# บทที่ 2
# ความรู้พื้นฐานและทฤษฎีที่เกี่ยวข้อง

ในบทนี้จะกล่าวถึงความรู้พื้นฐานและทฤษฎีที่จำเป็นสำหรับวิทยานิพนธ์ โดยจะแบ่งเนื้อหาออกเป็น 6 หัวข้อหลัก คือ ปริภูมิสีและขั้นตอนการแปลงปริภูมิสีที่จำเป็นในงานวิจัย แบบจำลองกล้องวีดิทัศน์และแบบจำลองการเคลื่อนที่ของกล้องวีดิทัศน์ ขั้นตอนวิธีการติดตามบุคคล ฟังก์ชันความเหมือนของ Bhattacharyya สำหรับการเปรียบเทียบฮิสโทแกรม ทฤษฎีเกม และทฤษฎีเกมกับการคัดเลือกกล้องวีดิทัศน์ที่มีมุมมองต่อบุคคลที่ติดตามที่ดีที่สุด

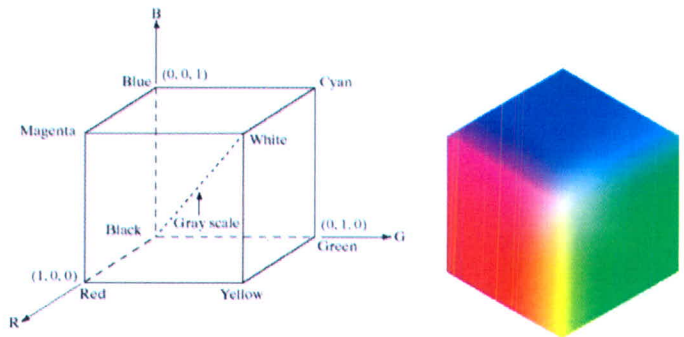
## 2.1 ปริภูมิสี

งานวิจัยทางด้านการประมวลผลภาพและวีดิทัศน์มักอาศัยข้อมูลสีเป็นลักษณะเด่น (Feature) เนื่องจากเป็นข้อมูลที่นำมาใช้ได้โดยตรง จึงสะดวกในการนำมาใช้สำหรับการประมวลผลภาพและวีดิทัศน์ ข้อมูลสีจะถูกเก็บอยู่ในรูปค่าสีซึ่งเป็นผลรวมขององค์ประกอบฐานสี (Color basis) เข้าเป็นสีต่าง ๆ ตามภาพที่รับได้ในหนึ่งจุดภาพ (Pixel) ดังนั้นจึงสามารถแทนค่าสีด้วยเวกเตอร์ค่าสีในปริภูมิสี (Color space) ได้หลายรูปแบบ โดยปริภูมิสีแบบ RGB เป็นปริภูมิสีที่ถูกนำมาใช้มากที่สุด เนื่องจากเป็นองค์ประกอบสีหลักในอุปกรณ์แสดงผลทั่ว ๆ ไป ทั้งนี้มนุษย์มีความสามารถในการมองเห็นความสว่างและสีได้ดีไม่เท่ากัน โดยสายตามนุษย์จะมีความไวต่อความสว่างมากกว่าสี จึงเกิดปริภูมิสีที่มีการแยกระหว่างความส่องสว่าง (Luminance) และค่าสี (Chrominance) ขึ้น เช่น ปริภูมิสี YUV YCbCr และHSL

### 2.1.1 ปริภูมิสี RGB

ปริภูมิสี RGB เป็นองค์ประกอบสีที่พบได้ในอุปกรณ์แสดงผลทั่วไปที่อาศัยการแผ่รังสีในย่านคลื่นแสงที่มองเห็นอันประกอบด้วยฐานสีแดง สีเขียว และสีน้ำเงิน การทำงานกับปริภูมิสี RGB จึงมีความสะดวกในการจัดการและการออกแบบระบบ ค่าสีที่เกิดจากปริภูมิสี RGB สามารถอธิบายได้จากแบบจำลองสี RGB ในลักษณะลูกบาศก์ในระบบพิกัดคาร์ทีเซียนดังแสดงในรูปที่ 2.1 โดยค่าสีแดง สีเขียว และสีน้ำเงินจะอยู่ที่แต่ละมุมหลักของลูกบาศก์ สีดำจะอยู่ที่จุดกำเนิด ส่วนสีขาวจะอยู่

ที่มุมตรงข้ามจุดกำเนิด หากลากเส้นเชื่อมต่อระหว่างสีดำและสีขาวจะได้ค่าระดับสีเทา (Gray level) ตามเส้นดังกล่าวที่มีค่า ที่อยู่




รูปที่ 2.1 แบบจำลองสี RGB

แต่ละค่าสีสามารถหาได้จากการพิจารณาเวกเตอร์ค่าสีซึ่งเป็นผลรวมเชิงเส้นขององค์ประกอบสีแดง สีเขียว และสีน้ำเงิน เวกเตอร์ค่าสี ณ จุดภาพหลักที่ $i$ แถวที่ $j$ ของลำดับภาพที่ $n$ ดังสมการที่ (2.1)

$$\mathbf{X}_{i,j}[n] = (X_{i,j}^{R}[n], X_{i,j}^{G}[n], X_{i,j}^{B}[n]) \tag{2.1}$$

โดย $X_{i,j}^{R}[n]$ เป็นค่าความเข้มขององค์ประกอบสีแดง ซึ่ง $0 \le X_{i,j}^{R}[n] \le D^{R}$

$X_{i,j}^{G}[n]$ เป็นค่าความเข้มขององค์ประกอบสีแดง ซึ่ง $0 \le X_{i,j}^{G}[n] \le D^{G}$

$X_{i,j}^{B}[n]$ เป็นค่าความเข้มขององค์ประกอบสีแดง ซึ่ง $0 \le X_{i,j}^{B}[n] \le D^{B}$

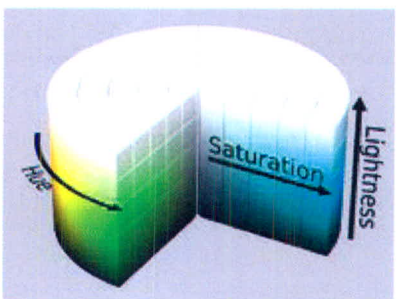
โดย $D^{R}$ $D^{G}$ และ $D^{B}$ เป็นค่าสีสูงสุดในองค์ประกอบสีแดง สีเขียว และสีน้ำเงินที่ขึ้นอยู่กับค่าความลึกของสี (Color depth) จากกระบวนการการแจงหน่วย ค่าความลึกของสีที่นิยมใช้ในการประมวลผลภาพและวีดิทัศน์ทั่วไปจะอยู่ในความลึกสี 24 บิตต่อจุดภาพ อันประกอบด้วยองค์ประกอบสีละ 8 บิตต่อจุดภาพ นั่นคือ $D^{R}$ $D^{G}$ และ $D^{B}$ มีค่าเป็น 255 ทั้งหมด

ปริภูมิสี RGB อาจไม่เหมาะสมในงานด้านการประมวลผลภาพและวีดิทัศน์ เนื่องจากตัวรับรู้สี (Color sensor) ที่อยู่ในอุปกรณ์รับภาพมีความไว (Sensitivity) ต่อแสงย่านสีแดง สีเขียว และสีน้ำเงินแตกต่างกัน ส่งผลให้เกิดความแตกต่างในรายละเอียดของระดับสีที่ได้รับ ทำให้ต้องปรับสมดุล

ค่าสี (Color equalization) ซึ่งจะทำให้สูญเสียข้อมูลในแถบสัญญาณสีที่มีความละเอียดสูง อีกทั้งในการอ่านค่าสีของปริภูมิ RGB สำหรับการประมวลผลภาพและวีดิทัศน์จะทำการอ่านค่าสีทั้ง 3 แถบสีพร้อมกันเสมอ ปัญหาที่สำคัญที่สุดสำหรับปริภูมิสี RGB คือ การรวมกันของค่าความสว่างและค่าสี ส่งผลให้ในการประมวลผลภาพและวีดิทัศน์ที่อาศัยค่าสีเป็นลักษณะเด่นเกิดความผิดพลาดเมื่อต้องทำงานกับสภาวะแวดล้อมที่มีปัจจัยทางด้านแสงสว่างเข้ามาเกี่ยวข้อง

