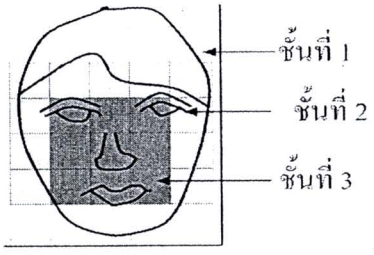
บทที่ 2
ปริทัศน์วรรณกรรมและงานวิจัยที่เกี่ยวข้อง

2.1 การทบทวนวรรณกรรมที่เกี่ยวข้อง

การค้นหาหน้าคนจากภาพนั้น เป็นงานวิจัยที่มีได้รับความสนใจศึกษากันอย่างกว้างขวาง เหตุผลที่ทำให้งานนี้ได้รับความสนใจเนื่องมาจากความท้าทายในการค้นหาใบหน้าบุคคล ที่มีความหลากหลายของใบหน้าที่ไม่ซ้ำกันในแต่ละบุคคล รวมถึงความเปลี่ยนแปลงไม่แน่นอนในการปรากฏของใบหน้า เช่น โครงสร้างทางกายภาพของใบหน้าที่แตกต่างกันไปตามเชื้อชาติ ตำแหน่งที่ตั้ง การวางท่า หรือการแสดงออกของสีหน้า เป็นต้น โดยในการวิจัยช่วงแรกจะพิจารณาเป็นภาพเดี่ยว สามารถแบ่งประเภทการตรวจจับออกได้เป็น 4 ประเภทหลัก ๆ คือ

2.1.1 วิธีเชิงความรู้ (Knowledge-based methods) เป็นวิธีการพิจารณาความสัมพันธ์ระหว่างลักษณะเด่นเฉพาะต่าง ๆ บนใบหน้า ที่มีตำแหน่งและองค์ประกอบหลักพื้นฐานตายตัวบนใบหน้า ตัวอย่างเช่น Yang and Huang (1994) ศึกษาวิธีเชิงความรู้แบบลำดับชั้น (hierarchical) ซึ่งพิจารณาแบ่งเป็น3 ลำดับชั้น ดังแสดงในรูปที่ 2.1



รูปที่ 2.1 ขอบเขตการพิจารณาของแต่ละลำดับชั้น

ชั้นแรก พิจารณาหาตำแหน่งความน่าจะเป็นของใบหน้าโดยรวม ชั้นที่ 2 พิจารณาโดยค่า ฮิสโทแกรม (histogram) ร่วมกับการหาเส้นขอบ ชั้นที่ 3 พิจารณาลักษณะเด่นภายในของใบหน้า เช่น ความสมมาตรกันของตาทั้ง 2 ข้าง ตำแหน่งของจมูกและปากในแนวตั้งและแนวนอน

ต่อมา Kotropoulos and Pitas (1997) ได้ศึกษาวิธีฐานกฎแบบกำหนดเขต (rule-based localization method) ซึ่งเป็นวิธีการที่คล้ายคลึงกับวิธีของ Yang and Huang แตกต่างกันที่พิจารณาค่าฮิสโทแกรมทั้งในแนวตั้งและแนวนอน แต่ยังไม่สามารถขจัดปัญหาของภาพที่มีพื้นหลังซับซ้อนและปัญหาความหลากหลายของใบหน้าบุคคล รวมทั้งการวางท่าที่แตกต่างกันได้

2.1.2 วิธีเชิงลักษณะ (Feature-based methods) เป็นการใช้อัลกอลิทึมพิจารณาลักษณะเด่นและโครงสร้างของใบหน้า รวมทั้งความเปลี่ยนแปลงขององค์ประกอบต่าง ๆ ของภาพ เช่น แสงเงาตัวอย่างเช่น Chetverikov and Lerch (1993) ใช้ความแตกต่างของแสงเงาและเส้นแนวเป็นแบบจำลองในการค้นหา

