

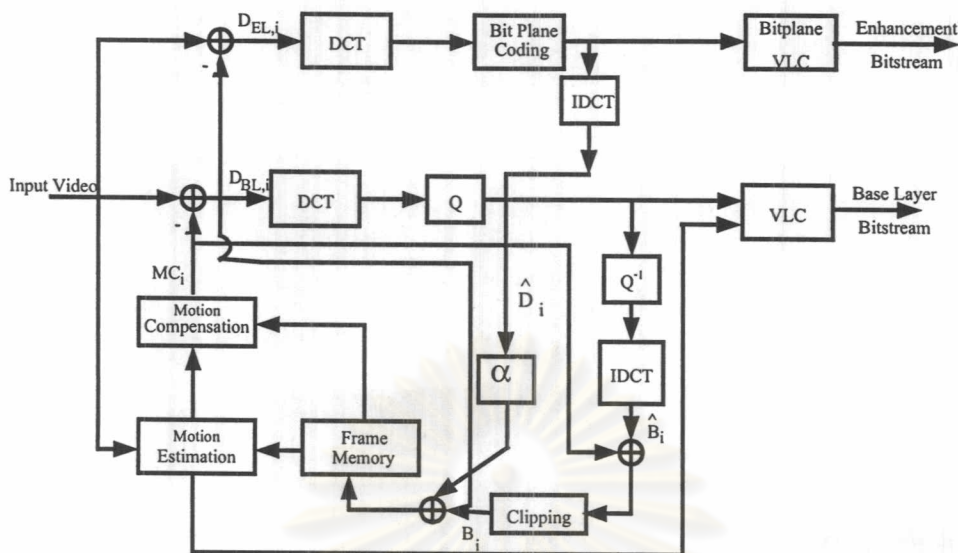
### บทที่ 3 กรรมวิธีที่นำเสนอ

วิทยานิพนธ์นี้นำเสนอการปรับปรุงวิธีการเข้ารหัสสัญญาณวิดีโอตามมาตรฐาน MPEG-4 เพื่อให้ได้คุณภาพของสัญญาณวิดีโออัตราบิตต่ำที่ดีขึ้น โดยมีรายละเอียดประกอบไปด้วยส่วนต่างๆ ดังนี้

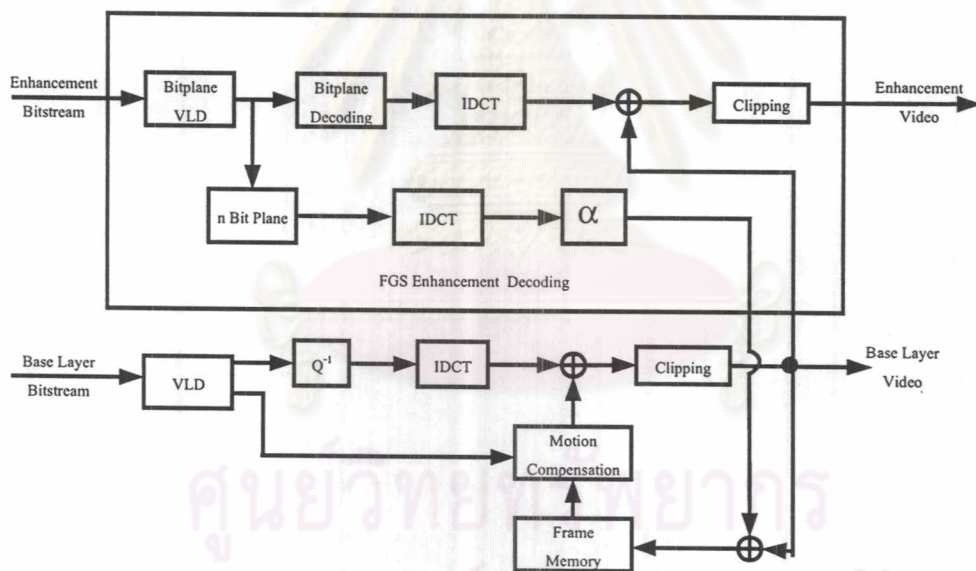
#### 3.1 การเข้ารหัสแบบไพน์แกรนูลาริตีสเกลลาบิลิตีโดยใช้การทำนายรั่ว (Fine granularity scalability using leaky prediction: FGS-LP)

