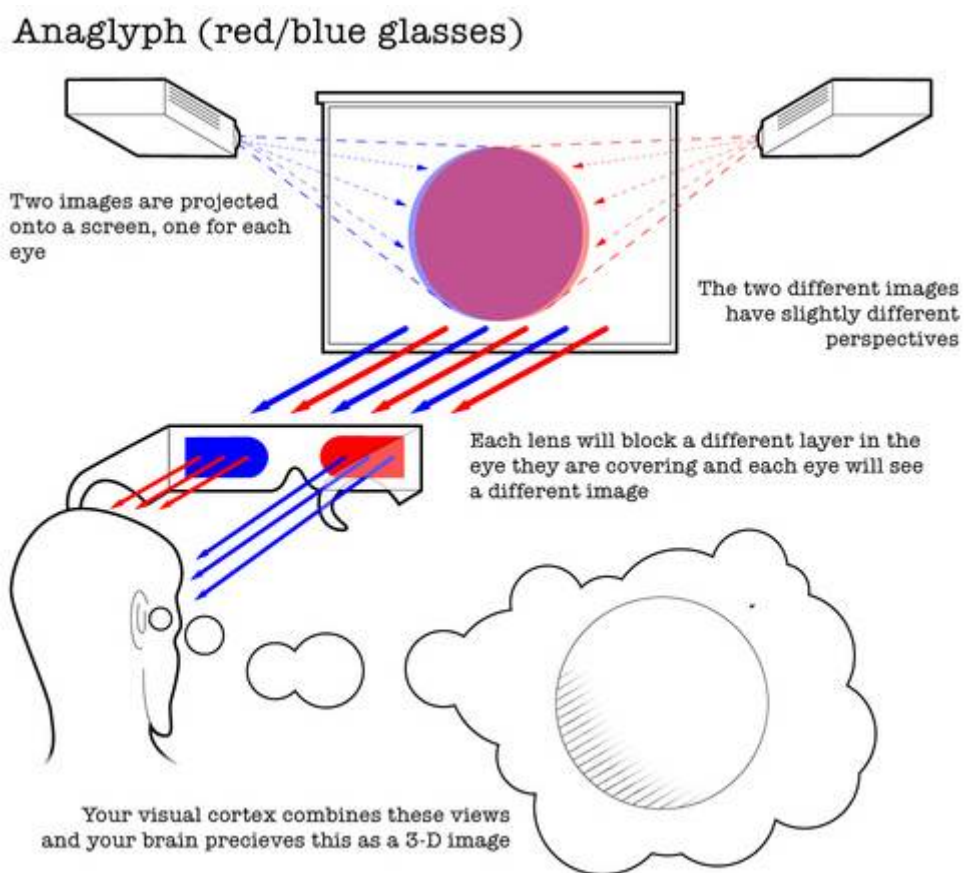


## บทที่ 2

### แนวคิดและทฤษฎีที่เกี่ยวข้อง

#### 2.1 3 มิติ <sup>[1]</sup>

เทคโนโลยี “ภาพ 3 มิติ” ได้รับความสนใจจากผู้บริโภคมานานแล้ว ในขณะเดียวกันมันก็มีเทคนิคที่ใช้หลอดสายตาให้เห็นภาพที่ฉายอยู่นั้นเกิดมิมิติตื้นลึกชัดเบลอขึ้นมามากมายจนน่าปวดหัว โดยแต่ละเทคนิคก็ยังคงใช้แว่นที่ไม่เหมือนกันอีกต่างหาก รูปภาพประกอบและคำอธิบายเพื่อช่วยเพิ่มความเข้าใจมากขึ้น



