โครงการวิจัยย่อยลำดับที่ 21

เรื่อง ระบบจำแนกแบบหลายตัวโดยกรรมวิธีการเข้ารหัสช่องสัญญาณและฐานหลักดิสครีมิแนนต์เฉพาะที่เสถียร

.1ผู้รับผิดชอบโครงการ อาจารย์ ดร .วิทยากร อัศดรวิเศษ

บทที่ 1
บทนำ

1.1 ความเป็นมาและเหตุผล

โดยทั่วไป ปัญหาในการรู้จำเป้าหมายอัตโนมัติ (Automatic Target Recognition: ATR) ประกอบด้วย การตรวจจับ (detecting) และการบ่งชี้ (identifying) และการประมาณค่าพารามิเตอร์ของเป้าหมายในภาพอาฌาบริเวณ(scene)โดยการใช้เซ็นเซอร์เพียงตัวเดียวหรือหลายตัว โดยที่เราสามารถตั้งสมมติฐานได้ว่าเป้าหมายได้ถูกตรวจจับและตำแหน่งที่ตรวจจับในอาฌาบริเวณภาพได้ถูกทราบแล้ว การรู้จำเป้าหมายของเรดาร์แบบช่องเปิดสังเคราะห์ (Synthetic Aperture Radar: SAR) นั้น หมายถึงเป้าหมายได้ถูกจัดกรอปจากภาพอาฌาบริเวณของ SAR แล้ว ภาพ SAR นั้นค่อนข้างอ่อนไหวต่อมุมการหมุนของเป้าหมาย เหตุผลที่สำคัญของลักษณะแตกต่าง (variability) อย่างมากมายของแต่ละภาพเป้าหมาย SAR ประกอบด้วย ผลของเงา การที่เป้าหมายมีผลต่อสิ่งแวดล้อม เช่น Clutter การฉาย (projection) ของภาพอาฌาบริเวณสามมิติไปบนระนาบ slant รวมทั้งลักษณะแตกต่างของสัญญาณเรดาร์เชิงซ้อนที่เกิดจากการรวมเชิงโคฮีเรนต์ของสัญญาณสะท้อนกลับจากตัวกระจาย (scatter) เมื่อเป้าหมายมีการเคลื่อนที่เป็นระยะห่างสัมพัทธ์ต่ออันดับ (order) ของความยาวคลื่นของสัญญาณส่ง (target speckle)

นอกจากนี้ภาพ SAR ของเป้าหมายที่ได้จากภาพอาฌาบริเวณมีมิติขนาดใหญ่ ในขณะเดียวกันจำนวนภาพตัวอย่างที่ได้จะมีขนาดน้อย จะเห็นได้ว่าปัญหาในการรู้จำเป้าหมายอัตโนมัติเป็นปัญหาในการรู้จำรูปแบบที่ท้าทายเป็นอย่างมาก องค์ความรู้ที่ได้จากการรู้จำเป้าหมายอัตโนมัติสามารถนำมาใช้กับปัญหาการรู้จำรูปแบบต่างเช่น การรู้จำใบหน้าที่จำเป็นอย่างยิ่งในการป้องกันประเทศ ทั้งในการตรวจสอบหนังสือเดินทาง การเดินทางทางเครื่องบิน และ การทำเหมืองข้อมูล (data mining) เป็นต้น

ในทางปฏิบัติเราสามารถสร้างตัวจำแนกเป้าหมายอัตโนมัติให้ทำการตัดสินใจชนิดของเป้าหมายได้หลายวิธี ตัวอย่างเช่น การสร้างตัวจำแนกให้เสถียรต่อการหมุนและลักษณะแตกต่างของเป้าหมาย หรือ โดยการสกัดลักษณะบ่งต่าง (features) ที่มีความเสถียรต่อการหมุนของเป้าหมายในภาพอาฌาบริเวณ และสัญญาณรบกวน หรือ การประมาณค่ามุมการหมุนของเป้าหมายในภาพอาฌาบริเวณ

การสร้างตัวจำแนกที่มีกำลังในการแยกแยะเป้าหมายโดยไม่คำนึงถึงมุมการหมุนของเป้าหมายมี ข้อดี คือ เป็นการทำการรู้จำกับพิกเซลดิบ (raw pixel) ของภาพเป้าหมาย จึงไม่มีข้อมูลที่สำคัญสูงเสีย แต่มีข้อเสีย

ที่สำคัญคือ ภาพมีขนาดมิติใหญ่ มีจำนวนภาพในการฝึกฝนน้อย จะทำให้เกิดปัญหาการฝึกฝนเกิน (overtraining) หรือ การจำ (memorizing) หรือ Curse of dimensionality ที่สามารถอธิบายได้โดยทฤษฎี Occam's razor ปัญหานี้สามารถแก้ไขได้โดยการระบบตัวจำแนกแบบหลายตัว (multiple classifier systems: MCS) อย่างไรก็ตามการใช้ MCS อาจยังไม่สามารถแก้ปัญหาได้ทั้งหมด เราสามารถเพิ่มประสิทธิภาพในการรู้จำให้เพิ่มขึ้นได้โดยการใช้การสกัดคุณลักษณะที่เสถียรต่อการหมุน

การสกัดคุณลักษณะที่เสถียรต่อการหมุนอาจสามารถแก้ปัญหา curse of dimensionality แต่ในปัจจุบันยังไม่สามารถแก้ปัญหาการรู้จำเป้าหมายอัตโนมัติได้ดีนัก แนวทางเลือกที่ดีอีกทางหนึ่งคือแบ่งการทำงานของระบบจำแนกออกเป็น 2 ส่วนตามแนวความคิด แบ่งแยกและปกครอง (divide and conquer) โดยที่ในส่วนแรกเราจะทำการประมาณมุมการหมุนก่อน และแยกการรู้จำให้ดำเนินการในแต่ละเซกเมนต์ของการหมุน ซึ่งจะทำให้ ขอบเขตการตัดสิน (boundary decision) มีโครงสร้างที่ง่าย

ไม่ว่าวิธีการรู้จำเป้าหมายอัตโนมัติที่เพิ่มประสิทธิภาพด้วยวิธีการใดๆตามวิธีข้างต้น การรู้จำเป้าหมายอัตโนมัติจำเป็นต้องมีการสกัดคุณลักษณะโดยการแปลง (transform) เพื่อให้ได้ระบบพิกัด (coordinate system) ที่เหมาะสมสำหรับการรู้จำ โดยปกติการแปลงอาจเป็นไปได้ทั้งแบบเชิงเส้น และไม่เป็นเชิงเส้นที่ใช้ การแปลงเชิงเส้นที่สำคัญ ได้แก่ การวิเคราะห์ส่วนประกอบมุขสำคัญ (Principal Component Analysis: PCA) อนึ่งการแปลง PCA ยังมีข้อเสียเปรียบซึ่งได้มีการแนะนำ การแปลงเชิงเส้นชนิดอื่น เพื่อเพิ่มประสิทธิภาพการรู้จำ เช่น Linear Discriminant Analysis หรือ Local Discriminant Bases เป็นต้น อย่างไรก็ตามการแปลงที่กล่าวมาข้างต้น เป็นการแปลงที่ไม่ได้ให้ความสนใจต่อโครงสร้าง (structure) ของภาพเป้าหมาย ในที่นี้เราจึงสนใจการแปลงระบบพิกัดสามารถสกัดคุณลักษณะที่ให้ความหมายทางโครงสร้างของเป้าหมาย

ในบทที่ 2 เราจะแนะนำระบบจำแนกแบบหลายตัว หรือ คณาจักรกล ที่สร้างบนฐานของการเข้ารหัสโดยเฉพาะการเข้ารหัสต่อกันของ คณาจักรกล ต่างชนิดกัน รวมทั้งผลการศึกษาและทดลองการรู้จำเป้าหมายที่มีการประมาณหมุนการหมุน ในบทที่ 3 เราจะพิจารณาและนำเสนอการแปลงเชิงเส้นที่ให้ความสนใจต่อโครงสร้างของภาพเป้าหมายและแก้ปัญหาเดิมของการแปลง PCA การทดสอบการแปลงเชิงเส้นที่นำเสนอขึ้นกับฐานข้อมูลใบหน้าหลายชนิดได้แสดงให้เห็นถึงการปรับปรุงประสิทธิภาพในการรู้จำที่ดีขึ้น บทที่ 4 กล่าวถึงงานที่ได้ทำและที่จะดำเนินต่อไปในอนาคต รวมทั้งสรุป

1.2 วัตถุประสงค์ของโครงงานวิจัย

1.2.1 เพื่อศึกษาความเป็นไปได้ในการกำหนดระบบจำแนกแบบหลายตัว (multiple classifier systems: MCS) ที่มีนัยทั่วไปบนฐานของการใช้การเข้ารหัสช่องสัญญาณที่มีนัยทั่วไป และฐานหลักดิสครีมิแนนต์เฉพาะที่เสถียร

1.2.2 เพื่อศึกษาความสัมพันธ์ของการขยายระบบจำแนกแบบหลายตัวที่มีนัยทั่วไปกับระบบจำแนกหลายตัวแบบอื่นๆเพื่อสร้างแบบแผนการรวม (combining scheme) และกลุ่มระบบพิกัด (coordinate system ensemble) ที่มีประสิทธิภาพ

1.3 ขอบเขตของโครงงานวิจัย

1.3.1 ศึกษาการประยุกต์ใช้การแปลงขั้ว (polar) และ แรดอน (radon transform) สำหรับฐานหลักดิสคริมิแนนต์ เพื่อใช้สร้างระบบพิกัดเสถียรสำหรับระบบจำแนกแบบหลายตัว

1.3.2 ทดสอบการประยุกต์ใช้และนัยทั่วไปของฐานหลักดิสครริมิแนนต์เฉพาะที่เสถียรสำหรับระบบจำแนกแบบหลายตัวด้วยเวฟเลตฟิลเตอร์ ชนิด Daubechies, Coiflet, and Battle-Lemarie

1.3.3 ศึกษาการใช้ระบบจำแนกแบบหลายตัวบนฐานของฐานหลักดิสครริมิแนนต์เฉพาะที่เสถียร (invariant local discriminant bases-based MCS) ร่วมกับระบบจำแนกแบบหลายตัวบนฐานของการเข้ารหัสป้องกันความผิดพลาดเอาท์พุต (ECOC-based MCS) ซึ่งสามารถเรียกอีกอย่างหนึ่งว่าระบบจำแนกแบบหลายตัวบนฐานการเข้ารหัสต่อกัน (concatenation code-domain MCS) ซึ่งเกี่ยวข้องกับแบบแผนการเข้ารหัสต่อกันของดอเมนสัญญาณ (signal domain) ร่วมกันกับดอเมนการเข้ารหัสช่องสัญญาณ (code domain) ของระบบจำแนกแบบหลายตัวเปรียบเทียบระบบจำแนกแบบหลายตัวที่พัฒนาขึ้นเพื่อเปรียบเทียบกับระบบจำแนกแบบหลายตัวที่ทันสมัยอื่นๆ

1.3.4 เปรียบเทียบระบบจำแนกแบบหลายตัวที่พัฒนาขึ้นเพื่อเปรียบเทียบกับระบบจำแนกแบบหลายตัวที่ทันสมัยอื่นๆ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 2

ระบบจำแนกแบบหลายตัวโดยวิธีการเข้ารหัสช่องสัญญาณ และฐานหลักดิสคริมิ แนนต์เฉพาะที่เสถียรในการประยุกต์ใช้งานในการรู้จำเป้าหมายอัตโนมัติ

Multiple Classifier Systems using Channel Coding and Invariant Local Discriminant Bases for Automatic Target Recognition