จุดเด่นของโครงสร้างการเข้ารหัสแบบ FGS คือ ทั้งชั้นฐาน และชั้นเอนแฮนส์เมนต์ในเฟรมปัจจุบัน ถูกทำนายการเคลื่อนที่จากชั้นฐานของเฟรมก่อนหน้าเท่านั้น ถึงแม้เกิดความผิดพลาดหรือเกิดการสูญหายของชั้นเอนแฮนส์เมนต์บางระนาบบิต ก็ไม่มีผลกับเฟรมที่ตามมา แต่อย่างไรก็ตาม การทำนายถูกทำนายจากชั้นฐานซึ่งมีคุณภาพต่ำที่สุด ประสิทธิภาพการเข้ารหัสที่ได้รับของโครงสร้างแบบ FGS จึงไม่ดีเท่าการเข้ารหัสแบบ SNR scalability [2] และจากวิธีการเข้ารหัสแบบ single-loop Motion-Compensation FGS (1-loop MC-FGS) [6] ซึ่งนำแนวคิดของการเข้ารหัสแบบ SNR scalability ที่ตัวเข้ารหัสและตัวถอดรหัสใช้ข้อมูลทั้งชั้นฐานและชั้นเอนแฮนส์เมนต์ในการทำนายการเคลื่อนที่เพื่อเพิ่มคุณภาพเฟรมอ้างอิงมาประยุกต์ใช้กับการเข้ารหัสแบบ FGS เพื่อปรับปรุงประสิทธิภาพการเข้ารหัส และเนื่องจากวิธีนี้มีการชดเชยการเคลื่อนที่ที่ชั้นฐานเท่านั้นจึงมีความซับซ้อนต่ำ แต่ดังที่กล่าวมาก่อนหน้า การเข้ารหัสที่ใช้ข้อมูลของชั้นเอนแฮนส์เมนต์ในการทำนายการเคลื่อนที่มีข้อเสีย คือ ถ้าข้อมูลที่ใช้ในการทำนายเฟรมถัดไปไม่สามารถส่งไปได้ครบ จะทำให้เกิดความผิดพลาดในการทำนาย (prediction drift หรือ drift error และ error propagation) ขึ้น

เพื่อป้องกันและลดทอนความผิดพลาด เนื่องจากการสูญหายของข้อมูลในช่องสัญญาณที่อัตราบิตไม่คงที่ จึงเสนอเทคนิคการทำนายรั่ว (leak prediction) [16] สำหรับวงรอบการเข้ารหัสระหว่างเฟรม (interframe loop) โดยค่าตัวประกอบการทำนายรั่ว (leakage factor :  $\alpha$ ) มีค่าระหว่าง 0 – 1 เทคนิคการทำนายรั่วช่วยเสริมความทนทานต่อความผิดพลาด (error resilience) แต่ประสิทธิภาพการเข้ารหัสลดลง เนื่องจากข้อมูลเอนแฮนส์เมนต์บางส่วนเท่านั้นที่ถูกนำมารวมกับข้อมูลที่ชั้นฐานเพื่อลดความซ้ำซ้อนทางเวลา ดังนั้นในวิทยานิพนธ์นี้จะนำเสนอโครงสร้างการเข้ารหัสดังแสดงในรูปที่ 3.1 ที่ยอมให้ตัวเข้ารหัสเลือกระหว่างปรับปรุงประสิทธิภาพการเข้ารหัสหรือรักษาคุณภาพของวิดีโอสำหรับช่องสัญญาณที่มีแบนด์วิดท์ไม่คงที่ และมีโครงสร้างการถอดรหัสดังแสดงในรูปที่



รูปที่ 3.1 โครงสร้างการเข้ารหัส FGS-LP



รูปที่ 3.2 โครงสร้างการถอดรหัส FGS-LP

จากรูปที่ 3.1 อธิบายการเข้ารหัสได้ดังนี้ คือ เฟรมปัจจุบันเฟรมที่  $i$  มีภาพต้นฉบับเป็น  $F_i$  เฟรมก่อนหน้าที่สร้างกลับคืนมา คือ  $B_{i-1}$  ส่วนเหลือจากการชดเชยการเคลื่อนที่ที่พื้นฐานเฟรมที่  $i$  คือ  $D_{BL,i}$  โดยค่าความแตกต่างที่ขึ้นแอนเซนส์เมนต์หลังจากถูกเข้ารหัสของเฟรมของก่อนหน้าแทนด้วย  $\hat{D}_{i-1}$  และมีค่าตัวประกอบการรัว คือ  $\alpha$  ภาพเฟรมปัจจุบันได้จากสมการ (3.1)

