

รายงานการศึกษา

เรื่อง

Articulated Part-based Model for Joint Object Detection and Pose Estimation

นาย พชร ชูเกียรติจิกร

รหัสนักศึกษา 560610556

รายงานนี้เป็นส่วนหนึ่งของวิชา 261458 (Computer Vision)

ภาควิชาวิศวกรรมคอมพิวเตอร์

คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเชียงใหม่

ภาคเรียนที่ 1 ปีการศึกษา 2559

## คำนำ

รายงานนี้เป็นส่วนหนึ่งของวิชา 261458 (Computer Vision) ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะ  
วิศวกรรมศาสตร์ มหาวิทยาลัยเชียงใหม่ เพื่อให้ได้ศึกษาความรู้ และเทคนิคทาง computer vision จากตัวอย่าง  
งานวิจัยในเรื่อง Articulated Part-based Model for Joint Object Detection and Pose Estimation

ผู้จัดทำมีความคาดหวังว่า รายงานเล่มนี้จะเป็นประโยชน์ต่อผู้อ่าน หรือผู้ที่สนใจในการศึกษา หาข้อมูล  
ในเรื่อง Articulated Part-based Model for Joint Object Detection and Pose Estimation รวมทั้งสามารถช่วยให้  
เข้าใจถึงความรู้ในเรื่องนี้ได้เป็นอย่างดี หากมีข้อเสนอแนะ หรือข้อผิดพลาดประการใด ผู้จัดทำขอน้อมรับไว้ และ  
ขออภัยมา ณ ที่นี้ด้วย

ผู้จัดทำ

นาย พชร ชูเกียรติขจร

## สารบัญ

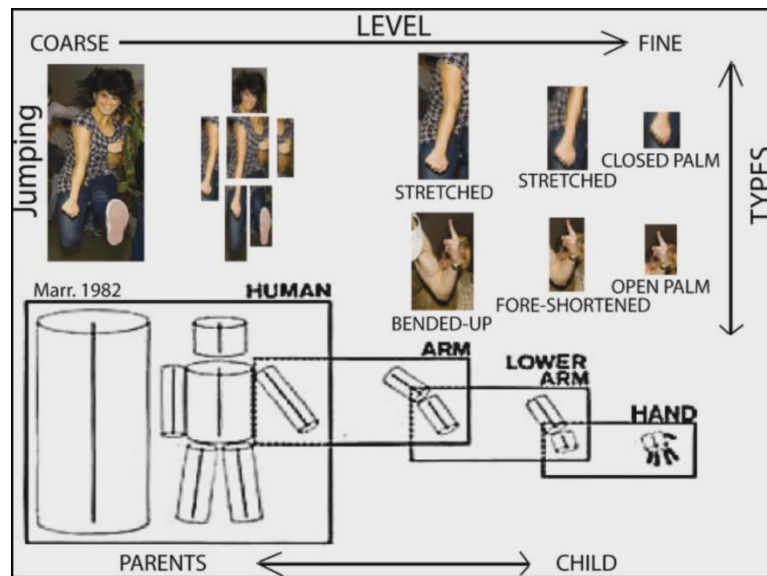
	หน้า
Articulated Part-based Model for Joint Object Detection and Pose Estimation	1-4
ตัวอย่างขั้นตอนวิธี APM	5
การทดลอง	6
ภาพผนวก	7

## Articulated Part-based Model for Joint Object Detection and Pose Estimation

(วิธีการแยกหาส่วนต่างๆของภาพ สำหรับการตรวจหาวัตถุร่วม และการคำนวณค่าตำแหน่งเชิง 3 มิติ)

วิธีการนี้ถูกคิดค้นขึ้นเพื่อแก้ปัญหาในการตรวจหาลักษณะ หรือท่าทางของสิ่งที่อยู่ในภาพ ที่การตรวจหานั้นยังไม่มีคามแม่นยำเพียงพอ และค่อนข้างจะเฉพาะเจาะจง

โดยวิธีการนี้จะทำการตรวจหาภาพด้วยโดยการแยกตรวจหาเพื่อตรวจสอบสิ่งที่สนใจ โดยจะตรวจหาลักษณะโดยรวมก่อน หรือหาภาพใหญ่ก่อน แล้วค่อยหาส่วนย่อยของส่วนนั้นๆในอีกทีหนึ่ง และหาข้อย่อยไปเรื่อยๆตามความเหมาะสม เพื่อความละเอียดในการตรวจหา