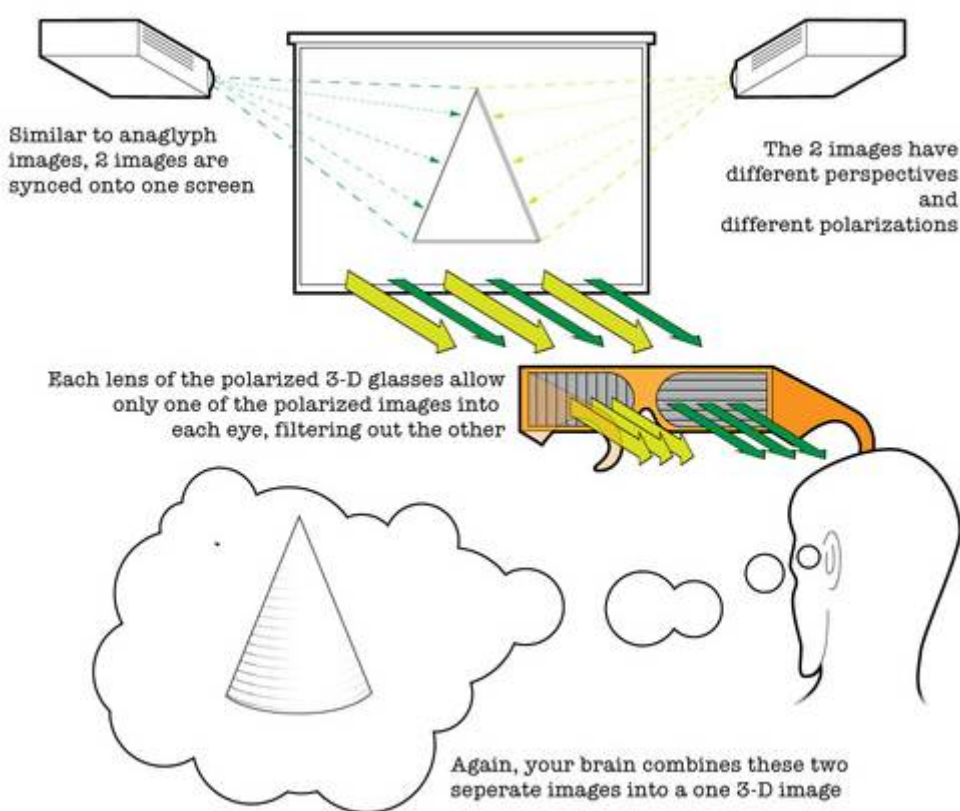
รูปที่ 2.1 แว่นตา Anaglyph

Anaglyph (แว่นตาน้ำเงิน/แดง) เทคนิคแรกนี้จะพบเห็นกันมาก และที่เป็นที่คุ้นเคยมากที่สุด ซึ่งหากจะอธิบายหลักการจากภาพที่เห็นข้างล่างนี้ก็คือ Anaglyph จะใช้กล้องฉายภาพ 2 ตัว ฉายภาพที่มีสีสลับ (น้ำเงินกับแดง) และมุมมองที่แตกต่างกัน (เหมือนกับเวลาเราปิดตาแล้วมองทีละข้าง ภาพที่เห็นจะมีมุมที่แตกต่างกันเล็กน้อย) ส่วนแว่นตาทำหน้าที่กรองภาพแต่ละสีออกไป เช่น แว่นตาสีแดงจะกรองภาพสีแดงออกไปให้เห็นแต่ภาพสีน้ำเงิน ส่วนแว่นตาสีน้ำเงินก็จะกรองภาพส่วนที่เป็นสี

แดงออกไป ทำให้ตาทั้งสองเห็นภาพที่แตกต่างกัน สมองจะตีความด้วยการรวมภาพที่มองเห็น  
 แตกต่างกันสองภาพ อีกทั้งมีมุมแตกต่างกันกลายเป็นภาพที่มีมิติขึ้นมา (อีกคำอธิบายหนึ่งที่น่าสนใจก็  
 คือ ภาพสีแดงจะตกหลังจอตา ส่วนภาพสีน้ำเงินจะตกกระทบบนจอตา ความแตกต่างกันของ  
 การตกกระทบบนจอตาทั้งสองภาพบนจอตา เมื่อมองเห็นพร้อมกันทำให้เกิดมิติคลึกลงที่ไม่เท่ากัน เลย  
 เห็นเป็นภาพลอยออกมาได้นั่นเอง)

