเทิร์นตารางหมากรุกตลอดเวลาที่ใช้งาน

6.) การประเมินผล

ในส่วนของการประเมินผลมีดังนี้

1. วัดความถูกต้องแม่นยำของการหาดำแหน่งศีรษะผู้ใช้ โดยคำนวณเปอร์เซ็นต์ความคลาดเคลื่อนเฉลี่ยระหว่างตำแหน่งของผู้ใช้จริง และตำแหน่งที่โปรแกรมคำนวณได้

- วัดความถูกต้องของโมเดลเมื่อมองไปยังบริเวณต่างๆ รอบตัวผู้ใช้ ทำการทดสอบโดยให้โมเดลมองไปยังจุดๆ หนึ่งก่อน แล้วให้ผู้ใช้สังเกตดวงตาของโมเดลแล้วระบุว่าโมเดลมองไปยังจุดใด คำนวณความคลาดเคลื่อนที่ได้ จากนั้นให้ผู้ใช้เคลื่อนศีรษะเพื่อมองจากมุมอื่นแล้วทำแบบทดสอบซ้ำ
- ให้ผู้ใช้ทดสอบใช้งานระบบ แล้วให้กรอกแบบสอบถามเพื่อวัดความพึงพอใจของผู้ใช้ โดยในแบบสอบถามจะมีช่องให้ผู้ใช้เลือกกรอกระดับความพึงพอใจพร้อมทั้งเหตุผล

7.) ขั้นตอนการวิจัย

- ศึกษาความรู้และงานวิจัยที่เกี่ยวข้องกับ Fish Tank VR ,Depth Perception และลักษณะของศีรษะมนุษย์
- พัฒนา Fish Tank VR โดยใช้โมเดลศีรษะมนุษย์
- เพิ่มความสามารถด้านการปฏิสัมพันธ์กับผู้ใช้ให้กับโมเดล
- ประเมินผล ตรวจสอบความเที่ยงตรงและความเสมือนจริง
- สรุป และเรียบเรียงวิทยานิพนธ์

7.1) สิ่งที่ได้ทำไปแล้ว

- ศึกษาความรู้และงานวิจัยที่เกี่ยวข้องกับ Fish Tank VR ,Depth Perception และลักษณะของศีรษะมนุษย์
- พัฒนาระบบ Fish Tank VR ตรวจสอบแพทเทิร์นโดยใช้ไลบรารี OpenCV และแสดงกราฟฟิกสามมิติโดยใช้ OpenGL
- ออกแบบโมเดลศีรษะมนุษย์โดยใช้โปรแกรม Poser, 3DS Max, Maya

7.2) สิ่งที่ต้องทำต่อ

- นำโมเดลศีรษะมนุษย์ที่ออกแบบไว้ไปรวมกับระบบ Fish Tank VR
- เพิ่มความสามารถด้านการปฏิสัมพันธ์กับผู้ใช้ให้โมเดล
- ประเมินผล ตรวจสอบความเที่ยงตรงและความเสมือนจริง
- สรุป และเรียบเรียงวิทยานิพนธ์

8.) ประโยชน์

งานวิจัยชิ้นนี้พัฒนาขึ้นเพื่อให้ผู้ใช้ได้รับความสะดวกและความรู้สึกเสมือนจริงในการติดต่อสื่อสารกับคอมพิวเตอร์ ผู้ใช้จะรู้สึกว่ามีโมเดลศีรษะและใบหน้ามนุษย์สามมิติ อยู่ในจอมอนิเตอร์จริงๆ ซึ่งจะแปรเปลี่ยนมุมมองไปตามตำแหน่งของผู้ใช้ และมีปฏิสัมพันธ์กับผู้ใช้ในรูปแบบต่างๆ กัน ทำให้การติดต่อกับคอมพิวเตอร์มีความน่าสนใจมากยิ่งขึ้น

9.) รายการอ้างอิง

- [1] Cynthia Breazeal, Brian Scassellati, Robot in Society: Friend of Appliance?, *In Agents99 Workshop on Emotion-based Agent Architectures*, Seattle, WA. 18-26, 1999.
- [2] Cynthia Breazeal, Brian Scassellati, How to build robots that make friends and influence people., *IROS99*, Kyonjiu, Korea, 1999.
- [3] Carolina Cruz-Neira, Daniel J. Sandin, Thomas A. DeFanti, Surround-Screen Projection-Based Virtual Reality: The Design and Implementation of the CAVE, *Proceedings of the 20th annual conference on Computer graphics and interactive techniques*, 1993
- [4] James E. Cutting, How the eye measures reality and virtual reality, *Behavior Research Methods, Instruments, & Computers* 1997
- [5] Cagatay Demiralp, David H. Laidlaw, Cullen Jackson, Daniel Keefe, and Song Zhang. Subjective usefulness of cave and fish tank vr display systems for a scientific visualization application. *In Poster Compendium IEEE Conference on Visualization*, Seattle, WA, 2003
- [6] Bill Freming, Darris Dobbs, Animating Facial Features & Expressions, *Charles River Media Inc.* 1999
- [7] Geoffrey S. Hubona, The Relative Contributions of Stereo, Lighting, and Background Scenes in Promoting 3D Depth Visualization, *ACM Transactions on Computer-Human Interaction*, Vol 6, No. 3, September 1999, Pages 214-242
- [8] Intel Corporation, Open Source Computer Vision Library Reference Manual, 2001, "<http://sourceforge.net/projects/opencvlibrary/>"
- [9] Frederic I. Parke, Keith Waters, Computer Facial Animation, *A K Peters Ltd.* 1996
- [10] Jun Rekimoto, A Vision-Based Head Tracker for Fish Tank Virtual Reality - VR without Head Gear -, *Proceedings of VRAIS'95*, 1995, Pages 94-100
- [11] Dave Shreiner, OpenGL(R) Reference Manual: The Official Reference Document to OpenGL, Version 1.2 (3rd Edition)
- [12] Ivan Sutherland, A head-mounted three dimensional display. *Proceedings of the Fall Joint Computer Conference, AFIPS Conference Proceedings*, 33 (1968), 757-764
- [13] Michael Treadgolf, Kevin Novins, Geoff Wyvill, and Brian Niven, What do you think you're doing? Measuring perception in Fish Tank Virtual Reality, *Computer Graphics International Conference* 2001
- [14] Colin Ware, Kevin Arthur, Kellogg S. Booth, Fish Tank Virtual Reality, *Proceedings of the SIGCHI conference on Human factors in computing systems*, 1993
- [15] Web-Content, "SimVis Head-mounted display", <http://www2.dcs.hull.ac.uk/simmod/Technology/HMD.htm>
- [16] Web-Content, "Welcome to Vislab", <http://www.ccom.unh.edu/vislab/FishTankVR.html>

[17] Web-Content, "DocShow-VR", <http://www.dfn-expo.de/Technologie/DocShow-VR/stereo.html>

[18] Web-Content, "Virtual Skiing Environment ", <http://www-personal.umich.edu/~galtland/skiVR/video/tech.htm>