2.1. บทนำ

ในรอบหลายปีที่ผ่านมานี้ *คณาจักรกล* (committee machine) หรือ วงหรือชุดของตัวประมาณ ได้พิสูจน์ว่าให้ผลการทำนายที่แม่นยำกว่าใช้ตัวประมาณเพียงตัวเดียว แนวความคิดพื้นฐานของ คณาจักรกล คือการฝึกฝนคณะของตัวจำแนกและรวมการประมาณของตัวประมาณแต่ละตัวเข้าด้วยกัน เพื่อเพิ่มสมรรถนะของนัยทั่วไป สมมติฐานหนึ่งของ คณาจักรกล คือ การที่ตัวประมาณแต่ละตัวควรเป็นอิสระต่อกัน ด้วยเหตุผลที่ว่าความผิดพลาดของการประมาณของตัวประมาณแต่ละตัวควรจะไม่มีค่าสหสัมพันธ์ (correlation) ต่อกันและกันสูงมากนัก ซึ่งจะทำให้การรวมการประมาณของตัวจำแนกหลายตัวสามารถขจัดความผิดพลาดออกไปได้ดี

ปัจจุบันนี้ นักวิจัยได้มีการนำ คณาจักรกล มาใช้ในงานหลายชนิด เช่นใช้ในการเฉลี่ยแบบจำลองทั้งชุด (ensemble-averaging model) ที่สร้างด้วยข่ายวงจรประสาทเพื่อใช้ในการวิเคราะห์ทางการเงิน การรู้จำตัวอักษรด้วยการใช้ แบกกิ้ง (bagging) บูสทิ้ง (boosting) และชุดของข่ายวงจรประสาท การใช้ข่ายเกทติ้งในการทำชัด (deblur) และการรู้จำใบหน้า

โดยปกติโครงสร้างของ คณาจักรกล นั้นสามารถแบ่งตามโครงสร้างออกได้เป็น 2 แบบ คือ แบบสถิต (static) และ แบบจลน์ (dynamic) ในทำนองเดียวกัน คณาจักรกล ก็สามารถแบ่งตาม กระบวนการเล็งเลิศออกได้เป็น 2 แบบ คือ การครอบคลุมที่เหมาะที่สุด (coverage optimization) และ การทำนายที่เหมาะที่สุด (prediction optimization)

2.1.1. คณาจักรกลแบ่งตามโครงสร้าง

2.1.1.1. โครงสร้างแบบสถิต

เป็นกระบวนการสร้าง คณาจักรกล ที่เกิดจากการใช้ ตัวจำแนกต่างชนิด หรือการเข้าถึงตัวสกัดคุณลักษณะ (feature) ที่แตกต่างกันของตัวจำแนกชนิดเดียวกัน หรือการฝึกฝนตัวจำแนกชนิดเดียวกันด้วยชุดฝึกฝนที่ต่างกัน (different training set) การฝึกฝนตัวจำแนกที่มีการป้อนสัญญาณรบกวนให้แต่ละตัวจำแนก

ตัวอย่างของ คณาจักรกล ที่มีโครงสร้างแบบสถิตนี้ได้แก่ การสร้างข่ายวงจรประสาทที่มีค่าเริ่มต้นของค่าถ่วงน้ำหนักที่ต่างกัน (different initialization weights) การทำการสุ่มซ้ำ (resampling) ชุดฝึกฝน เทคนิคนี้ถูกเรียกอีกอย่างหนึ่งว่า *แบกกิ้ง*

(Bagging) และเทคนิคการสุ่มซ้ำทางสถิติโดยที่ตัวอย่าง (sample) แต่ละตัวของชุดฝึกฝนที่ถูกเลือกมาใช้ในการฝึกตัวจำแนกในปัจจุบัน จะขึ้นอยู่กับสมรรถนะของตัวจำแนกที่ถูกสร้างขึ้นก่อนหน้านี้ เทคนิคนี้ถูกเรียกอีกอย่างหนึ่งว่า *บูสติ้ง* (Boosting)

2.1.1.2. โครงสร้างแบบจลน์

เป็นกระบวนการสร้าง คณาจักรกล ที่เกี่ยวข้องโดยตรงกับอินพุท ในกระบวนการรวม (combining) เทคนิคของการสร้างตัวจำแนกแบบหลายตัวนี้ รู้จักกันดีในอีกชื่อหนึ่ง คือ การรวมผสมผู้เชี่ยวชาญ (mixture of experts) โดยที่เอาท์พุทของตัวจำแนกตัวหนึ่งตัวใดจะถูกส่งผ่านโดยข่ายเกทติ้ง ที่จริงเกทติ้งจะเป็นตัวตัดสินว่าผู้เชี่ยวชาญตัวใดที่เกี่ยวข้องกับชุดย่อย (subsets) ที่แตกต่างกันของปริภูมิอินพุท (input space) ในที่นี้ผู้เชี่ยวชาญและข่ายเกทติ้งอาจเป็นข่ายวงจรประสาทก็ได้

2.1.2. คณาจักรกลแบ่งตามกระบวนการเล็งเลิศ

คณาจักรกล ที่จริงประกอบด้วย ชุดของตัวจำแนก และ กระบวนการในการรวบรวมการตัดสินใจของกลุ่มของตัวจำแนก โดยปกติ คณาจักรกลที่ได้รับความนิยม จะมีลักษณะแบบขนานซึ่งสามารถแบ่งออกได้ เป็น 2 แบบ คือ

1. ให้มีการรวบรวมการทำนายที่มีการกำหนดไว้ก่อน แต่เน้นที่การสร้างตัวจำแนกให้มีความหลากหลายเพื่อให้ผลของการจำแนกมีความถูกต้องแม่นยำที่สุด
2. ออกแบบและสร้างตัวจำแนกจำนวนหนึ่ง แต่เน้นที่การรวมการตัดสินใจของชุดของตัวจำแนก

2.1.2.1. การครอบคลุมที่เหมาะที่สุด (coverage optimization)

คณาจักรกล แบบที่ 1 ที่กล่าวถึงก่อนหน้านี้ สามารถเรียกอีกชื่อหนึ่งว่า *วิธีการครอบคลุมที่เหมาะสมที่สุด* (coverage optimization method) ซึ่งกรรมวิธีย่อยต่างๆ ที่อยู่ในการแบ่งแยกแบบนี้ ได้แสดงไว้ในตารางที่ 1 ดังต่อไปนี้

2.1.2.2. การทำนายที่เหมาะที่สุด (prediction optimization)

คณาจักรกล แบบที่ 2 ที่กล่าวถึงก่อนหน้านี้ สามารถเรียกอีกชื่อหนึ่งว่า *วิธีการทำนายที่เหมาะที่สุด* (prediction optimization method) ซึ่งกรรมวิธีย่อยต่างๆที่อยู่ในการแบ่งแยกแบบนี้ ได้แสดงไว้ในตารางที่ 2 ดังต่อไปนี้

งานวิจัยในส่วนนี้เน้นวิธีแบบขนานตามรูปแบบของวิธีการครอบคลุมที่เหมาะที่สุด โดยผสม คณาจักรกล จำนวน 3 แบบ เพื่อสร้าง คณาจักรกลแบบใหม่ คณาจักรกล สามแบบ ที่นำมาพิจารณา ได้แก่

- วิธี Adaboost ที่ทำการสร้างคณาจักรกล โดยการสุ่มซ้ำชุดการฝึกฝนที่มีการถ่วงน้ำหนักที่มีการเปลี่ยนแปลงตามความถูกต้อง
- วิธีรหัสแก้ไขผิดพลาดเอาท์พุต ที่มีการบังคับให้มีการฝึกฝน ขอบการตัดสินใจบางส่วน (partial decision boundary)
- วิธีที่ใช้ชุดการฝึกฝนหลายชุด (multiple training sets)

ตาราง 1 วิธี coverage optimization

Method	Training mechanism for introducing complementariness	Physical level
noise injection	inject noise to the training set	input
stacking	train classifiers by nonoverlapping subsamples of training set	input distribution $\mathcal{D}$
bagging	resample the training set by bootstrap replicates	input distribution $\mathcal{D}$
boosting	resample the training set by weights evolving with accuracy	input distribution $\mathcal{D}$
random subspace	project training set to random chosen subspaces	input representation
stochastic discrimination	generate random kernels to measure coverage of training set	classifier
perturbation	vary initial conditions or parameters of training process	classifier
error correction output coding	force training on partial decision boundaries	output
multiple feature sets	train classifiers using different feature subsets	input representation

ตาราง 2 วิธี prediction optimization

Trainable	binary or one of N decisions	ranked lists of classes	continuous prob. estimates or belief scores
No	majority	order statistics	sum, product rules
Yes	weighted vote	logistic regression	Bayes rules

โดยหลักการและเหตุผลที่ว่า การรู้จำโดยวิธี Adaboost มีความละม้ายคล้ายกับการรู้จำโดยการเข้ารหัสช่องสัญญาณชนิดที่มีการส่งข้อมูลซ้ำ (repetition coding) (รายละเอียดของวิธี Adaboost จะได้

กล่าวถึงใน บทที่ 2.2) ส่วนการรู้จำโดยวิธีรหัสแก้ไขผิดพลาดเอาท์พุต เป็นการรู้จำที่เกี่ยวข้องกับการเข้ารหัสช่องสัญญาณโดยตรง เป็นการเพิ่มความซ้ำซ้อน (redundancy) เพื่อเพิ่มประสิทธิ์ภาพในการรู้จำ โดยพิจารณาความละม้ายเหมือนของการรู้จำกับการรับส่งข้อมูลผ่านช่องสัญญาณที่มีการรบกวน (noisy channel) รายละเอียดของวิธีรหัสแก้ไขผิดพลาดเอาท์พุต จะได้กล่าวถึงใน บทที่ 2.3 ในปัจจุบันนั้นมีการศึกษาการเข้ารหัสที่เหมาะสมกับการส่งข้อมูลภาพและสัญญาณวิดีโอทางโครงข่าย (network) หรือช่องสัญญาณที่มีการรบกวน ที่เรียกว่า *แบบจำลองการเข้ารหัสแบบหลายการพรรณนา* (multiple description coding model) ซึ่งให้ความสัมพันธ์กับการเข้ารหัสทางด้านอินพุทก่อนมีการบีบอัด หรือ การสร้างความซ้ำซ้อนของอินพุท ก่อนมีการควอนไทเซชัน แล้วทำการส่งเข้าสู่ช่องสัญญาณ ซึ่งโดยปกติตามกรรมวิธีเดิม จะทำการบีบอัดภาพก่อนแล้วจึงค่อยทำการเข้ารหัส เทคนิคที่สำคัญที่ได้ผลดีเทคนิคหนึ่งในการทำการเข้ารหัสด้วยแบบจำลองการเข้ารหัสแบบหลายการพรรณนา โดยการใช้การแปลงเวฟเลตแบบซ้ำซ้อนในการประยุกต์ใช้ การเข้ารหัสด้วย แบบจำลองการเข้ารหัสแบบหลายการพรรณนา กับการรู้จำแบบหลายตัวจำแนกนั้น ในทางกายภาพการใช้งานจะมีความละม้ายเหมือนกับการสร้าง คณาจักรกล ด้วย การชุดฝึกฝนหลายชุด ซึ่งรายละเอียดของวิธีรหัสด้วย แบบจำลองการเข้ารหัสแบบหลายการพรรณนา ด้วยการใช้เวฟเลตแบบซ้ำซ้อน จะได้กล่าวถึงใน บทที่ 2.4