แบบที่สองนั้นก็คือ แว่น 3 มิติ ชนิด Polarized 3-D Glasses

### Polarized 3-D Glasses



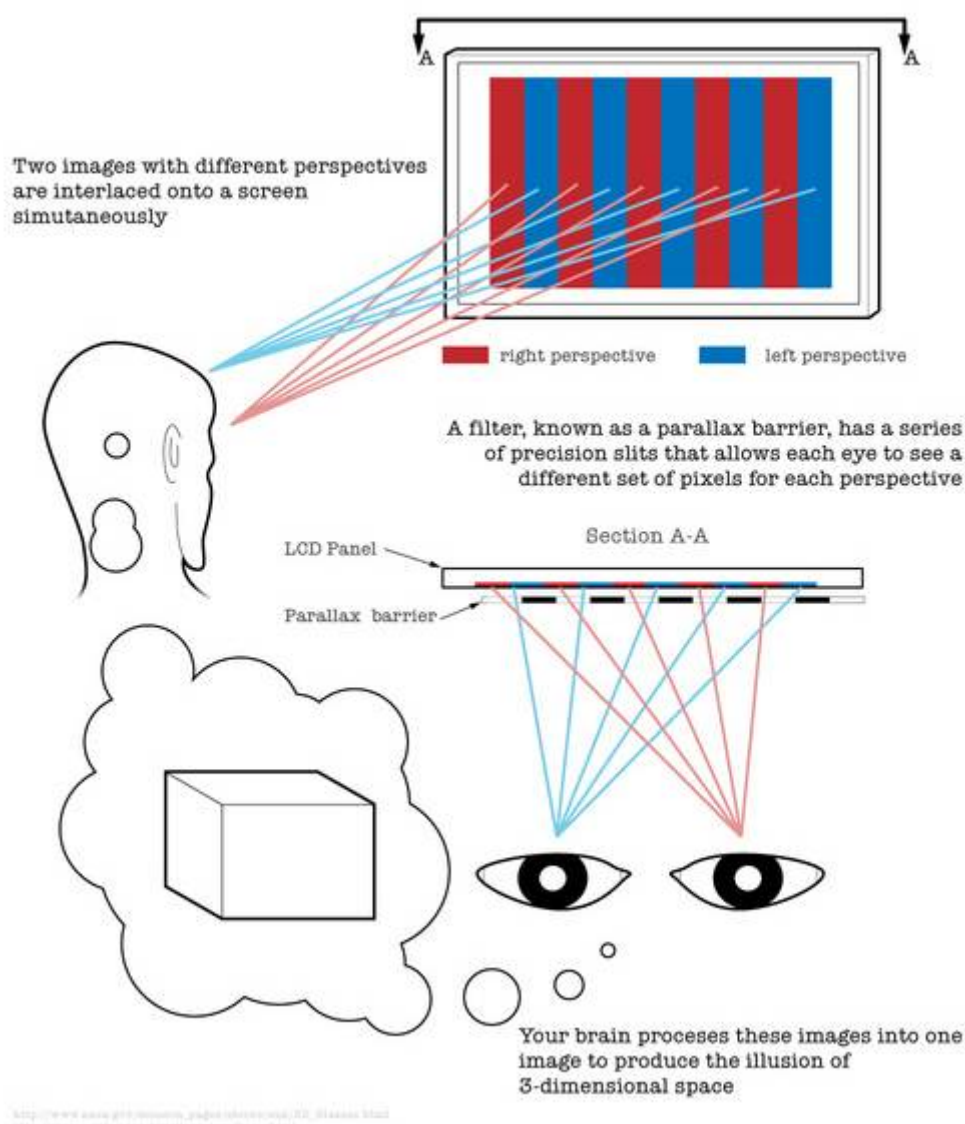
รูปที่ 2.2 แว่นตา Polarized 3-D Glasses

Polarized 3-D Glasses หลักการจะคล้ายกับ Anaglyph โดยเฉพาะการฉายภาพจากสองดวงด้วย  
 ภาพที่แตกต่างกัน แต่เทคนิคนี้ไม่ได้ใช้สีเป็นตัวแบ่งภาพที่ต่างกัน แต่จะใช้แนวการวางตัวของช่อง  
 การมองเห็นแต่ละภาพที่ฉายซ้อนกันอยู่ เช่น จากในภาพแว่นตาข้างซ้ายจะเห็นมองเป็นภาพที่ผ่าน  
 ช่องในแนวตั้ง ส่วนตาขวาจะมองเห็นภาพที่ช่องในแนวนอน ซึ่งทั้งสองภาพมีมุมมองที่แตกต่างกัน

ดังนั้นมันก็จะเข้าหลักการเดิม นั่นก็คือ การทำให้ตาแต่ละข้างของเรามองเห็นภาพที่ไม่เหมือนกัน เมื่อสมองพยายามรวมภาพทั้งสองที่มีความแตกต่างของมุมมอง ภาพที่เห็นจึงเกิดเป็น 3 มิติขึ้นมา

สำหรับหน้าจออีกประเภทหนึ่งชนิด Parallax Barrier ซึ่งหน้าจอประเภทนี้ไม่ต้องสวมแว่นตา

### Parallax Barrier (no glasses)



รูปที่ 2.3 หน้าจอแบบ Parallax Barrier

Parallax Barrier เทคนิคนี้ผู้ชมไม่จำเป็นต้องใส่แว่นตา ซึ่งในงาน Commart ครั้งหนึ่งได้มีการนำกล้อง Fujifilm ที่สามารถมองเห็นภาพถ่ายบน LCD ด้านหลังกล้องเป็น 3D ได้ โดยไม่ต้องใส่

แว่นตา นี้คือตัวอย่างของเทคนิค Parallax ซึ่งหากจะอธิบายจากภาพที่เห็นก็คือ มันจะใช้วิธีแบ่งภาพที่มีมุมมองต่างกันเป็นแท่งๆ วางตัวสลับกัน (เหมือนเส้นสแกนในทีวี แต่ทีวีจะใช้ภาพมุมมองเดียวกัน) โดยมี Parallax Barrier ที่เป็นชั้นกรองพิเศษสามารถแบ่งแต่ละส่วนของภาพให้ตาแต่ละข้างที่มองผ่านชั้นนี้มองเห็นภาพที่ไม่เหมือนกันได้พร้อมกัน เมื่อสมองพยายามรวมภาพที่มีมุมมองต่างกันให้เป็นภาพเดียว เราก็จะมองเห็นเป็นภาพสามมิตินั่นเอง

## 2.2 Virtual Reality <sup>[2]</sup>

ความจริงเสมือน (Virtual Reality : VR)

ความจริงเสมือน (Virtual Reality) หรือที่เรียกกันย่อ ๆ ว่า “ วีอาร์ ” (VR) เป็นกลุ่มเทคโนโลยีเชิงโต้ตอบที่ผลักดัน ให้ผู้ใช้เกิดความรู้สึกของการเข้าร่วมอยู่ในสิ่งแวดล้อมที่ไม่ได้มีอยู่จริงที่สร้างขึ้น โดยคอมพิวเตอร์ พัฒนาการของความเป็นจริงเสมือนได้รับอิทธิพลมาจากแนวความคิดง่าย ๆ แต่มีอำนาจมากเกี่ยวกับการที่จะเสนอสารสนเทศอย่างไรให้ดีที่สุด คือ ถ้าผู้ออกแบบสามารถให้ประสาทสัมผัสของมนุษย์มีความค่อยเป็นค่อยไปในปฏิสัมพันธ์กับโลกทางกายภาพซึ่งเป็นสิ่งที่อยู่ล้อมรอบตัวเราแล้ว มนุษย์ก็จะสามารถรับและเข้าใจสารสนเทศได้ง่ายขึ้น ถ้าสารสนเทศนั้นกระตุ้นการรับรู้สัมผัสของผู้รับ

เทคโนโลยีความเป็นจริงเสมือนสามารถเลียนการรับรู้สัมผัสของโลกทางกายภาพได้โดยสร้างการรับรู้หลายทางในสิ่งแวดล้อมสามมิติขึ้นมา ความเป็นจริงเสมือน ได้สร้างเนื้อหาสาระของสิ่งที่แสดงให้เห็น โดยการรับรู้ซึ่งเป็นผลลัพธ์ของคอมพิวเตอร์ เพื่อสนองต่อการเคลื่อนไหวทางกายภาพของผู้ใช้ที่สืบหาด้วยเครื่องรับรู้ของคอมพิวเตอร์

ประเภทของระบบ ( VR)

ระบบ VR แบ่งตามลักษณะตามพื้นฐานบนวิธีที่ติดต่อกับผู้ใช้.