### 2.1.2 ปริภูมิสี HSL

ปริภูมิสี HSL สร้างขึ้นตามพื้นฐานการมองเห็นสีด้วยสายตาของมนุษย์ ประกอบด้วยค่าสีสัน (Hue) ค่าความอิ่มตัวของสี (Saturation) และค่าความสว่างของสี (Lightness) โดยค่าสีสันเป็นค่าสีที่สะท้อนจากสีของวัตถุ ซึ่งแตกต่างกันตามความยาวของคลื่นแสงที่มากระทบวัตถุและสะท้อนกลับมาที่ตาของมนุษย์ ทำมุมได้ 0 ถึง 360 องศา โดยสีแดง สีเหลือง และสีเขียวจะมีค่าต่างกันสีละ 60 องศา สำหรับค่าความอิ่มตัวของสีสามารถเรียกได้ว่าเป็นค่าความเข้มของเนื้อสีหรือค่าความบริสุทธิ์ของสี โดยมีค่าเป็นเปอร์เซ็นต์ตั้งแต่ 0 ถึง 100 เปอร์เซ็นต์ กล่าวคือ สีของวัตถุจะมีความเข้มมากขึ้นเมื่อค่าอิ่มตัวของสีมีค่าเพิ่มขึ้น และจะอิ่มตัวเต็มที่เมื่อค่าความอิ่มตัวของสีมีค่าเท่ากับ 100 เปอร์เซ็นต์ ส่วนค่าความสว่างของสีเป็นค่าที่ใช้บอกปริมาณความเข้มของแสงในหน่วยเปอร์เซ็นต์ จาก 0 ถึง 100 ซึ่งเป็นค่าความสว่างของสีที่สว่างที่สุด ค่าสีที่เกิดจากปริภูมิสี HSL สามารถอธิบายได้จากแบบจำลองสี HSL ในลักษณะรูปทรงกระบอกในระบบพิกัดทรงกระบอก ดังแสดงในรูปที่ 2.2




รูปที่ 2.2 แบบจำลองสี HSL

ข้อดีในการใช้ปริภูมิสี HSL ในการประมวลผลภาพและวีดิทัศน์ คือ องค์ประกอบในปริภูมิมีการแยกองค์ประกอบเชิงความสว่างออกจากองค์ประกอบสีอย่างชัดเจน การวิเคราะห์ความ

เปลี่ยนแปลงของแต่ละองค์ประกอบจึงเป็นอิสระต่อกัน วิทยานิพนธ์นี้จึงเลือกใช้ปริภูมิสี HSL โดยพิจารณาเฉพาะช่วงค่าความอิ่มตัวของสีและค่าความสว่างของสีที่อยู่ในช่วง 10 ถึง 90 เปอร์เซ็นต์ เนื่องจากเป็นช่วงที่เหมาะสมในการประมวลผลวีดิทัศน์ซึ่งได้จากการทำการทดลอง

ทั้งนี้องค์ประกอบสีที่ได้จากกล้องวีดิทัศน์เป็นองค์ประกอบสีที่อยู่ในปริภูมิสี RGB จึงจำเป็นต้องทำการแปลงให้อยู่ในปริภูมิสี HSL โดยมีขั้นตอนดังนี้

1. นำข้อมูลสีแดง ($R$) สีเขียว ($G$) และสีน้ำเงิน ($B$) ที่ได้จากปริภูมิสี RGB มาทำการหาค่าแถบสีสูงสุด ($C_{MAX}$) ค่าแถบสีต่ำสุด ($C_{MIN}$) และผลต่างระหว่างค่าแถบสีสูงสุดต่ำสุด ($\Delta$) ดังสมการที่ (2.2) (2.3) และ (2.4) ตามลำดับ

$$C_{MAX} = \max(R', G', B') \tag{2.2}$$

$$C_{MIN} = \min(R', G', B') \tag{2.3}$$

$$\Delta = C_{MAX} - C_{MIN} \tag{2.4}$$

เมื่อ $R' = \frac{R}{255}$ , $G' = \frac{G}{255}$ และ $B' = \frac{B}{255}$

2. หาค่าสีสัน ($H$) ในหน่วยองศา ความอิ่มตัวของสี ($S$) ในหน่วยเปอร์เซ็นต์ และความสว่างของสี ($L$) ในหน่วยเปอร์เซ็นต์ ได้จากสมการที่ (2.5) (2.6) และ (2.7) ตามลำดับ

$$H = \begin{cases} 60° \times \left( \frac{G' - B'}{\Delta} \bmod 6 \right), & C_{MAX} = R' \\ 60° \times \left( \frac{B' - R'}{\Delta} + 2 \right), & C_{MAX} = G' \\ 60° \times \left( \frac{R' - G'}{\Delta} + 4 \right), & C_{MAX} = B' \end{cases} \tag{2.5}$$

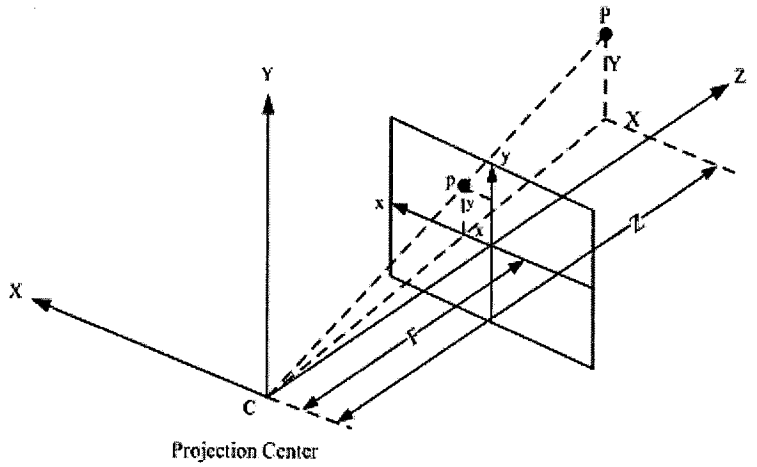
$$S = \begin{cases} 0, & \Delta = 0 \\ \dfrac{\Delta}{1 - \left| 2\left( \frac{C_{MAX} + C_{MIN}}{2} \right) - 1 \right|}, & \Delta \neq 0 \end{cases} \tag{2.6}$$

$$L = \left( \frac{C_{MAX} + C_{MIN}}{2} \right) \tag{2.7}$$

## 2.2 แบบจำลองกล้องวีดิทัศน์และแบบจำลองการเคลื่อนที่กล้องวีดิทัศน์

ทฤษฎีการทำงานเกี่ยวกับแบบจำลองกล้องมีความสำคัญในด้านการมโนภาพ (Imagine) และทำให้เข้าใจหลักการหลักการเกี่ยวกับการควบคุมกล้องวีดิทัศน์ โดยแบบจำลองกล้องวีดิทัศน์เป็นการจำลองรูปแบบการสร้างภาพที่ปรากฏอยู่บนจอภาพในพิกัดภาพ (Image coordinate) 2 มิติ จากวัตถุในฉากจริง (Real scene) ในพิกัด 3 มิติ มีแบบจำลองหลายรูปแบบที่ใช้บรรยายการสร้างภาพ เช่น แบบจำลองกล้องรูเข็ม (Pinhole camera model) และแบบจำลองกล้อง CAHV (CAHV Camera model) เป็นต้น โดยแบบจำลองกล้องวีดิทัศน์ที่สามารถแสดงความเข้าใจพื้นฐานเกี่ยวกับกล้องวีดิทัศน์ได้อย่างดี คือ แบบจำลองกล้องวีดิทัศน์แบบรูเข็ม

### 2.2.1 แบบจำลองกล้องวีดิทัศน์แบบรูเข็ม




รูปที่ 2.3 การฉายภาพแบบทัศนมิติ ของแบบจำลองกล้องรูเข็ม

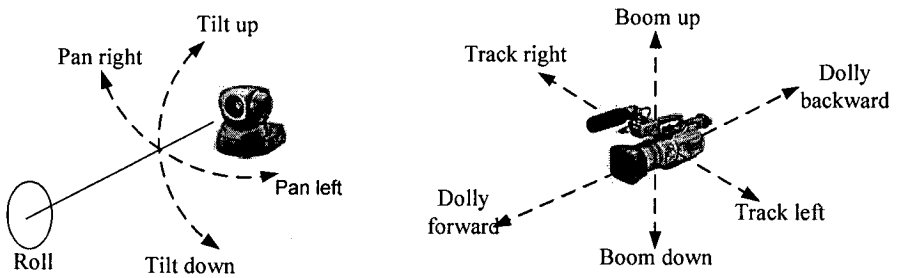
แบบจำลองกล้องรูเข็ม เป็นแบบจำลองที่นิยมใช้กันอย่างมากมายในการวิเคราะห์การสร้างภาพและงานประยุกต์ใน 3 มิติ แสดงดังรูปที่ 2.3 โดยอาศัยพื้นฐานการฉายภาพแบบทัศนมิติ (Perspective projection) จากวัตถุในพิกัด 3 มิติ ในการสร้างภาพพิกัด 2 มิติ