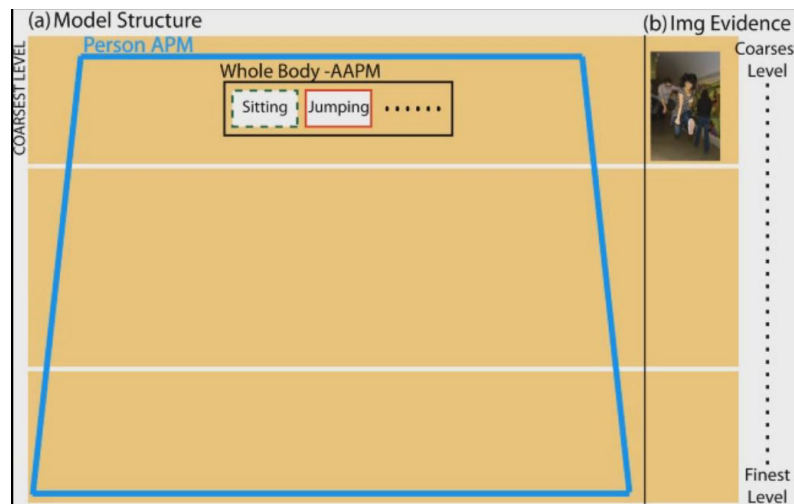


รูป 1

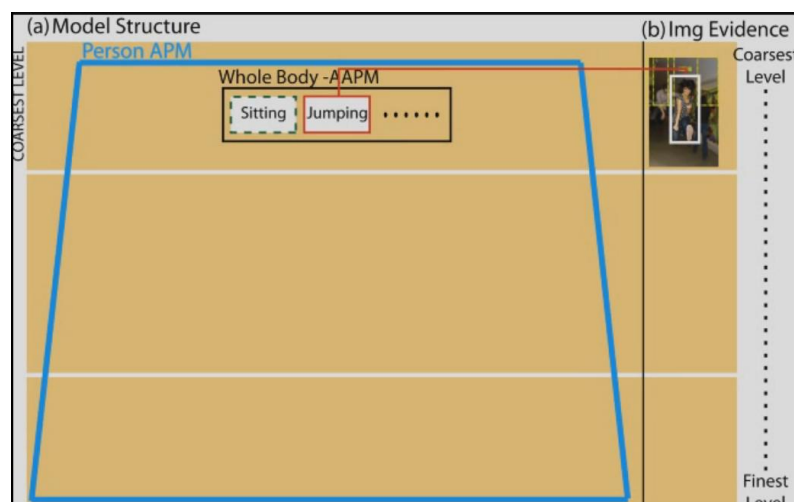
ในภาพมีรูปคนกระโดดด้วยท่าทางหนึ่งอยู่ ซึ่งวิธีนี้จะเริ่มหาจากลักษณะของภาพส่วนใหญ่มาก่อน แล้วแบ่งส่วนลงไปหาลักษณะของภาพส่วนย่อย แล้วหาลักษณะของภาพส่วนย่อยของแต่ละส่วนอีกที เช่นจากรูปนี้ การหาในขั้นแรกคือ การตรวจหาว่าสิ่งที่สนใจคืออะไร ซึ่งในรูปนี้คือ คนที่อยู่กลางภาพ จากนั้นก็ตรวจหาว่า คนที่ตรวจหาอยู่นั้นกำลังทำอะไร เช่น ยืน นั่ง นอน เป็นต้น ในรูปนี้คือ กระโดด ถ้าหาเพียงเท่านี้ความแม่นยำอาจไม่เพียงพอ เพื่อยืนยันสิ่งที่สนใจในภาพอย่างแม่นยำเพิ่มขึ้น ก็ต้องหาอย่างละเอียดมากขึ้น โดยการหาลักษณะของภาพย่อยที่แบ่งออกมาจากภาพหลักว่าภาพส่วนย่อยนั้นๆ มีส่วนประกอบ และลักษณะเป็นอย่างไรบ้าง อย่างเช่นขั้นตอนต่อมาของรูปข้างต้น ภาพหลัก(รูปคนกระโดด) ถูกแบ่งส่วนออกตามจำนวนลักษณะเด่นชัด