1. Desktop VR or Window onWorld Systems (WoW) : ระบบเหล่านี้ใช้จอภาพคอมพิวเตอร์ธรรมดาเพื่อแสดงโลกเสมือนจริงบางครั้งถูกเรียกว่า “Desktop VR” หรือ “Window on a World” (WoW).
2. Video Mapping : สิ่งนี้คือการเปลี่ยนแปลงของ WoW เข้าใกล้ / กลมกลืนกับสิ่งที่นำเข้าวิดีโอของรูปแบบตัวคนดำพื้นขาวของผู้ใช้กับ 2D /3D กราฟฟิคคอมพิวเตอร์ ผู้ใช้จะเห็นตนเองและการเปลี่ยนแปลงของตนจากจอภาพ ความคิดนี้ถูกใช้ในเกมส์คอมพิวเตอร์ที่เป็นที่ชื่นชอบ เช่น Doom and Quake.
3. Immersive Systems : เป็นระบบ VR ล่าสุดที่สมบูรณ์แบบสำหรับผู้ใช้ส่วนบุคคล immersive VR systems ถูกจัดรวมไว้ร่วมกับ Head Mounted Display (HMD). HMD คือ หมวกเหล็กหรือหน้ากากหน้าซึ่งจำลองภาพและการได้ยิน (auditory)

4. Telepresence : คือการเปลี่ยนแปลงของคอมพิวเตอร์ที่เสร็จสิ้นสมบูรณ์ทำให้มองเห็นโลกที่สร้างใหม่เทคโนโลยีนี้เชื่อมต่อตัวตรวจจับสัญญาณระยะไกลในโลกจริงกับไหวพริบของมนุษย์ตัวตรวจจับสัญญาณระยะไกลอาจจะถูกค้นหาที่ตั้งบนหุ่นยนต์ พนักงานดับเพลิงใช้พาหนะที่กระทำอย่างระยะไกลที่จะจัดการสภาวะอันตรายจำนวนหนึ่ง
5. Augmented/Mixed Reality Systems: การผสมกันกับของ Telepresence and Virtual Reality systems ให้ระบบคอมพิวเตอร์สิ่งที่น่าสนใจที่สร้างขึ้นถูกผสมกันกับสิ่งที่น่าสนใจ telepresence และ/หรือผู้ใช้ของโลกรจริง เช่น การดูของศัลยแพทย์ของศัลยกรรมสมอง

### Virtual Reality hardware

1. Head-mounted display (HMD) จอภาพสวมศีรษะ หรือที่รู้จักกันว่า “ชุดแว่นตา” (goggles) ประกอบด้วยแว่นตาที่บรรจุจอมอนิเตอร์ขนาดเล็กซึ่งทำด้วยกระจก 3 มิติ เรียกว่า “stereoscopic glasses” ทำมุมกว้างประมาณ 140 องศาเพื่อให้ผู้ใช้สามารถมองเห็นสิ่งที่ป็นนามธรรมในลักษณะ 3 มิติ ในโลกของความเป็นจริงเสมือนได้



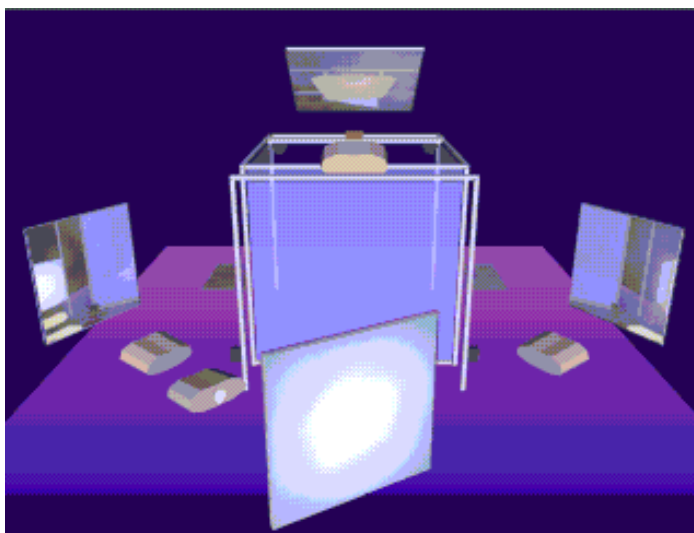
รูปที่ 2.4 ชุดแว่นตา (Goggles)

2. Data gloves ถุงมือรับรู้ (sensor glove) เป็นถุงมือขนาดเบาที่มีเส้นใยนำแสงเรียงเป็นแนวอยู่ตามนิ้วและข้อมือเพื่อเป็นเครื่องรับรู้การเคลื่อนที่และส่งสัญญาณไปยังคอมพิวเตอร์ เมื่อสวมถุงมือนี้แล้วจะทำให้ผู้ใช้เข้าถึงสิ่งแวดลอม 3 มิติ ถุงมือรับรู้จะทำให้ผู้ใช้จับต้องและรู้สึกวัดสิ่งของซึ่งไม่มีอยู่ที่นั่นจริงๆ