จากรูปที่ 2.3 หากพิจารณาที่จุด $P$ ในพิกัด 3 มิติ หรือจุด $P(X,Y,Z)$ สามารถฉายลงมายังจุด $P(x,y)$ ในระนาบภาพ (Image plane) 2 มิติ ได้ดังสมการที่ (2.8) โดยมีจุดตำแหน่ง $C$ เป็นจุดศูนย์กลางในการฉายภาพ (Center of projection) และ $F$ เป็นความยาวโฟกัส (Focal length)

$$x = F\left(\frac{X}{Z}\right)$$
$$y = F\left(\frac{Y}{Z}\right) \tag{2.8}$$

### 2.2.2 แบบจำลองการเคลื่อนที่กล้องวีดิทัศน์

การเคลื่อนที่ของกล้องโดยทั่วไปจะแบ่งออกได้เป็น 2 รูปแบบ คือ




(ก) ตำแหน่งศูนย์กลางไม่เปลี่ยนแปลง (ข) ตำแหน่งศูนย์กลางมีการเปลี่ยนแปลงตำแหน่ง

รูปที่ 2.4 การเคลื่อนที่ของกล้องวีดิทัศน์

#### 2.2.2.1 ตำแหน่งศูนย์กลางการฉายภาพไม่เปลี่ยนแปลง

การเคลื่อนที่ของกล้องวีดิทัศน์ในรูปแบบนี้จะประกอบด้วยการส่าย (Pan) เป็นลักษณะการเคลื่อนที่กล้องโดยยึดแกนตั้ง (Vertical axis) เป็นแกนหมุนรอบ การก้มเงย (Tilt) เป็นลักษณะการเคลื่อนที่กล้องโดยยึดแกนนอน (Horizontal axis) เป็นแกนหมุนรอบ และ การหมุน (Roll) เป็นลักษณะการเคลื่อนที่กล้องโดยยึดแกนเชิงแสง (Optical axis) เป็นแกนหมุนรอบ ดังแสดงในรูปที่ 2.4(ก) การเคลื่อนที่ของกล้องวีดิทัศน์ที่ตำแหน่งศูนย์กลางการฉายภาพไม่เปลี่ยนแปลงนี้ เป็นการเคลื่อนที่ของกล้องวีดิทัศน์สำหรับการสอดส่องดูแล (Surveillance) โดยทั่วไป ซึ่งเป็นรูปแบบเดียวกับที่ใช้ในวิทยานิพนธ์นี้

#### 2.2.2.2 ตำแหน่งศูนย์กลางการฉายภาพมีการเปลี่ยนแปลงตำแหน่ง

ประกอบด้วยการติดตาม (Track) ซึ่งเป็นลักษณะการเคลื่อนที่กล้องในลักษณะการเลื่อนขนาน (Translation) ตามแนวแกนนอนของระนาบภาพ การยก (Boom) เป็นลักษณะการเคลื่อนที่กล้องในลักษณะการเลื่อนขนานตามแนวแกนตั้งของระนาบภาพ และการดอลลี (Dolly) เป็นลักษณะการเคลื่อนที่กล้องในลักษณะการเลื่อนขนานตามแนวแกนเชิงแสง ดังแสดงในรูปที่ 2.4 (ข)

## 2.3 การติดตามวัตถุหรือบุคคลในลำดับภาพวีดิทัศน์

การติดตามวัตถุหรือบุคคลที่เคลื่อนที่ภายในลำดับภาพวีดิทัศน์ คือ การพยายามที่จะระบุตำแหน่งของบุคคลหรือวัตถุในลำดับภาพวีดิทัศน์ให้ถูกต้องและต่อเนื่องไปในทุก ๆ ลำดับภาพ หรือทำนายว่าทิศทางการเคลื่อนที่ของวัตถุหรือบุคคลนั้นควรเคลื่อนที่ไปในทิศทางใด โดยทั่วไปแล้วสามารถแบ่งรูปแบบการติดตามวัตถุหรือบุคคลที่เคลื่อนที่ได้ 4 ประเภท คือ การติดตามวัตถุหรือบุคคลด้วยโมเดลรูปร่าง (Model-based tracking) การติดตามวัตถุหรือบุคคลด้วยพื้นที่ (Region-based tracking) การติดตามวัตถุหรือบุคคลด้วยขอบ (Contour-based tracking) และการติดตามวัตถุหรือบุคคลโดยอาศัยลักษณะเด่น (Feature-based tracking) การติดตามวัตถุหรือบุคคลด้วยโมเดลรูปร่าง เป็นวิธีที่จะติดตามวัตถุที่มีรูปร่างลักษณะเข้ากับโมเดลรูปร่างที่เตรียมไว้ ซึ่งการติดตามวัตถุหรือบุคคลด้วยวิธีนี้ จำเป็นต้องสร้างโมเดลรูปร่างที่สนใจไว้ก่อน ดังนั้นวิธีการนี้จึงมีข้อจำกัด คือ ไม่สามารถทำงานกับวัตถุที่ไม่ทราบรูปร่างล่วงหน้าได้ การติดตามวัตถุหรือบุคคลด้วยพื้นที่ เป็นการดึงข้อมูลสำคัญออกมาเป็นลักษณะของพื้นที่ที่ประกอบไปด้วยข้อมูลใดๆ เช่น ข้อมูลของสี ข้อมูลลวดลาย เป็นต้น จากนั้นจะทำการติดตามพื้นที่เหล่านั้นโดยหาข้อมูลของสิ่งที่ดึงออกมาใช้นั่นเอง การติดตามวัตถุหรือบุคคลด้วยขอบ วิธีนี้จะขึ้นอยู่กับขอบของวัตถุหรือบุคคลที่มีการเคลื่อนที่ ไม่ได้ขึ้นกับตัววัตถุทั้งชิ้น ข้อมูลทางขอบวัตถุจะถูกดึงออกมาถูกปรับแต่งในแต่ละลำดับภาพที่เกิดขึ้น ดังนั้นการติดตามวัตถุที่เคลื่อนที่จะได้ผลดีหรือไม่ขึ้นกับขอบวัตถุเริ่มแรกที่ถูกดึงออกมาได้ สำหรับการติดตามวัตถุหรือบุคคลโดยอาศัยคุณลักษณะเฉพาะ มีจุดประสงค์ที่จะค้นหาและติดตามวัตถุที่เคลื่อนที่โดยใช้ลักษณะเด่นบางอย่างที่ได้จากการคำนวณในระดับสูงขึ้น เช่น การประมาณความเร็วของวัตถุ การหาค่าการกระจายตัวของข้อมูลของวัตถุ การประมาณความ

หนาแน่นของข้อมูลวัตถุ เป็นต้น ซึ่งวิธีการประเภทนี้จะมีข้อดี คือ วัตถุหรือบุคคลที่จะติดตามนั้นจะมีข้อมูลแยกออกมาจากสิ่งแวดล้อมได้อย่างดีกว่าวิธีอื่น ๆ