ใบหน้า โดยแบบจำลองนั้นประกอบด้วยจุดความสว่างน้อย (จุดมืด) 2 จุด เพื่อแสดงดวงตา และจุดความสว่างมาก (จุดอ่อน) 3 จุด เพื่อแสดงโหนกแก้มและจมูก แล้วหาความสัมพันธ์ของระยะห่างและ ตำแหน่งของจุดต่าง ๆ เพื่อคัดเลือกแบบหน้าที่เหมาะสม ระบบนี้มีข้อจำกัดอยู่ที่เมื่อแสงเงาของสภาวะแวดล้อมเปลี่ยนแปลงจะทำให้ประสิทธิภาพในการค้นหาเปลี่ยนไป ต่อมา Graf, Chen, Petajan, and Cosatto (1995) พัฒนาทฤษฎีนี้โดยใช้งานในระบบภาพระดับเทา และนำตัวกรองช่วงผ่าน (band pass filter) ตัวกรองเกาส์เซียน (Gaussian filter) และค่าฮิสโทแกรม มาใช้หาค่าจุดสูงสุดและต่ำสุดเพื่อกำหนดขอบเขตของใบหน้า

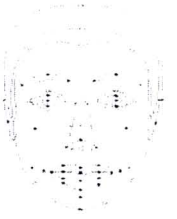
ลักษณะเฉพาะอีกประการหนึ่งที่นิยมนำมาพิจารณาหาใบหน้าบุคคล คือค่าสีผิวมนุษย์โดยมีการค้นหาในหลาย ๆ ปริภูมิสี โดย Crowley and Berard (1997) ทำการเก็บค่าเวกเตอร์ของสีผิวมนุษย์โดยแยกค่าสีออกเป็น R, G และ B เพื่อพิจารณาค่าฮิสโทแกรมในรูปของสมการ $h(r,g)$ □□ โดยค่า คือค่าเริ่มเปลี่ยน (threshold) ของค่าสีผิวมนุษย์ที่ได้จากการสำรวจข้อมูลกลุ่มตัวอย่าง ต่อมา Saxe and Fould (1996) ประยุกต์ใช้วิธีการดังกล่าวกับปริภูมิสี HSV และใช้แนวคิดค่าฮิสโทแกรมตัดผ่าน (histogram intersection) เปรียบเทียบค่าฮิสโทแกรมควบคุมกับค่าฮิสโทแกรมปัจจุบัน ซึ่งทำให้ประสิทธิภาพดีกว่าการใช้ค่าเริ่มเปลี่ยน

Chai and Ngan (1998) พิจารณาค่าสีผิวในปริภูมิสี YCrCb โดยพิจารณาในรูปของสมการ $Cr1$□Cr□$Cr2$ และ $Cb1$□Cb□$Cb2$ ซึ่งค่า [$Cr1$, $Cr2$] และ [$Cb1$, $Cb2$] เป็นค่าเริ่มเปลี่ยนของค่าสีผิวมนุษย์ที่ได้จากการสำรวจข้อมูลกลุ่มตัวอย่าง ซึ่งต้องระมัดระวังในการเลือกใช้ให้เหมาะสมกับภาพที่ทำการทดลอง และเป็นข้อเสียของการคัดแยกสีผิวมนุษย์ด้วยค่าเริ่มเปลี่ยน ต่อมามีการนำการคัดแยกสีผิวมาใช้ร่วมกับวิธีการอื่น ๆ ในการค้นหาใบหน้า ตัวอย่างเช่น Wu, Yokoyama, Pramadihanto, and Yachida, (1996) ใช้ทฤษฎีคลุมเครือ (fuzzy theory) ในการคัดแยกสีผิวมนุษย์และสีผมในปริภูมิสี CIE XYZ โดยทำการแปลงจากปริภูมิสี RGB นอกจากนี้ยังมีการนำเอาลักษณะเด่นของใบหน้าอื่น ๆ มาใช้ในการค้นหาตัวอย่างเช่น Yokoo and Hagiwara (1996) ใช้คุณสมบัติของโครงหน้าทั่วไปของมนุษย์ที่เป็นรูปไข่ สร้างแบบจำลองรูปวงรีเปรียบเทียบพารามิเตอร์ร่วมกับการค้นหาแบบจีนเนติกอัลกอลิทึม (genetic algorithm) จากการศึกษาพบว่าวิธีการนี้มีข้อจำกัดอยู่ที่รายละเอียดของใบหน้าที่ทำการค้นหาต้องมีความชัดเจนและมีขนาดใหญ่เพียงพอต่อการค้นหา

2.1.3 วิธีเทียบเคียงแผ่นแบบ (Template matching methods) เป็นการเปรียบเทียบภาพที่ต้องการค้นหากับโครงสร้างแบบจำลองของใบหน้ามาตรฐาน โดยเก็บข้อมูลความสัมพันธ์อย่างอิสระของส่วนต่าง ๆ บนใบหน้า ได้แก่ โครงสร้างใบหน้า ตา จมูก และปาก ในท่าหน้าตรง ตัวอย่างเช่น Sakai, Nagao and Fujibayashi, (1969) เสนอโครงสร้างแบบจำลองแผ่นแบบย่อย (subtemplates model) โดยใช้ตัวกรองโซเบล (sobel filter) หาเส้นขอบ เพื่อหาตำแหน่งความน่าจะเป็นของส่วนย่อยต่าง ๆ บนใบหน้า ที่สามารถเข้ากันได้ดีที่สุดกับแบบจำลองแผ่นแบบย่อยให้เป็นตำแหน่งของใบหน้าที่ต้องการต่อมา Tsukamoto, Lee, and Tsuji (1994) เสนอแบบจำลองคุณภาพสำหรับรูปแบบหน้า (qualitative model for face pattern : QMF) โดยใช้

พารามิเตอร์ของแสงสว่างและเส้นขอบใบหน้าเป็นแบบจำลองของใบหน้า และเทคนิคที่นิยมใช้อีกอันหนึ่งคือการวิเคราะห์องค์ประกอบหลัก (principal component analysis: PCA) โดย Jung, Lee C.W., and Lee Y.C (2002) ศึกษาการค้นหาตำแหน่งที่แน่นอนของใบหน้าโดยใช้หลักการพื้นฐานทางด้านเรขาคณิต วิเคราะห์องค์ประกอบของใบหน้าให้อยู่ในรูปของสมการทางคณิตศาสตร์

ต่อมามีการพัฒนาแบบจำลองให้มีความสามารถปรับตัวได้เรียกว่า แผ่นแบบเปลี่ยนรูปได้ (deformable templates) โดยแผ่นแบบสามารถปรับตัวให้รองรับความยืดหยุ่นของโครงสร้างใบหน้าที่เปลี่ยนแปลงไปและลดปัญหาที่มีในแผ่นแบบคงตัว ตัวอย่างเช่น Malciu and Preteux (2002) นำแบบจำลองหน้าจากฐานข้อมูล MPEG-4 มาสร้างเป็นแบบจำลองแผ่นแบบพารามิเตอร์ (template parameterization) และพิจารณาความยืดหยุ่นของโครงสร้างใบหน้าเปรียบเทียบเป็นความยืดหยุ่นของสปริง เพื่อหาค่าพลังงานที่สปริงใช้น้อยที่สุดในการจะปรับตัวเพื่อให้ภาพที่หาเส้นขอบนั้นเข้ากับแบบจำลองมากที่สุด โดยทำการพิจารณาในส่วนของตาและปาก แสดงในรูปที่ 2.2

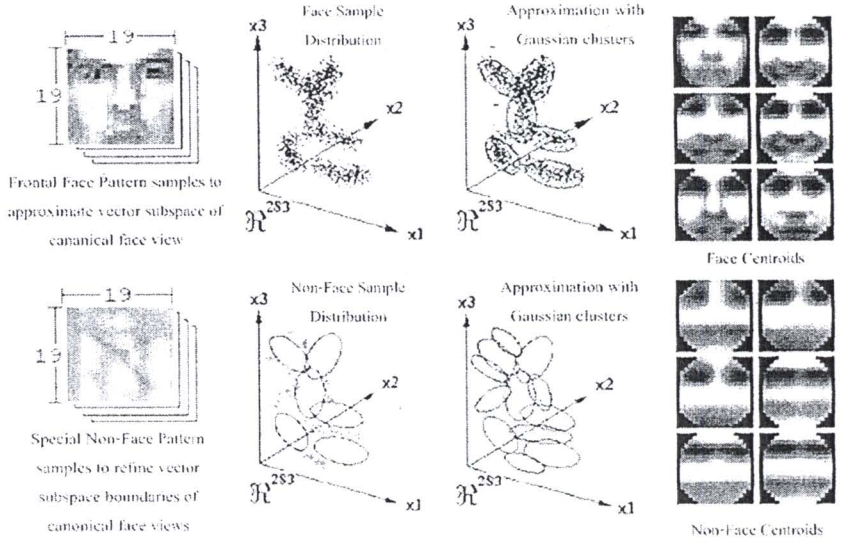


รูปที่ 2.2 แบบจำลองใบหน้ามนุษย์จากฐานข้อมูล MPEG-4

2.1.4 วิธีเชิงลักษณะปรากฏ (Appearance-based methods) เป็นการเปรียบเทียบภาพที่ต้องการค้นหากับโครงสร้างแบบจำลองของใบหน้าที่ทำการเรียนรู้และฝึกสอนให้ระบบจดจำ และนำความรู้ในฐานข้อมูลมาใช้ในการพิจารณา โดยสามารถแบ่งย่อยได้หลายวิธีด้วยกัน ในที่นี้จะทำการเสนอเพียงวิธีการที่ได้รับความนิยมใช้งานกันอย่างกว้างขวาง 4 วิธีดังนี้

2.1.4.1 วิธีหน้าลักษณะเฉพาะ (Eigenface Methods) เป็นการวิเคราะห์องค์ประกอบพื้นฐานของใบหน้าด้วยกระบวนการทางสถิติของใบหน้าที่หลากหลาย โดยหน้าลักษณะเฉพาะคือเซตของเวกเตอร์ลักษณะเฉพาะ (eigen vector) ที่สามารถหาได้จากเมทริกซ์ความแปรปรวนร่วมเกี่ยว (covariance matrix) สร้างเป็นแบบจำลองของใบหน้าที่รวมเอาลักษณะเด่นต่าง ๆ ของภาพใบหน้าตัวอย่างมาร่วมกันเพื่อหาค่าเฉพาะเจาะจงขององค์ประกอบบนใบหน้า ซึ่งในแต่ละบุคคลจะเป็นค่าเฉพาะของบุคคลนั้น ๆ ตัวอย่างเช่น Turk and Pentland (1991) นำภาพระดับเทามาแปลงเป็นเวกเตอร์เพื่อหาค่าลักษณะเฉพาะ และนำค่าลักษณะเฉพาะของตัวอย่างภาพหน้าบุคคลมาสร้างเป็นแบบจำลองหน้าลักษณะเฉพาะเพื่อค้นหาตำแหน่งของใบหน้า

2.1.4.2 วิธีเชิงการกระจาย (Distribution-Base Methods) เป็นการแสดงการกระจายตัวของรูปแบบข้อมูลตัวอย่างที่มีความเป็นหน้าและความไม่เป็นหน้าเพื่อใช้เป็นบรรทัดฐานในการตัดสินใจ ตัวอย่างเช่น Sung and Poggio (1998) นำฟังก์ชันเกาส์เซียน (Gaussian function) มาประมาณกลุ่มการกระจายของค่าเฉลี่ยกลุ่มตัวอย่าง แสดงในรูปที่ 2.3



รูปที่ 2.3 การกระจายตัวของข้อมูลตัวอย่าง

2.1.4.3 โครงข่ายประสาทเทียม (Artificial Neural Networks) เป็นการฝึกสอนโครงข่ายด้วยโครงสร้างใบหน้าที่มีความซับซ้อน โดยทำการปรับค่าน้ำหนักประสาทเพื่อให้ค่าความผิดพลาดกำลังสองเฉลี่ยมีค่าลดลง ความถูกต้องของโครงข่ายประสาทเทียมขึ้นกับตัวอย่างที่ทำการฝึกสอน จำนวนชั้นนิวรอน และจำนวนนิวรอนที่เหมาะสม ตัวอย่างเช่น Rowley, Baluja, and Kanade (1998) ใช้โครงข่ายประสาทเทียมแบบแพร่กลับในการฝึกสอนข้อมูลตัวอย่าง โดยทำการแบ่งข้อมูลออกเป็นภาพตัวอย่างใบหน้าบุคคลและภาพที่ไม่ใช่ใบหน้าบุคคลอย่างละ 1,048 ภาพ และแบ่งชุดข้อมูลออกเป็น 15 ชุดข้อมูล รวมข้อมูลตัวอย่างทั้งหมดได้ 15,720 ภาพ ซึ่งการฝึกสอนที่มีจำนวนข้อมูลตัวอย่างที่ครอบคลุมและมากเพียงพอจะช่วยลดความผิดพลาดของระบบ

2.1.4.4 วิธีเวกเตอร์เกื้อหนุน (Support Vector Machines : SVMs) เป็นโครงข่ายแบบมีการฝึกสอนคล้ายคลึงกับโครงข่ายประสาทเทียม แต่มีหลักการที่แตกต่างกันตรงที่โครงข่ายประสาทเทียมนั้นจะทำการฝึกสอนเพื่อลดผลของค่าความผิดพลาด (empirical risk minimization) แต่ SVMs นั้นจะทำการฝึกสอนเพื่อลดของโครงสร้างความเสี่ยงให้น้อยที่สุด (structural risk minimization) และต้องการ ข้อมูล

ตัวอย่างที่มาก เพื่อลดความผิดพลาดของระบบ ตัวอย่างเช่น Osuna, Freund, and Girosi (1997) ทำการศึกษาเพื่อลดจำนวนในการฝึกสอน SVMs ให้มีปริมาณลดลงแต่ประสิทธิภาพในการทำงานไม่ลดลงไปด้วย

นอกจากนั้นยังมีการนำเอาวิธีการต่าง ๆ ข้างต้นมาใช้งานร่วมกัน โดยพัฒนาให้สามารถลดขอบเขตในการตรวจจับหรือเพิ่มประสิทธิภาพในการตรวจจับใบหน้าบุคคลให้ดียิ่งขึ้น ตัวอย่างเช่น Mitsukura, Fukumi and Akamatsu (2001) ใช้การค้นหาแบบจีนเนติกอัลกอลิทึมร่วมกับโครงข่ายประสาทเทียม LSDNN (Lip detection neural network and Skin Distinction Neural Network) โดยทำการค้นหาตำแหน่งของปากด้วย LSDNN และทำการค้นหาความสัมพันธ์ของระยะห่างระหว่างตำแหน่งของปากและสีผิวบริเวณโดยรอบด้วยการค้นหาแบบจีนเนติกอัลกอลิทึม และ Li, Gong, Sherrah, and Liddell (1999) นำเสนอวิธีเวกเตอร์เกื้อหนุนร่วมกับปริภูมิลักษณะเฉพาะ (eigenspace) ในการตรวจจับใบหน้าบุคคล โดยใช้ปริภูมิลักษณะเฉพาะในการแบ่งขอบเขตความเป็นหน้าบุคคลออกจากสิ่งแวดล้อม และทำการพิจารณาร่วมกับการตรวจจับรูปแบบใบหน้าด้วยวิธีเวกเตอร์เกื้อหนุน

2.2 สรุป

บทที่ 2 นำเสนอรายงานผลการสืบค้นวรรณกรรมวิจัยย้อนหลังที่เกี่ยวข้องกับงานวิจัยที่จะดำเนินการ จากฐานข้อมูล IEEE IEE Science Direct และอื่น ๆ ซึ่งทำให้ทราบถึงแนวทางการวิจัยที่เกี่ยวข้อง ระเบียบวิธีที่ผู้วิจัยอื่น ๆ ได้นำมาใช้ ผลการดำเนินงาน ข้อเสนอแนะต่าง ๆ จากคณะนักวิจัยตั้งแต่อดีตจนถึงปัจจุบัน จากการสืบค้นปริทัศน์วรรณกรรมและงานวิจัยที่เกี่ยวข้อง พบว่าเทคนิคการตรวจจับหน้าคนจากภาพสีส่วนใหญ่เป็นการวิจัยด้วยเทคนิคการประมวลผลภาพและคอมพิวเตอร์วิทัศน์จากข้อมูลภาพทั้งจากการเก็บตัวอย่างภาพด้วยตัวเองและจากฐานข้อมูลรูปต่าง ๆ ที่ถ่ายภาพภายใต้การควบคุมปัจจัยต่าง ๆ ได้แก่ แสงสว่าง ความซับซ้อนของภาพพื้นหลัง และ ลักษณะการวางท่า เป็นต้น ซึ่งเทคนิคในการตรวจจับหน้าคนแต่ละแบบนั้น ล้วนมีข้อจำกัดต่าง ๆ กันไป จากแนวคิดดังกล่าวงานวิจัยนี้จึงได้ทำการพัฒนาเพื่อลดข้อจำกัดของขนาดข้อมูลภาพอินพุต เพื่อให้ระบบมีขนาดเล็กลง และสามารถทำงานได้ดีขึ้น โดยที่ระบบมีการพัฒนาตัวเองเพื่อให้ระบบเกิดความเป็นปัจจุบันอยู่เสมอ ซึ่งเป็นข้อจำกัดในงานวิจัยที่ใช้โครงข่ายประสาทเทียมในการตรวจจับหน้าคน รายละเอียดต่างๆ จะได้กล่าวถึงในบทต่อไป