รูปที่ 2.5 ถุงมือรับรู้ (Sensor Glove)

3. CAVE เป็นห้องแสดงผลสามมิติที่ประกอบจากจอภาพจำนวน 4 จอดต่อกันเป็นรูปลูกบาศก์สำหรับจอด้านซ้าย ด้านหน้า ด้านขวา และที่พื้น สามารถมองเห็นวัตถุจำลองที่เตรียมขึ้นจากโปรแกรมคอมพิวเตอร์โดยอาศัยแว่นสามมิติชนิด LCD shutter glasses รวมทั้งสามารถติดต่อกับวัตถุจำลองสามมิติชิ้นต่างๆ ได้โดยอาศัยอุปกรณ์ติดต่ออย่าง ถุงมือ 3 มิติ (CyberGlove), เมาส์สามมิติ (3D mouse), และ กทาสามมิติ (Wanda)



รูปที่ 2.6 ห้องแสดงผล 3 มิติ

### ข้อจำกัดของความจริงเสมือน (VR Limitations)

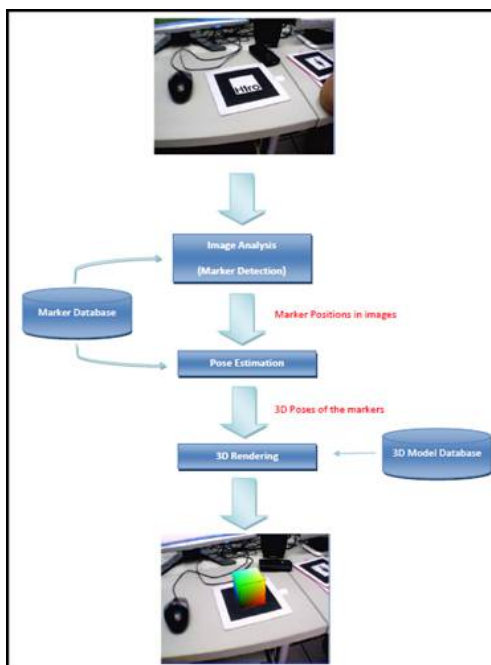
การใช้งานในการทำงานที่เสมือนจริงนั้น ดูเหมือนจะมีขีดจำกัดอย่างเดียว คือ ต้นทุนของเทคโนโลยี ตัวอย่างเช่น ผู้ใช้ VR สำหรับการพัฒนา Cybersickness เช่น Eyetrain และ Motion Sickness ที่ประกอบด้วยชุดครอบหัวที่ติดกับแว่นและหูฟัง ถุงมือส่งข้อมูลด้วยไฟเบอร์ออปติก อุปกรณ์การรับรู้ความรู้สึก และเครื่องสำหรับวิศวกรที่ทำงานร่วมกับซอฟต์แวร์การออกแบบโมเดลสามมิติที่ดีที่สุด ซึ่งมีราคาถึง 50,000 เหรียญสหรัฐ องค์กรหลายแห่ง เช่น นาซา กรมการป้องกันประเทศสหรัฐอเมริกา IBM, Lockheed, Matsushita Electric, Caterpillar และมหาวิทยาลัยหลายแห่ง ลงทุนกว่าล้านเหรียญสหรัฐสำหรับโครงการการวิจัยและพัฒนาที่เสมือนจริง (Virtual Reality R&D Projects) บางแห่งใช้ VRML (Virtual Reality Modeling Language) ในการพัฒนากราฟิกไฮเปอร์มีเดียสามมิติ (3-D Hypermedia Graphics) และผลิตภัณฑ์ภาพเคลื่อนไหว (Animation Products) ที่ทำงานร่วมกับอินเทอร์เน็ต ความล้าหน้าอย่างต่อเนื่องในเทคโนโลยีของ VR ที่ถูกคาดการณ์ว่าจะสร้างการทำงานเสมือนจริงที่มีประโยชน์สำหรับการทำงานในธุรกิจอีกมาก ตัวอย่าง VR ที่ Morgan Stanley แผนกความเสี่ยงทางการตลาดของบริษัท Morgan Stanley & Co. ได้ใช้ซอฟต์แวร์ Discovery Virtual Reality จาก Visible Decision ในการสร้างแบบจำลองความเสี่ยงสำหรับการลงทุนทางการเงินที่ประกอบด้วยเงื่อนไขมากมาย ที่แสดงผลออกมาในรูปแบบของสามมิติ