$$F_i = (B_{i-1} + \alpha \hat{D}_{i-1})_{mc} + D_{BL,i} \quad (3.1)$$

โดย  $(B_{i-1} + \alpha \hat{D}_{i-1})_{mc}$  หมายถึงการชดเชยการเคลื่อนที่ของ  $(B_{i-1} + \alpha \hat{D}_{i-1})$  และส่วนเหลือจากการชดเชยการเคลื่อนที่  $D_{BL,i}$  เมื่อถูกเข้ารหัสแล้วถูกแทนด้วย  $\hat{B}_i$  มีความผิดพลาดจากการควอนไทซ์เป็น  $D_{EL,i}$  เขียนแสดงความสัมพันธ์ได้ดังสมการ (3.2)

$$D_{BL,i} = \hat{B}_i + D_{EL,i} \quad (3.2)$$

แทนสมการ (3.2) ในสมการ (3.1) จะได้

$$\begin{aligned} F_i &= (B_{i-1} + \alpha \hat{D}_{i-1})_{mc} + \hat{B}_i + D_{EL,i} \\ &= (B_{i-1})_{mc} + \hat{B}_i + (\alpha \hat{D}_{i-1})_{mc} + D_{EL,i} \\ &= B_i + \hat{D}_i \end{aligned} \quad (3.3)$$

โดยที่

$$B_i = (B_{i-1})_{mc} + \hat{B}_i \quad (3.4)$$

$$\hat{D}_i = (\alpha \hat{D}_{i-1})_{mc} + D_{EL,i} \quad (3.5)$$

สัญญาณ  $B_i$  และ  $\hat{D}_i$  จะถูกใช้สำหรับการทำนายเฟรมต่อไป เมื่อข้อมูล  $D_{EL,i}$  ทั้งหมดถูกนำมาใช้ในวงรอบการทำนาย เราสามารถเขียน  $\hat{D}_i$  ในสมการที่ (3.5) ได้ว่า

$$\begin{aligned} \hat{D}_i &= (\alpha((\alpha \hat{D}_{i-2})_{mc} + D_{EL,i-1}))_{mc} + D_{EL,i} \\ &= (\alpha((\alpha((\alpha \hat{D}_{i-3})_{mc} + D_{EL,i-2}))_{mc} + D_{EL,i-1}))_{mc} + D_{EL,i} \end{aligned} \quad (3.6)$$

จากสมการที่ (3.6) จะสังเกตได้ว่าถ้าเกิดความผิดพลาดขึ้นใน  $\hat{D}_i$  ก็จะถูกลดทอนในโครงสร้างนี้ สมมติว่าเกิดความผิดพลาดที่ชั้นเอนแฮนส์เมนต์ในเฟรมที่  $i-2$  บิตสตรีมของชั้นเอนแฮนส์เมนต์ที่ได้รับ คือ  $\bar{D}_{i-2}$  และความผิดพลาดจากการส่ง คือ  $\Delta \hat{D}_{i-2}$  ดังนั้นจะได้ว่า

$$D_{EL,i-2} = \bar{D}_{i-2} + \Delta \hat{D}_{i-2} \quad (3.7)$$

และ  $\hat{D}_{i-2}$  ที่สร้างกลับคืนมาที่ตัวถอดรหัส คือ  $\tilde{D}_{i-2}$  ดังนั้น

$$\begin{aligned}\tilde{D}_{i-2} &= (\alpha \hat{D}_{i-3})_{mc} + \bar{D}_{i-2} \\ &= (\alpha \hat{D}_{i-3})_{mc} + D_{EL,i-2} - \Delta \hat{D}_{i-2}\end{aligned}\quad (3.8)$$