สำหรับวิทยานิพนธ์นี้ได้เลือกใช้ขั้นตอนวิธีการย้ายตามค่าเฉลี่ยแบบปรับตัวได้อย่างต่อเนื่อง (Continuously Adaptive Mean-Shift) ซึ่งเป็นขั้นตอนวิธีการติดตามวัตถุโดยอาศัยลักษณะเด่นของความหนาแน่นของข้อมูลวัตถุมาประยุกต์ใช้ในขั้นตอนการติดตามบุคคล ซึ่งมีข้อดี คือ สามารถทำการติดตามวัตถุที่มีการเปลี่ยนแปลงทางรูปร่างและขนาดได้ดี และไม่พิจารณาข้อมูลที่อยู่ภายนอกหน้าต่างการค้นหาทำให้สัญญาณภาพรบกวนและลักษณะเด่นของข้อมูลอื่นที่อยู่ใกล้เคียงกับวัตถุไม่ส่งผลต่อการติดตาม ทั้งนี้ในขั้นตอนการเริ่มต้นของการติดตามวัตถุหรือบุคคลในลำดับภาพวีดิทัศน์จำเป็นต้องมีการกำหนดจุดเริ่มต้นหรือพื้นที่เริ่มต้นก่อนซึ่งเรียกว่าการตรวจจับวัตถุ (Object detection) โดยพื้นฐานมีวิธีที่นิยม 3 วิธี ได้แก่ วิธีการลบฉากหลัง (Background Subtraction) ที่ทำการเปรียบเทียบภาพในปัจจุบันกับภาพฉากหลังที่สร้างไว้ วิธีการหาความแตกต่างของภาพทางเวลา (Frame Differencing) เป็นการหาความแตกต่างของภาพที่อยู่ติดกัน วัตถุที่เคลื่อนที่จะถูกแสดงความแตกต่างให้อยู่ในรูปผลต่างของทั้งสองภาพ เป็นวิธีที่สามารถทำงานได้อย่างรวดเร็ว แต่ผลการตรวจจับวัตถุจะไม่สมบูรณ์เนื่องจากได้จากเพียงผลต่างของภาพที่อยู่ติดกัน และวิธีการหาการไหลของวัตถุที่เคลื่อนที่ในภาพ (Optical Flow) ซึ่งสามารถนำไปประยุกต์ใช้งานได้กับกล้องวีดิทัศน์ที่อยู่นิ่ง และกล้องวีดิทัศน์ที่มีการเคลื่อนที่ไปด้วยได้ แต่การคำนวณจะมีความซับซ้อน สำหรับวิทยานิพนธ์นี้ให้ความสนใจกับการตรวจจับบุคคลขณะที่กล้องวีดิทัศน์ยังไม่มีการเคลื่อนที่ ดังนั้นวิธีการตรวจจับบุคคลโดยใช้วิธีการลบฉากหลังจึงเป็นวิธีที่เหมาะสม ซึ่งเป็นวิธีที่นิยมและสามารถทำงานได้อย่างรวดเร็ว

### 2.3.1 การตรวจจับวัตถุหรือบุคคลที่เคลื่อนที่ด้วยวิธีการลบภาพฉากหลัง

วิธีการลบฉากหลัง (Background Subtraction) เป็นวิธีการที่นำเอาภาพที่รับเข้ามาในปัจจุบันมาเปรียบเทียบกับภาพอ้างอิงที่ถูกสร้างไว้เพื่อใช้ในการตรวจจับวัตถุที่เคลื่อนที่อยู่ในลำดับภาพวีดิทัศน์ วิธีการลบฉากหลังแบบพื้นฐานที่สุด คือ การหาผลต่างระหว่างภาพปัจจุบันที่รับเข้ามากับแบบจำลองฉากหลังสามารถอธิบายได้ดังสมการที่ (2.9)

$$\left|I(x,y,t)-B(x,y,t')\right|>Th \tag{2.9}$$

เมื่อ $I(x,y,t)$ คือ ลำดับภาพปัจจุบันที่รับเข้ามา ณ เวลา $t$ ใด ๆ

$B(x,y,t')$ คือ แบบจำลองฉากหลังที่ถูกสร้างเมื่อเวลา $t'$

และ $Th$ คือ ค่าขีดเริ่มเปลี่ยน (Threshold) ซึ่งเป็นตัวกำหนดขนาดผลต่างที่ยอมรับได้

แบบจำลองฉากหลังอาจถูกสร้างจากลำดับภาพแรกที่รับเข้ามา หรือเกิดจากการประมาณจากแบบจำลองต่าง ๆ ขึ้นมา ทั้งนี้องค์ประกอบต่าง ๆ ในลำดับภาพวีดิทัศน์อาจทำให้การตรวจหาวัตถุหรือบุคคลจากวิธีการลบฉากหลังมีความผิดพลาดได้ เช่น การเปลี่ยนแปลงของฉากหลังจากความสว่าง อัตราลำดับภาพ การเคลื่อนที่ของบุคคลและกล้องวีดิทัศน์ วิทยานิพนธ์ฉบับนี้จึงเลือกใช้วิธีการลบฉากหลังแบบปรับตัวได้อย่างง่าย ด้วยการกำหนดจำนวนลำดับภาพที่ใช้สำหรับการปรับเปลี่ยนฉากหลัง

### 2.3.2 ขั้นตอนวิธีการย้ายตามค่าเฉลี่ยแบบปรับตัวได้อย่างต่อเนื่อง

วิธีการย้ายตามค่าเฉลี่ยแบบปรับตัวได้อย่างต่อเนื่อง (Continuously Adaptive Mean-Shift, CAMShift) เป็นวิธีการติดตามวัตถุที่นำเอาพื้นฐานของการติดตามแบบการย้ายตามค่าเฉลี่ย (Mean-Shift) มาทำงานร่วมกับขั้นตอนการปรับเปลี่ยนขนาดของหน้าต่างการค้นหา ซึ่งมีข้อดี คือ สามารถทำการติดตามวัตถุที่มีการเปลี่ยนแปลงทางรูปร่างและขนาดได้ดี และไม่พิจารณาข้อมูลที่อยู่ภายนอกหน้าต่างการค้นหาทำให้สัญญาณภาพรบกวนและลักษณะของข้อมูลอื่นที่ใกล้เคียงกับวัตถุไม่ส่งผลต่อการติดตาม อีกทั้งตราบใดที่วัตถุไม่ถูกบดบังทั้งหมดจากสิ่งกีดขวาง CAMShift จะยังคงมีแนวโน้มติดตามวัตถุได้จากข้อมูลที่เหลืออยู่ภายในหน้าต่างการค้นหา

จาก [13] ได้นำเสนอวิธีการย้ายตามค่าเฉลี่ยแบบปรับตัวได้อย่างต่อเนื่องมาใช้สำหรับการติดตามวัตถุด้วยการคำนวณจากภาพวีดิทัศน์แบบ 2 มิติ โดยการทำงานของ CAMShift [13] จะพิจารณาจากฟังก์ชั่นการแจกแจงความน่าจะเป็นของสีของวัตถุโดยพัฒนาต่อยอดจากขั้นตอนวิธีการติดตามวัตถุแบบการย้ายตามค่าเฉลี่ย (Mean-Shift) [21] ด้วยการให้มีการปรับเปลี่ยนขนาดของกรอบการติดตามวัตถุให้สามารถลดหรือขยายขนาดจากเดิมได้ และการแจกแจงความน่าจะเป็น

ของสีของวัตถุที่ติดตามก็จะถูกปรับเปลี่ยนไปด้วย ขั้นตอนการทำงานของ CAMShift สามารถอธิบายได้ดังนี้

1. กำหนดตำแหน่งตั้งต้นของกรอบการติดตามวัตถุ

2. ทำการค้นหาตำแหน่งใหม่ของวัตถุที่ติดตามด้วยขั้นตอนวิธี Mean-Shift ด้วยการเริ่มต้นจากการคำนวณค่าโมเมนต์ลำดับที่ศูนย์ซึ่งเป็นการคำนวณทางด้านผลรวมของค่าความน่าจะเป็นของสีในกรอบการติดตามวัตถุนั้น สำหรับโมเมนต์ลำดับที่ศูนย์ ($M_{00}$) คำนวณได้ดังสมการที่ (2.10)

$$M_{00} = \sum_{x}\sum_{y} I(x,y) \tag{2.10}$$

เมื่อ $I(x,y)$ คือ ค่าความน่าจะเป็นของสีที่ตำแหน่ง $(x,y)$ ในภาพ

3. คำนวณตำแหน่งกรอบการติดตามวัตถุจากค่าโมเมนต์ลำดับที่ศูนย์ และค่าโมเมนต์ลำดับที่หนึ่งของภาพการแจกแจงความน่าจะเป็นของสีในกรอบการติดตาม การคำนวณโมเมนต์ลำดับที่หนึ่งในแนวแกน x และ y คำนวณได้จากสมการที่ (2.11) และ (2.12)

$$M_{10} = \sum_{x}\sum_{y} xI(x,y) \tag{2.11}$$

$$M_{01} = \sum_{x}\sum_{y} yI(x,y) \tag{2.12}$$

จากนั้นทำการคำนวณหาตำแหน่งจุดศูนย์กลางใหม่ $(x_c, y_c)$ ของกรอบการติดตามจากสมการที่ (2.13) และ (2.14)

$$x_c = \frac{M_{10}}{M_{00}} \tag{2.13}$$

$$y_c = \frac{M_{01}}{M_{00}} \tag{2.14}$$

4. วนซ้ำขั้นตอนที่ 2-3 จนกว่าตำแหน่งกรอบการติดตามวัตถุที่ได้ไม่เปลี่ยนตำแหน่งหรือวนครบจำนวนรอบที่กำหนด (จำนวนรอบปกติที่กำหนดใช้ คือ 10 ถึง 20 รอบ) โดยตำแหน่งที่ได้ล่าสุดจะถูกนำไปใช้เป็นตำแหน่งตั้งต้นของกรอบการติดตามในลำดับภาพถัดไป