### 2.3 Augmented Reality (AR) <sup>[3]</sup>

Augmented Reality หรือ AR เป็นเทคโนโลยีที่พัฒนาในรูปแบบ Human-Machine Interface ที่อาศัยเทคโนโลยีคอมพิวเตอร์ และระบบเสมือนจริง (Virtual Reality) โดยที่วัตถุเสมือนนั้น ๆ จะถูกสร้างมาผสมกับสภาพในโลกจริงในรูปแบบ 3D และแสดงผลแบบ Real Time โดยเทคโนโลยีนี้จะต้องประกอบด้วย 3 ระบบ คือ ระบบ Tracking ระบบแสดงผลระบบประมวลผลเพื่อสร้างวัตถุ 3D โดยระบบ Tracking (กล้อง) จะรับข้อมูลรูปภาพเข้าไป เช่น รูปแบบ ตำแหน่ง และทิศทาง จากนั้นระบบประมวลผลก็จะนำไปแปลความหมาย และแสดงผลภาพสามมิติออกมาในตำแหน่งและทิศทางเดียวกันกับภาพที่กล้องจับได้ เทคโนโลยี AR สามารถแบ่งประเภทตามส่วนวิเคราะห์ภาพ (Image Analysis) ได้ออกเป็น 2 ประเภทด้วยกัน ได้แก่ Marker based AR และ Marker-less Based AR โดยที่ Marker based AR นั้น เป็นการวิเคราะห์ภาพโดยอาศัย Marker (วัตถุสัญลักษณ์) เป็นหลักในการทำงาน ส่วน Marker-less Based AR เป็นการวิเคราะห์ภาพที่ใช้คุณลักษณะต่าง ๆ ที่อยู่ในภาพ (Natural Features) มาทำการวิเคราะห์เพื่อคำนวณหาตำแหน่งเชิง



3 มิติ (3D Pose) เพื่อนำไปใช้งานต่อไป ซึ่งขั้นตอนของ Marker based AR สามารถแบ่งได้เป็น 3 ขั้นตอน ได้แก่ Image Analysis, Pose Estimation และ 3D Graphic Rendering



รูปที่ 2.7 แผนภาพการทำงานของเทคโนโลยี Augmented Reality

โดยรวมแล้วกระบวนการ Image Analysis และ Pose Estimation จะถูกเรียกรวมกันว่าการ Visual Tracking

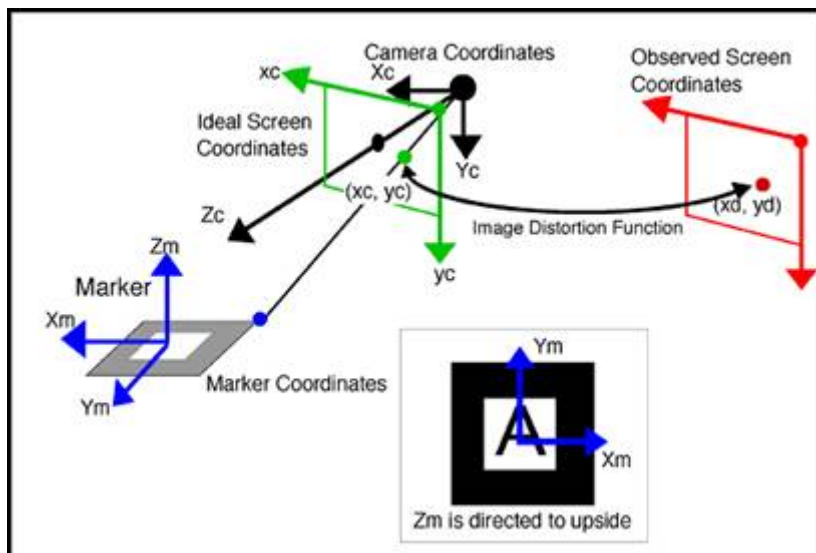
### Pose Estimation

Pose Estimation เป็นขั้นตอนของการคำนวณค่าตำแหน่งเชิง 3 มิติ (3D Pose) ของ Marker เมื่อเทียบกับกล้องวิดีโอ ค่านี้จะถูกแสดงในรูปเมตริกซ์ขนาด 4 x 4 (TCM) ที่ระบุความสัมพันธ์ระหว่าง Camera Coordinated Frame และ Marker Coordinated Frame แสดงดังสมการที่ (1)

$$\begin{bmatrix} X_C \\ Y_C \\ Z_C \\ 1 \end{bmatrix} = \begin{bmatrix} R_{11} & R_{12} & R_{13} & T_1 \\ R_{21} & R_{22} & R_{23} & T_2 \\ R_{31} & R_{32} & R_{33} & T_3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_M \\ Y_M \\ Z_M \\ 1 \end{bmatrix} = T_{CM} \begin{bmatrix} X_M \\ Y_M \\ Z_M \\ 1 \end{bmatrix}$$

รูปที่ 2.8 สมการที่ (1)

ซึ่ง Camera Coordinated Frame ก็คือ Coordinated Frame ที่ใช้อ้างอิงตำแหน่งใด ๆ ของกล้องวิดีโอ และ Marker Coordinated Frame ก็คือ Coordinated Frame ที่ใช้อ้างอิงตำแหน่งใดๆ ของ Marker ซึ่งสามารถแสดงดังรูปที่ 2.9