ของคนทั้งหมด 6 ส่วน คือ หัว, ช่วงตัว, แขนซ้าย-ขวา, ขาซ้าย-ขวา ซึ่งแต่ละส่วนที่ถูกแบ่งออกมานั้นก็จะถูกนำไปหาลักษณะของส่วนนั้นๆอีกที เช่น ภาพแขน ก็จะถูกแยกหาเป็น แขนช่วงบนอยู่ลักษณะไหน แขนช่วงล่างอยู่ลักษณะไหน และส่วนมืออยู่ลักษณะไหน เป็นต้น ซึ่งกำหนดภาพส่วนย่อยจะใช้วิธี Structured Support Vector Machine (SSVM) ในการจำแนกส่วนของภาพหลักเพื่อหาภาพส่วนย่อยตามที่ต้องการ

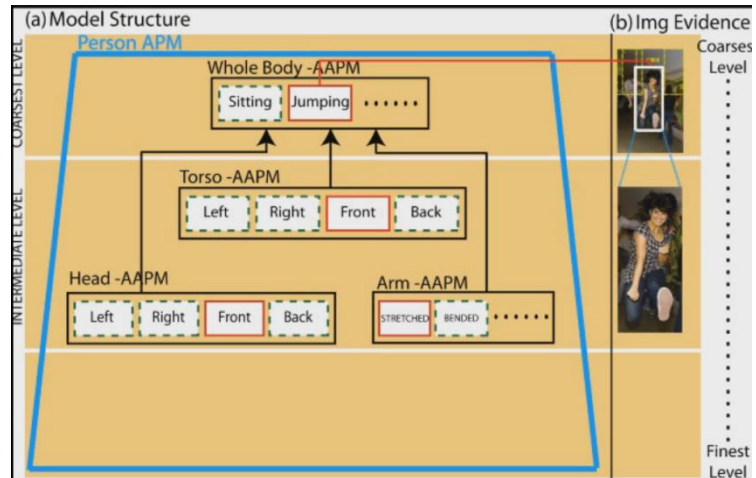
-วิธีที่กล่าวมาข้างต้น คือ APM (Articulated Part-based Model) ซึ่งมีวิธีการตามภาพดังนี้



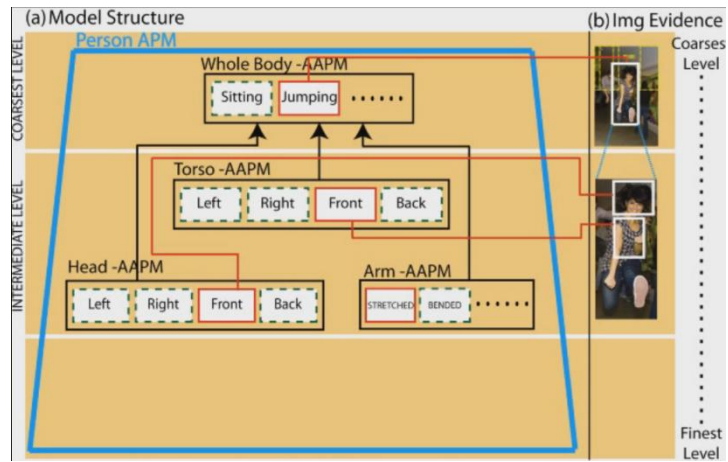
รูป 2 ตรวจสอบว่าเป็นสิ่งใด คนหรือวัตถุอื่นๆ



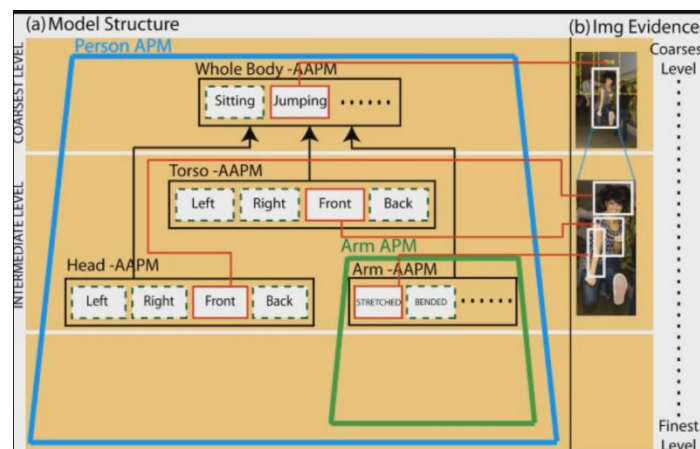
รูป 3 ตรวจสอบว่าเป็นคนที่กำลังกระโดด จากวิธีการAAPM