อนึ่งในทางปฏิบัติการเข้ารหัสช่องสัญญาณสามารถเพิ่มประสิทธิภาพและให้ผลดีได้โด*ยการเข้ารหัสต่อกัน* (concatenation code) ด้วยรหัสสั้นตั้งแต่สองรหัสขึ้นไป ซึ่งในหลายกรณีมีสมรรถนะดีกว่าการเข้ารหัสโดยใช้ รหัสยาว (long code) จะเห็นได้ว่าเราสามารถผสานกรรมวิธีการเข้าสร้าง คณาจักรกล โดยการผสาน คณาจักรกลชนิดต่างๆเข้าด้วยกันได้ ในการนี้การผสานคณาจักรกลเข้าด้วยกันจะมีความละม้ายเหมือนการเข้ารหัสต่อกัน ในบทที่ 2.5 จะอธิบายถึงการสร้างคณาจักรกล ด้วย คณาจักรกล 3 แบบ คือ Adaboost, วิธีรหัสแก้ไขผิดพลาดเอาท์พุต และ การใช้ชุดฝึกฝนหลายชุด ตลอดจนนำเทคนิคดังกล่าวมาใช้กับการรู้จำเป้าหมายอัตโนมัติ

ธรรมชาติของการรู้จำเป้าหมายอัตโนมัตินั้น จะอยู่ที่การที่ ภาพของเป้าหมายที่ถูกจับด้วยตัวตรวจจับทางภาพ (image sensor) อาจมีการหมุนของเป้าหมาย หากเป็นภาพที่ถ่ายจากทางอากาศของเป้าหมายที่อยู่บนดินเช่น รถถัง และ รถลำเลียง นั้น เราสามารถสมมติให้มีแต่การหมุนในแนวระนาบ (in-of-plane rotation) ส่วนภาพของวัตถุ เช่น ใบหน้า สามารถมีการหมุนนอกแนวระนาบ (out-of-plane rotation) นอกจากนี้ยังมีการเลื่อนทางสเปซ (translation หรือ position) และการเปลี่ยนสัดส่วน (scaling) การรู้จำที่คงทนต่อการเปลี่ยนแปลงทางการหมุน การเปลี่ยนสัดส่วน และ การเลื่อนทางสเปซ สามารถเรียกอีกอย่างหนึ่งว่า *การรู้จำรูปแบบเสถียร* (invariant pattern recognition หรือ position, rotation, and scaling invariance: prsi)

ในขั้นต้น การรู้จำเป้าหมายอัตโนมัติ จำเป็นที่ต้องคงทนต่อการหมุน ซึ่งสามารถทำได้โดยการแปลงขั้ว และแรดอน ในงานวิจัยนี้จึงเป็นการศึกษาการแปลง (transform) ที่เหมาะสำหรับใช้ในการสกัดคุณลักษณะที่เสถียรต่อการหมุน ในที่นี้การแปลงที่เหมาะสมคือการแปลงที่ให้สัมประสิทธิ์การแปลงที่เสถียรต่อการหมุนนั้น หรืออีกนัยหนึ่งคือเมื่อมีการหมุนของภาพอินพุทจะมีการเปลี่ยนแปลงของ

สัมประสิทธิ์การแปลงเพียงเล็กน้อย สมมติฐานของการศึกษาการแปลงที่เสถียรต่อการหมุนนั้นเพื่อให้การรู้จำมีความแม่นยำมากขึ้น อนึ่งการเพิ่มความแม่นยำในการรู้จำเป้าหมายอัตโนมัตินั้น นอกจากการใช้การแปลงที่เสถียรต่อการหมุนแล้ว เราสามารถใช้ทางเลือกอื่นในการเพิ่มอัตราการรู้จำได้ เช่น ใช้การประมวลก่อนหน้าในการประมาณมุม (pose) ในบทที่ 2.6 เราจะทำการเปรียบเทียบระเบียบวิธีและปัญหาของวิธีทั้งสอง ผลการทดลองการใช้รหัสช่องสัญญาณต่อกันในการรู้จำเป้าหมายอัตโนมัติ แสดงในบทที่ 2.7

2.2. Adaboost

Adaboost เป็นอัลกอริธึมแบบหนึ่งของวิธีบูสติ้ง ที่มีการเรียกใช้ตัวฝึกฝน (ตัวจำแนก) ชนิดอ่อน (weak classifier หรือ weak learner) หลายครั้ง โดยในแต่ละครั้งมีการปรับชุดฝึกฝนและการรวบรวมการตัดสินใจ ทีละเล็กละน้อยก่อนที่จะรวมกันเป็นสมมติฐาน (hypothesis) สุดท้าย ในการนี้การตัดสินใจสุดท้ายจะให้ผลการตัดสินใจที่แม่นยำกว่าการตัดสินใจของ ตัวฝึกฝนชนิดอ่อนแต่ละตัว ทั้งนี้ในทางทฤษฎีได้กำหนดไว้ว่า กรณีที่ตัวฝึกฝนชนิดอ่อนแต่ละตัวมีความถูกต้องเกินกว่าครึ่งหนึ่ง ผลรวมของการตัดสินใจจะลู่เข้าสู่การตัดสินใจที่เหมาะที่สุด แนวความคิดของ Adaboost คือการแจงค่าถ่วงน้ำหนักให้กับตัวอย่างแต่ละตัวของชุดฝึกฝน ในตอนต้นค่าถ่วงน้ำหนักจะมีค่าเท่ากัน แต่ในแต่ละรอบของการสร้างตัวฝึกฝนอ่อน ตัวฝึกฝนอ่อนจะให้ค่าสมมติฐาน โดยที่ค่าถ่วงน้ำหนักของตัวอย่างที่ตัดสินผิดพลาดจะมีค่าถ่วงน้ำหนักที่เพิ่ม ข้อดีคือการฝึกฝนจะถูกเน้นที่ชุดตัวอย่างที่ตัดสินยากของชุดฝึกฝน ส่วนข้อเสียของวิธีนี้ คือการฝึกฝนจะไม่คงทนต่อสัญญาณรบกวน เนื่องจากในบางกรณีอาจมีการเน้นไปที่ตัวอย่างและสัญญาณรบกวนมากกว่าที่ควรจะเน้นที่ตัวอย่างที่ยากในการฝึกฝน การตัดสินสุดท้ายจะเป็นการรวบรวมการตัดสินใจของสมมติฐานของแต่ละรอบการฝึกฝน หรือที่เรียกอีกอย่างหนึ่งว่า การถ่วงน้ำหนักของเสียงส่วนใหญ่ (weighted majority vote) โดยที่ สมมติฐานที่ความถูกต้องมากจะมีค่าถ่วงน้ำหนักมาก Adaboost ถูกพิจารณาว่าเป็น วิธี bias และ variance ไปพร้อมกัน อัลกอริธึมของ Adaboost แสดงดังในรูปที่ 1

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

The algorithm:

Let $D_t(i)$ denote the weight of example i in round t.
Initialization: Assign each example $(x_i, y_i) \in E$ the weight $D_1(i) = 1/n$.

- For t = 1 to T:
  - Call the weak learning algorithm with example set E and weight s given by D_t.
  - Get a weak hypothesis $h_t : X \to Y$.
  - Update the weights of all examples.
- Output the final hypothesis, generated from the hypotheses of rounds 1 to T.

Updating the weights in round t: $D_{t+1}(i) := \frac{D_t(i)exp(-\alpha_i y_i h_t(x_i))}{Z_t}$, where

- Z_t is chosen such, that $D_t + 1$ is a distribution.
- α_t is chosen according to the importance of hypothesis h_t.
  For $h_t : X \to \{-1, +1\}$ usually α_t is chosen as $\alpha_t := \frac{1}{2} \log \frac{(1-r_t)}{r_t}$,
  where $r_t = \sum_{i=1}^{l} D_t(i) y_i h_t$.

The final hypothesis $H : X \to \{-1, +1\}$ is chosen as

$$H(X) = sign(\sum_{t=1}^{T} \alpha_t h_t(x)).$$

รูปที่ 1 อัลกอริธึม Adaboost

2.3. รหัสแก้ไขผิดพลาดเอาท์พุต (error-correcting output code)

รหัสแก้ไขผิดพลาดเอาท์พุต (error-correcting output code: ECOC) เป็นวิธีสร้างและรวบรวมการตัดสินใจโดยการแยกองค์ประกอบเอาท์พุตจาก หลายคลาส เป็นชุดของปัญหาแบบทวิคลาส ใน ECOC นั้นเราแจงรหัส หรือ อักขระไบนารี ให้คลาสแต่ละคลาส โดยที่แต่ละบิตจะแทน ปัญหาทวิคลาสที่เกิดจากการแยกองค์ประกอบของคลาสหลายคลายให้เป็นชุดที่มีการแยกแยะเป็นสองส่วนอย่างชัดเจน ตัวอย่างเช่น ในกรณีของ 4 คลาส A, B, C และ D เราสามารถแบ่งเซตของคลาสสี่คลาสออกได้เป็นเซตสองเซต {A} และ {BCD} หรือ {B} และ {ACD} หรือ {AD} และ {BC} โดยในแต่ละค่าเราจะฝึกฝนตัวจำแนกที่ทำหน้าที่ดิสคริมิเนตระหว่างสองเซต ด้วยการใช้ตัวจำแนกแบบเข้ม (strong classifier)

ในการจำแนกคลาสของข้อมูลหรือรูปแบบ รูปแบบจะถูกป้อนเข้ากับตัวจำแนกทั้งหมด ลำดับที่ได้จากการตัดสินใจจะเป็นอักขระของแต่ละรูปแบบที่ยังไม่รู้คลาส ในกรณีที่การทำการแบ่งแยก (dichotomices) คลาสข้อมูลได้ถูกออกแบบจากเหมาะสม รหัสเอาท์พุตที่ได้จากลำดับการตัดสินใจของตัวจำแนกทั้งหมดจะทนต่อความผิดพลาด ในกรณีที่แตกต่างกับบิตแต่ละบิตของรหัสที่ได้ออกแบบมาก่อนหน้านี้ ด้วยเหตุนี้ความแม่นยำของการรู้จำโดยวิธีนี้ สามารถปรับปรุงให้ดีขึ้นได้ เมื่อมีการใช้วิธีการถอดรหัสที่มีการป้องกันความผิดพลาด (error correction) ด้วย ระยะห่างถูกต้อง (correction distance) ทำการถอดรหัส เช่น ระยะห่าง Hamming เป็นต้น

ข้อเสียเปรียบของการใช้ระบบตัวจำแนกแบบหลายตัวด้วยวิธี ECOC เนื่องจากการฝึกฝนใช้ ตัวจำแนกแบบเข้ม จะมีโอกาสค่อนข้างสูงที่ความผิดพลาดจากการตัดสินใจของแต่ละตัวจำแนกจะมีสหสัมพันธ์กันสูง โดยเฉพาะตัวจำแนกที่ใช้ข้อมูลเฉพาะที่ (local information) เช่น k-NN โดยปกติในการสร้างระบบตัวจำแนกหลายตัว ที่ใช้ตัวจำแนกชนิด k-NN นั้น เราจะทำการประมวลผลทางกายภาพ ผ่าน

ทางการพรรณนาอินพุต (ดูตารางที่ 1) หรือ การใช้ชุดฝึกฝนหลายชุด นั่นเอง รูปที่ 2 แสดงเมทริกซ์การเข้ารหัสหนึ่งที่เป็นไปได้ของปัญหาขนาด 10 คลาส

class	h_0	h_1	h_2	h_3	h_4	h_5	h_6	h_7	h_8	h_9	h_{10}	h_{11}	h_{12}	h_{13}	h_{14}
0	1	1	0	0	0	0	1	0	1	0	0	1	1	0	1
1	0	0	1	1	1	1	0	1	0	1	1	0	0	1	0
2	1	0	0	1	0	0	0	1	1	1	1	0	1	0	1
3	0	0	1	1	0	1	1	1	0	0	0	0	1	0	1
4	1	1	1	0	1	0	1	1	0	0	1	0	0	0	1
5	0	1	0	0	1	1	0	1	1	1	0	0	0	0	1
6	1	0	1	1	1	0	0	0	0	1	0	1	0	0	1
7	0	0	0	1	1	1	1	0	1	0	1	1	0	0	1
8	1	1	0	1	0	1	1	0	0	1	0	0	0	1	1
9	0	1	1	1	0	0	0	0	1	0	1	0	0	1	1