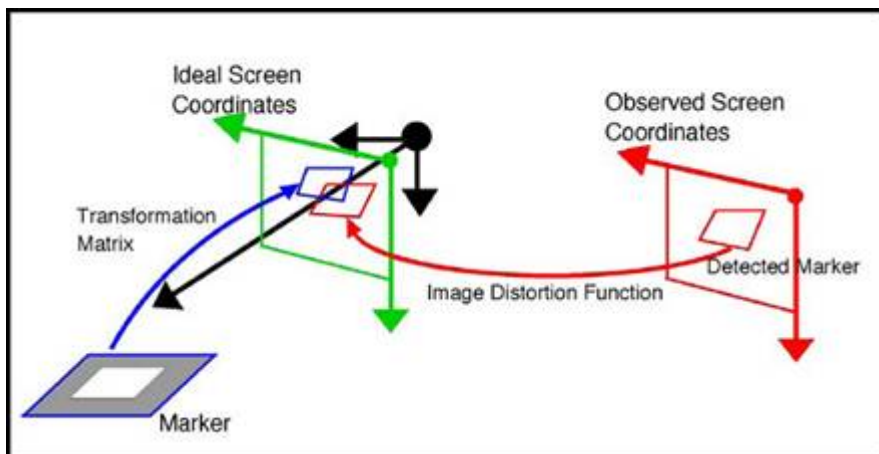
รูปที่ 2.9 ความสัมพันธ์ระหว่าง Camera Coordinated Frame และ Marker Coordinated Frame

จากรูปที่ 2.9 ความสัมพันธ์ระหว่างจุดใดๆ \$(X\_c, Y\_c, Z\_c)\$ บน Camera Coordinated Frame กับจุดที่ตรงกันใน Ideal Screen Coordinated Frame เป็นไปตาม Perspective Projection แสดงดังสมการที่ (2)

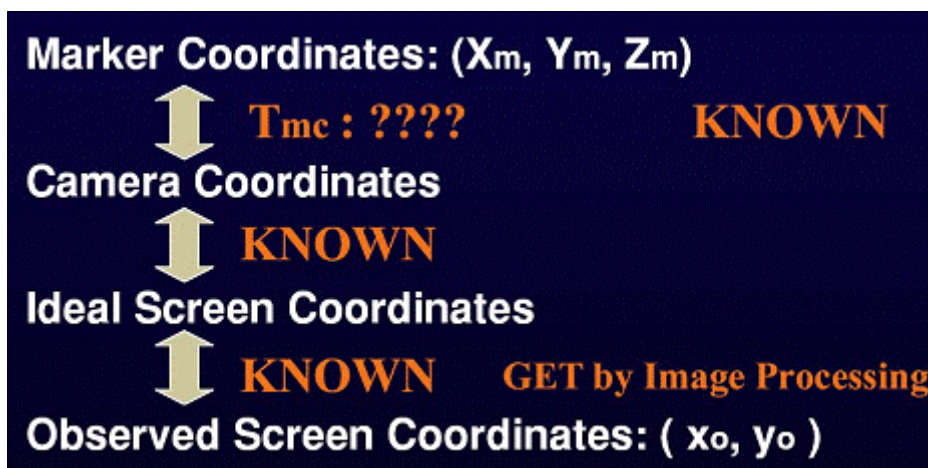
$$\begin{bmatrix} hx_I \\ hy_I \\ h \end{bmatrix} = \begin{bmatrix} sf_x & 0 & x_c & 0 \\ 0 & sf_y & y_c & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} = C \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix}$$

รูปที่ 2.10 สมการที่ (2)

โดยที่ \$C\$ ซึ่งเป็นเมตริกซ์ขนาด \$3 \times 4\$ ซึ่งประกอบไปด้วยค่า \$s, f\_x, f\_y, x\_c, y\_c\$ โดยทั่วไปค่าเหล่านี้รวมกันเรียกว่า Camera Parameters ซึ่งจะคำนวณได้มาจากขั้นตอน Camera Calibration ส่วนค่าความสัมพันธ์ระหว่างจุดใด ๆ บน Ideal Screen Coordinated Frame \$(X\_I, Y\_I)\$ กับ Observe Screen Coordinated Frame \$(X\_0, Y\_0)\$ ซึ่งเป็นจุดที่เราเห็นจริงๆ ในภาพ สามารถแสดงดังรูปที่ 2.11 และสามารถอธิบายได้แสดงดังสมการที่ (3)



รูปที่ 2.11 ความสัมพันธ์ระหว่าง Ideal Screen Coordinates และ Observed Screen Coordinates



รูปที่ 2.12 กระบวนการคำนวณค่า 3D Poses

จากรูปที่ 2.12 จะแสดงกระบวนการที่จะได้มาของค่า TCM เมื่อเรารู้ค่าตำแหน่งของ Marker ทั้ง 4 จุดบน Observed Screen Coordinates ในภาพที่ถ่ายจากกล้องวิดีโอ ซึ่งกล่าวโดยเฉพาะค่านี้สามารถหาได้จากการคำนวณหาคำตอบของฟังก์ชันค่าผิดพลาด (Error Function) แสดงดังสมการที่ (4) ซึ่งโดยทั่วไปแล้วเราจะใช้เทคนิคทางด้านกรหาค่าที่เหมาะสม (Optimization) ซึ่งเป็นกระบวนการแบบ Iterative

$$err = \frac{1}{4} \sum_{i=1,2,3,4} \{ (x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2 \}$$

รูปที่ 2.13 สมการที่ (4)