5. คำนวณขนาดกรอบการติดตามวัตถุจากค่าโมเมนต์ลำดับที่สองดังสมการที่ (2.15) (2.16) และ (2.17)

$$M_{20} = \sum_x \sum_y x^2 I(x,y) \tag{2.15}$$

$$M_{02} = \sum_x \sum_y y^2 I(x,y) \tag{2.16}$$

$$M_{11} = \sum_x \sum_y yx I(x,y) \tag{2.17}$$

จากนั้นจึงคำนวณพารามิเตอร์ a b และ c สำหรับใช้ในการคำนวณหาความยาวในแนวตั้ง ($l_1$) และแกนนอน ($l_2$) จากจุดศูนย์กลางของกรอบการติดตาม ดังสมการที่ (2.18) (2.19) (2.20) (2.21) และ (2.22) ตามลำดับ

$$a = \frac{M_{20}}{M_{00}} - x_c^2 \tag{2.18}$$

$$b = 2\left(\frac{M_{11}}{M_{00}} - x_c y_c\right) \tag{2.19}$$

$$c = \frac{M_{02}}{M_{00}} - y_c^2 \tag{2.20}$$

$$l_1 = \sqrt{\frac{(a+c) + \sqrt{b^2 + (a-c)^2}}{2}} \tag{2.21}$$

$$l_2 = \sqrt{\frac{(a+c) - \sqrt{b^2 + (a-c)^2}}{2}} \tag{2.22}$$

ผลลัพธ์ของการติดตามวัตถุที่ได้นี้อาจถูกแสดงด้วยภาพของกรอบการติดตามวัตถุ ซึ่งอาจมีการใช้สีหรือระบุตัวเลขเข้าไปในบริเวณของกรอบการติดตาม เพื่อความสะดวกในการเปรียบเทียบผลการทดลองที่ได้จากการติดตามวัตถุนั่นเอง

## 2.4 ฟังก์ชั่นความเหมือนของ Bhattacharyya

การประมวลผลร่วมกันของระบบกล้องวีดิทัศน์เพื่อติดตามบุคคลคนเดียวกัน จำเป็นต้องมีกา ระบวนการวัดความเหมือนของข้อมูลคุณลักษณะสำคัญว่ามีความใกล้เคียงหรือเหมือนกันหรือไม่ ไม่ว่าลักษณะ (Feature) ใด ที่จะนำมาเป็นตัวเปรียบเทียบ เช่น ลักษณะสี รูปร่าง และขอบ เป็นต้น โดยวิธีการหนึ่งที่เป็นที่นิยมใช้สำหรับการเปรียบเทียบความเหมือน คือ กระบวนการวัดความคล้าย (Similarity measure) โดยอาศัยวิธีการวัดระยะทางระหว่างกลุ่มการกระจายตัวของข้อมูล 2 ชุดที่สนใจ ซึ่งมีวิธีการในการหาระยะทางนี้ได้หลายแบบ ในวิทยานิพนธ์นี้จะอาศัยเทคนิคการเปรียบเทียบความเหมือนฮิสโทแกรมด้วยฟังก์ชั่นความเหมือนของ Bhattacharyya (Bhattacharyya similarity function) ซึ่งประกอบไปด้วยตัวแปรสองสวน ได้แก่ระยะ Bhattacharyya (Bhattacharyya distance) และสัมประสิทธิ์ของ Bhattacharyya (Bhattacharyya coefficient)

ฟังก์ชั่นความเหมือนของ Bhattacharyya เป็นรูปแบบการเปรียบเทียบลักษณะหนึ่งในกระบวนการวัดความคล้าย ระหว่างชุดการกระจายตัวของข้อมูล 2 ชุด ซึ่งได้ถูกนำมาประยุกต์ใช้ในกระบวนการทางการประมวลผลภาพสำหรับการเปรียบเทียบฮิสโทแกรมของสองกลุ่มข้อมูลที่ไม่จำเป็นต้องเป็นค่าฮิสโทแกรมของทั้งภาพ อาจเป็นแค่บริเวณใดบริเวณหนึ่งที่สนใจเท่านั้น และไม่จำเป็นว่า ทั้งสองบริเวณจะต้องมีขนาดเท่ากันในหน่วยจุดภาพ จึงเป็นข้อดีที่สามารถรองรับ การเปรียบเทียบวัตถุที่มีการเปลี่ยนแปลงรูปร่างได้ทั้งในลักษณะการเคลื่อนที่เปลี่ยนมุมมองและการเปลี่ยนแปลงขนาด

สมมติให้ $\hat{p}$ เป็นฮิสโทแกรมสีของแบบจำลองเป้าหมายและ $\hat{q}$ เป็นฮิสโทแกรมสีของแบบจำลองรับเข้าที่ต้องการนำมาเปรียบเทียบ โดยที่ $\hat{p}_u = \{p_u\}_{u=1,2,\ldots,m}$ (เมื่อ $\sum_{u=1}^{m} p_u = 1$) และ $\hat{q}_u = \{q_u\}_{u=1,2,\ldots,m}$ (เมื่อ $\sum_{u=1}^{m} q_u = 1$) เมื่อ $m$ เป็นจำนวนถังข้อมูลของฮิสโทแกรม (Bin histogram) เปรียบเทียบ เช่น ถ้าในกรณีเลือกใช้ขนาดขององค์ประกอบสีแต่ละองค์ประกอบมีขนาด 4 บิต ฉะนั้นจะได้ขนาดถังสีเท่ากับ 16 ถัง ดังนั้นสามารถคำนวณหาสัมประสิทธิ์ Bhattacharyya ( $\rho$ ) ได้ดังสมการที่ (2.23)

$$\rho\left[\hat{p},\hat{q}\right]=\sum_{u=1}^{m}\sqrt{p_u\times q_u} \tag{2.23}$$

อธิบายความหมายทางเรขาคณิตตามสมการที่ (2.23) จะแสดงถึงค่าโคไซน์ (Cosine) ของมุมระหว่างค่าถังข้อมูลทั้ง $m$ ถังมิติในรูปเวกเตอร์หนึ่งหน่วย $\left(\sqrt{p_1},\ldots,\sqrt{p_m}\right)^T$ และ $\left(\sqrt{q_1},\ldots,\sqrt{q_m}\right)^T$ และจากสมการที่ 2.23 จะสามารถหาค่าระยะ Bhattacharyya ที่เป็นระยะระหว่างสองกลุ่มของการกระจายตัวของข้อมูล $\left(d_{BH}\right)$ดังสมการที่ (2.24)

$$d_{BH}=\sqrt{1-\rho\left[\hat{p},\hat{q}\right]} \tag{2.24}$$

ข้อดีของการเปรียบเทียบวัดความคล้ายด้วยฟังก์ชั่นความเหมือนของ Bhattacharyya คือ สามารถรองรับการเปรียบเทียบชุดข้อมูล 2 ชุดที่ไม่จำเป็นต้องมีขนาดหรือจำนวนข้อมูลที่เท่ากันก็ได้ ซึ่งนอกจากการเปรียบเทียบด้วยระยะ Bhattacharyya