รูปที่ 2 รหัสแก้ไขผิดพลาดเอาท์พุต

2.4. คณาจักรกลด้วยแบบจำลองการเข้ารหัสแบบหลายการพรรณนา (Multiple Description Coding Model Based Multiple Classifier Systems)

ในการรับส่งข้อมูลภาพที่มีการบีบอัดผ่านทางโครงข่ายนั้น แบบจำลองการเข้ารหัสแบบหลายการพรรณนา (multiple description coding model: MDC) สามารถใช้การแปลงเวฟเลตซ้ำซ้อน ในหลายเหตุการณ์ทางการสื่อสาร สมรรถนะของการสื่อสารด้วยแบบจำลองการเข้ารหัสแบบหลายการพรรณนาดีกว่า การเข้ารหัสช่องสัญญาณ

แบบจำลองการเข้ารหัสแบบหลายการพรรณนา มีแนวคิดที่แตกต่างจากการเข้ารหัสช่องสัญญาณดังนี้ คือ แบบจำลองการเข้ารหัสแบบหลายการพรรณนาด้วยการแปลงรหัส (multiple description transform coding) หรือ ด้วยการขยายเฟรม (frame expansion) จะทำการแยกข้อมูลโดยการแปลงเชิงเส้นซ้ำซ้อน ข้อมูลที่ถูกขยายแล้วมีองค์ประกอบที่ไม่เป็นอิสระต่อกันเชิงเส้น ซึ่งก็คือการแยกองค์ประกอบแบบเฟรม นั่นเอง หลังจากการแยกองค์ประกอบ สัมประสิทธิ์การแปลงที่ซ้ำซ้อนแต่ละตัวจะถูกควอนไทซ์และส่งไปในช่องสัญญาณ ส่วนในการสื่อสารแบบดั้งเดิม ข้อมูลจะถูกควอนไทซ์ ด้วยหนังสือรหัสควอนไทซ์ (quantization codebook) แล้วจึงใช้ linear block code เข้าแปลี่ยนข้อมูลไบนารี ที่ได้จากการควอนไทซ์ จะเห็นได้ว่า linear block code ก็เป็น linear system แบบหนึ่ง ดังนั้น ความแตกต่างระหว่าง MDC กับการเข้ารหัสช่องสัญญาณคือ การสลับที่โดยใช้การแปลงเชิงเส้น ของขบวนการแปลงกับการควอนไทซ์ นั่นเอง สำหรับการนำเอา MDC มาใช้ในกรอบของการรู้จำนั้น เราต้องอาศัย ทฤษฎีของ Komogolov มาอธิบายถึง การหาค่าควอนไทเซชัน ของฟังก์ชันการจำแนก ทฤษฎีของ Komogolov อธิบายถึงการหา covering ต่อฟังก์ชัน

2.4.1. ฐานหลักดิสคริมิแนนต์เฉพาะที่ซ้ำซ้อน (Redundant Local Discriminant bases)

ในอีกทางหนึ่งการรู้จำรูปแบบดั้งเดิมนั้น ระบบพิกัดที่เหมาะสมสามารถหาได้ด้วยการแปลงเชิงเส้น เช่น การแปลง Karhunen-Loeve และการวิเคราะห์การแบ่งแยกเชิงเส้น (linear discriminant analysis: LDA) เร็วนี้ๆ ฐานหลักดิสคริมิแนนต์เฉพาะที่ (local discriminant bases: LDB) ได้ถูกแนะนำเป็นระบบพิกัดที่เหมาะสมสำหรับงานรู้จำรูปแบบ ดังที่แสดงใน [Saito] LDB คงทนต่อ outlier และการรบกวนกว่า การแปลง Karhunen-Loeve และ LDA นอกจากนี้ LDB ยังสามารถตรวจจับคุณลักษณะเฉพาะที่ โดยมีการคำนวนที่รวดเร็ว และสามารถใช้ในการตีความการรู้จำในเชิงโครงสร้างของข้อมูลได้ดีกว่า

การแปลงฐานหลักดิสคริมิแนนต์เฉพาะที่ ได้รับแรงบันดาลใจจากการแปลงเวฟเลตแพกเก็ต (wavelet packet) และการเลือกฐานหลักที่เหมาะสมที่สุด (Best Basis Selection) การหาฐานหลักที่เหมาะสมที่สุดสามารถหาได้จากอัลกอริธึมโปรแกรมมิ่งจลน์ (dynamic programming algorithm) ความแตกต่างระหว่างการใช้งานการแปลงเวฟเลตแพกเก็ต กับ ฐานหลักดิสคริมิแนนต์เฉพาะที่ คือ การแปลงเวฟเลตแพกเก็ตจะนิยมใช้กับการพรรณนาและการบีบอัดภาพ ฟังก์ชันราคา (cost function) จึงเป็น ฟังก์ชันราคาทาง เอนโทรปี ที่เหมาะกับการบีบอัดข้อมูล ในขณะที่ฐานหลักดิสคริมิแนนต์เฉพาะที่ จะใช้ ฟังก์ชันราคาที่เหมาะกับการรู้จำข้อมูล เช่น ระยะห่างคู่ (pairwise distance) ระหว่างรูปแบบ 2 รูปแบบ

ในงานวิจัยเราเป็นคณะวิจัยแรกที่ขยายการใช้งาน ฐานหลักดิสคริมิแนนต์เฉพาะที่ ให้เกิดความซ้ำซ้อนเพื่อให้มีลักษณะที่เหมาะสมกับการสกัดคุณลักษณะ เพื่อสร้างคณาจักรกลแบบใช้หลายชุดฝึกฝน หรือ คณาจักรกลด้วยแบบจำลองการเข้ารหัสแบบหลายการพรรณนา นั่นเอง

2.5. การเข้ารหัสช่องสัญญาณต่อกันและการประยุกต์ใช้สำหรับระบบตัวจำแนกแบบหลายตัวสำหรับการรู้จำเป้าหมายอัตโนมัติ

แนวความคิดของการเข้ารหัสต่อกัน เริ่มต้นเสนอโดย Forney โดยที่มีการกำหนด *รหัสใน* (inner code) กับช่องสัญญาณรวมกันเป็น *ช่องสัญญาณยวดยิ่ง* (superchannel) โดยมีรหัสที่สอง ที่เรียกว่า *รหัสนอก* (outer code) ที่ใช้สำหรับเข้ารหัสข้อมูลก่อนส่งเข้าไปในช่องสัญญาณยวดยิ่ง การเข้ารหัสนี้เรียกอีกอย่างหนึ่งว่า การ

ในโครงการนี้เราจะนำเอาวิธีการทาง คณาจักรกล 3 แบบ คือ Adaboost, ECOC และ การใช้ชุดฝึกฝนหลายชุด มาทำงานต่อกันตามแนวความคิดการเข้ารหัสช่องสัญญาณต่อกัน ข้อดีของการสร้างคณาจักรกลแบบหลายคณาจักรกลที่ต่อกันตามลำดับ สามารถลำดับได้ดังนี้ คือ การใช้ชุดฝึกฝนแบบหลายชุด สามารถทนต่อการสัญญาณรบกวน ซึ่งเป็นข้อเสียเปรียบของวิธี Adaboost ในขณะเดียวกัน Adaboost ก็สามารถแก้ปัญหาการเกิดสหสัมพันธ์ของการทำนายของ ECOC ได้ดี ในขณะเดียวกัน ECOC ก็สามารถลดความซับซ้อนและเวลาในการฝึกฝนของ Adaboost เป็นต้น

2.6. การแปลงที่เสถียรต่อการหมุน เปรียบเทียบ กับการประมาณมุมการหมุน สำหรับการรู้จำเป้าหมายอัตโนมัติ

ดังที่ทราบ ว่าสัมประสิทธิ์ที่ได้จากการแปลงเวฟเลตนั้นไม่เสถียรต่อการเลื่อน (translation) การหมุน (rotation) และ การเปลี่ยนสัดส่วน (scaling) ได้มีการแก้ปัญหาดังกล่าวบ้างแล้ว เช่น การใช้การแปลง log-polar สำหรับการรู้จำลวดลายเสถียร (texture) อนึ่งนอกเหนือจากวิธีที่ใช้ในงานวิจัยข้างต้นนี้ มีงานวิจัยทางด้านการรู้จำรูปแบบเสถียร ที่ใช้การแปลงรูปแบบอื่น เช่น การแปลงแรดอน การแปลงแรดอนเป็นการแปลงที่มีประสิทธิภาพในการแปลงฟูริเยร์ ของพิกัดขั้ว ในทางปฏิบัติคุณลักษณะที่เสถียรต่อการหมุนสามารถหาได้โดยการประมวลผลตามหลังต่อค่าสัมประสิทธิ์ของการแปลงแรดอน เช่น การแยกองค์ประกอบค่าเอกฐาน (singular value decomposition) หรือ การขยายฟังก์ชันฮาร์โมนิกเชิงวงกลม (circular harmonic function expansion) อนึ่งการวิจัยในส่วนนี้ยังไม่เหมาะสมกับการใช้ในการรู้จำเป้าหมาย เท่าที่ศึกษาจะมีการนำไปใช้กับการรู้จำเป้าหมายด้วยขอบ (boundary target recognition) หรือ การรู้จำลวดลาย (texture reconition) ซึ่งเป็นการรู้จำอย่างง่าย

ในงานวิจัยนี้ เราได้ศึกษาการประมาณมุมการหมุน (pose estimation) ซึ่งเป็นวิธีที่นิยมใช้กันมากในการรู้จำเป้าหมายอัตโนมัติ ในขั้นต้นการศึกษานี้เป็นเพื่อการสร้างตัวเปรียบเทียบ โดยที่คุณลักษณะอาจไม่จำเป็นต้องใช้ LDB แต่เราจะใช้ การวิเคราะห์ส่วนประกอบมุขสำคัญ (PCA) แทนในขั้นต้น

2.6.1. ผลการศึกษาการประมาณมุมการหมุน

ตัวอย่างของภาพ SAR ที่เป็นเอเรย์ของพิกเซลขนาด m x n นั้นเราสามารถเปลี่ยนเป็นเวกเตอร์แถวขนาด mn ได้ เราสามารถพิจารณาให้สัญญาณที่เราสนใจนั้นอยู่ในแบบจำลองทางสถิติที่ประกอบด้วย พารามิเตอร์ 2 ชนิด คือ ชนิดของเป้าหมายและทิศทางการหมุน โดยในที่นี้เราสามารถพิจารณาให้ขบวนการของสัญญาณเป็นขบวนการแบบสุ่มเกาเสียนที่มีสัญญาณรบกวน

เป้าหมายหรือสัญญาณที่เป็นขบวนการแบบสุ่มเกาเสียนจะมีค่าเฉลี่ยเงื่อนไข (conditional mean) และ ค่าโคแวเรียนซ์เงื่อนไข (conditional covariance) โดยที่สัญญาณรบกวนจะมีค่าเฉลี่ยเป็นศูนย์ และ โควาแวเรียนซ์ เท่ากับ $N_0 I$