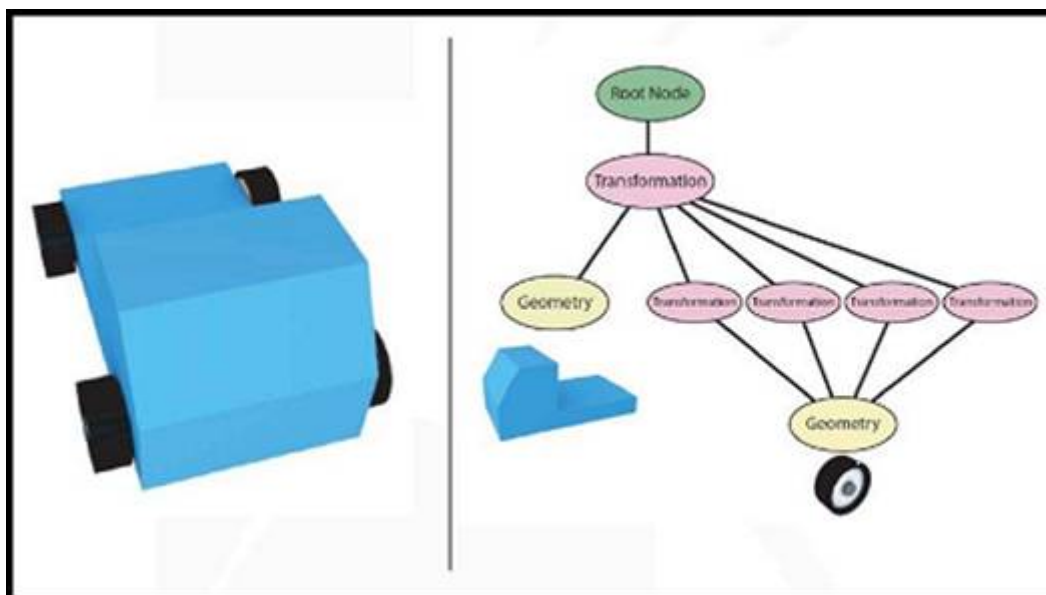
โดยที่  $\hat{x}_i$   $\hat{y}_i$  แสดงได้โดย

$$\begin{bmatrix} h\hat{x}_i \\ h\hat{y}_i \\ h \end{bmatrix} = \mathbf{C} \cdot \mathbf{T}_{CM} \begin{bmatrix} X_{Mi} \\ Y_{Mi} \\ Z_{Mi} \\ 1 \end{bmatrix}, i = 1,2,3,4$$

รูปที่ 2.14 สมการที่ (5)

### 3D rendering <sup>[3]</sup>

3D rendering เป็นส่วนสุดท้ายที่จะทำให้กระบวนการ AR ครบถ้วนสมบูรณ์ ซึ่งก็คือ การเพิ่ม (Augment) ข้อมูลที่เราต้องการซึ่งโดยทั่วไปแล้วจะเป็นโมเดล 3 มิติ (3D Model) ลงไปในภาพที่ได้จากกล้องวิดีโอ ณ ตำแหน่งของ Marker ที่ตรวจพบจากขั้นตอน Image Analysis โดยใช้ค่าตำแหน่งเชิง 3 มิติที่คำนวณได้จากขั้นตอน Pose Estimation กล่าวโดยทั่วไปแล้ว 3D Rendering หมายถึง กระบวนการที่ทำการสร้างภาพ 2 มิติ จากโมเดล 3 มิติ ซึ่งโมเดล 3 มิติ นี้ จะอธิบายวัตถุหรือสิ่งแวดล้อมหนึ่ง ๆ ที่เราต้องการสร้างภาพนั้น เนื่องจากเทคนิคทางด้าน 3D Rendering นั้นมีหลากหลาย จึงยกตัวอย่างเทคนิค 3D Rendering โดยใช้หลักการ Scene Graph (โดยเลือกตามเครื่องมือที่เราจะใช้ซึ่งคือ Open Scene Graph) ดังนั้นจึงขออธิบายหลักการโดยคร่าว ๆ ของ Scene Graph



รูปที่ 2.16 กระบวนการคำนวณค่า 3D Poses

จากภาพที่ 2.16 จะเห็นได้ว่ามีอยู่สองส่วนด้วยกัน ส่วนทางซ้ายคือ ส่วนของโมเดล 3 มิติ ที่ได้รับการ Render หรือ การแสดงออกมาเป็นภาพเรียบร้อยแล้ว ส่วนทางขวาก็คือ โครงสร้างของ Scene Graph ซึ่งเป็น Tree-like structure ซึ่งเมื่อ Render ตาม Tree นี้แล้วก็จะได้โมเดล 3 มิติตามที่ปรากฏในส่วนทางซ้ายออกมา เมื่อพิจารณากันที่ตัว Tree ที่อยู่ทางขวา ดังนี้ วิธีการแหวผ่านต้นไม้ (Tree Traversal) แบบ “การแหวผ่านแบบก่อนลำดับ” (Preorder Traversal) ก็จะสามารถอธิบายเป็น Node โดยเริ่มจาก RootNode ได้

การแปลงค่า (Transformation) Node นี้จะเป็น Node ที่เก็บ Matrix ขนาด  $4 \times 4$  ที่ได้จากขั้นตอน Pose estimation เอาไว้ โดยที่กล่าวได้คร่าว ๆ ว่า Node นี้ จะเป็น Node ที่เก็บค่าสำหรับการระบุตำแหน่งที่ต้องการแสดงภาพในส่วนต่าง ๆ ลงในหน้าจอ โดยใช้ Matrix ดังกล่าวนี เพื่อเป็นตัวแปลง พิกัดในเชิง 3 มิติของ Camera หรือ Viewer เพื่อให้อยู่ในรูปของ พิกัดในเชิง 3 มิติของ Object หรือ โมเดล 3 มิติ ที่ต้องการแสดงลงไปบนหน้าจอ