เปรียบเทียบสมการ (3.6) กับ (3.8) จะได้ค่าความแตกต่าง คือ  $\Delta \hat{D}_{i-2}$

ถ้าเรากลับไปดูที่เฟรมที่  $i-1$  และสมมติว่าไม่เกิดความผิดพลาดที่ชั้นเอนแฮนส์เมนต์ในเฟรมนี้ ซึ่งบิตสตรีมของชั้นเอนแฮนส์เมนต์ที่ได้รับของเฟรม  $i-1$  คือ  $D_{EL,i-1}$  ดังนั้นบิตสตรีมของชั้นเอนแฮนส์เมนต์ที่ตัวถอดรหัสของเฟรมที่  $i-1$  คือ

$$\begin{aligned}\tilde{D}_{i-1} &= (\alpha \tilde{D}_{i-2})_{mc} + D_{EL,i-1} \\ &= (\alpha(\alpha \hat{D}_{i-3})_{mc} + D_{EL,i-2} - \Delta \hat{D}_{i-2})_{mc} + D_{EL,i-1}\end{aligned}\quad (3.9)$$

ความแตกต่างระหว่าง  $\hat{D}_{i-1}$  กับ  $\tilde{D}_{i-1}$  คือ  $\alpha(\Delta \hat{D}_{i-2})$

ถ้าเรากลับไปดูที่เฟรมที่  $i$  และสมมติว่าไม่เกิดความผิดพลาดที่ชั้นเอนแฮนส์เมนต์ในเฟรมนี้ ซึ่งบิตสตรีมของชั้นเอนแฮนส์เมนต์ที่ได้รับของเฟรม  $i$  คือ  $D_{EL,i}$  ดังนั้นบิตสตรีมของชั้นเอนแฮนส์เมนต์ที่ตัวถอดรหัสของเฟรมที่  $i$  คือ

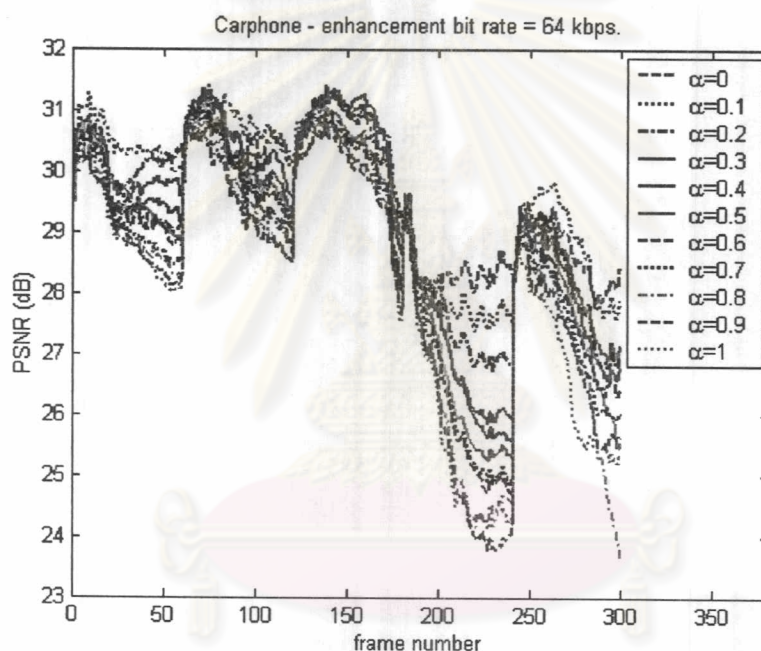
$$\begin{aligned}\tilde{D}_i &= (\alpha \tilde{D}_{i-1})_{mc} + D_{EL,i} \\ &= (\alpha((\alpha(\alpha \hat{D}_{i-3})_{mc} + D_{EL,i-2} - \Delta \hat{D}_{i-2})_{mc} + D_{EL,i-1}))_{mc} + D_{EL,i}\end{aligned}\quad (3.10)$$

ความแตกต่างระหว่าง  $\hat{D}_i$  กับ  $\tilde{D}_i$  คือ  $\alpha^2(\Delta \hat{D}_{i-2})$

จากที่กล่าวมาข้างต้นจะเห็นว่า ความผิดพลาดที่เกิดในบิตสตรีมที่ถูกถอดรหัสที่ชั้นเอนแฮนส์เมนต์นั้น จะถูกลดทอนโดยค่าตัวประกอบการรั้วที่เฟรมต่อไป สำหรับค่าตัวประกอบการรั้วที่น้อยกว่าหนึ่ง