เราสามารถสมมติโดยไม่มีการสูญเสียทางนัยสำคัญ (without loss of generality) ให้สัญญาณรบกวนมีความเป็นอิสระต่อกันระหว่างพิกเซลต่อพิกเซล และให้สัญญาณรบกวนมีความเป็นอิสระต่อกันกับสัญญาณ นอกจากนั้นจากการทดลอง เราสามารถจำลองให้สัญญาณมีความเป็นอิสระต่อกันในขณะเดียวกันมีค่าเฉลี่ยเท่ากันศูนย์ ทำให้เราสามารถคำนวนหาค่าฟังก์ชัน log-likelihood ได้โดยไม่ต้องคำนวนเมตริกซ์เพียวแต่คำนวนหาค่าผลบวกของค่า log ของพิกเซลทั้งหมดเท่านั้น

ในที่นี้ค่าไม่ทราบของแบบจำลองของสัญญาณเพื่อใช้ในการหาค่าฟังก์ชัน log-likelihood คือค่า แวเรียนซ์ของพิกเซลแต่ละพิกเซล ดังนั้นการฝึกฝนแบบจำลองของสัญญาณคือ

การประมาณค่า แวเรียนซ์ ค่าแวเรียนซ์ที่ประมาณได้นี้สามารถนำมาใช้ในการประมาณทิศการหมุนของเป้าหมายและรู้จำเป้าหมายอัตโนมัติ

เราแบ่งฐานข้อมูลออกเป็น N_w หน้าต่าง โดยที่แต่ละหน้าต่างจะมีการฝึกฝนเป็นช่วงช่วงละ d เรเดียน ฟังก์ชันแวเรียนซ์ของพิกเซลที่ I จะได้จากการเฉลี่ยของผลบวกของพิกเซลของภาพฝึกฝนที่อยู่ในช่วงหมุนที่กำหนดของเป้าหมายแต่ละเป้าหมาย ในงานที่ทดลองทางด้านนี้อาจมีภาพแวรียนซ์ 72 ภาพ (หน้าต่าง) ซึ่งแต่ละหน้าต่างมีความกว้างขนาด 10 องศา

การมีการประมาณการหมุนเป็นกระบวนการหนึ่งของการรู้จำเป้าหมายจะทำให้การลดจำนวนของการเปรียบเทียบตัวอย่างของสมมติฐานการหมุน ดังนั้นเราสามารถที่จะนำกำลังการคำนวนมาเพิ่มประสิทธิภาพในการรู้จำ

การรู้จำเป้าหมายอัตโนมัติสามารถทำไปพร้อมกับการประมาณมุม โดยที่การประมาณการหมุนสามารถเป็นไปตามการประมาณแบบ Hilbert-Schmidt หรือ แบบ MAP (maximum a posteriori) ในทำนองเดียวกัน โดยที่เราตั้งสมมติฐานให้ ฟังก์ชันความหนาแน่น posteriori probability ที่คงที่ แบบpiecewise ของมุมในการหมุนและตำแหน่ง นอกจากนี้เนื่องจากเราไม่รู้ค่า priori probability ของมุมในการหมุนและตำแหน่ง โดยไม่มีผลทางนัยสำคัญเราสามารถกำหนดค่า priori probability ของมุมในการหมุนและตำแหน่งให้มีการกระจายเป็นแบบสม่ำเสมอ (uniform)

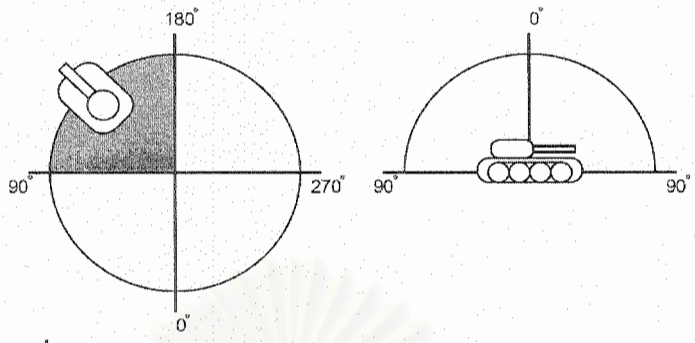
สำหรับคลาสของเป้าหมายสามารถตัดสินได้โดยการพิจารณาการกระจายความหนาแน่นฟังก์ชัน posteriori ของค่ามุม ตำแหน่ง และคลาสของเป้าหมาย ต่อ สัญญาณที่ได้รับ ในทำนองเดียวกันเราสามารถเพิ่มค่าเทรสโฮลด์ต่อค่า log ของ posteriori probability ของสัญญาณที่ได้รับ หรือ การทดสอบสัดส่วนใกล้เคียงที่มีนัยทั่วไป (generalized likelihood ratio test)

ในการทดลองเกี่ยวกับการประมาณมุมของการรู้จำเป้าหมายอัตโนมัติของฐานข้อมูล MSTAR นั้น ดารประมาณผลมีความถูกต้องสูงมาก เราจึงสามารถเพิ่มประสิทธิภาพในการรู้จำได้โดยไม่ต้องมีการแปลงที่เสถียรต่อการหมุน

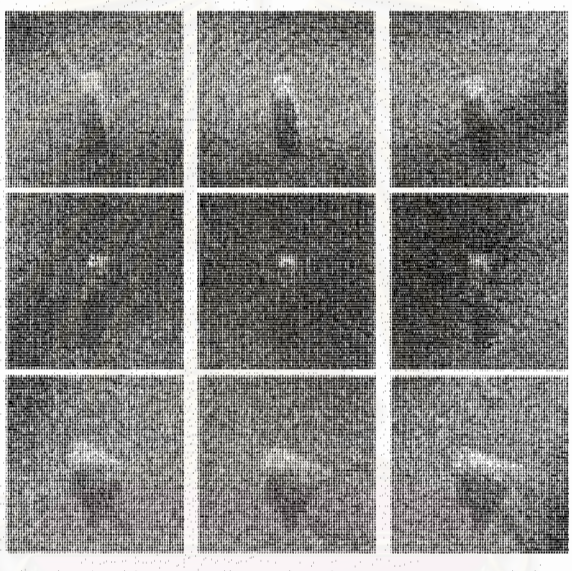
2.7. ผลการทดลอง

ฐานข้อมูล MSTAR ประกอบด้วยชุดข้อมูล Synthetic Aperture Radar (SAR) ความละเอียดสูง รวบรวมโดยห้องวิจัยโปรแกรม DARPA/Wright laboratory Moving and Stationary Target Acquisition and Recognition (MSTAR) ชุดข้อมลประกอบด้วยภาพ SAR ขนาด 128 x 128 ของรถทางทหาร 3 แบบ ได้แก่ BMP2 armored personal carriers (APCs), BTR70 APCs และ รถถัง T72 ซึ่งรูป SAR ดังกล่าวขึ้นกับการวางตัวของรถเป้าหมายที่ต่างกัน ดังนั้นการตรวจรู้ทิศทางการวางตัวของรถก่อน (มุม Aspect) ดังรูปที่ 3 สามารถเพิ่มประสิทธิผลการรู้จำได้มากขึ้นซึ่งในการทดลองของเราจะทำด้วยมือ ตัวอย่างรูปภาพจากฐานข้อมูล MSTAR แสดงดังรูปที่ 4 โดยภาพทั้งหมดถูกตัดเอาเฉพาะบริเวณกลางภาพขนาด 32 x 32, 48 x 48, 36 x 36 และ 80 x 80 พิกเซล ตารางที่ 3 แสดงชุดฝึกฝน และตารางที่ 4 แสดงชุดทดสอบ ตารางที่ 5 แสดงผลการทดลองการรู้จำเป้าหมายอัตโนมัติด้วยคณาจักรกลแบบเข้ารหัสต่อกัน โดยทำการเปรียบเทียบกับวิธีชั้นนำที่ให้ผลการรู้จำที่ดีที่สุดในขณะนี้ ตารางที่ 6 แสดงผลการรู้จำเป้าหมายอัตโนมัติแบบใช้การ

ประมาณมุมการหมุนและการหาส่วนประกอบมุขสำคัญแทนการใช้การแปลงเชิงเส้นที่เสถียรต่อการหมุน (รายละเอียดวิธีการหาส่วนประกอบมุขสำคัญที่โครงการนำเสนอแสดงในบทถัดไป)



รูปที่ 3 ภาพซ้ายคือมุม Aspect และภาพขวาคือมุม Depression



รูปที่ 4 ตัวอย่างภาพ SAR จากฐานข้อมูล MSTAR โดยแถวบนเป็นรถ BMP2 armored personal carriers (APCs) แถวกลางเป็นรถ BTR70 APCs และแถวล่างเป็นรถถัง T72

ตารางที่ 3 ชุดฝึกฝนของฐานข้อมูล MSTAR

ชนิดของรถ	หมายเลขรถ	หมายเลขซีเรียล	มุม Depression (องศา)	จำนวนภาพ
BMP2	1	9563	17	233
	2	9566		231
	3	C21		233
BTR70	1	C71	17	233

T72	1	132	17	232
	2	812		231
	3	S7		228
รวม	-	-	17	1,621

ตารางที่ 4 ชุดทดสอบของฐานข้อมูล MSTAR

ชนิดของรถ	หมายเลขรถ	หมายเลขซีเรียล	มุม Depression (องศา)	จำนวนภาพ
BMP2	1	9563	15	195
	2	9566		196
	3	C21		196
BTR70	1	C71	15	196
T72	1	132	15	196
	2	812		195
	3	S7		191
รวม	-	-	15	1,365

ตารางที่ 5 แสดงผลการทดลองการรู้จำเป้าหมายอัตโนมัติด้วยคณาจักรกลแบบเข้ารหัสต่อกัน (เปรียบเทียบวิธีต่างๆ ในเปอร์เซ็นต์ของการรู้จำภาพทั้งหมดในฟังก์ชันของขนาดภาพ)

Method/Image Size	32 x 32	48 x 48	64 x 64	80 x 80
MDBC+NN	75.88	N/A	N/A	N/A
CGSM	N/A	N/A	N/A	**98.53**
ECOC-SVM (raw)	84.46	90.16	91.76	92.70
ECOC-SVM (MDB)	85.42	90.81	92.51	92.97
Adaboost (raw)	88.24	93.35	93.48	93.68
Adaboost (MDB)	89.66	93.16	94.26	93.97
MDC-Adabost-ECOC	91.21	94.95	94.80	95.24

ตารางที่ 6 การเปรียบเทียบความถูกต้องในการรู้จำของวิธีการ 2DPCA B2DPCA 2DPCA+FSS และ B2DPCA+FSS ที่ช่วงมุมต่าง ๆ บนฐานข้อมูล MSTAR

วิธีการ	ช่วงมุม (องศา)	ความถูกต้อง (%)	d	q
2DPCA	5	98.28	5	-
	10	98.35	5	-
	15	98.73	5	-
	20	98.57	5	-
	30	98.66	5	-
	40	98.24	5	-
B2DPCA	5	98.24	5	10
	10	98.14	5	10
	15	98.14	5	10
	20	98.15	5	10
	30	98.35	5	10
	40	97.62	5	10
2DPCA+FSS	5	93.10	5	-
	10	94.57	5	-
	15	94.22	5	-
	20	93.14	5	-
	30	90.22	5	-
	40	86.62	5	-
B2DPCA+FSS	5	94.54	5	10
	10	95.93	5	10
	15	95.60	5	10
	20	93.80	5	10
	30	91.82	5	10
	40	88.16	5	10

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 3
การรู้จำใบหน้าและเป้าหมายอัตโนมัติโดยวิธีการวิเคราะห์ส่วนประกอบมุขสำคัญแบบสองมิติและปริภูมิย่อยใบหน้าเฉพาะ