## 2.5 ทฤษฎีเกม

ทฤษฎีเกมเป็นแนวทางที่เกี่ยวกับการตัดสินใจในสถานการณ์เดียวกันของผู้เล่นสองฝ่ายหรือมากกว่า ซึ่งการตัดสินใจของแต่ละฝ่ายนั้นจะส่งผลต่อสถานการณ์ดังกล่าวโดยตรง ดังนั้นผู้เล่นจึงจำเป็นต้องพิจารณาองค์ประกอบต่าง ๆ ในสถานการณ์นั้น รวมถึงความเป็นไปได้ในการเลือกตัดสินใจของฝ่ายอื่นที่จะมีผลกระทบต่อตนเอง และเลือกใช้กลยุทธ์ที่เหมาะสม โดยมีจุดประสงค์เพื่อให้ตนเองได้รับประโยชน์สูงสุด หรือได้รับความเสียหายน้อยที่สุด ทฤษฎีเกมจึงเป็นทฤษฎีทางคณิตศาสตร์ที่เกี่ยวข้องกับความขัดแย้งและความร่วมมือ ที่ต้องมีผู้เกี่ยวข้องอย่างน้อยสองคน เรียกว่า ผู้เล่น (Players) ซึ่งอาจเป็นมนุษย์หรือไม่ก็ได้ โดยผู้เล่นแต่ละคนจะมีแนวทางและกลยุทธ์ (Strategies) ต่าง ๆ ที่ถูกนำมาใช้เป็นสิ่งตัดสินผลลัพธ์ (Outcome) ของเกม

นอกจากนี้ทฤษฎีเกมยังเป็นการศึกษาวิธีการเล่นเกมของผู้เล่นอย่างมีเหตุผล (Rational player) กล่าวคือ ผู้เล่นแต่ละคนจะพยายามเล่นเกมให้ได้ผลลัพธ์ของตนมากที่สุดเท่าที่จะเป็นไปได้ โดยผู้เล่นสามารถควบคุมผลการเล่นได้บางส่วนจากแนวทางและกลยุทธ์ที่เลือก แต่ผลลัพธ์ของเกมไม่ได้ถูกตัดสินจากการเลือกกลยุทธ์ใด ๆ ของผู้เล่นเพียงคนเดียว แต่เป็นผลลัพธ์ที่เกิดจากการเล่นของผู้เล่นทุกคนที่อาจทำให้เกิดความขัดแย้งหรือความร่วมมือ โดยความขัดแย้งเกิดจากการที่ผู้เล่น

ได้ผลลัพธ์ที่ต่างกัน และความร่วมมือเกิดจากผู้เล่นที่ได้ประโยชน์ร่วมกัน ความพึงพอใจของผู้เล่นของผู้เล่นแต่ละคนที่มีต่อเกมจึงเป็นสิ่งสำคัญ เพื่อความสะดวกในการเปรียบเทียบปริมาณความพึงพอใจนี้ จึงมีการนิยามฟังก์ชันอรรถประโยชน์ (Utility function) ในรูปแบบทางคณิตศาสตร์เพื่อใช้แทนปริมาณความพึงพอใจของผู้เล่นแต่ละคน

### 2.5.1 ฟังก์ชันอรรถประโยชน์ (Utility function)

ปริภูมิกลยุทธ์ (Strategy space) หรือกลยุทธ์ต่าง ๆ ของผู้เล่นแต่ละคน จะถูกจัดเรียงให้อยู่ในรูปแบบทางคณิตศาสตร์ซึ่งถูกเรียกว่าฟังก์ชันอรรถประโยชน์ ค่าของฟังก์ชันอรรถประโยชน์จะเป็นปริมาณที่แสดงความพึงพอใจของผู้เล่นที่มีต่อผลลัพธ์ของเกม หากกำหนดให้ $A$ และ $B$ เป็นชุดของกลยุทธ์ที่อยู่ในปริภูมิกลยุทธ์ และให้ $u_i(.)$ เป็นฟังก์ชันอรรถประโยชน์ของผู้เล่นคนที่ $i$ โดยเครื่องหมาย $\prec$ แสดงการเปรียบเทียบค่าความพึงพอใจ หรือความสนใจของผู้เล่นที่มีต่อผลลัพธ์ที่ได้จากเกมดังสมการที่ (2.25)

$$u_i(A) < u_i(B) \Leftrightarrow A \prec B \tag{2.25}$$

พบว่าค่าอรรถประโยชน์สำหรับผลลัพธ์ที่เกิดจากชุดของกลยุทธ์ $A$ ของผู้เล่น $i$ จะมีค่าน้อยกว่าอรรถประโยชน์สำหรับผลลัพธ์ที่เกิดจากชุดของกลยุทธ์ $B$ ของผู้เล่น $i$ ก็ต่อเมื่อผู้เล่น $i$ พอใจผลลัพธ์ที่เกิดจากชุดของกลยุทธ์ $A$ น้อยกว่าผลลัพธ์ที่เกิดจากชุดของกลยุทธ์ $B$

ผู้เล่นแต่ละคนสามารถนิยามฟังก์ชันอรรถประโยชน์ เพื่อนำไปใช้กับเกมในรูปแบบต่าง ๆ ชนิดของเกมที่เป็นที่รู้จักอย่างกว้างขวาง เช่น เกมร่วมมือและเกมไม่ร่วมมือ เกมสมมาตรและเกมไม่สมมาตร เกมผลรวมศูนย์และเกมผลรวมไม่เป็นศูนย์ เกมร่วมมือและเกมไม่ร่วมมือเป็นชนิดของเกมที่ผู้เล่นแต่ละฝ่ายคำนึงถึงกลยุทธ์เพื่อให้ฝ่ายตนเองได้รับผลประโยชน์สูงสุด เกมร่วมมือเป็นเกมที่ผู้เล่นแต่ละฝ่ายสามารถตกลงกันได้เพื่อให้ได้รับผลตอบแทนรวมที่ดีที่สุด โดยจะถือว่าผู้เล่นที่ร่วมมือกันจะเป็นผู้เล่นฝ่ายเดียวกันและจะปฏิบัติตามข้อตกลงที่ได้ตกลงกันไว้ ซึ่งแตกต่างจากเกมไม่ร่วมมือที่ผู้เล่นแต่ละฝ่ายไม่สามารถตกลงผลตอบแทนกันได้ ทำให้ต้องตัดสินใจโดยใช้ผลตอบแทนของตนเป็นหลักเท่านั้น

### 2.5.2 เกมไม่ร่วมมือ (Non-cooperative game)

เกมไม่ร่วมมือเป็นชนิดของเกมที่ผู้เล่นทุกคนจะไม่สามารถสื่อสารหรือตกลงผลประโยชน์ร่วมกันได้ แต่ละผู้เล่นจึงเลือกกลยุทธ์ของตนเองอย่างอิสระเพื่อผลประโยชน์สูงสุดแก่ฝ่ายตนเองโดยไม่คำนึงถึงผลประโยชน์ของส่วนรวม รูปแบบเกมที่ใช้นิยามเกมไม่ร่วมมือนี้ คือ เกมรูปแบบครอบคลุมและรูปแบบปกติ

#### 2.5.2.1 เกมรูปแบบครอบคลุม (Extensive form game)

เป็นเกมในลักษณะที่มีลำดับในการเล่นหรือการเลือกทางเลือกของผู้เล่นแต่ละคนกำหนดไว้อย่างชัดเจน ผู้เล่นจะทราบถึงการตัดสินใจของผู้เล่นอีกฝ่ายในรอบก่อนหน้าสามารถเขียนเกมประเภทนี้ได้ในรูปแผนภาพต้นไม้ โดยตั้งต้นที่จุดเริ่มแรกและจบที่จุดสิ้นสุดของเกมซึ่งสามารถมีได้หลายจุด มีการใช้จุดยอดแทนสถานะที่มีทางเลือกในการตัดสินใจของผู้เล่น และใช้เส้นแทนทางเลือกของผู้เล่นในรอบถัดไป ตัวอย่างเกมรูปแบบครอบคลุม ได้แก่ เกมหมากรุก เกมทิก-แทก-โท

#### 2.5.2.2 เกมรูปแบบปกติ (Normal form game)

เกมรูปแบบปกติเป็นเกมที่ผู้เล่นไม่ทราบถึงการตัดสินใจของผู้เล่นคนอื่น นิยมเขียนแสดงเกมในรูปแบบตารางซึ่งมักจะใช้ในกรณีที่มีผู้เล่น 2 คน โดยผู้เล่นคนหนึ่งจะแทนการตัดสินใจด้วยแถวต่าง ๆ และผู้เล่นอีกคนหนึ่งแทนการตัดสินใจด้วยคอลัมน์ต่าง ๆ ผู้เล่นในเกมจะเลือกกลยุทธ์ของตนพร้อม ๆ กัน

### 2.5.3 เกมร่วมมือ (Cooperative game)

เกมร่วมมือเป็นเกมที่ผู้เล่นมีอิสระในการสื่อสารก่อนการเล่นเกม ผู้เล่นสามารถต่อรองและมีข้อตกลงเพื่อผลประโยชน์ร่วมกันได้ ความขัดแย้งระหว่างผู้เล่นจะมีการตัดสินอย่างยุติธรรมจากผู้ตัดสิน เกมร่วมมือส่วนใหญ่อยู่ในรูปแบบลักษณะเฉพาะ (Characteristic function form) กล่าวคือไม่มีรูปแบบเกมที่ตายตัว เกมร่วมมือจึงมีความเหมาะสมในการนำมาประยุกต์ใช้กับกระบวนการส่งต่อบุคคลระหว่างกล้องวีดิทัศน์และการเลือกกล้องที่มีมุมมองต่อบุคคลที่ดีที่สุด ด้วย

การตกลงใช้เงื่อนไขจากกฎเกณฑ์เดียวกันภายใต้การตัดสินจากศูนย์การควบคุมส่วนกลางให้มีความเท่าเทียมกันและมีความเหมาะสมในการเลือกกล้องวีดิทัศน์สำหรับการติดตามบุคคลซึ่งเป็นขั้นตอนที่สำคัญในวิทยานิพนธ์นี้ ผลเฉลยของเกมร่วมมือสามารถหาได้จากกลไกการต่อรองหรือการเปรียบเทียบค่าอรรถประโยชน์ระหว่างผู้เล่น ซึ่งขึ้นอยู่กับรูปแบบกลยุทธ์ที่ผู้เล่นนำมาใช้ด้วย

### 2.5.4 รูปแบบกลยุทธ์

ลักษณะของกลยุทธ์ที่ผู้เล่นสามารถเลือกนำมาใช้แบ่งออกเป็น 2 รูปแบบ คือ

#### 2.5.4.1 กลยุทธ์แท้ (Pure strategy)

ผู้เล่นแต่ละคนจะเลือกกลยุทธ์เด่น (Dominant strategy) ที่ตนมีมาใช้กับทุก ๆ รอบของการแข่งขัน โดยกลยุทธ์เด่น คือ กลยุทธ์ที่ให้ผลตอบแทนแก่ผู้เล่นมากกว่ากลยุทธ์อื่น ๆ ที่อยู่ในปริภูมิกลยุทธ์ของผู้เล่นนั้น กลยุทธ์เด่นสามารถแบ่งออกเป็นสองรูปแบบ คือ กลยุทธ์เด่นอย่างชัดเจน (Strictly dominant strategy) และกลยุทธ์ไม่ด้อยกว่า (Weakly dominant strategy)

#### 2.5.4.2 กลยุทธ์สุ่ม (Randomized strategy)

ผู้เล่นแต่ละคนจะใช้เครื่องมือสุ่ม เช่น การโยนเหรียญ ในการตัดสินใจเลือกกลยุทธ์ที่ตนมีในแต่ละรอบของการแข่งขัน กลยุทธ์แบบนี้จึงขึ้นอยู่กับความน่าจะเป็นของผลลัพธ์ในการสุ่มเลือก

การเลือกรูปแบบกลยุทธ์จะพิจารณาจากผลลัพธ์ของเกมที่จะเกิดขึ้นด้วยการหาดุลยภาพของเกม ดุลยภาพของเกมรูปแบบปกติที่เป็นที่นิยมและเปิดโอกาสให้วิเคราะห์เกมการแข่งขันโดยทั่วไป คือ ดุลยภาพของแนช

## 2.6 ทฤษฎีเกมกับการคัดเลือกกล้องวีดิทัศน์ที่มีมุมมองต่อบุคคลที่ดีที่สุด [20]

การติดตามบุคคลด้วยกล้องวีดิทัศน์หลายตัวมีความเป็นไปได้ที่บุคคลเดียวกันจะถูกติดตามด้วยกล้องวีดิทัศน์มากกว่าหนึ่งตัว การคัดเลือกกล้องวีดิทัศน์ที่มีมุมมองต่อบุคคลดังกล่าวที่ดีที่สุดจึงต้องอาศัยข้อมูลของแต่ละบุคคลที่ได้จากกล้องวีดิทัศน์ทุกตัว ด้วยการนำข้อมูลต่าง ๆ ที่มีส่วน

เกี่ยวข้องมาร่วมใช้ในการพิจารณา จากที่ได้กล่าวไว้ในหัวข้อที่ 2.5 ทฤษฎีเกมจึงเป็นแนวทางที่สามารถนำมาใช้สำหรับการตัดสินใจในสถานการณ์ที่ต้องการคัดเลือกกล้องวีดิทัศน์ได้ จาก [20] ได้กำหนดให้กล้องวีดิทัศน์แต่ละตัวคือผู้เล่นในเกม ซึ่งผู้เล่นที่ชนะในเกมคือกล้องวีดิทัศน์ที่มีมุมมองต่อบุคคลที่ดีที่สุด การเล่นเกมจะนำข้อมูลที่เกี่ยวข้องมาพิจารณาให้อยู่ในรูปของสมการคณิตศาสตร์ ซึ่งประกอบด้วยฟังก์ชันอรรถประโยชน์ทั้งหมด (Global utility) ที่ให้ผลลัพธ์เป็นค่าที่แสดงปริมาณความพึงพอใจทั้งหมดของประสิทธิภาพในการติดตาม ฟังก์ชันอรรถประโยชน์ของกล้องวีดิทัศน์ (Camera utility) ที่ให้ผลลัพธ์เป็นค่าที่แสดงปริมาณความเหมาะสมที่กล้องวีดิทัศน์แต่ละตัวจะทำการติดตามบุคคลที่กำหนดให้ต่อไป และฟังก์ชันอรรถประโยชน์ของกล้องวีดิทัศน์ (Camera utility) ที่ให้ค่าที่แสดงปริมาณเหมาะสมที่แต่ละบุคคลสมควรถูกติดตามโดยกล้องวีดิทัศน์บางตัวในระบบ ดังนั้นค่าอรรถประโยชน์ของบุคคลจึงเป็นข้อมูลสำคัญที่นำไปสู่การแก้ปัญหาในการเลือกกล้องวีดิทัศน์ที่มีมุมมองต่อบุคคลที่ดีที่สุด ฟังก์ชันอรรถประโยชน์ทั้งหมดถูกกำหนดไว้ล่วงหน้าและใช้ร่วมกัน ดังนั้นชนิดของเกมที่ [20] นำมาใช้นี้จึงเป็นเกมร่วมมือที่จะมีการตัดสินผู้เล่นอย่างยุติธรรม

จาก [20] พบว่าค่าอรรถประโยชน์ที่ได้จากฟังก์ชันอรรถประโยชน์ของบุคคลเป็นข้อมูลสำคัญที่นำไปสู่ผลเฉลยของเกมสำหรับการแก้ปัญหาในการเลือกกล้องวีดิทัศน์ที่มีมุมมองต่อบุคคลที่ดีที่สุด วิทยานิพนธ์นี้จึงนำเสนอเฉพาะส่วนที่นำไปสู่ผลเฉลยของเกมดังขั้นตอนต่อไปนี้

1. คำนวณหาค่าอรรถประโยชน์ของบุคคลจาก 3 กฎเกณฑ์ที่แต่ละลำดับภาพ ซึ่งเป็นกฎเกณฑ์สำหรับการชี้วัดว่ากล้องวีดิทัศน์ตัวใดในระบบเป็นกล้องที่มีมุมมองต่อบุคคลที่ทำการติดตามที่ดีที่สุด ดังต่อไปนี้

1.1 กฎเกณฑ์ที่หนึ่งขนาดของบุคคลที่ถูกติดตาม ( $Crt^1$ ) เป็นค่าที่ได้จากอัตราส่วนของจำนวนจุดภาพภายในกรอบการติดตามบุคคล ( $N_{Tracked\,blob}$ ) และจำนวนจุดภาพทั้งหมดของวีดิทัศน์ที่พิจารณา ( $N_{image}$ ) แสดงการคำนวณดังสมการที่ (2.26)

$$r = \frac{N_{Tracked\,blob}}{N_{image}} \tag{2.26}$$

เพื่อความสะดวกในการสังเกตผลลัพธ์ของการติดตามบุคคล กรอบการติดตามที่ได้ไม่ควรใหญ่หรือเล็กจนเกินไป ค่าขีดเริ่มเปลี่ยนของขนาดบุคคลที่ถูกติดตามที่ดีที่สุด ($\lambda$) ถูกกำหนดขึ้นมาให้มีขนาดของกรอบการติดตามบุคคลมีค่าเท่ากับ 1 ต่อ 15 ส่วนของขนาดภาพวีดิทัศน์ $Crt^1$ จึงคำนวณได้จากสมการที่ (2.27)

$$Crt^1 = \begin{cases} \frac{1}{\lambda} r, & r < \lambda \\ \frac{1-r}{1-\lambda}, & r \geq \lambda \end{cases} \tag{2.27}$$

1.2 กฎเกณฑ์ที่สองตำแหน่งของบุคคลที่ได้จากมุมมองการรับภาพของกล้องวีดิทัศน์ ($Crt^2$) เป็นค่าที่วัดได้จากระยะทางยุคลิด (Euclidean distance) ที่บุคคลอยู่ห่างจากจุดศูนย์กลางของภาพวีดิทัศน์ สามารถคำนวณได้จากสมการที่ (2.28)

$$Crt^2 = \frac{\sqrt{(x-x_c)^2 + (y-y_c)^2}}{\frac{1}{2}\sqrt{x_c^2 + y_c^2}} \tag{2.28}$$

เมื่อ $(x, y)$ คือ จุดศูนย์กลางของกรอบการติดตามบุคคล

$(x_c, y_c)$ คือ จุดศูนย์กลางของภาพวีดิทัศน์

1.3 กฎเกณฑ์ที่สามใบหน้าของบุคคล ($Crt^3$) เป็นค่าที่คำนวณได้จากอัตราส่วนของจำนวนจุดภาพของใบหน้าที่ตรวจจับได้ขณะที่ทำการติดตาม ($N_{face}$) กับจำนวนจุดภาพทั้งหมดของกรอบการติดตามบุคคลขณะนั้น ($N_{tracked\ blob}$) แสดงการคำนวณดังสมการที่ (2.29)

$$R = \frac{N_{face}}{N_{tracked\ blob}} \tag{2.29}$$

เช่นเดียวกับกฎเกณฑ์ที่หนึ่งขนาดของบุคคลที่ถูกติดตาม เพื่อกำหนดอัตราส่วนที่เหมาะสมของกฎเกณฑ์ที่สาม ค่าขีดเริ่มเปลี่ยนของอัตราส่วนสำหรับใบหน้าและกรอบการติดตามที่ดีที่สุด ($\xi$) จึงถูกกำหนดขึ้นมาให้มีค่าเท่ากับ 1 ต่อ 6 แสดงการคำนวณหากฎเกณฑ์ที่สามได้ดังสมการที่ (2.30)

$$Crt^3 = \begin{cases} \frac{1}{\xi}R, & R < \xi \\ \frac{1-R}{1-\xi}, & R \geq \xi \end{cases} \tag{2.30}$$

เนื่องจากใบหน้าของบุคคลเป็นสิ่งที่สามารถบ่งบอกความเป็นมนุษย์ได้ดีที่สุด [20] จึงให้นัยสำคัญของข้อมูลที่ได้ไม่เท่ากันด้วยการกำหนดค่าแบ่งน้ำหนักของกฎเกณฑ์ทั้งสาม ดังนี้

ให้ $w_1$ คือ ค่าแบ่งน้ำหนักของกฎเกณฑ์ที่หนึ่งซึ่งมีค่าเท่ากับ 0.2

$w_2$ คือ ค่าแบ่งน้ำหนักของกฎเกณฑ์ที่สองซึ่งมีค่าเท่ากับ 0.1

$w_3$ คือ ค่าแบ่งน้ำหนักของกฎเกณฑ์ที่สามซึ่งมีค่าเท่ากับ 0.7

จะเห็นได้ว่าน้ำหนักของกฎเกณฑ์ที่สองมีนับสำคัญน้อยที่สุด เนื่องจากการวัดระยะทางของบุคคลด้วยการอ้างอิงเพียงพิกัดในภาพ 2 มิติ จะทำให้เกิดความผิดพลาดจากการอ้างอิงตำแหน่งที่ไม่ถูกต้องได้ จากกฎเกณฑ์ทั้งสามและค่าแบ่งน้ำหนักที่ได้สามารถนำมาเขียนให้เป็นกฎเกณฑ์ที่สี่ ($Crt^4$) ดังแสดงในสมการที่ (2.31)

$$Crt^4 = w_1 Crt_{s1} + w_2 Crt_{s2} + w_3 Crt_{s3} \tag{2.31}$$

จากกฎเกณฑ์ที่สี่สามารถนำมานิยามฟังก์ชันอรรถประโยชน์ของบุคคลได้ดังสมการที่ (2.32)

$$U_{P_i}(a) = \sum_{i=1}^{n_p} Crt_{P_i}^4 - \sum_{\substack{i=1 \\ i \neq P_i}}^{n_p} Crt_{P_i}^4 \tag{2.32}$$

เมื่อ $U_{P_i}(a)$ คือ ค่าอรรถประโยชน์ของบุคคล $P_i$ สำหรับใช้ในการเลือกกล้องวีดิทัศน์ทุกตัวที่มีอยู่ในระบบ

$n_p$ คือ จำนวนของบุคคลที่ติดตามได้ในกล้องวีดิทัศน์ที่พิจารณา

$Crt_{P_i}^4$ คือ ค่าของกฎเกณฑ์ที่สี่สำหรับบุคคล $P_i$ ในกล้องวีดิทัศน์ที่พิจารณา

2. การหาผลเฉลยของเกม ผลเฉลยของเกมสำหรับ [20] คือ การเลือกกล้องวีดิทัศน์ที่เหมาะสมที่สุดในการติดตามบุคคล กล่าวคือ ณ เวลาใด ๆ หากบุคคลที่ต้องการติดตามปรากฏให้เห็นมากกว่าหนึ่งกล้องวีดิทัศน์ จะมีเพียงกล้องวีดิทัศน์เพียงตัวเดียวที่เหมาะสมสำหรับติดตามบุคคลดังกล่าว การเลือกกล้องวีดิทัศน์ที่เหมาะสมที่สุดนี้เกิดจากวิธีการต่อรองค่าอรรถประโยชน์ของแต่ละบุคคลที่ได้จากกล้องวีดิทัศน์ทุกตัวในระบบ ผลเฉลยของเกมจะอยู่ในรูปของค่าความน่าจะเป็นที่สูงที่สุดซึ่งกล้องวีดิทัศน์แต่ละตัวจะได้รับจากบุคคลที่ได้ทำการติดตาม สามารถทำการคำนวณได้ดังสมการที่ (2.33)

$$p_i^l(k) = \frac{e^{\frac{1}{\tau}\overline{U}_{P_i}^l(k)}}{e^{\frac{1}{\tau}\overline{U}_{P_i}^1(k)} + \ldots + e^{\frac{1}{\tau}\overline{U}_{P_i}^{n_C}(k)}} \tag{2.33}$$

เมื่อ $p_i^l(k)$ คือ ค่าความน่าจะเป็นของบุคคลที่ $i$ ของกล้องวีดิทัศน์ตัวที่ $l$ ณ ลำดับการคำนวณที่ k

$l$ คือ ลำดับกล้องวีดิทัศน์ที่สามารถทำการติดตามบุคคลที่ $i$ ได้ ดังนั้นเซตของกล้องวีดิทัศน์ในระบบจึงมีค่าเท่ากับ $l = \{1,2,\ldots,n_C\}$

$n_C$ คือ จำนวนกล้องวีดิทัศน์ทั้งหมดที่สามารถทำการติดตามบุคคลที่ $i$ ได้

$\overline{U}_{P_i}^l(k)$ คือ ค่าอรรถประโยชน์ของการทำนายบุคคล $i$ เพื่อใช้สำหรับการคำนวณในลำดับที่ $k$

ค่าอรรถประโยชน์ของการทำนายบุคคล $i$ จะถูกใช้เป็นข้อมูลในการทำนายค่าความน่าจะเป็น $p_i^l(k)$ ในลำดับการคำนวณรอบถัดไป โดยการคำนวณจะสิ้นสุดลงและให้ผลเฉลยของเกมเมื่อ $p_i^l(k)$ ให้ผลเฉลยที่ลู่เข้าสู่ผลลัพธ์