### 3.2 กรรมวิธีปรับค่าตัวประกอบการรั้ว ( $\alpha$ ) ของการทำนายรั้ว (FGS-ALP)

การทำนายรั้วเป็นเทคนิคที่รู้จักกันดีว่าเพิ่มความทนทานต่อความผิดพลาด แต่ต้องแลกกับประสิทธิภาพการเข้ารหัส [7] ในโครงสร้างที่นำเสนอ สามารถบดบังแรกของข้อมูลชั้นเอนแฮนซ์เมนต์ที่ถูกแปลงกลับไปโดเมนปริภูมิ (the spatial domain) โดย IDCT และถูกถอดทอนด้วยตัวประกอบการรั้ว  $\alpha$  ก่อนนำไปรวมในหน่วยความจำภาพ (frame memory) ดังนั้น ค่าความแตกต่างระหว่างตัวเข้ารหัสและตัวถอดรหัส (drift error) จะถูกถอดทอน และถ้าค่าตัวประกอบการรั้วถูกกำหนดเป็นศูนย์หมายความว่าไม่ใช่ข้อมูลชั้นเอนแฮนซ์เมนต์ในการประมาณการเคลื่อนที่ ซึ่งจะทำให้ drift ถูกกำจัดอย่างสมบูรณ์



รูปที่ 3.3 ค่า PSNR ของลำดับภาพ Carphone ที่แบนด์วิธเอนแฮนซ์เมนต์เท่ากับ 64 kbps.

การกำหนดค่าตัวประกอบการรั้วเท่ากับศูนย์ เปรียบได้กับการเข้ารหัสแบบ FGS เดิม ซึ่งจะทนทานต่อความผิดพลาดมากที่สุด ส่วนการกำหนดค่าตัวประกอบการรั้วเท่ากับหนึ่ง ก็เปรียบได้กับการเข้ารหัสแบบ 1-loop MC-FGS ซึ่งมีความทนทานต่อความผิดพลาดต่ำ อย่างไรก็ตาม การเลือกค่าตัวประกอบการรั้วที่เหมาะสมเป็นสิ่งสำคัญ ซึ่งค่าตัวประกอบการรั้วถูกแสดงให้เห็นว่าปรับเปลี่ยนตามภาพวิดีโอและแบนด์วิธของช่องสัญญาณที่ต่างกัน เมื่อเราปรับค่าตัวประกอบการรั้วในช่วง 0-1 (0, 0.1, 0.2, ..., 1) ในลำดับภาพ Carphone ผลที่ได้แสดงดังรูปที่ 3.3 โดยสังเกตได้ว่าการกำหนดค่าตัวประกอบการรั้วของแต่ละเฟรม ควรปรับตามค่าการเปลี่ยนแปลงทางพื้นที่ของเฟรมนั้น นอกจากนี้ การปรับค่าตัวประกอบการรั้วจาก 0-1 ของหลายๆลำดับภาพที่อัตราบิต

กัน จะได้ความสัมพันธ์ระหว่างแบนด์วิดท์เอนแฮนส์เม้นท์กับค่าตัวประกอบการรบกวนที่แสดงในสมการที่ (3.11)

$$\alpha_{initial} = 0.1 \times \log_2 \left( \frac{BW_{enh}}{64} \right) : 0 < BW_{enh} \leq 256 \quad (3.11)$$

กรรมวิธีปรับค่าตัวประกอบการรบกวนนี้ ถูกเสนอเพื่อกำหนดค่าตัวประกอบการรบกวนที่เหมาะสมสำหรับแบนด์วิดท์ของช่องสัญญาณ และลำดับภาพต่างๆ โดยกรรมวิธีที่นำเสนอจะปรับค่าตัวประกอบการรบกวนทุกเฟรมที่เข้ารหัส ขั้นตอนการปรับค่ามีดังนี้

**ขั้นตอนที่ 1 :** หาค่าการเปลี่ยนแปลงทางพื้นที่โดยเฉลี่ย (Average activity measure : avg\_act) ซึ่งเป็นค่าเฉลี่ยของค่าการเปลี่ยนแปลงทางพื้นที่ (Spatial Activity Measure : act) ในแต่ละเฟรม ซึ่งค่าการเปลี่ยนแปลงทางพื้นที่คำนวณที่ระดับมาโครบล็อก โดยค่าการเปลี่ยนแปลงทางพื้นที่ของมาโครบล็อกที่  $j$  หากจากบล็อกย่อยของความเข้มแสงที่จัดเรียงข้อมูลแบบเฟรมทั้งสี่บล็อก (four luminance frame-organized sub-blocks)  $n=1, \dots, 4$  และบล็อกย่อยของความเข้มแสงที่จัดเรียงข้อมูลแบบฟิลด์ทั้งสี่บล็อก (four luminance field-organized sub-blocks)  $n=5, \dots, 8$  โดยใช้ค่าพิกเซลต้นฉบับ ดังนี้

$$act_j = 1 + \min(vblk_1, vblk_2, \dots, vblk_8) \quad (3.12)$$

โดยค่าความแปรปรวนของบล็อกขนาด  $8 \times 8$  พิกเซล (the variance of  $8 \times 8$  block) หากจาก

$$vblk_n = \frac{1}{64} \times \sum_{k=1}^{64} (P_k^n - P_{-mean_n})^2 \quad (3.13)$$

และ

$$P_{-mean_n} = \frac{1}{64} \times \sum_{k=1}^{64} P_k^n \quad (3.14)$$

และ  $P_k^n$  คือ ค่าตัวอย่างในบล็อกต้นฉบับขนาด  $8 \times 8$  พิกเซล

**ขั้นตอนที่ 2 :** หาค่าเฉลี่ยของ avg\_act ตั้งแต่เฟรมแรกถึงเฟรมปัจจุบันที่ถูกเข้ารหัส แทนค่าดังกล่าวด้วยตัวแปร Tot\_avg\_act

ขั้นตอนที่ 3 : หาค่าตัวประกอบการรั่วที่เหมาะสม โดยมีแนวคิดในการตัดสินใจ คือ ถ้าค่าการเปลี่ยนแปลงทางพื้นที่เฉลี่ยของเฟรมที่เข้ารหัสมีค่าสูงกว่า Tot\_avg\_act มากๆ จะกำหนดค่าตัวประกอบการรั่วต่ำกว่า  $\alpha_{initial}$  และถ้าค่าการเปลี่ยนแปลงทางพื้นที่เฉลี่ยของเฟรมที่เข้ารหัสมีค่าต่ำกว่า Tot\_avg\_act มากๆ จะกำหนดค่าตัวประกอบการรั่วสูงกว่า  $\alpha_{initial}$  แต่ด้านนอกเหนือจากสองกรณีดังกล่าว จะกำหนดค่าตัวประกอบการรั่วเท่ากับ  $\alpha_{initial}$  ซึ่งมีกระบวนการเป็นดังนี้

```

iff((avg_act - Tot_avg_act) > Th)
     $\alpha = \alpha_{initial} - 0.1;$ 
else iff((avg_act - Tot_avg_act) < -Th)
     $\alpha = \alpha_{initial} + 0.1;$ 
else
     $\alpha = \alpha_{initial} ;$ 

```

### 3.3 วิธีการทดสอบ

วิธีการทดสอบประสิทธิภาพของการเข้ารหัส มีขั้นตอนดังนี้

1. ทดสอบประสิทธิภาพของโครงสร้างที่นำเสนอ (FGS-LP) โดยเปรียบเทียบผลการทดสอบกับโครงสร้างอื่นๆ
2. ทดสอบการลดทอนความผิดพลาดของโครงสร้างที่นำเสนอ โดยเปรียบเทียบผลการทดสอบกับโครงสร้าง 1-loop MC-FGS
3. เปรียบเทียบผลการทดสอบเมื่อปรับค่าตัวประกอบการรั่วในช่วง 0-1 (0, 0.1, 0.2, ..., 1) เพื่อศึกษาวิธีการกำหนดค่าตัวประกอบการรั่วว่าควรกำหนดค่ามากหรือน้อยเมื่อไร
4. ทดสอบประสิทธิภาพของกรณวิธีปรับค่าตัวประกอบการรั่วของการทำนายรั่ว (FGS-ALP) โดยเปรียบเทียบผลการทดสอบกับโครงสร้างอื่นๆ
5. จำลองช่องสัญญาณที่มีแบนด์วิดท์ไม่คงที่ และทดสอบกับวิธีการเข้ารหัสต่างๆ