Face and Automatic Target Recognitions using

Two-Dimensional Principal Component Analysis and Face-Specific Subspace

3.1. บทนำ

การวิเคราะห์ส่วนประกอบมุขสำคัญ (principal component analysis: PCA) เกี่ยวข้องกับสถิติอันดับที่สองของภาพอินพุท และการหาการพรรณนา (reconstruction) ที่มีค่าเหมาะสมที่สุดโดยการหาค่าผิดพลาดของการพรรณนาที่น้อยที่สุดตามกระบวนการหาค่า *กำลังสองต่ำสุด* (least square) ค่าเวกเตอร์เจาะจง (eigenvector) ที่คำนวนจากเมตริกซ์โควาเรียนซ์ของใบหน้า ก็คือ ค่าเจาะจงใบหน้า นั่นเอง ขนาดมิติของปริภูมิคุณลักษณะของใบหน้า (face feature space) สามารถลดให้มีขนาดเล็กลงได้โดยเลือก ไอเกนเวกเตอร์ เฉพาะที่มี ค่าเจาะจง (eigenvalue) มากที่สุด

เมื่อเราสามารถสร้างปริภูมิของใบหน้าได้แล้ว ใบหน้าของภาพที่จะถูกทดสอบจะถูกโพรเจกไปบนปริภูมิใบหน้าเพื่อให้ได้เวกเตอร์คุณลักษณะซึ่งก็คือ ค่าสัมประสิทธิ์ของการพรรณนา (representation coefficients) ของปริภูมิใบหน้า ตัวจำแนกสามารถตัดสินเลือกคลาสของภาพอินพุท โดยเปรียบเทียบคะแนนความเหมือน (similarity score) ระหว่าง เวกเตอร์คุณลักษณะ ของภาพทดสอบ กับ ค่า เวกเตอร์คุณลักษณะ PCA ของแต่ละบุคคลที่อยู่ในฐานข้อมูล

ปัญหาแรกของปัญหาหลักของการใช้ PCA เกิดจากการประมาณค่าที่อาจไม่แม่นยำของเมตริกซ์โควาเรียนซ์ ที่มีมิติขนาดใหญ่และมีจำนวนตัวอย่างน้อย เช่น ในการประยุกต์การรู้จำใบหน้า หรือ ที่รู้จักกันดีในชื่อ ว่า ขนาดจำนวนตัวอย่างน้อย (small sample size: SSS) หรือ คำสาปของมิติ (curse of dimensionality) ปัญหาที่สอง คือ โครงสร้างทางสเปซ ของภาพที่หายไปเนื่องจากการทำ เวกเตอร์ไรซ์ ปัญหาสุดท้าย คือ PCA คำนึงถึงการกระจายของข้อมูลเป็นหลัก หรือ ข้อมูลภายในคลาส (within-class) ซึ่งทำด้วยการนอมอลไรซ์ข้อมูลตามค่าแวเรียนซ์ ซึ่งข้อมูลของ ระหว่างคลาส (between-class) ได้ถูกละเลย ปัญหานี้สามารถแก้ไขได้โดย การวิเคราะห์การแบ่งแยกเชิงเส้น (Linear Discriminant Analysis: LDA)

3.2. การวิเคราะห์ส่วนประกอบมุขสำคัญแบบสองมิติ

เป็นที่ทราบกันดีแล้วว่า การสกัดคุณลักษณะที่ใช้ในการรู้จำภาพในด้านต่างๆ นั้นมีเทคนิคที่น่าสนใจได้แก่ Principal Component Analysis (PCA) และ Linear Discriminant Analysis (LDA) ซึ่งมีการใช้กันอย่างแพร่ภาพในงานการรู้จำใบหน้าและเป้าหมายอัตโนมัติ (Face and Automatic Target Recognitions) เป็นที่ทราบกันดีว่าเทคนิคในการสกัดคุณลักษณะทั้งสองวิธีนั้นต่างก็มีข้อเสียอันเกิดจากการที่มิติของภาพที่จะทำการรู้จำมีขนาดมากกว่าจำนวนตัวอย่างในการฝึกฝนมากๆ ทำให้เกิดปัญหาที่เรียกว่า Small Sample Size (SSS) ปัจจุบันมีการเสนอเทคนิค Two-Dimensional Principal Component Analysis (2DPCA)

ซึ่งพัฒนาบนหลักการของ PCA ซึ่งสามารถแก้ปัญหาดังกล่าวได้ โดยนิยาม เมตริกซ์ความแปรปรวนร่วมเกี่ยว ขึ้นมาใหม่โดยไม่ต้องมีการแปลงรูปภาพจากสองมิติซึ่งเป็นเมตริกซ์ไปเป็นเวกเตอร์หนึ่งมิติ ทำให้ *เมตริกซ์ความแปรปรวนร่วมเกี่ยว* (covariance matrix) ที่ได้มีขนาดเล็กลงและเต็มขั้น ในทำนองเดียวกันพบว่า LDA ก็สามารถพัฒนาเป็นแบบสองมิติได้เช่นกัน ในที่นี้เราจึงนำเสนอเทคนิคขึ้นใหม่เรียกว่า Two-Dimensional Linear Discriminant Analysis (2DLDA) ขึ้นโดยใช้หลักการพื้นฐานของ 2DPCA โดยขั้นแรกได้ทำมาทดลองบนฐานข้อมูลใบหน้า 2 ฐานข้อมูล อันได้แก่ ORL face database ของ AT&T และ Yale face database ของมหาวิทยาลัย Yale ก่อนที่จะทำการทดลองกับการรู้จำเป้าหมายอัตโนมัติกับภาพเรดาร์ช่องเปิดสังเคราะห์ (Synthetic Aperture Radar: SAR) (ผลการทดลองของการรู้จำด้วยวิธี 2DPCA และ 2DLDA จะกล่าวถึงในบทย่อยการทดลอง ของวิธีนี้กับวิธี *ปริภูมิย่อยของใบหน้าเฉพาะ* (Face-Specific Subspace)) ทั้งนี้พบว่าประสิทธิภาพการรู้จำของวิธี 2DLDA ลดลง เนื่องจากคุณลักษณะที่สกัดได้มีข้อมูลที่ไม่เหมาะสมในการคัดแยกใบหน้ารวมอยู่ด้วย ตัวอย่างเช่น ฉากหลัง แสง (illumination) ทำให้ 2DLDA สกัดคุณลักษณะในส่วนฉากหลังและแสงดังกล่าวไปเป็นส่วนแยกแยะ ซึ่งเมื่อใบหน้าของบุคคลเดียวกันแต่อยู่บนฉากหลังต่างกันจะถูก 2DLDA จำแนกเป็นบุคคลละคนบุคคลกัน

เพื่อเป็นการแก้ไขและปรับปรุงการรู้จำใบหน้าและเป้าหมายอัตโนมัติของวิธี 2DLDA ให้มีความแม่นยำในการรู้จำสูงขึ้น เราจึงนำเสนอเทคนิคใหม่ที่มีโดยการประยุกต์ใช้ 2DPCA เพื่อสกัดเอา *ปริภูมิย่อยของใบ* (Face Subspace) ไว้เป็นคุณลักษณะสำคัญก่อนแล้วจึงนำปริภูมิย่อยที่ได้มาทำการสกัดคุณลักษณะเฉพาะของใบหน้าอีกครั้งหนึ่งด้วย 2DLDA ในที่นี้เราสามารถเรียก *ปริภูมิย่อยของใบหน้า* ที่สกัดด้วยวิธี 2DPCA อีกอย่างหนึ่งว่า "Principal Component Vector"

จากการเปรียบเทียบกันของวิธี 3 วิธี คือ

1. 2DPCA
2. 2DLDA และ
3. วิธีผสมระหว่าง 2DPCA และ 2DLDA หรือ 2DLDA of 2DPCA (Two-dimensional Linear Discriminant Analysis of two-dimensional Principal Vectors)

ซึ่งจากการวิเคราะห์ส่วนประกอบมุขสำคัญด้วยวิธีที่เสนอขึ้นนี้ ทำให้ประสิทธิภาพในการการรู้จำสูงขึ้น ทั้งนี้เนื่องมาจากเดิมระบบที่ใช้วิธี 2DPCA เพียงอย่างเดียวนั้นมีความเหมาะสมกับการพรรณนาหน้า (Face representation) และ การสร้างคืนหน้า (Face reconstruction) มากกว่าเพราะด้วยหลักการเดียวกันกับ PCA ซึ่งมีการนอร์มอลไลซ์ข้อมูลเพียงค่าวาเรียนซ์ของข้อมูลทั้งหมดเท่านั้น แต่ความสัมพันธ์ระหว่างคลาส (class) ของข้อมูลไม่ได้นำมาพิจารณาด้วย โดยทั่วไปแล้วการจำแนกคลาสข้อมูลขึ้นกับทั้งความสัมพันธ์ระหว่างคลาส (between-class) และภายในคลาส (within-class) ของข้อมูล ซึ่งวิธี LDA นั้นมีการพิจารณาในส่วนนี้ ดังนั้นการประยุกต์ใช้ PCA ร่วมกับ LDA จึงสามารถแก้ไขข้อบกพร่องในแต่ละวิธีได้ กล่าวคือ PCA สามารถขจัดข้อมูลที่มีผลทำให้ LDA แยกแยะผิดออกไป ในขณะเดียวกัน LDA ก็สามารถใช้ความสัมพันธ์ของข้อมูลทั้งระหว่างและภายในคลาสของข้อมูลได้ ส่งผลในประสิทธิภาพในการรู้จำของระบบรู้จำดีขึ้น เราได้นำการหลักการนี้มาประยุกต์ใช้ในการประยุกต์ใช้ 2DPCA และ 2DLDA เข้าด้วยกัน

โดยใช้ 2DPCA เป็นตัวสกัดคุณลักษณะให้กับ 2DLDA ทั้งนี้เหตุผลที่เราไม่ประยุกต์ใช้ LDA กับ 2DPCA โดยตรงนั้นเนื่องจากขนาดมิติของเมตริกซ์คุณลักษณะที่ได้มาจาก 2DPCA นั้นมีขนาดใหญ่อยู่เพราะ 2DPCA เป็นการแปลงเชิงเส้นเพียงแนวแถวเท่านั้น หากใช้ LDA ประยุกต์เข้าไปโดยตรงจะทำให้เกิดปัญหา SSS ได้ การประยุกต์ใช้ 2DLDA กับ 2DPCA ไม่เพียงแต่ลดขนาดมิติของเมตริกซ์คุณลักษณะซึ่งทำให้ใช้เวลาในการประมวลผลในการทดสอบลดลง และยังสามารถเพิ่มประสิทธิภาพในการรู้จำอีกด้วย ซึ่งจากการทดลองกับฐานข้อมูลใบหน้าคนมาตรฐานทั้งของ ORL และ Yale พบว่าสามารถเพิ่มประสิทธิภาพการรู้จำได้เมื่อเทียบกับทั้งวิธี 2DPCA เพียงอย่างเดียว และวิธีที่ใช้ 2DLDA เพียงอย่างเดียวในทุกกรณี

3.3. การวิเคราะห์ส่วนประกอบมุขสำคัญแบบสองมิติและวิธีปริภูมิย่อยของใบหน้าเฉพาะ

จากการศึกษา PCA พบว่าเราสามารถเพิ่มข้อมูลเกี่ยวกับคลาสข้อมูลได้ด้วยหลักการ *คลาสเฉพาะ* (class-specific) สำหรับการรู้จำใบหน้า หรือที่สามารถเรียกว่า ปริภูมิย่อยใบหน้าเฉพาะ (Face-Specific Subspace (FSS)) โดยที่ในการฝึกฝนนั้นจะแยกการฝึกฝนเป็นของแต่ละคลาสข้อมูล ไม่ฝึกฝนรวมหมดทุกคลาสข้อมูลพร้อมกันภายในครั้งเดียวตามแบบหลักการของ PCA เดิม ซึ่งเป็นประมาณค่า เมตริกซ์ความแปรปรวนร่วมเกี่ยว (covariance matrix)ของแต่ละเฉพาะคลาสข้อมูล จากเดิมที่มีเพียง เมตริกซ์ความแปรปรวนร่วมเกี่ยวรวมของทุกคลาสข้อมูลเพียงเมตริกซ์เดียว ทำให้ได้จำนวนปริภูมิย่อยและจำนวนการแปลงเชิงเส้นเท่ากับจำนวนคลาสข้อมูล ที่จริงการใช้หลักการคลาสเฉพาะในการรู้จำใบหน้านั้นเป็นการพรรณนาหน้าเฉพาะของใบหน้าแต่ละคลาส หรืออีกนัยหนึ่งเกี่ยวข้องกับ *วิธีการจำแนกคลาสด้วยการสร้าง* (generative-based classification method) ที่ใช้ในการหาฟังก์ชันการกระจายทางสถิติของข้อมูล (probability distribution function)