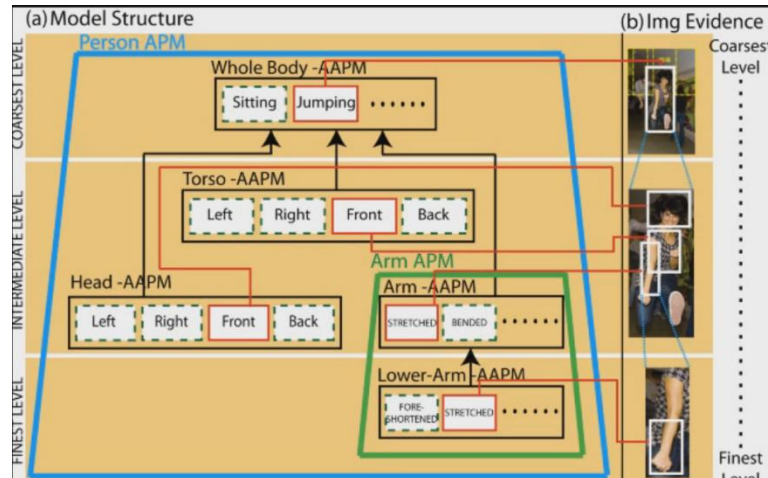
รูป 4 ตรวจสอบลักษณะของส่วนย่อยอื่น(หัว, ลำตัว, แขน)ด้วยวิธีAPM



รูป 5 ได้ลักษณะของส่วนย่อยอื่น(หัว, ลำตัว, แขน)ด้วยวิธีAPM



รูป 6 ตรวจสอบลักษณะของส่วนย่อยของส่วนย่อยอีกที(ช่วงล่างของแขน)ด้วยวิธีAPM



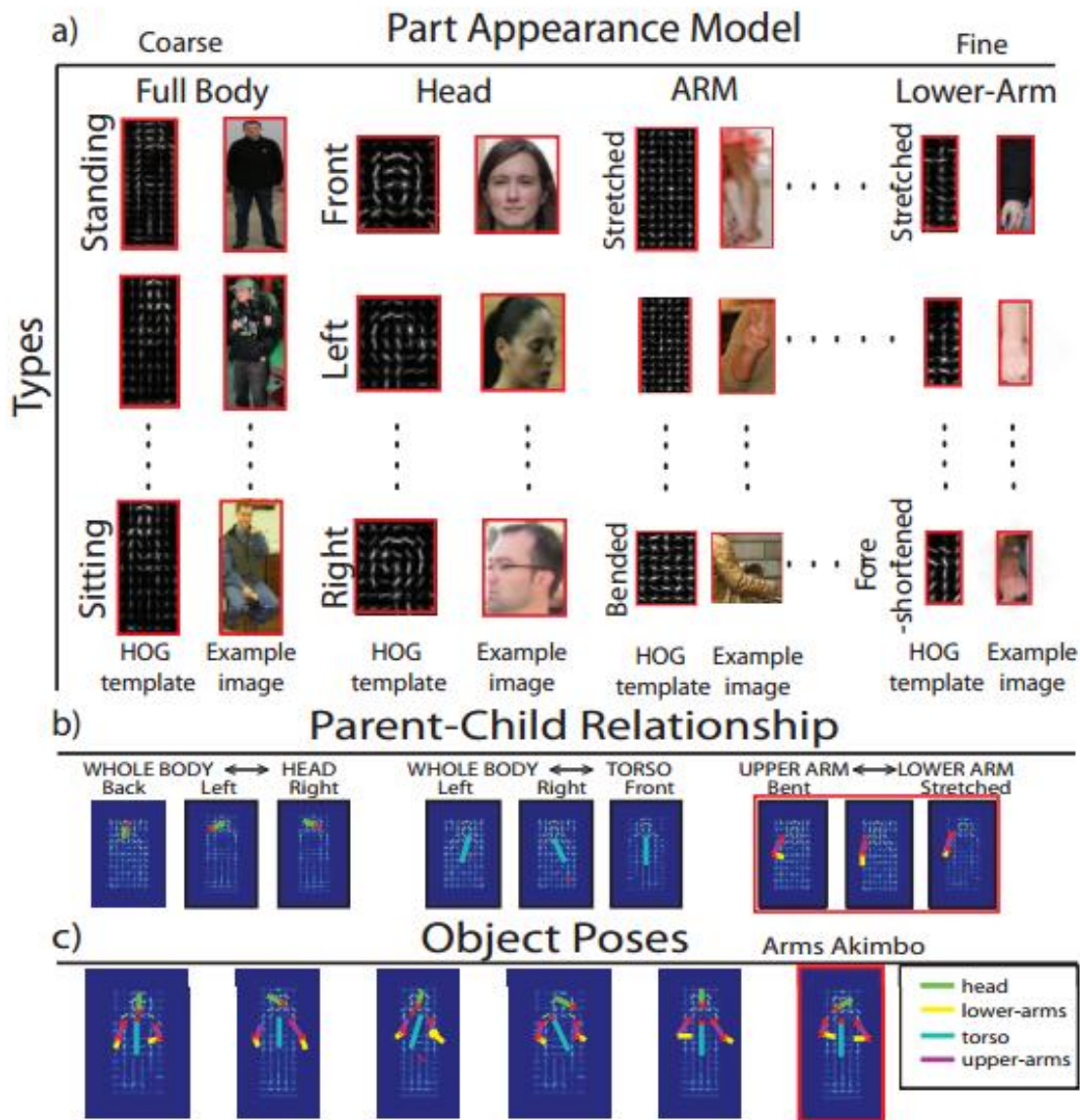
รูป 7 ได้ลักษณะของส่วนย่อยทั้งหมด

โครงสร้างของ APM จะแบ่งเป็น Number of parts (จำนวนของส่วนย่อยที่แบ่งส่วนภาพหลักออก เช่น ภาพคนก็จะแบ่งออกเป็น คนทั้งตัว, หัว, ลำตัว, แขน, ขา เป็นต้น), Part-types (ชนิดของภาพแต่ละส่วนที่ถูกแบ่งออก เช่น หัวก็จะแบ่งชนิดได้เป็น หัวด้านหน้า, หัวด้านข้าง, หัวด้านหลัง เป็นต้น) และ Parents-Child relationships (ความเกี่ยวข้องกันของแต่ละส่วนที่ถูกแบ่งออกกับส่วนหลักของภาพ) ที่อาจแสดงถึงการเคลื่อนไหวของวัตถุในภาพ โดยมีจุดมุ่งหมายที่จะเรียนรู้การรวมรูปแบบลักษณะที่ปรากฏ ของทุกๆ ชนิดส่วนและความสัมพันธ์ทางกายภาพระหว่างส่วนหลัก และส่วนย่อย

การรู้จำหรือ Recognition สำหรับ APM มีจุดประสงค์เพื่อกำหนดส่วนที่ดีที่สุดของภาพ โดยใช้ประโยชน์จากการทำซ้ำของโครงสร้าง APM ที่มีการใช้ top-down search strategy สำหรับการสำรวจสมมติฐานองค์ประกอบของลักษณะท่าทางที่เป็นไปได้ทั้งหมดของภาพ พร้อมกับใช้ Bottom-up matching score assignment สำหรับการให้ Matching score เพื่อช่วยเลือกสมมติฐานองค์ประกอบของลักษณะท่าทางที่ดีที่สุด

คุณสมบัติของ APM ที่ดีกว่าวิธีอื่นก็คือ APM รองรับทั้งการตรวจหาตั้งแต่ระดับแบบหยาดจนถึงแบบละเอียด, รองรับการแยกชนิดส่วนที่หลากหลาย, รองรับความซับซ้อนของแบบจำลองที่อาจเพิ่มขึ้นตามจำนวนขององค์ประกอบหรือลักษณะท่าทางในภาพ และได้ข้อสรุปแบบธรรมดาที่แน่นอน

## ตัวอย่างขั้นตอนวิธี APM

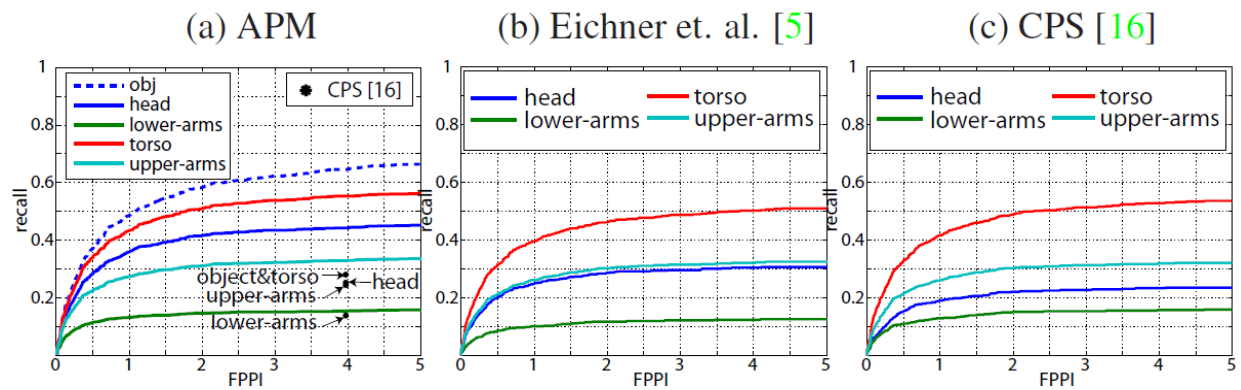


รูป 8 (a)แสดงการใช้ Histogram of Oriented-Gradient (HOG) ที่สอดคล้องกับภาพตัวอย่างสำหรับส่วนย่อยในแต่ละชนิด (b)แสดง Parent-Child geometric relationships ของแบบจำลอง โดยในส่วนย่อยที่ต่างกันจะใช้สีของ stick ที่ใช้แทนส่วนย่อยนั้นๆ ที่ต่างกัน (c)แสดงตัวอย่างลักษณะท่าทางของวัตถุที่แตกต่างกันของชนิดส่วนที่ได้จากวิธี APM

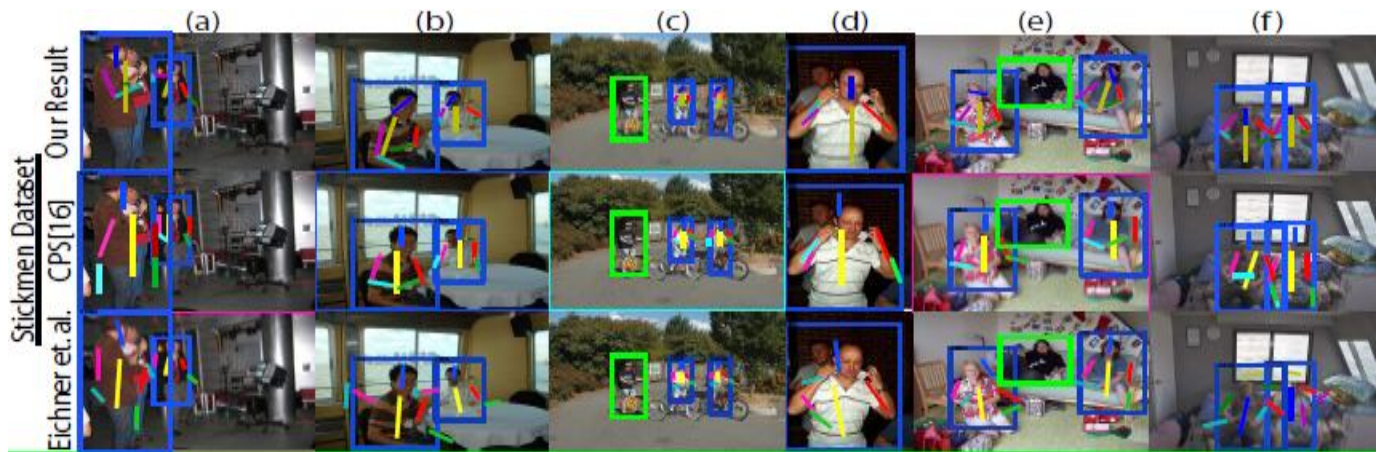


## การทดลอง

ในทดลองจะมีการคำนวณเปรียบเทียบประสิทธิภาพระหว่างวิธี APM และวิธีอื่นๆ โดยจะเทียบกับ Recall vs False Positive Per Image (FPPI) ดังรูปที่ 9 ซึ่งจะสังเกตเห็นได้ชัดว่าในช่วงค่า FPPI เดียวกันนั้นวิธี APM จะมีความแม่นยำในการตรวจพบมากกว่า หรือตรวจพบสิ่งที่ต้องการสูงกว่าวิธีอื่น ดูได้จากแกน recall ในกราฟ



รูปที่ 9 การเปรียบเทียบด้วย Recall vs False Positive Per Image (FPPI) ระหว่างวิธี (a)APM และวิธีอื่น(b,c)



รูปที่ 10 ผลลัพธ์ในการใช้งานวิธี APM และวิธีอื่นๆ

## ภาคผนวก

ไฟล์งานวิจัยที่ศึกษา

<https://drive.google.com/file/d/0B9oWgOT0D6NVdlk5Yy1HTEFlczA/view?usp=sharing>