อนึ่งในการคำนวนหาค่าระยะห่าง (Distance) ของการค่าสัมประสิทธิ์ของ PCA ระหว่างคลาสข้อมูลสองคลาสนั้นไม่สามารถใช้วิธีการหาค่าระยะทางแบบเดิมได้ เพราะเป็นการแปลงไปบนปริภูมิย่อยคนละปริภูมิกันที่เกิดขึ้นของการแปลงเชิงเส้นของแต่ละคลาสข้อมูล ดังนั้นวิธีการแยกแยะโดยใช้ระยะห่างจึงต้องมีการปรับปรุงขึ้นใหม่โดยอาศัยหลักการที่ในการหาระยะทางของการสร้างภาพสร้างคืน (Reconstruction Image) โดยใช้ตัวแปลงเชิงเส้นของแต่ละคลาสข้อมูลในการหาค่าสัมประสิทธิ์ของ PCA แล้วนำค่าสัมประสิทธิ์นั้นมาสร้างภาพคืน ภาพที่สร้างคืนที่มีระยะห่างจากภาพต้นแบบน้อยที่สุด ควรอยู่ในคลาสข้อมูลของปริภูมิย่อยชุดนั้น

หลักการคลาสเฉพาะ (class-specific) นี้มีข้อดีที่สามารถเพิ่มความแม่นยำในการรู้จำของการรู้จำใบหน้าที่มีตัวอย่างของใบหน้าของแต่ละบุคคลจำนวนมาก ในกรณีที่มีจำนวนตัวอย่างน้อยเราสามารถใช้วิธีเพิ่มจำนวนตัวอย่างเพื่อให้การประมาณ เมตริกซ์ความแปรปรวนร่วมเกี่ยว มีความถูกต้องมากขึ้นได้

เมื่อนำหลักการ FSS ที่กล่าวมามาประยุกต์ใช้กับ 2DPCA นั้นพบว่าสามารถทำได้ แต่อย่างไรก็ตามระยะห่างที่ใช้ในวิธี PCA+FSS นั้นเป็นระหว่างเวกเตอร์สองเวกเตอร์ใดๆ แต่ข้อมูลบนปริภูมิย่อยของ 2DPCA มีลักษณะเป็นเมตริกซ์ จึงต้องใช้ระยะห่างต่างคลาสกัน จากการทดลองพบว่าระยะห่างที่นิยามด้วยสมการที่ (3.1) เป็นระยะห่างที่มีนัยทางการแยกแยะสูงที่สุด

$$d(\mathbf{B}_i, \mathbf{B}_j) = \sqrt{\sum_{m=1}^{q}\sum_{n=1}^{d}(\mathbf{b}_{(m,n)_i} - \mathbf{b}_{(m,n)_j})^2} \qquad (3.1)$$

โดย $\mathbf{b}_{(m,n)}$ เป็นอีลีเมนต์ที่ (m,n) ในเมตริกซ์ $\mathbf{B}$ และ d,q เป็นจำนวนแถวและหลักของเมตริกซ์คุณลักษณะตามลำดับ

จากนั้นได้ลองประยุกต์ใช้หลักการของ FSS กับ Bilateral-projection-based 2DPCA (B2DPCA) ซึ่งเป็น 2DPCA ที่มีการแปลงทั้งแนวและแถว

3.4. ผลการทดลอง

การทดลองการรู้จำใบหน้าที่ดำเนินการทดลองในโครงการนี้มีเพื่อสัตถุประสงค์เพื่อเทียบเคียงผลและประยุกต์ใช้ ในที่นี้เราทำการทดสอบกับฐานข้อมูลใบหน้า Yale AR และ ORL โดยทำการทดสอบ 2 ส่วน คือ การหาส่วนประกอบมุขสำคัญสองมิติ และ ส่วนประกอบมุขสำคัญสองมิติบนปริภูมิย่อยเฉพาะของใบหน้า โดยมีรายละเอียดดังนี้

3.4.1. ส่วนประกอบมุขสำคัญสองมิติ

ตารางที่ 7 และตารางที่ แสดงการเปรียบเทียบวิธีการ 2DPCA 2DLDA และ 2DPCA+2DLDA ณ จุดที่มีประสิทธิภาพการรู้จำมากที่สุดบนฐานข้อมูล Yale และ ORL โดยพบว่าเมตริกซ์คุณลักษณะของ 2DPCA มีขนาด 100 x 23 เมตริกซ์คุณลักษณะของ 2DLDA มีขนาด 100 x 14 และเมตริกซ์คุณลักษณะของ 2DPCA+2DLDA มีขนาด 100 x 7 ซึ่งวิธี 2DPCA+2DLDA มีมิติของเมตริกซ์คุณลักษณะขนาดเล็กที่สุด และใช้เวลาทั้งฝึกฝนและทดสอบน้อยที่สุดบนฐานข้อมูล Yale และยังให้ความถูกต้องในการรู้จำมากที่สุด ในขณะเดียวกัน 2DPCA+2DLDA ที่ทดสอบบนฐานข้อมูล ORL ก็ให้ความถูกต้องในการรู้จากที่สุดเช่นกัน โดยมีมิติของเมตริกซ์คุณลักษณะเท่ากับของ 2DPCA แต่ใช้เวลามากกว่า 2DPCA เพียงไม่มาก อัตราการรู้จำของ 2DLDA บนฐานข้อมูล ORL มีค่าน้อยลงตามสมมติฐานที่ว่า 2DLDA จะนำข้อมูลที่ไม่มีนัยในการแยกแยะกลุ่มมาใช้ในการแยกแยะ แต่เมื่อนำ 2DPCA มาเป็นตัวสกัดคุณลักษณะก่อนป้อนเข้า 2DLDA พบว่าสามารถเพิ่มอัตราการรู้จำได้ดีขึ้นจริง

ตารางที่ 7 การเปรียบเทียบความถูกต้องในการรู้จำของวิธีการ 2DPCA 2DLDA และ 2DPCA+2DLDA บนฐานข้อมูล Yale

วิธีการ	ความถูกต้อง (%)	d	q	มิติ	เวลา (วินาที)		
					ฝึกฝน	ทดสอบ	ทั้งหมด
2DPCA	87.78	23	-	100 x 23	0.17 (1)	3.14 (1)	3.31 (1)
2DLDA	88.89	-	14	100 x 14	0.16 (0.94)	1.94 (0.62)	2.10 (0.63)
2DPCA+2DLDA	90.00	21	7	100 x 7	0.24 (1.14)	1.01 (0.32)	1.25 (0.38)

ตารางที่ 8 การเปรียบเทียบความถูกต้องในการรู้จำของวิธีการ 2DPCA 2DLDA และ 2DPCA+2DLDA บนฐานข้อมูล ORL

วิธีการ	ความถูกต้อง (%)	*D*	*q*	มิติ	เวลา (วินาที)		
					ฝึกฝน	ทดสอบ	ทั้งหมด
2DPCA	91.50	5	-	112 x 5	0.33 (1)	4.33 (1)	4.69 (1)
2DLDA	90.50	-	3	112 x 3	0.50 (1.51)	2.75 (0.64)	3.28 (0.7)
2DPCA+2DLDA	93.50	14	5	112 x 5	0.64 (1.94)	4.33 (1)	5.22 (1.07)

3.4.2. ส่วนประกอบมุขสำคัญสองมิติบนปริภูมิย่อยเฉพาะของใบหน้า

ประยุกต์ใช้วิธีปริภูมิย่อยเฉพาะของใบหน้ากับ 2DPCA และ B2DPCA บนฐานข้อมูล Yale และ ARได้ผลดังตารางที่ 99 และ ตารางที่ 10 โดย d และ q คือจำนวนค่าเจาะจงที่ใช้ในแนวแถวและหลักตามลำดับ

ตารางที่ 9 การเปรียบเทียบความถูกต้องในการรู้จำของวิธีการ 2DPCA B2DPCA 2DPCA+FSS และ B2DPCA+FSS บนฐานข้อมูล Yale

วิธีการ	ความถูกต้อง (%)	*d*	*q*	มิติ
2DPCA	87.78	5	-	100 x 5
B2DPCA	92.22	23	22	22 x 23
2DPCA+FSS	92.22	5	-	100 x 5
B2DPCA+FSS	94.44	1	1	1 x 1

ตารางที่ 10 การเปรียบเทียบความถูกต้องในการรู้จำของวิธีการ 2DPCA B2DPCA 2DPCA+FSS และ B2DPCA+FSS บนฐานข้อมูล AR

วิธีการ	ความถูกต้อง (%)	*d*	*q*	มิติ
2DPCA	54.49	5	-	100 x 5
B2DPCA	55.45	23	22	22 x 23
2DPCA+FSS	56.41	5	-	100 x 5
B2DPCA+FSS	59.46	1	1	1 x 1

จากผลการทดลองพบว่าวิธี FSS สามารถเพิ่มประสิทธิภาพในการรู้จำของวิธีเดิมได้ โดย เมื่อประยุกต์ใช้กับ 2DPCA สามารถเพิ่มอัตราการรู้จำได้ถึงร้อยละ 4.44 และเมื่อนำมาประยุกต์ใช้กับ B2DPCA พบว่าอัตราการรู้จำสูงขึ้นร้อยละ 2.22 ซึ่งวิธีการ B2DPCA บน FSS เป็นวิธีการที่มีอัตราการรู้จำสูงสุดคือร้อยละ 94.44

บทที่ 4
สรุปและงานต่อไป

4.1. สรุป

โครงงานนี้นำเสนอการรู้จำเป้าหมายอัตโนมัติและรู้จำใบหน้า ในส่วนของการรู้จำเป้าหมายอัตโนมัติได้นำเสนอเทคนิคการจำแนกด้วยคณาจักรกลแบบเข้ารหัสช่องสัญญาณต่อกัน และใช้วิธีการประมาณการหมุน แทนที่การแปลงเชิงเส้นให้เสถียรต่อการหมุน นอกจากนี้ได้มีการศึกษาเพิ่มเติมเทคนิคที่ใช้ในการรู้จำใบหน้า ซึ่งเมื่อพัฒนาปรับปรุงแล้วจะสามารถนำมาใช้เพื่อเพิ่มประสิทธิภาพในการรู้จำเป้าหมายอัตโนมัติในอนาคต โดยเฉพาะคณาจักรกลสำหรับการรู้จำเป้าหมายอัตโนมัติและรู้จำใบหน้าซึ่งกำลังเป็นแนวทางที่มีประสิทธิภาพมากที่สุดในขณะนี้ โดยได้นำเสนอ 2DLDA of 2DPCA และ การพรรณาใบหน้าด้วย B2DFSS (Bilateral 2D Face Specific Representation) รายละอียดส่วนงานที่ได้ดำเนินการไปแล้วและส่วนที่จะดำเนินการต่อไปแสดงดังต่อไปนี้

4.2. ส่วนงานที่ได้ดำเนินการไปแล้ว

1. พัฒนาระบบจำแนกแบบหลายตัวบนฐานการเข้ารหัสต่อกัน (concatenation code-domain MCS) ซึ่งประกอบด้วยระบบจำแนกแบบหลายตัว 3 แบบ คือ วิธีฐานหลักดิสคริมิแนนต์ ระบบ Adaboost และ การเข้ารหัสช่องสัญญาณ ซึ่งได้มีการทดสอบกับการรู้จำเป้าหมายอัตโนมัติ (automatic target recognition) แล้ว
2. ศึกษาการวิเคราะห์ส่วนประกอบหลักแบบสองปริภูมิ (two dimension principal component analysis: 2DPCA) (เมทริกซ์) ซึ่งต่างจากวิธีเดิมที่เป็นคลาสปริภูมิเดียว (เวกเตอร์) ในขณะนี้ได้ศึกษาพบว่าในทางปฏิบัติ 2DPCA มีการประมาณค่า เมทริกซ์โคแวเรียนต์ (covariance matrix) ที่แม่นยำกว่า PCA เดิม
3. ศึกษาการการแปลงขั้ว (polar) และ แรดอน (radon transform) สำหรับฐานหลักดิสคริมิแนนต์ เพื่อใช้สร้างระบบพิกัดเสถียรสำหรับระบบจำแนกแบบหลายตัว โดยมีการร่วมใช้การแยกค่าเชิงเดียว (single value decomposition)
4. ศึกษาการประมาณการหมุน (pose estimation)

5. ศึกษาพื้นฐานทฤษฎีข้อมูลควอนตัม (quantum information theory) ถึงความเป็นไปได้ในการนำ การพรรณนาเกินบริบูรณ์ (overcomplete representation) และการเข้ารหัสควอนตัม (quantum error correcting code) เพื่อนำมาใช้กับระบบตัวจำแนกแบบหลายตัว ในขณะนี้ยังอยู่ในการศึกษาขั้นต้นเพื่อหาแนวทาง

4.3. ส่วนงานต่อไป

1. ศึกษาทฤษฎีเพื่อสนับสนุนว่า 2DPCA มีการประมาณค่า เมทริกซ์โคแวเรียนต์ (covariance matrix) ที่แม่นยำกว่า PCA เดิม เท่าที่ทราบยังไม่มีนักวิจัยอื่นได้ทำการพิสูจน์มาก่อน โดยจะศึกษาทฤษฎีเกี่ยวกับ แบบจำลองความถูกต้องเชิงประมาณที่เป็นไปได้ (probably approximately correct (PAC) model) เพื่อเปรียบเทียบจำนวนตัวอย่างที่จำเป็นต้องใช้สำหรับ การประมาณค่า เมทริกซ์โคแวเรียนต์ ของ 2DPCA และ PCA
2. พัฒนาวิธี 2DPCA แบบหลายระดับความละเอียด (multiresolution) ที่มีการนำ 2DPCA ที่กระทำตามแนวตั้ง แถว และ ทแยงมุม ของแบนด์ย่อยละเอียดทางแนวนอน แนวตั้ง และทแยงมุม (detail subbands: horizontal, vertical, and diagonal) รวมถึงการพัฒนาการพรรณนาเกินบริบูรณ์ที่เหมาะสำหรับการจำแนกแบบหลายตัว
3. พัฒนาเทคนิค 2DPCA กับวิธีอื่นๆที่ใช้ในการรู้จำใบหน้า เช่น วิธีแก่น (kernel method), null space, และ direct discriminant subspace เป็นต้น
4. พัฒนาวิธีการหาทิศ (pose) ของเป้าหมาย เพื่อใช้ในการรู้จำเป้าหมายอัตโนมัติ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

เอกสารอ้างอิง

1) T. Dietterich, "Machine learning research: Four current directions," AI Magazine, vol.18, no.4, pp.97--136, 1997.

2) A. Sharkey, Combining Artificial Neural Nets: Ensemble and Modular Multi-Net Systems, Springer Verlag, 1999.

3) R. Neal, "Probabilistic inference using Markov chain Monte Carlo methods," Tech. Rep., Dept. of Computer Science, University of Toronto, Tech. Rep. CRG-TR-93-1, 1993.

4) Y. Freund and R. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," Journal of Computer and System Sciences, 1995.

5) T. Dietterich and G. Bakiri, "Error-correcting output codes: A general method for improving multiclass inductive learning programs," Proc. of the Ninth AAAI, pp.572--577, 1991.

6) T. Ho, "The random subspace method for constructing decision forests," IEEE Trans. Pattern Anal. and Machine Intell., vol.20, no.8, pp.832--844, 1998.

7) C. Nikunj and K. Tumer, "Input decimation ensembles: Decorrelation through dimensionality reduction," in Lecture Notes in Computer Science, ed. J. Kittler and F. Roli, Multiple Classifier Systems, vol.2096, pp.238--247, Springer-Verlag, Berlin Heidelberg New York, 2001.

8) R. Nock, "Inducing interpretable voting classifiers without trading accuracy for simplicity: Theoretical results, approximation algorithm, and experiments," Journal of Aritificial Intelligent Research, pp.138--170, 2002.

9) R. Coifman and N. Saito, "Constructions of local orthonormal bases for classification and regression," C. R. Acad. Sci. Paris, vol.319 S'erie I, pp. 191--196, 1994.

10) R. Coifman and M. Mickerhauser, "Entropy-based algorithm for best basis selection," IEEE Trans. Inform. Theory, vol.38, no.2, pp.713--718, 1992.

11) N. Saito, Local Feature Extraction and Its Applications Using a Library of Bases, Ph.D. thesis, Dept. of Mathematics, Yale University, 1994.

12) J. O'Sullivan, M. DeVore, V. Kedia, and M. Miller, "SAR ATR performance using a conditionally Gaussian model," IEEE Trans. Aerospace and Electronic Systems, vol.37, no.1, pp.91--108, 2001.

13) C. Nilabul, O. Pham, R. Mersereau, M. Smith, and M. Clements, "Hidden Markov modeling for SAR automatic target recognition," ICASSP, pp.1061--1064, May 1998.

14) A.Y. Ng and M. Jordan, "On discriminative vs. generative classifiers: A comparison of logistic regression and naive Bayes," Advances in Neural Information Processing Systems, Cambridge, MIT Press, 2002.

15) Q. Zhao, D. Xu, and J. Principe, "Pose estimation for SAR automatic target recognition, Image Understanding Workshop '98, Pacific Grove, Canada, pp.827--831, 1998.

16) Q. Zhao and J. Principe, "Support vector machines for synthetic aperture radar automatic target recognition," Automatic target recognition IX, SPIE Conf., Orlando, FL, pp.101--107, 1999.

17) Q. Zhao, J.C. Principe, V. Brennan, D. Xu, and Z. Wang, "Synthetic aperture radar automatic target recognition with three strategies of learning and representation," Optical Engr., special issue on ATR, vol.39, no.5, pp.1230--1244, 2000.

18) E.P. Simoncelli, W.T. Freeman and D. Heeger, "Shiftable multiscale transforms," IEEE Trans. Inform. Theory, vol.38, no.2, pp.587--607, March 1992.

19) C.M. Pun and M.C. Lee, "Log-polar wavelet energy signatures for rotation and scale invariant texture classification," IEEE Trans. Pattern Anal. and Machine Intell., vol.25, no.5, pp.590--603, May 2003.

20) P.C. Teo, Theory and Applications of Steerable Functions, Ph.D. thesis, Stanford University, March 1998.

21) J. Wood, "Invariant pattern recognition: A review," Pattern Recognition, vol.29, no.1, pp.1--17, 1996.

22) E.P. Simoncelli, W.T. Freeman and D. Heeger, "Symmetries and discriminability in feedforward network architecture," IEEE Trans. Inform. Theory, vol.38, no.2, pp.816--826, March 1992.

23) O.K. Al-Shaykh and J.K. Doherty, "Invariant image analysis based on radon transform and SVD," IEEE Trans. Circuits and Systems, vol.43, no.2, pp.123--133, February 1996.

24) H.J. Dohse and A. Jain, "Object classification and registration by radon transform based invariants," Proc. of Int. Conf. on Acoustics, Speech, and Signal Processing, pp.225--228, April 1987.

25) E. Tapia, J. Gonzalez, and J. Villena, "A generalized class of boosting algorithms based on recursive decoding models," in Lecture Notes in Computer Science, ed. T. Windeatt and F. Roli, Multiple Classifier Systems, vol.2096, Springer-Verlag, Berlin Heidelberg New York, 2001.

26) S. Hashem, "Optimal linear combinations of neural networks," Neural Networks, vol.10, no.4, pp.599--614, 1997.

27) L. Sirovich and M. Kirby, Low-dimensional procedure for characterization of human faces," J. Optical Soc. Am., vol.4, pp.519-524, 1987.

28) M. Turk and A. Pentland, Eigenfaces for recognition," J. of Cognitive Neuroscience, vol.3, no.1, pp.71-86, 1991.

29) P.N. Belhumeur, J.P. Hespanha, and D.J. Kriegman, Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection," IEEE Trans. Pattern Anal. and Mach. Intell., vol.19, pp.711-720, July 1997.

30) W. Zhao, R. Chellappa, and A. Krishnaswamy, Discriminant analysis of principle components for face recognition," Japan, IEEE 3rd Inter. Conf. on Automatic Face and Gesture Recognition, 1998.

31) W. Zhao, R. Chellappa, and N. Nandhakumar, Empirical performance analysis of linear discriminant classifiers," CVPR, pp.164-171, IEEE Computer Society, 1998.

32) K. Fukunaga, Introduction to Statistical Pattern Recognition, second ed., Academic Press, 1990.

33) L. Chen, H. Liao, M. Ko, J. Lin, and G. Yu, A new LDA based face recognition system which can solve the small sample size problem," Pattern Recognition, vol.33, no.10, pp.1713-1726, 2000.

34) J. Lu, K.N. Plataniotis, and A.N. Venetsanopoulos, Regularized discriminant analysis for the small sample size problem in face recognition," Pattern Recogn. Lett., vol.24, no.16, pp.3079-3087, 2003.

35) R. Huang, Q. Liu, H. Lu, and S. Ma, Solving the small sample size problem of LDA," Pattern Recognition, vol.3, pp.29-32, 2002.

36) J. Yang, D. Zhang, A.F. Frangi, and J. yu Yang, Two-dimensional PCA: A new approach to appearance-based face representation and recognition," IEEE Trans. Pattern Anal. and Mach. Intell., vol.26, pp.131{137, Jan. 2004.

37) J. Ye, R. Janardan, and Q. Li, Two-dimensional linear discriminant analysis," in Advances in Neural Information Processing Systems 17, ed. L.K. Saul, Y. Weiss, and L. Bottou, pp.1569{1576, MIT Press, Cambridge, MA, 2005.

38) J. Yang and J.Y. Yang, From image vector to matrix: A straightforward image projection technique IMPCA vs. PCA," Pattern Recognition, vol.35, no.9, pp.1997-1999, 2002.

39) Yale, The Yale face database," 1997. Available from http://cvc.yale.edu/projects/yalefaces/yalefaces.html.

40) ORL, The ORL face database at the AT&T (Olivetti) research laboratory," 1992. Available from http://www.cl.cam.ac.uk/Research/DTG/attarchive/facedatabase.html.

41) D. Zhang and Z.H. Zhou, $(2D)^2$PCA: 2-directional 2-dimensional PCA for efficient face representation and recognition," Neurocomputing, 2005.

42) H. Kong, X. Li, L. Wang, E.K. Teoh, J.G. Wang, and R. Venkateswarlu, Generalized 2D principal component analysis," IEEE International Joint Conference on Neural Networks (IJCNN), 2005.

43) S. Shan, W. Gao, and D. Zhao, Face recognition based on face-specific subspace," International Journal of Imaging Systems and Technology, vol.13

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย