รายงานการวิจัย

เรื่อง

การค้นคืนความหมายภาพด้วยโครงสร้างกราฟลำดับขั้นบนพื้นฐาน WordNet

Deep Retrieval Similar Semantic Images with Hierarchical Graph base on WordNet

โดย

นัศพ์ชาณัณ ชินปัญช์ธนะ

มหาวิทยาลัยธุรกิจบัณฑิตย์

รายงานผลการวิจัยนี้ได้รับทุนอุดหนุนจากมหาวิทยาลัยธุรกิจบัณฑิตย์

พ.ศ. 2559

DPU

ชื่อเรื่อง : การคืนค้นคืนความหมายภาพด้วยโครงสร้างกราฟลำดับขั้นบนพื้นฐาน WordNet
ผู้วิจัย : นัศพ์ชาณัฐ ชินปัญช์ธนะ **สถาบัน** : มหาวิทยาลัยธุรกิจบัณฑิตย์
ปีที่พิมพ์ : พุทธศักราช 2559 **สถานที่พิมพ์** : มหาวิทยาลัยธุรกิจบัณฑิตย์
แหล่งที่เก็บรายงานการวิจัยฉบับสมบูรณ์ : ศูนย์วิจัยมหาวิทยาลัยธุรกิจบัณฑิตย์
จำนวนหน้าวิจัย : 60 หน้า **ลิขสิทธิ์** : มหาวิทยาลัยธุรกิจบัณฑิตย์
คำสำคัญ : การประมวลผลภาพ, กราฟ, ภาพความหมาย, WordNet

บทคัดย่อ

การค้นคืนความหมายภาพเป็นหัวข้องานวิจัยที่ยังคงท้าทายอย่างมากในสาขาการประมวลผลภาพ มีนักวิจัยหลายกลุ่มพยายามปรบปรุงวิธีการเพื่อแก้ไขปัญหาของการแทนความหมายภาพด้วยโมเดลที่ซับซ้อน และพยายามนำรูปแบบการแทนคำศัพท์ระดับสูงเข้ามาใช้ในการแท็ก ทำให้โมเดลมีความซับซ้อนแต่อย่างไรก็ตามยังไม่สามารถแทนความหมายของภาพได้ ดังนั้นในงานวิจัยนี้จึงนำรูปแบบของ WordNet เพื่อแทนความหมายภาพในรูปแบบของกราฟลำดับขั้น โดยจะทำการสร้างความสัมพันธ์ระหว่างวัตถุบนภาพ จะประกอบด้วยขั้นตอนดังต่อไปนี้ (1) ขั้นตอนการเตรียมข้อมูล (2) ขั้นตอนการประมวลผล (3) ขั้นตอนการวัดประสิทธิภาพความคล้ายกันของภาพ จากการทดลองสามารถเปรียบเทียบความคล้ายกันได้ค่าความถูกต้องสูงถึง 81.2%

Title : Deep Retrieval Similar Semantic Images with Hierachical Graph base on WordNet

Researcher : Nutchanun Chinpanthana **Institution** : Dhurakijpundit University.

Year of Publication : 2016 **Publisher** : Dhurakijpundit University.

Sources : Dhurakijpundit University Research Center.

Number of Pages : 60 Pages **Copyright** : Dhurakijpundit University.

Keyword : Image processing, Graph, Semantic images, WordNet

Abstract

Semantic image retrieval is an active problem in multimedia personal photo collections. Many researchers have attempted to improve semantic models such as developing more sophisticated models, or generating intermediate representations by high-level concept based on keyword annotation. The model is rather rudimentary and it does not specific enough for representing the meaning of images. Therefore, we are using WordNet concept based on hierarchical graph to represent the semantic meaning. This concept is formulates on a graph which is captured the relationships among objects in the images. The approach is composed of three main phases: (1) data preprocessing (2) data processing and (3) evaluation of similarity measurement. The experimental results indicate that our proposed approach offers significant performance improvements in the interpretation of semantic images, compared, with the maximum of 81.2% accuracy.

กิตติกรรมประกาศ

งานวิจัยชิ้นนี้สำเร็จลุล่วงได้ทั้งนี้เพราะได้รับความอนุเคราะห์ การสนับสนุน และแรงผลักดัน และอีกหลายฝ่ายที่ได้ให้ความช่วยเหลือเอื้อเฟื้อข้อมูลที่เป็นประโยชน์ในการทดลองโปรแกรม บันทึกผลการทดลอง และงานวิจัยนี้จะไม่สมบูรณ์ได้ หากไม่ได้ข้อเสนอแนะที่เป็นประโยชน์จากท่านอาจารย์ผู้ทรงคุณวุฒิ ผู้ซึ่งให้คำแนะนำที่ดีและมุมมองที่ผู้วิจัยได้นำมาปรับปรุงให้สมบูรณ์ยิ่งขึ้น

นอกจากนี้ใคร่ขอขอบคุณมหาวิทยาลัยธุรกิจบัณฑิตย์ ผู้ให้ทุนสนับสนุนงานวิจัยชิ้นนี้ หากงานวิจัยเล่มนี้มีข้อผิดพลาด ประการใดขออภัยไว้ ณ ที่นี้ด้วย

นัศพ์ชาณัณ ชินปัญช์ธนะ

เมษายน 2560

DPU

สารบัญ

	หน้า
บทคัดย่อภาษาไทย	ก
บทคัดย่อภาษาอังกฤษ	ข
กิตติกรรมประกาศ	ค
สารบัญ	(1)
สารบัญตาราง	(3)
สารบัญรูปภาพ	(4)
บทที่ 1 บทนำ	1
1.1 ความเป็นมาของปัญหา	1
1.2 วัตถุประสงค์	3
1.3 สมมติฐาน	3
1.4 ขอบเขตของการวิจัย	3
1.5 ประโยชน์ที่คาดว่าจะได้รับ	4
1.6 นิยามคำศัพท์	4
บทที่ 2 ทฤษฎีและงานวิจัย	5
2.1 ทฤษฎีเบื้องต้น	5
2.2 การเปรียบเทียบความคล้ายของกราฟ	14
2.3 การวัดประสิทธิภาพ	16
บทที่ 3 ขั้นตอนวิธีการดำเนินงานวิจัย	18
3.1 ขั้นตอนการเตรียมข้อมูล	18
3.2 ขั้นตอนการประมวลผล	37
3.3 ขั้นตอนการวัดประสิทธิภาพ	40
บทที่ 4 ผลการทดลอง	41
4.1 การกำหนดข้อมูลภาพ	41
4.2 ผลการทดลอง	44

DPU

บทที่ 5 สรุปผลการทดลอง 49
5.1 สรุปและอภิปรายผลการวิจัย 49
5.2 ข้อเสนอแนะ 50

บรรณานุกรม 52
ประวัติผู้วิจัย 56

DPU

สารบัญตาราง

ตารางที่หน้า

2.1 การวัดประสิทธิภาพ 16

3.1 แนวคิด Lexical เมตริกซ์ 21

3.2 ตัวอย่างความสัมพันธ์ของคำคุณศัพท์ 25

4.1 ผลลัพธ์การเปรียบเทียบความคล้ายกันของความหมายภาพ 45

DPU

สารบัญรูปภาพ

รูปภาพที่ หน้า

2.1 การประมวลผลแบบค้นคืนด้วยคุณลักษณะฟีเจอร์ระดับต่ำ 6

2.2 ตัวอย่างการประมวลผลแบบค้นคืนด้วยสีและลวดลาย 7

2.3 ตัวอย่างโครงสร้างแท็กคำหลักของภาพ 9

2.4 ตัวอย่างระดับคำอธิบายภาพ 10

2.5 ตัวอย่างงานวิจัยที่ใช้คำศัพท์แท็กลงในภาพ 11

2.6 ตัวอย่างแนวคิดเครือข่ายความสัมพันธ์ระหว่างคำอธิบายความหมายบนภาพ 12

3.1 ขั้นตอนการวัดความคล้ายกันของกลุ่มความหมายภาพ 18

3.2 ตัวอย่างที่ถูกแท็กด้วยคำหลัก 20

3.3 ตัวอย่างจำลอง Synset ในอนุกรมวิธาน WordNet 23

3.4 รายการคำหลักเริ่มต้นบน WordNet 23

3.5 รายการแสดงความสัมพันธ์ระหว่างหัวข้อหลักบน WordNet 24

3.6 ตัวอย่างการแสดงความสัมพันธ์ของคำหลักที่มีความหมายเหมือนและต่างกัน 25

3.7 ตัวอย่างรูปแบบ 2 Entailment 26

3.8 ตัวอย่างรูปแบบของ 3 Entailment 27

3.9 ตัวอย่างรูปแบบของ 4 Entailment 27

3.10 โปรแกรม LabelMe บนเบราว์เซอร์ 29

3.11 การเลือกสัดส่วนของวัตถุบนภาพ 29

3.12 ตัวอย่างวัตถุที่ถูกแท็กด้วยโปรแกรม LabelMe 30

DPU

สารบัญรูปภาพ

รูปภาพที่ หน้า

3.12 ตัวอย่างภาพที่ถูกแท็กคำหลักบนภาพ ด้วยโปรแกรม LabelMe 30

3.13 ตัวอย่างภาพที่ถูกแท็กคำหลักบนภาพ ด้วยโปรแกรม LabelMe 30

3.14 ตัวอย่างของคำหลักที่พบบ่อยในการให้ความหมายวัตถุ 31

3.15 ตัวอย่างของคำหลักในการให้ความหมายวัตถุบนภาพ 32

3.16 ตัวอย่างความสัมพันธ์ของวัตถุด้วยแนวคิดกราฟ 33

3.17 ตัวอย่างการแทนความสัมพันธ์ตามแนวคิดกราฟ 36

3.18 ตัวอย่างความสัมพันธ์กราฟด้วยตามระดับและเส้นเชื่อม 37

4.1 ตัวอย่างกลุ่มภาพทำงานในสำนักงาน และภาพครอบครัว 43

4.2 ตัวอย่างกลุ่มภาพพักผ่อน และ ภาพพิธีการ 44

4.3 กราฟแสดงการเปรียบเทียบความคล้ายกันด้วยค่าความถูกต้องระหว่างกลุ่มภาพ 46

4.4 ตัวอย่างผลลัพธ์ของการเปรียบเทียบความคล้ายกันของความหมายภาพ 47

DPU

บทที่ 1
บทนำ

1.1 ความเป็นมาของปัญหา

การประมวลผลภาพวิวัฒนาการของเครื่องมือและอุปกรณ์อิเล็คทรอนิคสำหรับการถ่ายภาพดิจิทัลได้พัฒนาอย่างรวดเร็วจนทำให้ภาพถ่ายภาพดิจิทัลมีจำนวนมากมาย และแทรกอยู่ทั่วไปไม่ว่าจะอยู่ใน สังคมออนไลน์อย่างเฟสบุ้ค (Facebook), กูเกิ้ล (Google) หรือ อินสตารแกรม (Instagram) เป็นต้น ทำให้เกิดปัญหาที่ตามมาก็คือการจัดเก็บข้อมูลภาพที่เพิ่มขึ้น การจัดเก็บอย่างไรให้มีระบบที่ดีสามารถสืบค้นข้อมูลภาพได้อย่างง่ายและจำแนกข้อมูลภาพให้ตรงตามความหมายของภาพที่ต้องการของผู้ใช้มากที่สุด ทำให้งานวิจัยในปัจจุบันที่เกี่ยวข้องกับการค้นคืนรวมทั้งการจัดกลุ่มภาพให้ตรงกับความต้องการได้รับความสนใจจากนักวิจัยหลายกลุ่มเป็นงานด้านการประมวลผลภาพ (image processing) ด้านการค้นคืนสารสนเทศ (image retrieval) และการจำแนกประเภทข้อมูลภาพ (image classification) เพื่อคัดเลือกภาพ ให้ตรงตามความต้องการของผู้ใช้งาน สำหรับงานวิจัยทางการประมวลผลภาพในการค้นคืนสารสนเทศ จะมีการค้นคืนตามคุณลักษณะพื้นฐานของภาพที่ถูกสกัดคุณลักษณด้วยอัลกอริทึมต่างๆ เช่น สี (color) ลวดลาย (texture) รูปทรง (shape) เป็นต้น [Smeulders, A.W.M, 2000] [Ying Liua 2007] [Kevin Lin 2015] งานวิจัยของ SIMPLIcity [Jia Li,2003] เป็นการใช้คุณลักษณะระดับต่ำด้วยสี ลวดลาย และตำแหน่งของพื้นที่ของภาพ จากผลลัพธ์จะสังเกตว่าผลลัพธ์ของภาพเป็นภาพที่มีโทนสีคล้ายกันเป็นหลัก แต่มีลักษณะวัตถุที่แตกต่างกันอย่างเห็นได้ชัดเจนการประมวลผลภาพระดับต่ำนั้นค่อนข้างยากที่จะจัดให้หมวดหมู่เดียวกัน แต่มีกลุ่มนักวิจัยที่พยายามปรับปรุงอัลกอริทึม เพื่อทำการค้นคืนภาพที่มีลักษณะฟีเจอร์ที่ใกล้เคียงกับภาพที่ต้องการมาก การปรับปรุงเทคนิค ด้วยการนำวิธีการมาผสมผสานกันระหว่างคุณลักษณะเพื่อให้สามารถวิเคราะห์ในรูปแบบที่ซับซ้อน แต่อย่างไรก็ตามในการค้นคืนที่ใช้คุณลักษณะฟีเจอร์ระดับต่ำเพียงอย่างเดียว ทำให้ได้ผลลัพธ์ส่วนใหญ่ตรงกับคุณลักษณะของฟีเจอร์ที่สกัดมาแต่ไม่ได้ตรงกับความหมายที่เกิดภายในภาพที่ต้องการอย่างแท้จริง

งานวิจัยอีกกลุ่มที่พยายามจะใช้เทคนิคของการเข้าใจความหมายของภาพแทน การสืบค้นแบบข้างต้น งานวิจัยในกลุ่มนี้พยายามที่จะมองข้อมูลบนภาพเป็นวัตถุ (object) [Galleguillos C., 2010] ที่มีความหมายและแทนวัตถุนั้นๆ ด้วยคำหลัก (keyword) บนภาพ เรียกการแทนข้อมูลบน

ภาพแบบนี้ว่า Keyword-based image indexing [Tie Hua Zhou 2015] เป็นการให้ความหมายของวัตถุบนภาพเป็น ชื่อวัตถุ หรือคำศัพท์ ที่สอดคล้องกัน เช่น “grass” , “plant” , “boat”, “sky” เป็นต้น หรือแสดงเนื้อหาภาพเป็นคำอธิบาย สิ่งต่าง ๆที่เกี่ยวข้องกับภาพ เช่น ชื่อเรื่อง, ผู้ถ่ายภาพ, รายละเอียดภาพ เป็นต้น ส่วนใหญ่ในการให้ความหมายหรือการเขียนคำอธิบายภาพจะเป็นลักษณะของอัติโนมัติบนภาพหรือใช้การคัดเลือกคำศัพท์ด้วยมนุษย์ ขึ้นกับระบบที่เลือกใช้งาน ในปัจจุบันระบบที่ใช้ในการให้ความหมายภาพ (annotation) จะสามารถให้ความหมายวัตถุบนภาพได้อย่างง่ายดายและรวดเร็ว ตัวอย่างระบบ เช่น Flickr [Flickr], LabelMe [Russell, B.C 2008], Peekaboom [L. Von Ahn 2006] เป็นต้น การให้ความหมายของภาพด้วยคำศัพท์ต่าง ๆเพื่อทำให้การสืบค้นข้อมูลได้ผลดีขึ้นจากวิธีแรกที่ใช้เพียงคุณลักษณะต่ำเพียงอย่างเดียว แต่ยังคงมีปัญหาที่ตามมาจากผลของการค้นคืนเพื่อให้ถูกต้องและตรงตามความหมายมากยิ่งขึ้น การใช้หลายคำศัพท์ (multiple label) [K.Narnard 2003][M. Johnson 2005] เป็นอีกวิธีการหนึ่งที่ใช้ช่วยการค้นคืนแต่ก็ยังถูกจำกัดด้วยการค้นหาเพื่อให้ได้ภาพที่ตรงกับความหมายภาพที่เพิ่มขึ้น นักวิจัยบางกลุ่ม [Venkatesh N. Murthy 2015] พยายามใช้มิติของความสัมพันธ์ของคำศัพท์เพื่อหาความสัมพันธ์ของความหมายของการแทนคำ ร่วมกับวิธีการ Kernel Canonical Correlation Analysis เพื่อสร้างความสัมพันธ์ของคุณลักษณะคำศัพท์ที่เกิดขึ้น M. Johnson and R. Cipolla [M. Johnson 2005] ได้พยายามจัดกลุ่มภาพที่มีวัตถุเหมือนกันอยู่ในกลุ่มเดียวกันแทนที่จะเป็นการแยกทีละภาพ I. Simon และคณะ [I. Simon 2007] ได้พยายามจัดกลุ่มภาพในลักษณะเดียวกันโดยที่จะใช้ข้อมูลสถานที่ (location) ของภาพช่วยในการจัดกลุ่ม และการจัดกลุ่มตามชื่อเรื่อง (topic) หรือภาพที่เกิดขึ้นเป็นภาพธรรมชาติ อาจจะมีการจัดกลุ่มตาม เหตุการณ์ของธรรมชาติ (natural event) แต่บางงานวิจัยจะมีการใช้วิธีการเรียนรู้ตามแม่แบบ (learning concept templates) [Y. Wu 2007] เพื่อจัดกลุ่มภาพและเพิ่มส่วนของการตอบสนองกลับ (relevance feedback) [Y. Wu 2006] เพื่อช่วยการการแก้ไขกลุ่มของภาพในการค้นคืนให้ได้ตรงตามกลุ่มเป้าหมายที่ดีขึ้น Query-Frequency Pair [Tie Hua Zhou2015] เป็นอีกวิธีการหนึ่งที่พยายามค้นคืนภาพด้วย query ที่ซับซ้อนที่ประกอบด้วยคำศัพท์ที่มีการบันทึกไว้ ผลลัพธ์ที่ได้นั้นจะได้ผลที่ดีเมื่อการค้นคืนใช้รูปแบบของคำศัพท์ในแบบเดียวกัน แต่เมื่อมีความแตกต่างของการให้ความหมายคำศัพท์ผลที่ได้จะผิดเพี้ยนไป

เพื่อให้ผลลัพธ์ที่ได้มีความถูกต้องมากยิ่งขึ้น การค้นหาภาพด้วยเทคนิคที่กล่าวมานี้จะได้ผลลัพธ์ที่ขึ้นกับคำศัพท์ที่ถูกให้ความหมายไว้บนภาพยิ่งมีการแท็กข้อมูลบนภาพมากยิ่งสามารถหาความเหมือนกันบนภาพมากขึ้นเท่านั้น แต่เมื่อการให้ความหมายของวัตถุบนภาพมาจากหลายที่

หลายแหล่งข้อมูล หรือหลายผู้บันทึกข้อมูลทำให้คำศัพท์บางคำมีความหมายที่เหมือนกันแต่ต่างคำกัน เช่น “sky” กับ “trees”, “car” กับ“vehicle” เป็นต้น หรือคำศัพท์บางคำที่ให้ความหมายได้ลึกแตกต่างกัน เช่น “mango” กับ “fruit”, “mobile phone” กับ “electronic device” เป็นต้น การให้ความหมายเพียงผิวเผิน หรือความหมายที่ตรงกับวัตถุจริงหรือเพียงกลุ่มของวัตถุเท่านั้นจะมีผลต่อการค้นคืนภาพรวมทั้งมีผลต่อการแปลความหมายภาพด้วยเช่นกัน ดังนั้นในงานวิจัยนี้จึงนำเสนอวิธีการจัดการกลุ่มคำศัพท์และแทนข้อมูลความสัมพันธ์ภาพนั้นด้วยรูปแบบออนโทโลจี (Ontology) [F. Giunchiglia,2004][J. Euzenat,2013] ซึ่งออนโทโลจีมีโครงสร้างแบบดัชนีหลายลำดับ (multilevel index) สามารถนำมาแทนกลุ่มความหมายคำศัพท์ได้ และใช้วิธีการเปรียบเทียบความเหมือนกันของความหมายภาพด้วยการหาความเหมือนของภาพด้วย รูปแบบการวัดความเหมือนของโครงสร้างกราฟ (Similarity Graph Matching) และ การจัดกลุ่มแบบลำดับชั้น(Hierarchical Clustering Algorithm)

1.2 วัตถุประสงค์

1. เพื่อเปรียบเทียบเทคนิคที่นำเสนอ Ontology สามารถนำมาใช้ในการค้นหาความเหมือนกันของภาพดิจิทัล
2. เพื่อสร้างแบบจำลองวิเคราะห์การค้นหาความหมายภาพในรูปแบบของโครงสร้างลำดับขั้น
3. เพื่อพัฒนาเทคนิคในการหาความเหมือนกันของภาพดิจิทัล ด้วยการเปรียบการวัดความเหมือนของโครงสร้างกราฟ (Similarity Graph Matching) และ การจัดกลุ่มแบบลำดับชั้น (Hierarchical Clustering Algorithm)

1.3 สมมติฐาน

งานวิจัยนี้นำเสนอวิธีการแทนข้อมูลคำศัพท์ด้วยโครงสร้างออนโทโลจีเพื่อนำมาใช้ช่วยในการค้นคืนความหมายภาพได้ดียิ่งขึ้น โดยจะทำการแทนข้อมูลวัตถุบนภาพออกเป็นหมวดหมู่ตามลักษณะของออนโทโลยี

1.4 ขอบเขตของการวิจัย

1. ข้อมูลที่นำมาวิเคราะห์เป็นข้อมูลภาพดิจิทัลที่เห็นวัตถุชัดเจน

2. ข้อมูลคำหลักบนภาพหรือวัตถุบนภาพที่เป็นอินพุตบนภาพจะถูกแท็ก เป็นคำศัพท์เข้ามาก่อนเริ่มการประมวลผล
3. การเก็บผลข้อมูลเบื้องต้นของความหมายของภาพ ใช้กลุ่มนักศึกษา และบุคคลทั่วไปเป็นกลุ่มบุคคลที่ตัดสิน ความหมายของภาพสำหรับการทดลองในคลังภาพที่กล่าวมา

1.5 ประโยชน์ที่คาดว่าจะได้รับ

1) เพื่อนำกระบวนการที่นำเสนอเข้ามาประยุกต์ใช้ หรือปรับเปลี่ยน Module ให้เข้ากับการค้นหาภาพบนห้องสมุดดิจิทัลได้

2) เพื่อได้แนวทางการแปลและการตีความหมายภาพในรูปแบบของคำศัพท์

3) ช่วยให้การใช้คำศัพท์ในการค้นหาข้อมูลภาพได้ผลตามความหมายของภาพมากขึ้น

4) สร้างชื่อเสียงให้กับมหาวิทยาลัย เมื่อมีบทความวิจัยลงในวารสารต่างประเทศ บทความวิจัยที่นำเสนอในการประชุมระดับชาติ และนานาชาติในระดับที่ดี

1.6 นิยามคำศัพท์

1. การประมวลผลภาพ (Image processing) หมายถึง การนำภาพมาผ่านกระบวนการเพื่อประมวลผลสัญญาณบนสัญญาณ 2 มิติ เช่น ภาพนิ่ง (ภาพถ่าย) หรือภาพวีดิทัศน์(วีดิโอ) และนำมาใช้งาน
2. การจำแนกประเภทข้อมูลภาพ (Image Classification) หมายถึง การแยกแยะภาพที่มีคุณลักษณะเดียวกันออกเป็นกลุ่ม ๆ
3. การจำแนกความหมายของภาพ (Semantic Image Classification) หมายถึง การนำภาพมาแบ่งเป็นกลุ่มโดยพิจารณาจากความหมายโดยรวมของภาพ
4. การค้นคืนรูปภาพ (Image Retrieval) หมายถึง การค้นหาข้อมูลของภาพจากระบบฐานข้อมูล
5. วัตถุ (Object) หมายถึง ส่วนของวัตถุบนภาพ ยกตัวอย่างเช่น คน, ต้นไม้ เป็นต้น
6. คำหลัก (Keyword) หมายถึง คำที่มีความหมายได้ใจความใช้แทนวัตถุบนภาพ
7. แท็ก (Tag) หมายถึง การกำหนดคำ หรือ คำศัพท์บนภาพ หรือเรียกว่าการ label
8. การแทนที่ด้วยกราฟ (Graph representation) คือการแทนความหมายของภาพโดยใช้กราฟ

DPU

9. คุณลักษณะข้อมูล หรือ ฟีเจอร์ (Feature) หรือตัวแปร ที่ถูกสกัดออกมาจากภาพ เช่น สี (color) ลวดลาย (texture) หรือ รูปทรง (shape) รวมทั้ง วัตถุ ที่ปรากฏบนภาพเพื่อนำมาใช้ในการสืบค้นข้อมูลต่อไป
10. คิวรี (query) หมายถึงรูปแบบการสอบถามข้อมูลด้วยคุณลักษณะที่กำหนด

DPU

บทที่ 2
ทฤษฎีและงานวิจัย

2.1 ทฤษฎีเบื้องต้น

วิวัฒนาการของเครื่องมือและอุปกรณ์ถ่ายภาพดิจิทัลได้พัฒนาอย่างรวดเร็วจนทำให้ภาพถ่ายภาพดิจิทัล มีจำนวนเพิ่มมากขึ้น ปัญหาที่ตามมาก็คือการจัดเก็บข้อมูลภาพที่เพิ่มมากขึ้นอย่างไร้ขีดจำกัดนี้จะทำอย่างไรจึงจะสามารถจัดเก็บอย่างมีระบบและสามารถสืบค้นข้อมูลภาพ และจำแนกข้อมูลภาพให้ตรงตามความหมายของภาพที่ต้องการของผู้ใช้มากที่สุด ทำให้งานวิจัยในปัจจุบันที่เกี่ยวข้องกับการค้นคืนรวมทั้งการจัดกลุ่มภาพให้ตรงกับความต้องการเพิ่มมากขึ้น รวมไปถึงระยะเวลาในการสืบค้นที่น้อยลงกับปริมาณของภาพที่เพิ่มทวีคูณ ดังนั้นปัญหาดังกล่าวมาข้างต้นนั้นจึงได้รับความสนใจจากนักวิจัยหลายกลุ่ม ซึ่งเป็นงานด้านการประมวลผลภาพ (image processing) ด้านการค้นคืนข้อมูลภาพ (image retrieval) เป็นอีกหนทางหนึ่งในการแก้ปัญหาดังกล่าว เนื่องจากสามารถช่วยให้การค้นหาข้อมูลกระทำได้โดยสะดวกยิ่งขึ้นตรงตามความต้องการของผู้ใช้งาน

ภาพต้นฉบับ ภาพต้นฉบับที่ถูกแปลงเป็นฟีเจอร์

ภาพที่ 2.1 ภาพต้นฉบับสำหรับการค้นคืนที่ถูกแปลงเป็นฟีเจอร์ระดับต่ำ

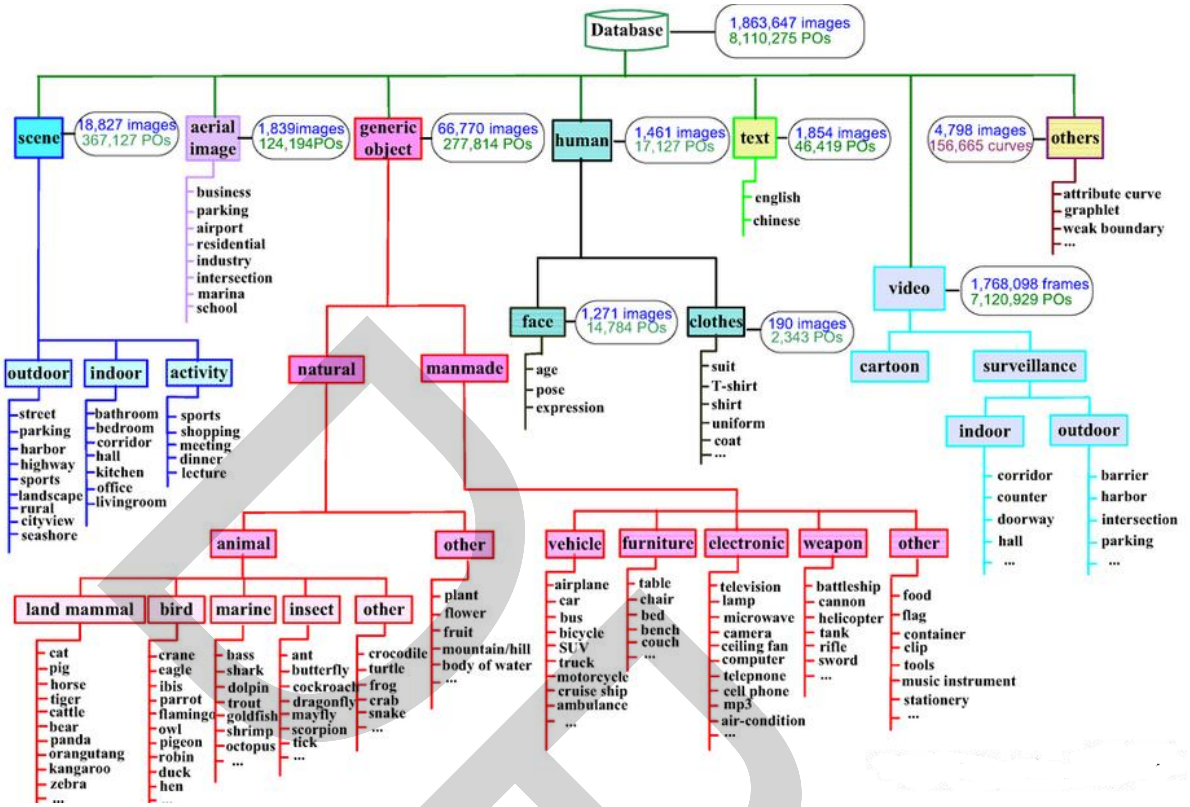
สำหรับงานวิจัยทางการประมวลผลภาพในการค้นคืนข้อมูลภาพกลุ่มแรกจะมีการค้นคืนตามคุณลักษณะพื้นฐานของภาพที่ถูกสกัดคุณลักษณะด้วยอัลกอริทึมต่างๆ ยกตัวอย่างเช่น สี (color) ลวดลาย (texture) รูปทรง (shape) เป็นต้น กระบวนการนี้ถูกเรียกว่า การประมวลผลภาพระดับต่ำ (low-level image processing) [Shaoqing R. 2016] [Sun Ting, 2016] [Jain A.K, 1996];[Smeulders, A. W. M, 2000] กระบวนการนี้สามารถค้นหาภาพได้ตามคุณลักษณะพื้นฐานเมื่อนำมาจำแนกด้วยการประมวลผลภาพระดับต่ำ (low-level feature) ภาพผลลัพธ์ส่วนใหญ่เป็นภาพที่มีคุณลักษณะไม่ซับซ้อน มีรูปทรงที่แตกต่างกันอย่างเด่นชัด ด้วยคุณลักษณะฟีเจอร์ระดับต่ำ

ด้วยสี ลวดลาย และตำแหน่งของพื้นที่ของภาพ ภาพที่ 2.1 จากผลลัพธ์จะสังเกตว่าผลลัพธ์ของภาพเป็นภาพที่มีโทนสีคล้ายกันเป็นหลัก แต่มีลักษณะวัตถุภายในภาพที่ได้ที่แตกต่างกัน

ภาพที่ 2.2 ตัวอย่างการแท็กองค์ประกอบของภาพ[1]

แต่อย่างไรก็ตามได้มีงานวิจัยอีกกลุ่ม ที่พยายามจะใช้เทคนิคของการเข้าใจความหมายของภาพแทน การสืบค้นแบบข้างต้น เรียกว่า การประมวลผลภาพระดับสูง (high level image processing) งานวิจัยในกลุ่มนี้พยายามที่จะมองข้อมูลบนภาพเป็นวัตถุ (object) ที่มีความหมาย [Benitez A.B, 2002];[Galleguillos C., 2010] [Wan, Ji, 2014] และแทนวัตถุนั้น ๆ ด้วยคำหลัก (keyword) บนภาพ เรียกว่า การแท็ก (tag) หรือการให้ความหมายของวัตถุบนภาพเป็น ชื่อวัตถุหรือคำศัพท์ ที่สอดคล้องกันเช่น “grass” , “plant” , “boat”, “sky” เป็นต้น ดังแสดงในภาพที่ 2.2 [Galleguillos C., 2010] และใช้ความหมายหรือคำศัพท์นั้นเพื่อทำการสืบค้นข้อมูลแทน ซึ่งเป็นการใช้ความหมายของคำศัพท์ที่มีความสอดคล้องกันด้วยความหมายตามพจนานุกรม หรือในลักษณะใช้ความสัมพันธ์ของความหมายที่เหมือนกันของคำหลัก (synonym) [Zhao T., 2001];[Benitez, A.B., 2002];[Kobus B., 2001]; เข้ามาใช้ในการค้นคืนข้อมูลภาพเพื่อให้ผลลัพธ์ตรงกับความต้องการของผู้ใช้ ยกตัวอย่างเช่น “stone” มีความหมายสอดคล้องกันกับ “rock” เป็นต้น จะได้ผลที่ค่อนข้างดีกว่า แต่ขึ้นอยู่กับว่าอัลกอริทึมที่ถูกนำมาใช้นั้นจะเป็นลักษณะใด Benjamin Yao [Benjamin Yao, 2009] ได้สร้างความสัมพันธ์ของกลุ่มคำศัพท์โดยจัดหมวดหมู่ของคำศัพท์ที่มีความเกี่ยวข้องเข้าด้วยกันดังแสดงในภาพที่ 2.3 และจะได้ผลที่ค่อนข้างดีกว่า แต่ขึ้นอยู่กับว่าอัลกอริทึมที่ถูกนำมาใช้นั้นจะเป็นลักษณะใด

[1] http://vision.ucsd.edu/project/context-based-object-categorization



ภาพที่ 2.3 ตัวอย่างโครงสร้างแท็กคำหลักของภาพ[2]

นักวิจัยบางกลุ่มใช้เทคนิคเพื่อสร้างความสัมพันธ์ (relationship) ของแท็กชื่อวัตถุบนภาพ [Benitez A.B, 2001];[R. Zhao, 2002];[Philippe M., 2002] ด้วยโมเดลเหตุการณ์ (event model) [Joo-Hwee L., 2003] เช่น การเชื่อมวัตถุด้วยคำต่าง ๆ เช่น “touch” “on” “top” เป็นต้น บางงานวิจัยได้พยายามใส่ข้อความ (context) เพื่อบรรยายความหมายของภาพ [Mathias Lux, 2003];[Mathias Lux, 2009] ในหัวข้อที่สอดคล้องกับภาพนั้น ๆ เช่น “birthday party of uncle Adam” หรือ “a picture showing a barking dog” แต่ภาพที่นำมาใช้จัดเก็บนั้นมักจะเป็นภาพส่วนตัว (personal images) ทำให้การค้นคืนจำกัดเพราะคำบรรยายส่วนใหญ่จะเป็นคำเฉพาะเจาะจง จึงไม่นิยมเท่าที่ควร บางงานวิจัยพยายามที่จะใส่ข้อมูลเหตุการณ์ต่าง ๆจนครบถ้วนในรูปแบบของ ใคร ทำอะไร ที่ไหน เมื่อไหร่ (who, what, when, where) ดังนั้นในการใส่ข้อมูลบางครั้ง เป็นข้อมูลที่นอกเหนือ หรือเกินความจำเป็นโดยใช่เหตุ ข้อมูลเหล่านี้อาจจะไม่มีความจำเป็นเลยสำหรับการสืบค้นข้อมูล อาจจะเป็นข้อมูลที่มีความเป็นส่วนตัวจนเกินไป และใช้คำศัพท์ซ้ำซ้อน ฟุ่มเฟือยฟุ่มเฟือย ทำ

---

[2] อ้างอิง http://www.stat.ucla.edu/~zyyao/projects/I2T.htm ค้นคืนเมื่อวันที่ 12 ส.ค. 2558

ให้การค้นคืนยิ่งเกิดความสับสนปนเปกันของภาพผลลัพธ์ที่ได้มา มีการนำคำศัพท์หลักที่ถูกแท็กลงในภาพเป็นข้อมูลที่ใช้ในการค้นหาซึ่งผลลัพธ์ของข้อมูลที่ได้มาจะมีความสอดคล้องกันของคำศัพท์ที่ถูกแท็กลงในแต่ละภาพ การค้นหาภาพด้วยเทคนิคนี้จะได้ผลลัพธ์ที่ขึ้นกับคำศัพท์ที่ถูกแท็กไว้บนภาพยิ่งมีการแท็กข้อมูลบนภาพมากยิ่งสามารถหาความเหมือนกันบนภาพมากขึ้นเท่านั้น แต่ในความเป็นจริงแล้ว การแท็กข้อมูลบนภาพในปัจจุบันนั้นเป็นเพียงการหาคำศัพท์ที่ต้องการบนภาพ แต่ไม่ได้ให้ความหมายภาพโดยรวม ความหมายของภาพคือการนำวัตถุที่ปรากฏบนภาพมารวมกันเพื่อวิเคราะห์จากความคิดของมนุษย์เพื่อให้ได้ผลลัพธ์คือคำศัพท์ใหม่ที่แทนความหมายของภาพทั้งภาพ

ปัจจุบันการพัฒนาซอพต์แวร์เพื่อให้เกิดความสะดวกสบายในการแท็กข้อมูล และข้อมูลที่ถูกนำมาเป็นคำศัพท์นั้นมีแบบแผนและโครงสร้างที่แน่นอน ยกตัวอย่าง Caliph & Emir [Mathias Lux, 2009], Annosearch [Xin-Jing, 2008], CAMEL [Apostol Paul N., 2001] เป็นต้น ส่วนใหญ่จะมีความสนใจในการแทนค่าความหมายภาพ (Semantic image representation) โดยขึ้นกับภาษาและโครงสร้างการแทนค่าข้อมูลเพื่อทำให้มนุษย์และเครื่องคอมพิวเตอร์เกิดความเข้าใจในการประมวลผลโดยมีคำอธิบายรายละเอียดของภาพด้านขวา คำศัพท์ที่จะมีเพิ่มลงในภาพ ขึ้นอยู่กับการผสมระดับคำอธิบาย สามารถแทนคำอธิบายภาพด้วย “Formula One”, “event”, หรือ “a kind of sport” เป็นต้น แต่อย่างไรก็ตามการเพิ่มระดับการแท็กลงบนภาพเพื่อให้ได้ความหมาย เช่น “a Formula-1 car in a street”, “buildings in the background”, “barriers delineating the circuit” และ “people” เป็นต้น บางข้อมูลสามารถได้มาจากภาพ แต่ข้อมูลบางอย่างไม่สามารถหาได้จากภาพ เช่น “Mika Hakkinen has been F1 world champion twice”, “the race he takes part in is part of an anti-drink driving campaign” หรือ “the photo has been taken on March 30, 2008” ซึ่งในภาพจะไม่สามารถปรากฏข้อมูลดังกล่าวบนภาพ S. Shatford Layne [S. Shatford Layne, 1994] ได้กำหนดวิธีการเก็บคำอธิบายรูปภาพ โดยแบ่งออกเป็น 4 ส่วนดังนี้ (1) แอทริบิวต์ชีวประวัติ เพื่อบอกสถานที่ เวลาของภาพ (2) แอทริบิวต์หัวเรื่อง เพื่ออธิบายเนื้อหาของภาพ (3) แอทริบิวต์ Exemplified เพื่อบอกประเภทของภาพถ่าย ภาพการ์ตูน หรือภาพร่าง เป็นต้น (4) แอทริบิวต์ความสัมพันธ์ เพื่อแสดงความสัมพันธ์ของภาพ

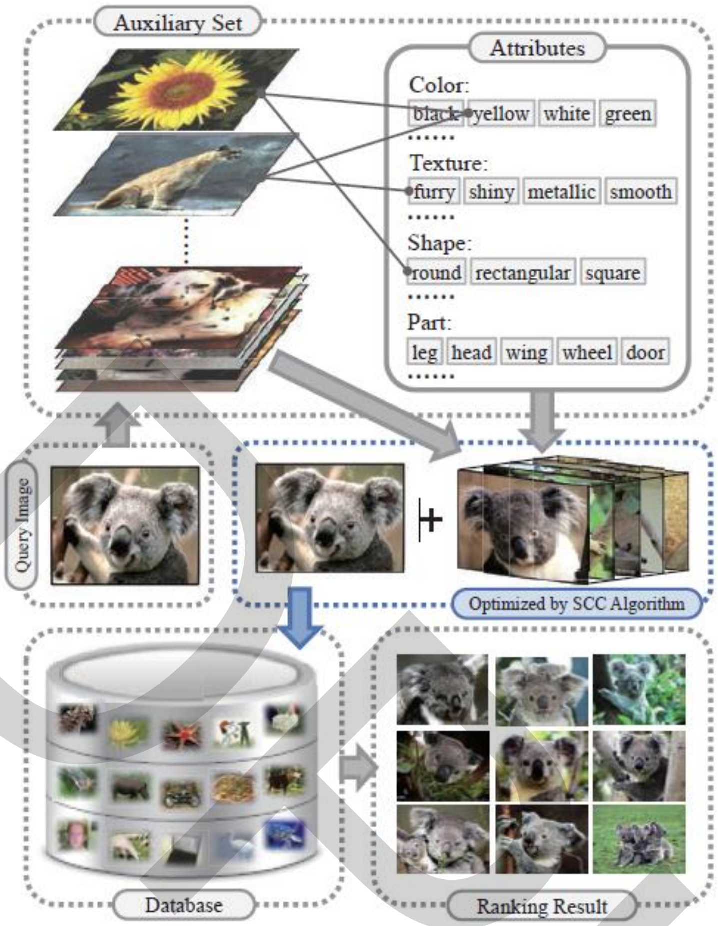
The description levels suggested by Jaimes and Chang

1	Type, technique
2	Global distribution
3	Local structure
4	Global composition
5	Generic objects
6	Generic scene
7	Specific objects
8	Specific scene
9	Abstract objects
10	Abstract scene

1.	photo
2.	(histogram)
4.	(segmentation)
5.	flower, leaf, water
6.	nature
7.	water lily and its leaves
8.	pond
10.	stillness, coldness

ภาพที่ 2.4 ตัวอย่างระดับคำอธิบายภาพ [A. Jaimes, 2000]

กลุ่มนักวิจัย [A. Jaimes, 2000];[L. Hollink, 2004] ได้สร้างโครงสร้างคำศัพท์ภาพแบบ 10 ระดับ เพื่อจัดเก็บข้อมูลภาพ ดังภาพที่ 2.4 แสดงตารางระดับคำอธิบายภาพโดยที่ 4 ลำดับแรกจะอ้างอิงถึง perceptual aspect หรือ syntax และ 6 ลำดับท้ายจะสัมพันธ์กับความหมายและแนวคิด ดังนั้นเมื่อทำการอธิบายความหมายของภาพตามระดับแล้วสามารถที่จะเปรียบเทียบภาพได้จากข้อมูลแต่ละระดับ ดังแสดงตัวอย่างการใส่คำอธิบายภาพตามลำดับ L. Hollink [L. Hollink, 2004] ได้มีการใช้ Unified Modeling Language (UML) ในการอธิบายลำดับการแทนคำอธิบายภาพเพื่อให้เกิดความเข้าใจวัตถุที่เกิดขึ้นบนภาพง่ายขึ้น Hare et al. [J.S. Hare, 2006] ได้นำเสนอโครงสร้างคำศัพท์เริ่มต้นจากข้อมูลภาพและมีการแปลในรูปแบบของคำอธิบายวัตถุเป็นส่วนตามชื่อของวัตถุซึ่งมีความหมายในระดับสูง คล้ายวิธีการที่ Jaimes and Chang [A. Jaimes, 2000] ได้นำเสนอในส่วนต้น จากภาพที่ 2.5 มีการนำคำศัพท์หลักที่ถูกแท็กลงในภาพเป็นข้อมูลที่ใช้ในการค้นหาซึ่งผลลัพธ์ของข้อมูลที่ได้มาจะมีความสอดคล้องกันของคำศัพท์ที่ถูกแท็กลงในแต่ละภาพ การค้นหาภาพด้วยเทคนิคนี้จะได้ผลลัพธ์ที่ขึ้นกับคำศัพท์ที่ถูกแท็กไว้บนภาพยิ่งมีการแท็กข้อมูลบนภาพมากยิ่งสามารถหาความเหมือนกันบนภาพมากขึ้นเท่านั้น แต่ในความเป็นจริงแล้ว การแท็กข้อมูลบนภาพในปัจจุบันนั้นเป็นเพียงการหาคำศัพท์ที่ต้องการบนภาพ แต่ไม่ได้ให้ความหมายภาพโดยรวม ความหมายของภาพคือการนำวัตถุที่ปรากฏบนภาพมารวมกันเพื่อวิเคราะห์จากความคิดของมนุษย์เพื่อให้ได้ผลลัพธ์คือคำศัพท์ใหม่ที่แทนความหมายของภาพทั้งภาพ



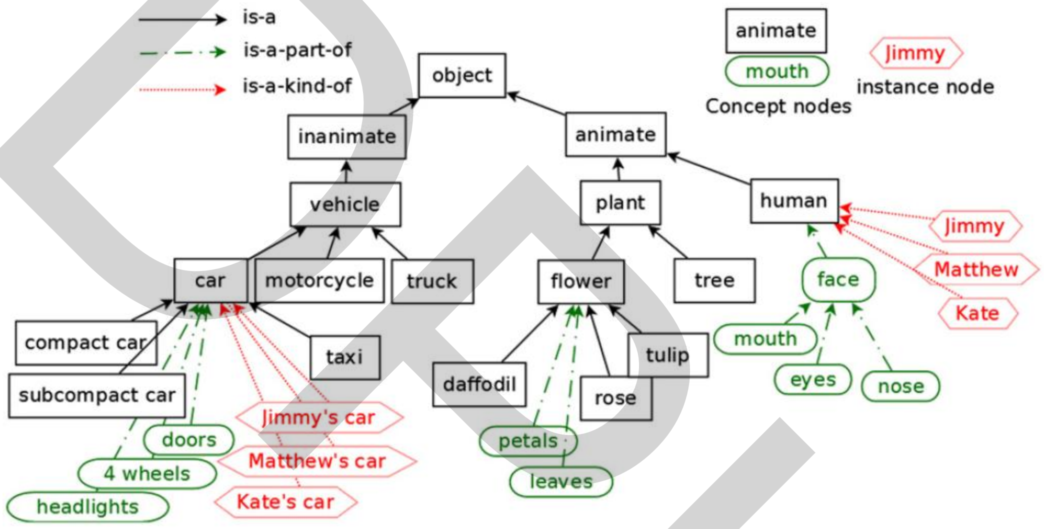
ภาพที่ 2.5 ตัวอย่างงานวิจัยที่ใช้คำศัพท์แท็กลงในภาพ [3]

ออนโทโลยีเป็นแนวคิดแบบจำลองสำหรับรูปแบบคำอธิบาย ซึ่งจะมีการกำหนดประเภทวัตถุ ความสัมพันธ์ระหว่างวัตถุ รวมทั้งความสัมพันธ์ระหว่างประเภท ซึ่งจะมีการนำส่วนของเครือข่ายความสัมพันธ์เข้ามาใช้บนออนโทโลยีดังแสดงในภาพที่ 2.6 ได้แสดงเครือข่ายความหมายที่ถูกแทนด้วยประเภทของความสัมพันธ์ โดยที่มีโหนด 2 ประเภท ที่เรียกว่า โหนดคอนเซปต์ (concept) จะแทนกลุ่มของวัตถุ เช่น car เป็นต้น และ โหนดอินสแตน (instance) แทนวัตถุ เช่น Jimmy’s car เป็นต้น จะเห็นว่า การใช้โครงสร้างความหมาย (semantic structure) จะแทนความสัมพันธ์แบบระดับขั้นระหว่างคอนเซปต์ แต่อย่างไรก็ตามการวิเคราะห์คำบรรยายเพื่อให้มาถึงความหมายของภาพนั้น อาจจะไม่สามารถควบคุมการรู้จำวัตถุในบางกรณีได้เนื่องจาก ขนาดของวัตถุเล็กเกินไป และคำศัพท์ที่แทนนั้นยากในการควบคุม ทุกปัญหาที่เกิดขึ้นจะมีความสัมพันธ์ที่เกี่ยวข้องกันกับ อินสแตน

---

3 http://lxiongh.com/2014/06/02/ImageRetrieval/ ค้นคืนเมื่อวันที่ 12 ส.ค. 2559

และกลุ่มของวัตถุสำหรับการใส่คำอธิบาย มีนักวิจัยหลายกลุ่มพยายามที่จะแก้ไข และใช้หลากหลายวิธีเพื่อที่จะเชื่อมโยงให้ได้ความหมายของภาพอย่างแท้จริง แต่อย่างไรก็ตามซอฟต์แวร์ส่วนใหญ่นั้น จะทำการสืบค้นภาพในรูปแบบของการเทียบคำหลักเป็นคำต่อคำตามการเรียนรู้ของเครื่องจักร (machine learning) ที่เข้ามาช่วยในการหาความหมายของภาพ หรือตามระดับขั้นของการเรียนรู้ตามที่มีการเก็บข้อมูลไว้เท่านั้น เพราะฉะนั้นคำตอบที่ได้จะไม่ได้ขึ้นกับความหมายของภาพ แต่ขึ้นกับคำศัพท์ที่มีการเก็บข้อมูลลงไปเท่านั้น เพราะฉะนั้นผลลัพธ์ของกลุ่มภาพที่ได้จะไม่ได้ขึ้นกับความหมายของภาพอย่างแท้จริง



ภาพที่ 2.6 ตัวอย่างแนวคิดเครือข่ายความสัมพันธ์ระหว่างคำอธิบายความหมายบนภาพ

แต่มีกลุ่มนักวิจัยที่พยายามปรับปรุงอัลกอริทึม เพื่อทำการค้นคืนภาพที่มีลักษณะฟีเจอร์ที่ใกล้เคียงกับภาพที่ต้องการมาก การปรับปรุงเทคนิคด้วยการนำวิธีการมาผสมผสานกันระหว่างคุณลักษณะเพื่อให้สามารถวิเคราะห์ในรูปแบบที่ซับซ้อน ส่วนใหญ่ในการให้ความหมายหรือการเขียนคำอธิบายภาพจะเป็นลักษณะของอัติโนมัติบนภาพหรือใช้การคัดเลือกคำศัพท์ด้วยมนุษย์ ขึ้นกับระบบที่เลือกใช้งาน ในปัจจุบันระบบที่ใช้ในการให้ความหมายภาพ (annotation) จะสามารถให้ความหมายวัตถุบนภาพได้อย่างง่ายดายและรวดเร็ว ตัวอย่างระบบ เช่น LabelMe [Russell B.C, 2008] Peekaboom [L. Von Ahn, 2006] เป็นต้น การให้ความหมายของภาพด้วยคำศัพท์ต่าง ๆเพื่อทำให้การสืบค้นข้อมูลได้ผลดีขึ้นจากวิธีแรกที่ใช้เพียงคุณลักษณะต่ำเพียงอย่างเดียว แต่ยังคงมีปัญหาที่ตามมาจากผลของการค้นคืนเพื่อให้ถูกต้องและตรงตามความหมายมากยิ่งขึ้น การใช้หลายคำศัพท์

(multiple label) [K. Barnard, 2003][M. Johnson, 2015] เป็นอีกวิธีการหนึ่งที่ใช้ช่วยการค้นคืน แต่ก็ยังถูกจำกัดด้วยการค้นหาเพื่อให้ได้ภาพที่ตรงกับความหมายภาพที่เพิ่มขึ้น นักวิจัยบางกลุ่ม [11] พยายามใช้มิติของความสัมพันธ์ของคำศัพท์เพื่อหาแตกต่างกัน เช่น “mango” กับ “fruit” “mobile phone” กับ “electronic device” เป็นต้น การให้ความหมายเพียงผิวเผิน หรือความหมายที่ตรง กับวัตถุจริงหรือเพียงกลุ่มของวัตถุเท่านั้นจะมีผลต่อการค้นคืนภาพรวมทั้งมีผลต่อการแปลความหมายภาพด้วยเช่นกัน

ดังนั้นในงานวิจัยนี้จึงนำเสนอวิธีการค้นคืนกลุ่มภาพและแทนข้อมูลความสัมพันธ์ภาพนั้นด้วยรูปแบบอนุกรมวิธาน WordNet [Patwardhan and Pedersen, 2006][Miller 1990] [Junshi Huang, 2015] ซึ่งเป็นฐานข้อมูลพจนานุกรม “good”, ภาษาอังกฤษที่มีโครงสร้างแบบลำดับขั้น โดยจะมีการจัดลำดับความสัมพันธ์ตามความหมายเป็นกลุ่มที่เรียกว่า Synset ตัวอย่างเช่น ความสัมพันธ์ของความหมายของการแทนคำ ร่วมกับวิธีการ Kernel Canonical Correlation Analysis เพื่อสร้างความสัมพันธ์ของคุณลักษณะคำศัพท์ที่เกิดขึ้น M. Johnson and R. Cipolla [M. Johnson, 2005] ได้พยายามจัดกลุ่มภาพที่มีวัตถุเหมือนกันอยู่ในกลุ่มเดียวกันแทนที่จะเป็นการแยกทีละภาพ I. Simon และคณะ [I. Simon, 2007] ทำการจัดกลุ่มภาพในลักษณะเดียวกันโดยที่จะใช้ข้อมูลสถานที่ (location) ของภาพ และการจัดกลุ่มตามชื่อเรื่อง (topic) หรือภาพที่เกิดขึ้นเป็นภาพธรรมชาติ จัดกลุ่มตาม เหตุการณ์ของธรรมชาติ (natural event) แต่บางงานวิจัยจะมีการใช้วิธีการเรียนรู้ตามแม่แบบ (learning concept templates) [Y. Wu, 2007] จัดกลุ่มภาพและเพิ่มส่วนของการตอบสนองกลับ (relevance feedback) เพื่อช่วยการการแก้ไขกลุ่มของภาพในการค้นคืนให้ได้ตรงตามกลุ่มเป้าหมายที่ดีขึ้น Query-Frequency Pair [Galleguillos C, 2010] เป็นอีกวิธีการหนึ่งที่พยายามค้นคืนภาพด้วยคิวรี (query) ที่มีความซับซ้อนประกอบด้วยคำศัพท์ที่มีการบันทึกไว้ ผลลัพธ์จะได้ผลเมื่อการค้นคืนใช้รูปแบบของคำศัพท์ในแบบเดียวกัน แต่เมื่อมีความแตกต่างของการให้ความหมายคำศัพท์ผลที่ได้จะผิดไปเพื่อให้ผลลัพธ์ที่ได้มีความถูกต้องมากยิ่งขึ้น และยังมีบางกลุ่มนักวิจัยการใช้วิธีการเรียนรู้เข้ามาช่วยเพื่อทำให้ผลการทดลองนั้นดียิ่งขึ้น[Albert Gordo 2017] [Junshi Huang, 2015]

2.3 การเปรียบเทียบความคล้ายของกราฟ

การเปรียบเทียบความคล้ายกันของกราฟ (Graph Similarity) เป็นการนำข้อมูลที่มีความสัมพันธ์ในรูปแบบกราฟทั้งหมดที่เก็บรวบรวม ผ่านกระบวนการขั้นตอนการเปรียบเทียบความคล้ายกันของกราฟแยกแยะข้อมูลลงในแต่ละกลุ่มที่จัดไว้ โดยในแต่ละกลุ่มของข้อมูลนั้นจะมีคุณลักษณะเด่นของแต่ละกลุ่มที่แตกต่างกัน ขึ้นกับข้อมูลที่เก็บรวบรวมมาได้ การเชื่อมโยงความสัมพันธ์ข้อมูลภาพ ด้วยกราฟลำดับขั้น ข้อมูลภาพได้จากการแท็กของผู้ใช้ออนไลน์ ที่มีผู้ใช้ที่แตกต่างกันทำให้เกิดความแตกต่างของคำศัพท์บ้างเล็กน้อย กับการแท็กวัตถุประเภทเดียวกัน แต่มีการแท็กคำที่ต่างกัน (เช่น คำศัพท์ “stone” หรือ “rock” มีความหมายใกล้เคียงกัน) เมื่อมีการค้นคืนด้วยคำศัพท์จะถูกสร้างความแตกต่างของภาพ แต่ในความเป็นจริงแล้วมีความหมายที่ใกล้เคียงกัน ดังนั้นจึงทำการลดความซ้ำซ้อนหรือกำกวมของคำศัพท์ที่เกิดขึ้นด้วยการใช้รูปแบบความสัมพันธ์ของคำศัพท์ WordNet [Patwardhan and Pedersen 2006] ดังนั้นเมื่อคำศัพท์สองคำหรือมากกว่ามีความหมายเหมือนกัน (synonym) จะถูกแทนค่าวัดความเหมือนได้ด้วยความหมายเดียวกัน ในรูปแบบความสัมพันธ์ Synset ตามอนุกรมวิธาน (taxonomy) ของ WordNet และขั้นตอนสุดท้ายขั้นตอนการประมวลผล (data processing) ดังนั้นวิธีการเปรียบเทียบความคล้ายข้อมูล ซึ่งในงานวิจัยนี้ได้ใช้วิธีการเปรียบเทียบกันทั้งหมด 4 วิธี

- การวัดความคล้ายด้วยอนุกรมวิธาน
- การวัดความคล้ายด้วยคิวรี่คำศัพท์
- การวัดความคล้ายด้วยคิวรี่ย้อนกลับ
- การวัดความคล้ายด้วยโครงสร้างกราฟตามความสัมพันธ์ลำดับขั้น

2.3.1 การวัดความคล้ายด้วยอนุกรมวิธาน

การวัดความคล้ายด้วยอนุกรมวิธาน (WordNet Similarity: WS) ลำดับแรกจะมีการกำหนดความเหมือนกันของคำศัพท์ที่อยู่บนพื้นฐานของอนุกรมวิธาน WordNet ที่มีความสัมพันธ์ของคำศัพท์แบบลำดับขั้น โดยที่จะมีการเลือกใช้ในส่วนเฉพาะโครงสร้างของ WS ที่มีสัดส่วนของเฉพาะข้อมูลคำศัพท์ที่มีในฐานข้อมูลภาพเท่านั้นทำให้มีการเปรียบเทียบที่จำกัดลง และการวัดความคล้ายที่ถูกนำมาใช้เป็นรูปแบบของความสัมพันธ์ระหว่างคำ (interword relationships) ที่มีการสอดคล้องกับ Synset ดังนั้นกำหนดให้ การวัดความคล้ายของ s_1 และ s_2 เป็น Synset และมีความ

ลึกของโหนดบรรพบุรุษของ Synset เป็น s_a แล้วสามารถเขียนสมการแสดงการวัดความคล้ายกันของความลึกระหว่างของ s_1 และ s_2 ได้ดังแสดงในสมการดังนี้

$$Sim_WS(s_1, s_2) = depth(s_a)/depth_{max}$$

2.3.2 การวัดความคล้ายด้วยคิวรี่คำศัพท์

การวัดความคล้ายด้วยคิวรี่คำศัพท์ (Keyword-based Query Similarity: KQS) เป็นการวัดความคล้ายแบบคำต่อคำที่มีการกำหนดค่าถ่วงน้ำหนักภายในคำศัพท์ที่ถูกแท็กเพื่อใช้เป็นเครื่องมือวัดรูปแบบหนึ่งระหว่างคิวรี่และภาพ โดยที่เมื่อกำหนดให้ $\rho = \{< k_1, w_{k_1} >, < k_2, w_{k_2} >, \ldots, < k_m, w_{k_m} >\}$ ให้ k_i แทนคำศัพท์หรือ Synset ที่ i ที่มีการแท็กบนภาพด้วยค่าถ่วงน้ำหนักมีค่าเป็น w_{k_i} ที่ i คิวรี่ของการวัดความคล้ายสามารถเขียนได้ $Q = \{< q_1, w_{q_1} >, < q_2, w_{q_2} >, \ldots, < q_m, w_{q_m} >\}$ ทุกคำศัพท์เมื่อผู้ใช้งานมีการค้นหาจะมีค่าถ่วงน้ำหนักเริ่มต้นคงเดิม แต่เมื่อมีการโต้ตอบของผู้ใช้งานเกิดขึ้นจะมีการปรับค่าถ่วงน้ำหนักไป ซึ่งสมการของการปรับเปลี่ยนค่าถ่วงน้ำหนักสามารถเขียนได้ดังนี้

$$Sim_KQS(Q, \rho) = \sum_{i=1}^{n} \max_{j=1 \ldots m} [sim(q_i, k_j) \cdot w_{q_i} \cdot w_{k_j}] / n$$

เมื่อกำหนดให้ $sim_KQS(q_i, k_j)$ เป็นการวัดความคล้ายระหว่าง q_i และ k_j เป็นการค้นหาคำศัพท์ของแต่ละคิวรี่และมีการคำนวณค่าเฉลี่ยที่ได้จากค่าสูงสุดของการวัดค่าความคล้ายระหว่างคิวรี่และภาพค่าความคล้ายกันของค่าน้ำหนักของคำศัพท์ w_{q_i} และ w_{k_j} ค่ามากจะแสดงถึงความเหมือนกันของคำศัพท์

2.3.3 การวัดความคล้ายด้วยคิวรี่ย้อนกลับ

การวัดความคล้ายด้วยคิวรี่คำศัพท์จากการสร้างแบบจำลองความสัมพันธ์โครงสร้างกราฟลำดับขั้น เมื่อมีการเพิ่มคำศัพท์ลงบนโครงสร้างจะเป็นการเพิ่มโหนดตามแนวคิดของ Synset บน WordNet ลงบนกราฟที่กำหนดขึ้น และจะสร้างตำแหน่งและที่ตั้งที่มีความสอดคล้องกันบนโครงสร้าง WS ดังนั้นในการใช้ถ้ามีการใช้การวัดความคล้ายด้วยคิวรี่คำศัพท์เพื่อค้นหาภาพที่คล้ายกัน ผลที่ได้จากการค้นคืนด้วยคิวรี่ จะได้ผลย้อนกลับ (relevance feedback) ที่เป็นทั้งภาพคล้ายกันและภาพต่างกันตามระดับลักษณะของภาพ ดังนั้นในการนำผลที่ได้จากการค้นคืนด้วยคิวรี่หลายครั้งมารวมกันเพื่อให้ได้ผลลัพธ์ที่ใกล้เคียงมากที่สุด เรียกวิธีการนี้ว่าการวัดความคล้ายด้วยคิวรี่ย้อนกลับ (Relevance Feedback based on Query Similarity: RFQS) เมื่อทุกคำศัพท์ที่ถูกแท็กบนตัวอย่างที่ถูกต้องจะเป็นคิวรี่บวก (positive query: Q_p) ด้วยค่าน้ำหนักที่เกิดขึ้นจากแต่ละการแท็ก และQ_n

เป็นคิวรี่ที่ถูกสร้างจากการย้อนกลับแบบลบ (negative query) ดังนั้นการสร้างค่าความคล้ายสุดท้ายของ S_RFQS_i เมื่อ i เป็นลำดับภาพบนฐานข้อมูลสามารถเขียนเป็นสมการได้ดังนี้

$$S_\mathrm{RFQS}_i = sim(Q,\rho_i) + \delta\left(1 + sim(Q_p,\rho_i)\right)\sum_{k\epsilon Np} S_{ik} + \beta(1 + 1 + sim(Q_n,\rho_i))\sum_{k\epsilon Nn} S_{ik}$$

เมื่อกำหนดให้ δ และ β เป็นค่าคงที่ และ S_{ik} เป็นค่าความคล้ายระหว่างภาพที่ i กับคิวรี่ตัวอย่างย้อนกลับที่ k, N_p และN_n เป็นจำนวนรวมของคิวรี่ย้อนกลับทั้งบวกและลบ

2.4 การวัดประสิทธิภาพ

การวัดประสิทธิภาพ (evaluation) เป็นขั้นตอนสุดท้าย เพื่อทำการตรวจสอบวิธีการที่ทำการทดลองมาข้างต้นว่ามีประสิทธิภาพมากหรือน้อยเพียงใดเมื่อนำมาใช้งานจริง จะเป็นขั้นตอนที่สำคัญ เพราะการนำวิธีการที่นำเสนอไปข้างต้นมาใช้งานได้นั้นจะต้องสอดคล้องกับความต้องการ จึงต้องมีการทดสอบศักยภาพการนำไปใช้ สถาปัตยกรรมที่ใช้ตัววัดความสำเร็จหลังการนำไปใช้หากนำไปใช้แล้วไม่ประสบผลสำเร็จต้องย้อนกลับไปเริ่มกระบวนการแรกใหม่ จึงต้องมีการประเมินผลก่อนการใช้งาน ในการประเมินนั้นกระทำได้โดยการวัดประสิทธิภาพของการจัดกลุ่มภาพมักจะถูกพิจารณาเป็นค่าของความถูกต้องของแต่ละกลุ่มข้อมูลซึ่งจะประกอบด้วย การวัดค่าความแม่นยำ ค่าความระลึก ค่าความถูกต้อง และ F-measure โดยยกตัวอย่างของค่าที่เกิดขึ้นจากตารางที่ 2.1

		ค่าทำนาย (predicted)	
		ปฏิเสธ (false/negative)	ยอมรับ (true/positive)
ค่าความจริง (actual)	ปฏิเสธ (negative)	a	b
	ยอมรับ (positive)	c	d
ค่าความถูกต้อง (accuracy)		Acc	

ตารางที่ 2.1 การวัดประสิทธิภาพ

- ค่าความแม่นยำ (false positive rate / Precision: Pr) เป็นอัตราส่วนของการค้นพบภาพที่ถูกต้องจากจำนวนภาพทั้งหมดที่ทำการค้นหามาได้

$$\Pr = \frac{a}{(a+b)}, a+b > 0$$

- ค่าความระลึก (true positive rate / Recall: Re) เป็นอัตราส่วนของการค้นพบภาพที่ถูกต้องจากจำนวนภาพที่ถูกต้องทั้งหมด

$$\mathrm{Re} = \frac{a}{(a+c)}, a+c > 0$$

- ค่าความถูกต้อง (accuracy: Acc) เป็นอัตราส่วนของการค้นพบภาพที่ถูกต้องทั้งหมดจากจำนวนภาพที่มีอยู่

$$Acc = \frac{(a+d)}{(a+b+c+d)}$$

- ค่า F-measure เป็นการวัดค่าความสัมพันธ์ระหว่างค่าความระลึกและค่าความแม่นยำในเชิงฮาร์โมนิค (harmonic) เหมาะสำหรับฐานข้อมูลสารสนเทศที่มีขนาดใหญ่มาก และมักจะไม่ทราบว่าข้อมูลภาพที่ถูกต้องทั้งหมดมีอยู่เท่าใด ทำให้ต้องทำการประมาณโดยใช้การสุ่มตัวอย่าง (sampling) ตามหลักทางสถิติหรือด้วยวิธีอื่นด้วย โดยทั่วไปจะเป็นการหาค่า F-measure ซึ่งแสดงสูตรได้ดังนี้

$$F = \frac{2(\Pr \cdot \mathrm{Re})}{(\Pr + \mathrm{Re})}$$

บทที่ 3
ขั้นตอนวิธีการดำเนินงานวิจัย

ในการศึกษางานวิจัยในครั้งนี้ได้มีการเก็บรวบรวมข้อมูลภาพและคัดเลือกภาพที่เหมาะสมเพื่อเตรียมเป็นข้อมูลภาพเบื้องต้น ดังนั้นข้อมูลภาพที่เตรียมพร้อมจะสามารถเข้าสู่กระบวนการประมวลผลภาพ โดยในงานวิจัยจะมีการแบ่งขั้นตอนวิธีการดำเนินงานวิจัยโดยทั่วไปจะแบ่งออกเป็น 3 ส่วนหลักดังนี้

3.1 ขั้นตอนการเตรียมข้อมูล (data preprocessing) เป็นการทำงานในส่วนของการนำข้อมูลเข้าด้วยเครื่องมือ (image annotation tool) และการแทนวัตถุลงในกราฟ (object representation into graph)

3.2 ขั้นตอนการประมวลผล (data processing) ทำการนำข้อมูลที่ได้ทั้งหมดมาทำงานบนกราฟแบบลำดับขั้น

3.3 ขั้นตอนการวัดประสิทธิภาพความคล้ายกันของภาพ (evaluation of similarity measurement) เป็นการเปรียบเทียบการทำงานของวิธีการที่นำเสนอโดยขั้นตอนทั้งหมดจะสามารถอธิบายได้ดังภาพที่ 3.1

ภาพที่ 3.1 ขั้นตอนการวัดความคล้ายกันของกลุ่มความหมายภาพ

3.1 ขั้นตอนการเตรียมข้อมูล

ขั้นตอนการเตรียมข้อมูล (data preprocessing) โดยการทำการแยกและคัดเลือกข้อมูลภาพดิจิทัลที่มีวัตถุบนภาพที่เด่นชัด มีวัตถุภาพพื้นหลัง และภาพที่คัดเลือกเข้ามานั้น

สามารถให้มนุษย์แปลความหมายภาพนั้นได้อย่างสมบูรณ์ สำหรับภาพบางภาพจะไม่นำเข้ามาทำการทดลองนั้นจะเป็นภาพที่มีความหมายกำกวม ภาพไม่มีความหมาย แปลความหมายไม่ได้หรือภาพที่มนุษย์แปลได้หลายความหมาย ภาพที่มีการโฟกัสระยะใกล้ ข้อมูลภาพที่มีความซ้ำซ้อน หรือไม่สอดคล้องกันจะถูกคัดเลือกภาพนั้นออกไป และทำการรวบรวมข้อมูลภาพที่ต้องการที่มาจากหลายฐานข้อมูลจุดประสงค์ก็เพื่อทำให้มั่นใจว่าคุณภาพของข้อมูลที่ถูกเลือกนั้นเหมาะสม ดังนั้นกระบวนการทั้งหมดนี้จะประกอบด้วย 2 กระบวนการดังนี้

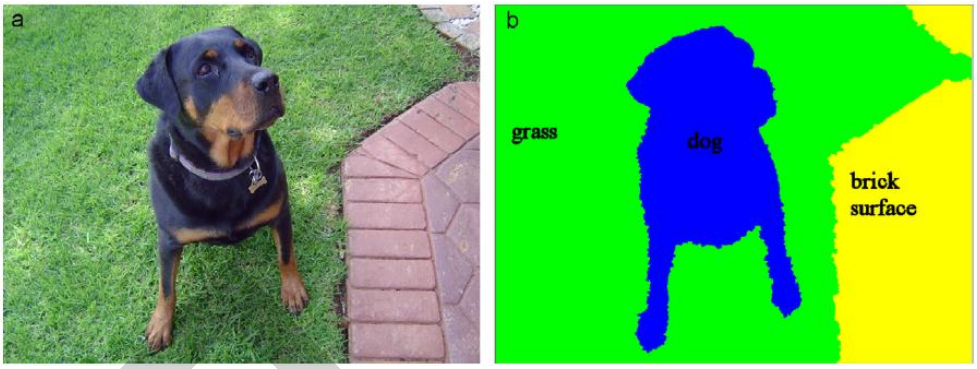
3.1.1 ฐานข้อมูลรูปภาพ

แหล่งข้อมูลภาพมีหลายแหล่งข้อมูลที่ได้รับการยอมรับและสามารถนำมาใช้เป็นฐานข้อมูลรูปภาพได้ เช่น Pascal[2] Fotosearch stock[3] The Cobis Stock[4] The Corel Corporation[5] เป็นต้น Pascal เป็นทีมงานวิจัยหนึ่งที่พยายามจะสกัดข้อมูลภาพที่ได้เป็นชื่อของวัตถุที่ถูกกำหนดไว้ซึ่งในงานวิจัยได้มีการทดลองถึงวัตถุหลายอย่าง เช่น person, bird, cat, cow, dog, bottle, chair เป็นต้น สำหรับแหล่งข้อมูลภาพที่หลากหลาย อาจจะมีภาพที่ไม่เหมาะสมกับการทดลองที่นำเสนอเนื่องจาก ภาพบางภาพมีลักษณะผิดปกติ (outlier) หรือ คุณลักษณะวัตถุ (object characteristic) ไม่ชัดคลุมเครือ มีขนาดวัตถุขนาดเล็กเกินไปไม่สามารถ บ่งชี้ชื่อวัตถุได้ ภาพถ่ายระยะใกล้ (close up) ภาพบางภาพอาจจะไม่สามารถแปลความหมาย หรือภาพมีความหมายกำกวมจนทำให้ไม่สามารถหาความหมายภาพได้ ทำให้ต้องมีการคัดเลือกภาพออกไปไม่นำมาใช้ในการทดลอง ดังนั้นในการหาแหล่งข้อมูลของการนำภาพเข้ามาใช้จึงจำเป็นต้องสมบูรณ์ที่สุด ภาพจะต้องความเหมาะสมกับงานที่จะนำมาใช้เพื่อตอบสนองกับความต้องการของการทดลองมากที่สุด

สำหรับการทดลองโดยทั่วไปในสาขาคอมพิวเตอร์วิชัน (computer vision) ในส่วนของการประมวลผลภาพระดับสูง รูปภาพที่นำเข้ามาทดลองจะเป็นภาพที่วัตถุถูกแท็ก หรือกระทำการให้ความหมายมาล่วงหน้าก่อน จะเรียกภาพจำพวกนี้ว่า annotated images [Feichao Wang, 2011] ดังแสดงในภาพที่ 3.2 แสดงรูปภาพถูกแท็กด้วยคำหลัก[Allan Hanbury,2008] ดังนั้นในการทดลองจะต้องมีฐานข้อมูลภาพที่สมบูรณ์เพียงพอที่จะสามารถนำส่วนของคำหลักที่ถูกแท็กมาใช้งานได้โดยไม่มีผลข้างเคียงต่อกระบวกการที่นำเสนอ ส่วนใหญ่แหล่งข้อมูลภาพจะทำการคัดเลือกบริเวณ (region) ที่เหมาะสมสำหรับการแท็กเป็นคำหลักหรือคำสำคัญ (keyword) เพื่อ

[2] Pascal: http://host.robots.ox.ac.uk/pascal/VOC/voc2012/
[3] Fotosearch Stock: http://www.fotosearch.com
[4] The corbis Stock: http://pro.corbis.com
[5] The Corel Corporation: http://www.corel.com/

ใช้สำหรับการสืบค้นข้อมูล ดังนั้นบางแหล่งข้อมูลจะใช้การแบ่งหมวดหมู่วัตถุ (object categories) เข้ามาช่วยในการจัดกลุ่มของคำที่จะนำมาทดแทนส่วน บริเวณนั้น ๆ



ภาพที่ 3.2 ตัวอย่างที่ถูกแท็กด้วยคำหลัก [Allan Hanbury,2008]

3.1.2 ฐานข้อมูลคำหลัก WordNet

WordNet [Miller G.A., 1990] [Zinger S., 2005] คือฐานข้อมูลคำศัพท์ภาษาอังกฤษขนาดใหญ่ของ คำนามคำกริยาคำคุณศัพท์และกริยาวิเศษณ์มีการจัดกลุ่มเป็นชุดของคำพ้องความรู้ความเข้าใจ (synsets) แต่ละการแสดงแนวคิดที่แตกต่างกัน Synsets จะเชื่อมโยงกันโดยวิธีการของความสัมพันธ์ของแนวคิดและความหมายคำศัพท์ เครือข่ายที่เกิดจากคำที่เกี่ยวข้องกับความหมายและแนวคิดสามารถสำรวจที่มีเบราว์เซอร์ WordNet ยังเป็นอิสระและเปิดเผยต่อสาธารณชนพร้อมสำหรับการดาวน์โหลด โครงสร้างมโนทัศน์ทำให้เครื่องมือที่มีประโยชน์สำหรับภาษาศาสตร์คอมพิวเตอร์และการประมวลผลภาษาธรรมชาติ

WordNet เป็นกลุ่มคำที่มันขึ้นอยู่ร่วมกันตามความหมาย WordNet เป็นฐานข้อมูลเกี่ยวกับคำศัพท์ที่ใช้ภาษาอังกฤษที่ประกอบด้วยคำนาม คำกริยา คำคุณศัพท์ และคำต่าง ๆถูกจัดไว้เป็นหมวดหมู่ซึ่งเป็นงานวิจัยที่ได้รับความนิยมมากมีนักวิจัยหลายกลุ่มที่อ้างอิงการใช้ฐานข้อมูลของคำหลักที่แท็กบนภาพ [Adrian Popescu, 2008][Javier Álvez, 2008] โดย WordNet จะใช้คำนาม (noun) ถูกเลือกจากพจนานุกรม (Dictionary) เป็นคำหลัก โดยทั่วไปแล้วมาตรฐานอักขระจะมีการจัดการข้อมูล lexical เพื่อให้มีความสอดคล้องกันทั้งการสะกดคำ การกระจายความหมายเหมือนของคำ ความสัมพันธ์ของคำ มีการสร้างฐานข้อมูลคำศัพท์ออนไลน์ด้วยคอมพิวเตอร์ที่เรียกว่า lexical database เพื่อค้นหาแต่ละตัวอักษรได้อย่างรวดเร็ว WordNet เป็นส่วนของโปรแกรมที่พยายามรวมข้อมูล lexicographic ที่มีการคำนวณอย่างรวดเร็ว ซึ่งจะมีการจัดการข้อมูล lexical ในเทอมของความหมายจนถึงรูปแบบของคำศัพท์โดยจะใช้ Lexical

Matrix ช่วยในการจัดการกฎไวยากรณ์ของคำและลดรูปแบบของคำ (word form) คำกำกวมความหมายของคำ (word meaning) โดยจะใช้รูปแบบของการ mapping ระหว่าง form และความหมาย [Miller,1986] ซึ่งจะสนใจเรื่องของกลุ่มไวยากรณ์คำโดยจะใช้การ mapping ที่มีรูปแบบต่างกัน จากตารางที่ 3.1 แสดงรูปแบบของ lexical เมตริกซ์โดยส่วนแนวคอลัมแสดงถึงรูปแบบของคำศัพท์ (Word Forms: $F_1,F_2,\ldots,F_n$) และแนวนอนจะแทนความหมายของคำศัพท์ (Word Meanings: $M_1,M_2,\ldots,M_m$) ผลลัพธ์ที่ได้คือ Expression ($E_{1,2},E_{1,3},\ldots,E_{n,m}$)ที่เป็นไปได้ การใช้แค่ lexical เมตริกซ์ยังทำให้เกิดความหมายที่ผิดพลาดอยู่ ตัวอย่างเช่น ต้นไม้“trees” มี ราก“roots”ที่ประกอบไปด้วยเซลต่าง ๆเพื่อใช้ในการดำรงอยู่ แต่ถ้ามองอีกแง่หนึ่งในเทอมของต้นไม้ “plant” เป็นอีกความหมายหนึ่งที่สามารถใช้ได้เช่นเดียวกัน ถ้าผู้อ่านไม่ได้ตีความเป็นอย่างอื่น เมื่อมีรูปประโยคที่ใช้ความหมายของคำที่มีความสอดคล้องกัน แต่ถ้ามอง “plant” เป็นสถานที่ใดที่หนึ่ง “place”ของโรงงานซึ่งสามารถใช้ได้เช่นเดียวกัน ในกรณีของคำศัพท์ ต้นไม้ “tree” ซึ่งต้นไม้จะมี เปลือกไม้ “bark” และ กิ่ง “twigs” ต้นไม้เจริญเติบโตจากเมล็ดพันธ์ ต้นไม้ใหญ่จะมีความสูงมากกว่ามนุษย์ เมื่อมีจำนวนต้นไม้เยอะ ๆจะถูกเรียกต้นไม้ที่อยู่รวม ๆเป็นป่า “forest” ไปโดยปริยายเนื้อไม้ “wood” ถูกแปรรูปเป็นน้ำมันเชื้อเพลิงหรืออื่น ๆได้มากมาย ดังนั้นข้อมูลเกี่ยวกันต้นไม้ที่มีความสัมพันธ์กันเหล่านี้จะต้องถูกเชื่อมโยงอยู่ในความหมายของต้นไม้ “tree” เช่นเดียวกันดังนั้นจึงต้องมีระบบของการสืบทอด lexical (Lexical inheritance systems)

ตารางที่ 3.1 แนวคิด Lexical เมตริกซ์

Word Meanings	Word Forms F_1	F_2	F_3	. . .	F_n
M_1	$E_{1,1}$	$E_{1,2}$			
M_2		$E_{2,2}$			
M_3			$E_{3,3}$		
⋮				. . .	
M_m					$E_{m,n}$

ความสัมพันธ์ของความหมายบน WordNet [G. Varelas 2005] มีดังนี้

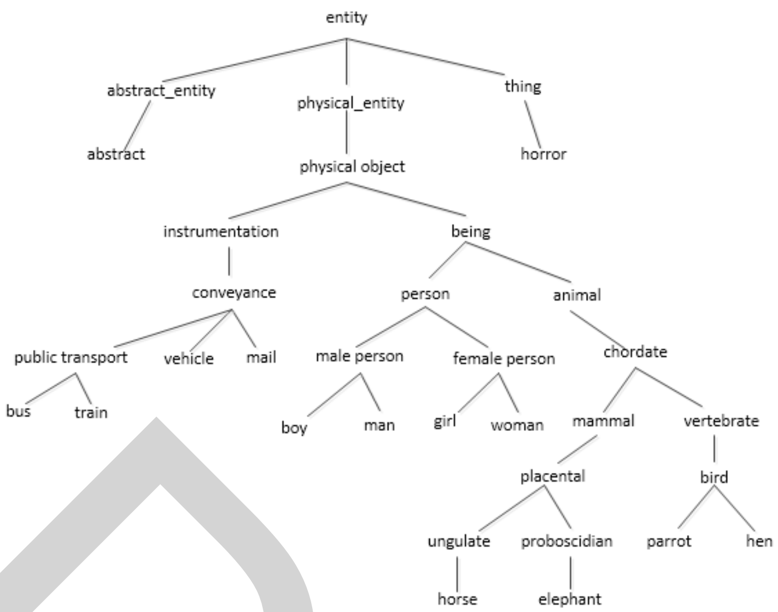
- Synonymy คือ เวิร์ดเน็ตใช้เซ็ตของคำที่มีความหมายคล้ายคลึงกัน(synsets, synonym sets) ในการแทนคำที่มีความหมาย เช่น pipe, tube เป็นต้น
- Antonymy คือคำที่มีความหมายตรงกันข้ามกันอย่างสมมาตรซึ่งเน้นที่คำ ประเภทคำคุณศัพท์และคำกริยาวิเศษ เช่น wet, dry เป็นต้น

- Hyponymy และ Hypernymy คือความสัมพันธ์ของคอนเซ็พที่มีลักษณะของ concept และ super concept โดย Hyponym คือคอนเซ็พที่อยู่เหนือคอนเซ็พอื่น ส่วน Hypernym คือคอนเซ็พที่อยู่ภายใต้ คอนเซ็พอื่น เช่น เซ็ตของคำที่แสดงลำดับชั้น {robin,redbreast} @{bird} @{animal,animate_being} @ {organism,life_form,living_thing}@ ชี้ไปยังคำที่อยู่ในลำดับที่สูงกว่า
- Meronymy และ Holonymy คือความสัมพันธ์แบบบางส่วนและทุกส่วน ตัวอย่าง แขน ขา มือ เป็นส่วนหนึ่งของร่างกาย หรือ พวงมาลัย เบรก ล้อ เป็นส่วนหนึ่งของรถ เป็นต้น
- Troponymy คือคำกริยาเพื่อแสดงอาการหรือการกระทำ ซึ่งใช้สำหรับการแสดงการกระทำของคำนามในลำดับชั้นของ concept
- Entailment คือความสัมพันธ์ระหว่างคำกริยาซึ่งกำหนดโดยเวิร์ดเน็ต

ระบบการสืบทอด lexical

ระบบการสืบทอด lexical (Lexical inheritance systems) เป็นการใช้คำหลักเหล่านั้นมาจัดกลุ่ม (class) ตามลำดับชั้น (hierarchy) ถึง 73,733 กลุ่มซึ่งจะมีคำนามที่ถูกจัดกลุ่มอยู่ภายในกิ่งก้าน (leave) ของลำดับชั้นมากมายถึง 60,000 กิ่ง และมีคำนามที่ใช้ถึง 116,364 คำ ดังแสดงโครงสร้างคำหลักของ WordNet ในภาพที่ 3.3 แสดงกิ่งก้านลำดับชั้นของคำว่า “elephant” จะอยู่ในกิ่งลำดับชั้นหรือรากฐานมาจากคำหลัก “animal” จากภาพแสดงถึงความสัมพันธ์ของคำหลักกับโครงสร้างใน WordNet เพื่อไปเรียกใช้ โดย WordNet จะถูกใช้คู่กับโปรแกรมเพื่อทำการสืบค้นข้อมูล และยังสามารถกำหนดความสัมพันธ์ของคำด้วยโครงสร้างของ WordNet เอง และสามารถสร้างความหมายเหมือนกันได้

พื้นฐานการออกแบบของ lexicographer สามารถเขียนเป็นความหมายของคำนามได้ ตัวอย่าง ความหมายของ “tree” สามารถเขียนคุณสมบัติของ “tree” ในรูปแบบกราฟได้ดังนี้ oak@ →tree@ →plant@ →organism โดยกำหนดให้ “@ →” คือการแทนค่าความสัมพันธ์กันในลักษณะของความหมายที่มีความเกี่ยวพันธ์กันสามารถอ่าน “is a” หรือ “is a kind of” โดยที่จะมีความสัมพันธ์กันตามลำดับขั้นจากล่างขึ้นบน สามารถกล่าวได้ว่า $w_h@ \rightarrow w_s$ และถ้าเขียนกลับกันสามารถเขียนได้ $w_s\sim \rightarrow w_h$ และกล่าวได้ว่า w_s เป็นซุปเปอร์คลาสของ w_h



ภาพที่ 3.3 ตัวอย่างจำลอง Synset ในอนุกรมวิธาน WordNet

ความหมายของคำหลักที่ถูกนำมาใช้บน WordNet นั้นถูกจัดลงบนลำดับขั้นตามความหมายที่สามารถเป็นไปได้โดยที่ WordNet ได้กำหนดรายการหัวข้อหลักต่าง ๆตั้งต้นไว้ดังแสดงในภาพที่ 3.4 เป็นกลุ่มหัวข้อเริ่มต้นเป็นลำดับบนสุดของกราฟจะเป็น {entity} {object, thing} และ {idea} เป็นหัวข้อเริ่มต้นดังนั้นคำศัพท์ที่ถูกใช้จะอยู่ภายในหัวข้อต่าง ๆที่กำหนดซึ่งบางคำมีหลายความหมายสามารถอยู่ได้ในกลุ่มหลายกลุ่มตามความหมายที่เกิดขึ้น

{act, action, activity}
{animal, fauna}
{artifact}
{attribute, property}
{body, corpus}
{cognition, knowledge}
{communication}
{event, happening}
{feeling, emotion}
{food}
{group, collection}
{location, place}
{motive}
{natural object}
{natural phenomenon}
{person, human being}
{plant, flora}
{possession}
{process}
{quantity, amount}
{relation}
{shape}
{state, condition}
{substance}
{time}

ภาพที่ 3.4 รายการคำหลักเริ่มต้นบน WordNet

{thing, entity}
- {living thing, organism}
  - {plant, flora}
  - {animal, fauna}
  - {person, human being}
- {non-living thing, object}
  - {natural object}
  - {artifact}
  - {substance}
  - {food}

ภาพที่ 3.5 รายการแสดงความสัมพันธ์ระหว่างหัวข้อหลักบน WordNet

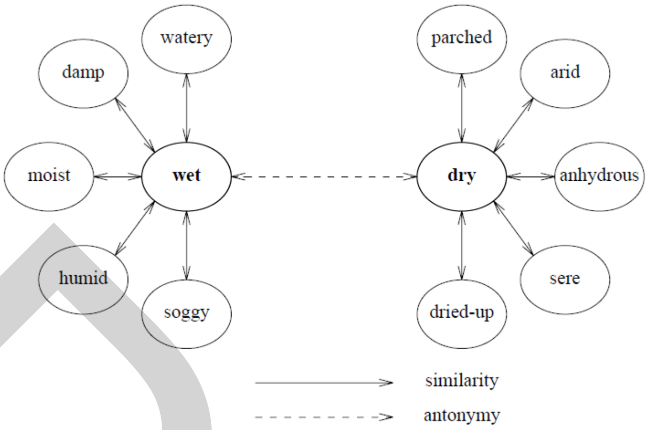
ดังนั้นความสัมพันธ์ที่เกิดขึ้นแต่ละวัตถุนั้นไม่จำเป็นต้องมีระดับความลึกหรือความสัมพันธ์ที่เท่ากันเสมอไปตัวอย่างเช่น Shetland pony เป็น pony, a horse, an equid, an odd-toed ungulate, a herbivore, a mammal, a vertebrate, และ an animal ตามลำดับจนถึงระดับสูงสุดคือ animal บางครั้งอาจจะมีระดับถึง 11 ก็เป็นไปได้หรือระดับของ person อาจจะเพียง 5-6 ระดับเช่น *roadster @→ car @→ motor vehicle @→ wheeled vehicle @→ vehicle @→ conveyance @→ artifact* หรือ *televangelist @→ evangelist @→ preacher @→ clergyman @→ spiritual leader @→ person* เป็นต้น

คำคุณศัพท์ในภาษาอังกฤษ

จะเห็นว่าความหมายของคำศัพท์จะต้องมีการแจกแจงคุณสมบัติของความหมายของคำเพื่อความชัดเจน คำศัพท์ในภาษาอังกฤษมีได้หลายความหมายดังที่กล่าวมาแล้ว ขึ้นกับความสัมพันธ์ของคำที่อยู่ใกล้เคียงกันด้วย เช่น “chair” ถ้าคำนี้อยู่โดด ๆ อาจจะหมายถึงเก้าอี้หรือ ผู้พูด (speaker) ก็เป็นได้แต่ถ้ามีคำคุณศัพท์เข้ามาเพิ่มสามารถที่จะเฉพาะเจาะจงได้มากขึ้นเช่น larger หรือ comfortable ตัวอย่างเช่น the creaking chair, the overstuffed chair หรือ armchair, barber chair เป็นต้น ในบางครั้งอาจจะเป็นวลีหรือกลุ่มคำนาม เช่น chair with green upholstery, my grandfather’s chair เป็นต้น ทำให้เห็นว่าความหมายของคำว่า chair สามารถแปรเปลี่ยนได้ตามคำศัพท์ที่อยู่ใกล้เคียง WordNet มีคำคุณศัพท์ประมาณ 19,500 คำ และถูกจัดเก็นตามความหมายของคำประมาณ 10,000คำ ซึ่งจะมีการจัดเก็บคำคุณศัพท์ เช่น big, interesting, possible และความสัมพันธ์ของคำคุณศัพท์ เช่น presidential และ nuclear เป็นต้น

ปัญหาที่เกิดขึ้นของคำศัพท์นั้นคือ จะหาความสัมพันธ์ของคำศัพท์ที่เกิดขึ้นได้อย่างไร ได้มีการแทนค่าความสัมพันธ์ที่เกิดขึ้นด้วย {heavy, weighty, ponderous} และ {light, weightless, airy} คนทั่วไปจะรู้ว่า heavy และ light เป็นคำที่มีความหมายต่างกัน (antonym) แต่ในบางครั้งบางคำจะมีความต่างเพียง suffix เช่น weighty กับ weightless หรือ heavy และ

weightless เป็นต้น เป็นการยากที่จะทำการกำหนดรูปแบบที่เกิดขึ้น ดังนั้นได้มีการกำหนดกลุ่มคำที่มีความหมายคล้ายกันด้วยความสัมพันธ์ ดังแสดงในภาพที่ 3.6 และตารางที่ 3.2 แสดงความสัมพันธ์ที่เกิดขึ้นระหว่างคำคุณศัพท์ในภาษาอังกฤษ



ภาพที่ 3.6 ตัวอย่างการแสดงความสัมพันธ์ของคำหลักที่มีความหมายเหมือนและต่างกัน[6]

ตารางที่ 3.2 ตัวอย่างความสัมพันธ์ของคำคุณศัพท์

SIZE	WHITENESS	AGE	VIRTUE	VALUE	WARMTH
astronomical	snowy	ancient	saintly	superb	torrid
huge	**white**	**old**	**good**	great	**hot**
large	ash-gray	middle-aged	worthy	**good**	warm
standard	gray	mature	ordinary	mediocre	tepid
small	charcoal	adolescent	unworthy	**bad**	cool
tiny	**black**	**young**	**evil**	awful	**cold**
infinitesimal	pitch-black	infantile	fiendish	atrocious	frigid

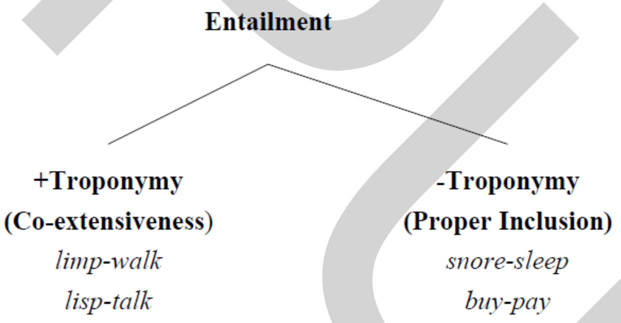
คำกริยาในภาษาอังกฤษ

คำกริยาเป็นคำที่เป็นความหมายหลักของประโยคเพื่อแสดงการกระทำที่เกิดขึ้นอย่างชัดแจ้ง ดังนั้นการแปลความหมายของภาพหรือประโยคจะไม่สามารถขาดสิ่งนี้ได้ ใน *Collins English Dictionary* ได้บรรจุคำนามที่มีความแตกต่างกันทั้งหมด 43,636 คำ และคำกริยา 14,190 คำที่แตกต่างกัน โดยส่วนใหญ่แล้วจะมีคำกริยาที่ถูกใช้บ่อย ๆ เช่น have, be, run, make, set, go และคำอื่น ๆ แต่ในคำศัพท์ที่เป็นตัวเดียวกันบางครั้งอาจมีความหมายที่แตกต่างกันได้ เช่น I have a Mercedes. และ I have a headache. กริยา have แต่มีความหมายที่ต่างกันตามคำที่อยู่รอบ ๆ และเช่นเดียวกันคำกริยาบางคำมีความหมายที่เหมือนกันเช่น {beat, strike, hit}, {beat, throb, pulse}, {beat, flatten} เป็นต้น จะเห็นว่า beat มีความหมายที่แตกต่างกันตามกลุ่มคำ ดังนั้นจะต้องมีการจัดการตามความหมายของคำกริยาไว้ เช่นเดียวกันกับความต่างของคำ

[6] รูปภาพ อ้างอิง http://www.corbisimages.com/ ค้นคืนเมื่อวันที่ 9 เมษายน 2560

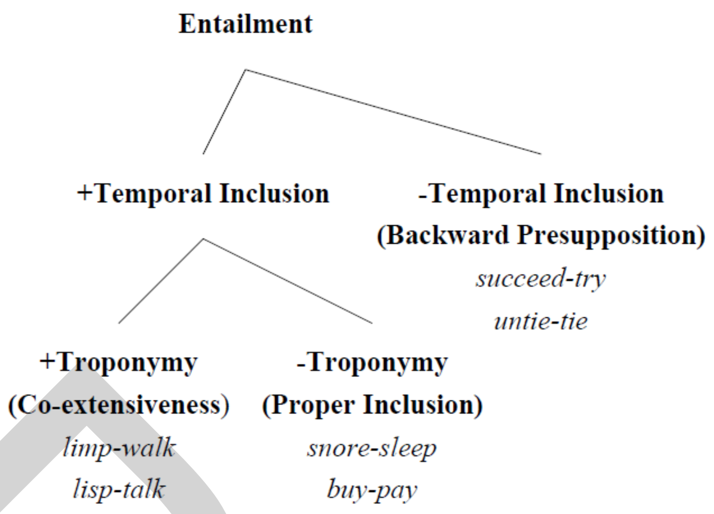
drive และ ride, snore และ sleep เป็นคู่ความสัมพันธ์ของคำกริยาที่เป็นกิจกรรมที่เกิดขึ้นซึ่งมีความเกี่ยวพันกัน อาจจะมีการตกทอดต่อมาหรือขยายจากความสัมพันธ์ของคำกริยาช่องที่ 1 หรือ 2 ก็ตาม เช่น buy และ pay , snore และ sleep เป็นคำกริยาที่มีความสัมพันธ์เกี่ยวพันกันด้วยกิจกรรม

ในบางครั้งคำศัพท์เดียวกันอาจจะมีความแตกต่างของคำกริยาที่เกิดขึ้นเมื่อมีการเปลี่ยนแปลงคำนามภายในประโยค เช่น “A horse is an animal หรือ A spade is a garden tool” และ “Ambling is walking หรือ Mumbling is talking” แสดงให้เห็นว่าคำกริยาเดียวกันมีความหมายที่แตกต่างกันตามคำนาม จะเห็นว่าคำนามบางคำที่ถูกจัดให้เป็นกลุ่ม เช่น bus, truck, bike เป็นต้น แต่อย่างไรก็ตามคำกริยาในภาษาอังกฤษนั้นมีระดับความแตกต่างของความหมายในเชิงของ ค่าน้ำหนัก แรง ความเร็ว เช่น run หรือ stroll, slide หรือ pull เป็นต้น จากภาพที่ 3.7 แสดงตัวอย่าง 2 กลุ่มคำที่มีความแตกต่างกันตามไดอะแกรม ซึ่งในบางครั้งความหมายที่เกิดขึ้นนั้นไม่สามารถแทนค่าลงบนโครงสร้างต้นไม้ได้ครบถ้วน เช่น เมื่อมี 2 โหนดความหมายดังนี้ {move, make a movement} และ {move, travel} ซึ่งความหมายอาจจะมีได้ถึง 3 ความหมายดังนี้ {give, transfer}, {take, receive} และ {have, hold} เป็นต้น

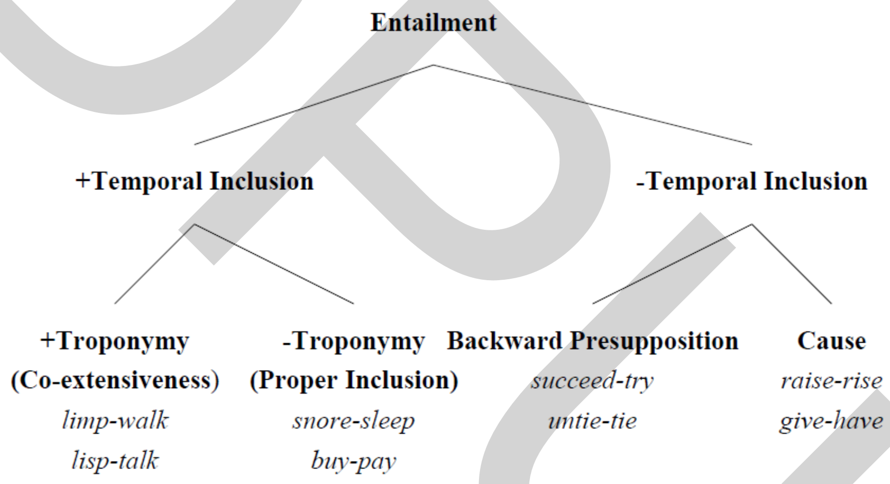


ภาพที่ 3.7 ตัวอย่างรูปแบบ 2 Entailment

การแสดงความหมายของคำกริยาจึงใช้การแทนของความหมายที่ตรงกันของ Entailment (+) และความที่แตกต่างกัน (Opposition manner (-))แทนเป็นการแทนความสัมพันธ์ที่เกิดขึ้นของคำศัพท์ที่แตกต่างกันดังแสดงในภาพที่ 3.8 เช่น “hit และ miss” ของ entail “aim” ซึ่งสามคำศัพท์นี้มีความสัมพันธ์กันเป็นลำดับ hit และ miss จะเกิดได้เมื่อมี “aim” ตามลำดับ เช่นเดียวกันกับคำศัพท์ “fail และ succeed” ที่จะต้องมี entail “try” มาก่อน หรือ “win และ lose” จะต้องมี entail ของ “play หรือ gamble” มาก่อนเช่นกัน ดังแสดงตัวอย่างของความสัมพันธ์ในภาพที่ 3.9



ภาพที่ 3.8 ตัวอย่างรูปแบบของ 3 Entailment



ภาพที่ 3.9 ตัวอย่างรูปแบบของ 4 Entailment

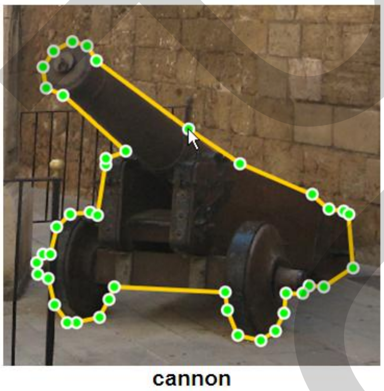
รูปแบบของ WordNet ที่มีการจัดเก็บความสัมพันธ์ของข้อมูลคำศัพท์ทั้งที่เกิดจากคำนาม คำคุณศัพท์ และคำกริยาที่พยายาให้เกิดความสอดคล้องความหมายภายในประโยคที่เกิดขึ้นเพื่อให้สื่อถึงความหมายที่แท้จริงโดยการใช้ลำดับขั้นของโครงสร้างต้นไม้ สำหรับในงานวิจัยนี้เป็นการนำคำศัพท์ที่ได้จากการแท็กบนภาพเข้ามาใช้เพื่อทำการแปลความหมายของภาพดังนั้นจะเลือกใช้รูปแบบของคำหลักที่เกิดจาก WordNet เพียงบางส่วนเท่านั้น ซึ่งคำคุณศัพท์จะไม่ได้ถูกนำมาใช้ในการทดลองในครั้งนี้

3.1.3 การแท็กคำหลักบนภาพ

สิ่งที่สำคัญถัดจากการเลือกใช้ฐานข้อมูลสำหรับกระบวนการของการประมวลผลภาพระดับสูง (high- level image processing) คือ การให้คำอธิบายภาพ (annotated images) โดยใช้ฐานข้อมูลคำหลักที่คัดเลือกมา การแท็กคำหลักบนภาพ โดยทั่วไป ภาพ (image) จะประกอบด้วย พื้นหลัง (background) และ พื้นหน้า (foreground) กล่าวคือภาพหนึ่งภาพจะประกอบด้วยวัตถุ (objects) หลายวัตถุ ทำให้ทุกวัตถุควรจะถูกแท็กด้วยคำหลักที่เหมาะสมจากฐานข้อมูล ดังนั้นในการแบ่งแยกวัตถุ (segmentation) เป็นหัวข้อที่มีการวิจัยอย่างต่อเนื่องในการประมวลผลภาพระดับต่ำ(low-level image processing) [Qian Huang, 1995][Vailaya A., 2001] ด้วยคุณลักษณะของภาพหลายคุณลักษณะ (features) เพื่อทำการแบ่งแยกให้ได้วัตถุที่สมบูรณ์แบบ แต่อย่างไรก็ตามการแบ่งแยกวัตถุที่มีรูปร่างลักษณะคล้ายคลึงกันหรือแตกต่างกันแต่มีความหมายเดียวกัน ยังคงเป็นเรื่องที่ค่อนข้างยากสำหรับการประมวลผลภาพระดับต่ำรวมถึงกระบวนการรู้จำรูปแบบวัตถุ (pattern recognition) เพื่อบอกความหมายของวัตถุหรือชื่อของวัตถุ แต่อย่างไรก็ตามในการวิเคราะห์ความหมายของภาพในงานวิจัยนี้ได้เฉพาะเจาะจงในส่วนการประมวลผลระดับสูง จึงได้ข้ามในส่วนของการแบ่งแยกวัตถุ และ การรู้จำรูปแบบวัตถุ ดังนั้นในส่วนที่กล่าวต่อไปสำหรับการประมวลผลภาพระดับสูงคือการให้คำอธิบายภาพด้วยการแท็ก (tag) [Ismail Haritaoglu, 1998] [Tele Tan, 2002] [Vasileios Mezaris, 2003] [R. Zhao, 2002] ข้อมูลภาพให้อยู่ในรูปของชื่อวัตถุ หรือคำหลักจากฐานข้อมูล

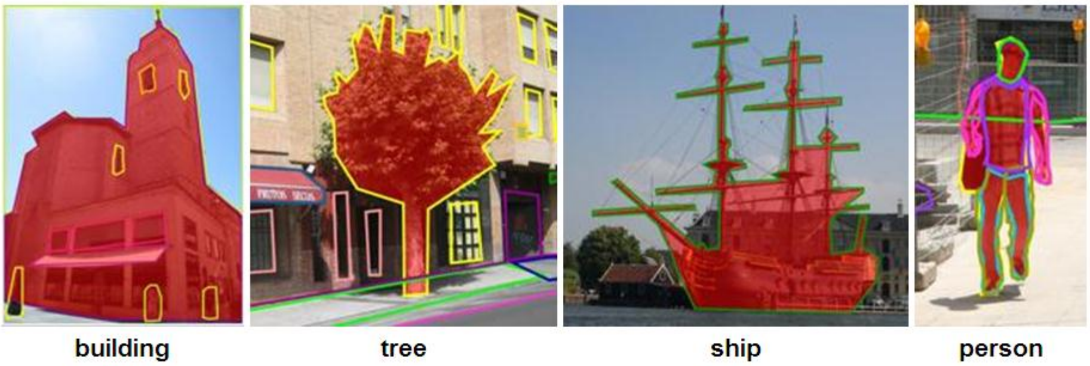
มีหลายกลุ่มงานวิจัยที่คิดค้นกระบวนการให้ความหมาย หรือแท็กวัตถุบนภาพด้วยวิธีการแตกต่างกัน จากกลุ่มเครื่องมือที่ให้ความหมายภาพที่นิยม [Jeroen Steggink, 2011] ในงานวิจัยนี้ได้เลือกเครื่องมือแบบออนไลน์ โดยใช้โปรแกรม LabelMe[9] [B. C. Russell,2008] [A. Torralba,2010] เป็นเครื่องมือที่ได้รับการยอมรับอย่างกว้างขวาง สำหรับงานวิจัยทางด้าน Computer Vision โดยแอพพลิเคชันนี้สามารถทำงานได้อย่างเต็มรูปแบบบนเว็บในลักษณะของเครื่องมือให้ความหมาย (Web-based annotation tools) เริ่มตั้งแต่ปี 2005 ปัจจุบันมีวัตถุบนภาพที่ถูกให้ความหมายรวมทั้งสิ้น 400,000 วัตถุ [Von Ahn and L. Dabbish, 2004.] [B. C. Russell, 2008] [A. Sorokin and D. Forsyth, 2008] [M. Spain and P. Perona, 2007] [D. G. Stork, 1999] ผู้ใช้สามารถเข้าถึงโปรแกรมผ่านทางเครือข่ายออนไลน์ได้ สามารถแท็กวัตถุผ่านทางออนไลน์ได้ดังแสดงในภาพที่ 3.10 แสดงหน้าเว็บไซต์ของโปรแกรม LabelMe โปรแกรมสามารถทำงานร่วมกันได้หลายเพลตฟอร์ม ซึ่งเป็นวัตถุประสงค์หลักของคณะผู้จัดทำโปรแกรมนี้ ทำให้ผู้ใช้งานที่เข้ามาให้ความหมายภาพมาได้มากมาย และมีพื้นฐานของการให้ความหมายที่แตกต่างกันตามความสามารถ ของแต่ละบุคคล

ภาพที่ 3.10 โปรแกรม LabelMe บนเบราว์เซอร์[7]



ภาพที่ 3.11 การเลือกสัดส่วนของวัตถุบนภาพ[9]

[7] LabelMe: http://labelme.csail.mit.edu/ ค้นคืนเมื่อวันที่ 9 เมษายน 2560

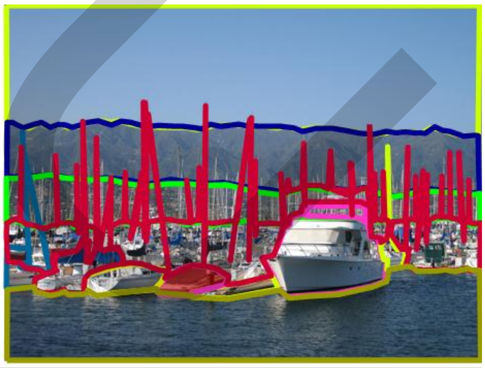


ภาพที่ 3.12 ตัวอย่างวัตถุที่ถูกแท็กด้วยโปรแกรม LabelMe[8]

คำหลัก: car traffic light car taxi building sky

คำหลัก: person woman sea person woman sky sea sand

คำหลัก: bus person wheel building



คำหลัก: boat sea water sky mountain wheel building

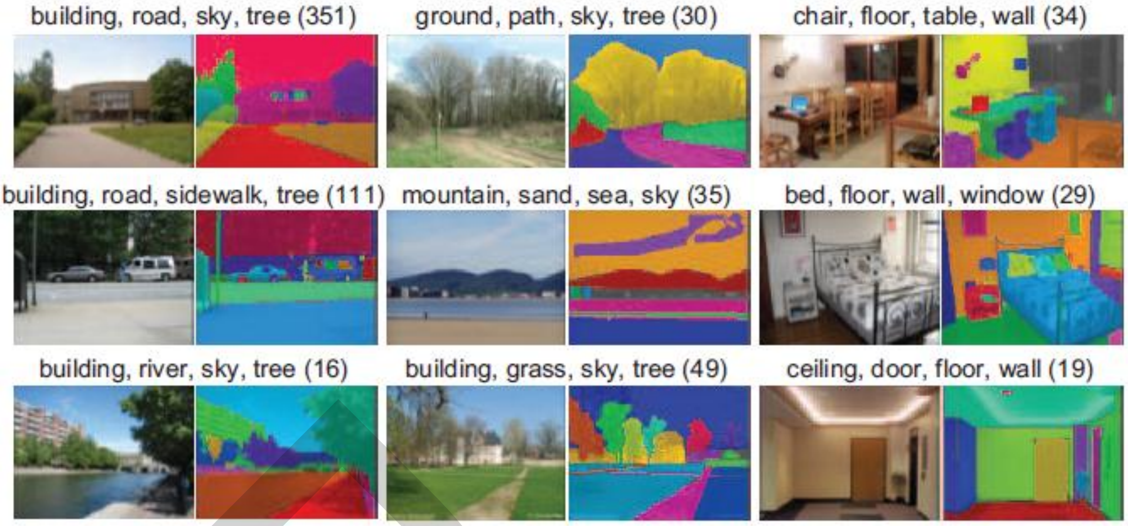
ภาพที่ 3.13 ตัวอย่างภาพที่ถูกแท็กคำหลักบนภาพ ด้วยโปรแกรม LabelMe[10]

[8] LabelMe: http://labelme.csail.mit.edu/ ค้นคืนเมื่อวันที่ 9 เมษายน 2560

รูปแบบของการให้ความหมายภาพ หรือแท็กข้อมูลบนโปรแกรม LabelMe จากภาพที่ 3.10 ประกอบด้วยเครื่องมือช่วยการแท็ก (ด้านบน) รูปภาพสามารถคัดเลือกจากฐานข้อมูลภายในโปรแกรม หรือ โหลดรูปภาพที่ต้องการจากเครื่องคอมพิวเตอร์เข้ามาในโปรแกรม LabelMe ได้ วิธีการแท็กวัตถุ แสดงในภาพที่ 3.12 เพื่อทำการแท็กคำหลักบนวัตถุ ด้วยการใช้เมาส์ลากตามสัดส่วนของวัตถุแบบ freehand บนรูปภาพ เมื่อลากตามสัดส่วนของวัตถุเสร็จสิ้น โปรแกรมจะให้ใส่คำหลักเพื่อให้ความหมายของวัตถุ ข้อมูลคำหลักจะถูกจัดเก็บลงบนฐานข้อมูลพร้อมกับรูปภาพ และ ในภาพที่ 3.11 แสดงตัวอย่างของภาพที่ถูกแท็ก จากภาพที่ 3.10 ข้อมูลคำหลักที่ถูกแท็กแล้วจะแสดงไว้ทางขวามือ จะได้ข้อมูลคำหลักของวัตถุบนภาพประกอบด้วย grass, snorkel, snorkel, kid, kid, ball, ball, flipper และ flipper

หลังจากที่ผู้ใช้งานแท็กภาพในโปรแกรม LabelMe หรือทำการค้นหารูปภาพที่ถูกแท็ก โปรแกรมจะมีการแสดงภาพที่ถูกแท็ก ดังแสดงในภาพที่ 3.12 แสดงภาพที่ถูกแท็ก เส้นแสดงขอบเขตของคำหลักแต่ละคำที่แท็กไว้ ดังนั้นการใช้ฐานข้อมูลในโปรแกรม สามารถดาวน์โหลดมาใช้ร่วมกันได้ และข้อมูลคำหลักที่ถูกนำมาใช้มีมากมายหลายฐาน ดังนั้นข้อมูลสำหรับการทดลองจะมีการใช้ฐานข้อมูลที่เก็บรวบรวมมาจาก LabelMe [B. C. Russell, 2008] โดยจะทำการให้ความหมายในลักษณะของรูปทรงแบบ polygon ที่วาดลงบนวัตถุในภาพดังแสดงตัวอย่างในภาพที่ 3.14 [A. Torralba, 2010] จากโปรแกรม LabelMe เป็นตัวอย่างคำหลักที่พบบ่อยหรือเกิดขึ้นบ่อยในการให้ความหมายของวัตถุบนภาพ ตัวเลขด้านข้างแทนจำนวนของการให้ความหมาย สำหรับในภาพที่ 3.15 [A. Torralba, 2010] แสดงตัวอย่างของข้อมูลคำหลักที่ถูกแท็กบนภาพในโปรแกรม LabelMe เช่นกัน

ภาพที่ 3.14 ตัวอย่างของคำหลักที่พบบ่อยในการให้ความหมายวัตถุ



ภาพที่ 3.15 ตัวอย่างของคำหลักในการให้ความหมายวัตถุบนภาพ

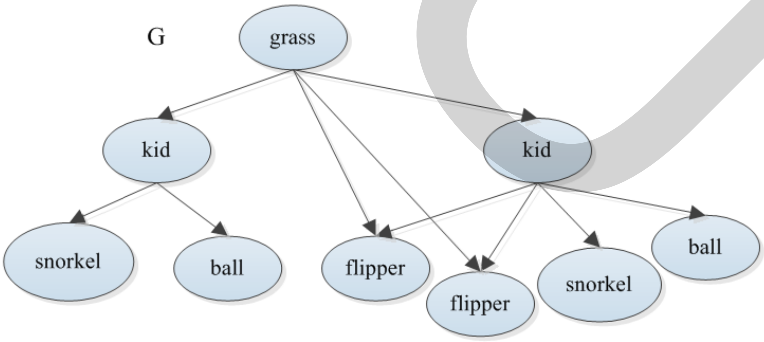
3.1.2 การแทนวัตถุลงในกราฟ

หลังจากผ่านขั้นตอนของการเตรียมข้อมูลภาพ ฐานข้อมูลคำหลักและฐานข้อมูลภาพได้ถูกคัดเลือกเข้ามา จะทำการจัดเก็บข้อมูลคำหลักที่ได้ลงในกราฟ ข้อมูลวัตถุภายในภาพที่ถูกจัดเก็บลงในเมตริกซ์จะถูกสร้างความสัมพันธ์ระหว่างวัตถุด้วยแนวคิดกราฟ (conceptual graph) กำหนดให้รูปภาพใดใด สามารถเขียนความสัมพันธ์ของแนวคิดกราฟ เป็นสเปเชียลเอ็นทิตี้ (spatial entities) ที่แทนด้วยเซตของ V คือจุด (vertex) หรือโหนด (node) เมื่อ $i \in \{1...N\}$ และความสัมพันธ์ของวัตถุภายในภาพใด เกิดจากโหนดสองโหนดที่เชื่อมต่อกันด้วย E เมื่อ $E \subset V \times V$, เมื่อ E คือความสัมพันธ์ระหว่างโหนดสองโหนด (edge) ความสัมพันธ์รูปแบบนี้ถูกเรียกว่า binary spatial relationship ดังนั้นสามารถแทนความสัมพันธ์ของสองจุดได้ด้วย $e_{ab} \equiv (v_a, v_b) \in E$ เมื่อกำหนดให้ $v_a, v_b \in V$ จากตัวอย่างภาพที่ 3.16 ก. แสดงภาพตัวอย่างจากโปรแกรม LabelMe ถูกให้ความหมายวัตถุตามคำหลัก ดังนี้ kid, kid, grass, ball, ball, snorkel, snorkel, flipper ดังแสดงในภาพที่ 3.16 ข. และสามารถแทนด้วยความสัมพันธ์ของวัตถุ ด้วยแนวคิดกราฟโหนดแต่ละโหนดบนกราฟถูกแทนด้วยคำหลักดังนั้น grass เป็น root node ที่มีความสัมพันธ์ร่วมกันกับ kid ทั้งสองและ โหนดของ kid จะเกิดความสัมพันธ์ต่อไปยังส่วนต่างๆ เช่น ball, snorkel และ flipper ดังนั้นการแสดงถึงความสัมพันธ์จะถูกเชื่อมเข้าด้วยกันเป็นค่าของ E

วัตถุ: kid kid grass ball ball snorkel snorkel flipper

ก. วัตถุที่ถูกแท็กบนภาพจากโปรแกรม LabelMe

ข. วัตถุที่ถูกแท็กบนภาพจากโปรแกรม LabelMe



ค.ความสัมพันธ์ของวัตถุที่ถูกแท็กด้วยแนวคิดกราฟ

ภาพที่ 3.16 ตัวอย่างความสัมพันธ์ของวัตถุด้วยแนวคิดกราฟ

จากภาพที่ 3.16 แสดงคำศัพท์ที่ถูกแท็กลงในภาพ ประกอบด้วยคำหลักดังนี้ grass kid kid ball ball snorkel snorkel flipper ในรูปแบบกราฟโดยที่ความสัมพันธ์ที่เกิดขึ้นภาพในภาพจะถูกแทนในรูปแบบกราฟแบบมีทิศทาง คำศัพท์ทั้งหมดจะถูกจัดกลุ่มให้อยู่ใน Synset ของ WordNet ตามกลุ่มที่มีการจัดเรียงไว้แล้วในภาพที่3.3 ดังนั้นการเปรียบเทียบความหมายทั้งหมดจะถูกแทนอยู่ในรูปแบบของกราฟ จะกล่าวต่อไปนี้

แนวคิดกราฟ

จากคุณสมบัติของคลาสกราฟสามารถเขียนเป็นทฤษฎีกราฟได้ดังนี้

กำหนดให้ $V \neq \phi$ เป็นเซตของเวคเตอร์ $H = (V, E), |V| < \infty$ เป็นกราฟจำกัดที่ระบุทิศทาง เมื่อ $E \subseteq \binom{V}{2}, V$ แทนจุดหรือโหนดและ E คือเซตของเส้นเชื่อมความสัมพันธ์ระหว่างจุด

กำหนดให้ $V \neq \phi$ เป็นเซตของเวคเตอร์ และเรียก $H = (V, E), E \subseteq V \times V.|V| < \infty$ เป็นจำนวนจำกัดสำหรับกราฟที่ไม่ระบุทิศทาง

กำหนดให้ $H = (V, E)$ เป็นจำนวนจำกัดสำหรับกราฟมีทิศทาง และ $G = (\hat{V}, \hat{E})$ เป็นกราฟย่อยของ H หมายถึง $\hat{V} \subseteq V$ และ $\hat{E} \subseteq E$ กำหนดให้กราฟย่อย G ของ H ด้วย $G \subset H$ ในกรณีของ $\ddot{E} = E \cap (\hat{V} \times \hat{V})$ เมื่อเรียก G เป็นกราฟย่อยของ $T = (V_T, E_T)$ เมื่อกำหนดให้ T แทนต้นไม้ใน G

กำหนดให้ $H = (V, E)$ เป็นจำนวนจำกัดสำหรับกราฟมีทิศทาง และกำหนดให้มีคุณลักษณะดังนี้

$\kappa_+ := \{\tilde{v} \in V \setminus \{v\} | (v, \tilde{v}) \in E\},$

$\kappa_- := \{\tilde{u} \in V \setminus \{v\} | (\tilde{u}, v) \in E\},$

$\varsigma_{out}(v) := |\kappa_+(v)|,$

$\varsigma_{in}(v) := |\kappa_-(v)|.$

เมื่อกราฟไม่ระบุทิศทางจะถูกเรียกว่าต้นไม้แบบไม่ระบุทิศทาง (undirected tree) ถ้าเป็นกราฟแบบเชื่อมต่อและวนรอบ (cycle) จะถูกเรียกว่า กราฟมีทิศทาง ต้นไม้มีทิศทาง แต่ถ้ากราฟนั้นไม่มีทิศทางจะกลายเป็นต้นไม้ (tree) สามารถเขียนเป็นคำจำกัดความได้ดังนี้ ต้นไม้ $T = (V_T, E_T)$ เป็นกราฟมีทิศทางที่มีเพียงหนึ่งโหนด $r \in V_T$ เมื่อ $\varsigma_{in}(r) = 0$ ทุก ๆโหนดใน T จะไม่ซ้ำกันและสามารถเข้าถึงจาก r ที่ถูกแทนเป็นโหนดราก

แนวคิดกราฟคลาส

แนวคิดของรูปแบบกราฟคลาสของกราฟมีทิศทาง โดยทั่วไปแล้วทฤษฎีของกราฟคลาสจะเริ่มต้นในรูปแบบของต้นไม้ (tree) ดังที่จะกล่าวต่อไปนี้

ต้นไม้ทั่วไป (Generalized trees)

ทฤษฎีของโครงสร้างต้นไม้ทั่วไป สามารถเขียน $T = (V, E_1)$ แทนรากต้นไม้ จุดกำเนิดสามารถเขียนเป็น

$V :=$

$\{v_{0,1}, v_{0,2}, \dots, v_{0,|V_0|}, v_{1,1}, v_{1,2}, \dots, v_{1,|V_1|} v_{2,1}, v_{2,2}, \dots, v_{2,|V_2|}, \dots, v_{d,1}, v_{d,2}, \dots, v_{d,|V_d|}\}$

โดยที่กำหนดให้ $|V| < \infty$. $|\kappa|$แทนค่าเป็นคาร์ดินอลลิตี้ (Cardinality) ของระดับ เซต κ และ d แทนความลึกของ T $|\kappa| = d + 1$ surjective mapping $L: V \to \kappa$ ถูกเรียกว่า ฟังกชันหลายระดับ (Multi level function) จะถูกกำหนดให้กับทุก ๆจุดในอิลิเมนต์ของระดับในเซต κ เมื่อกำหนดให้ $v_{i,j}$ เป็นจุดของตำแหน่งที่ j บนระดับ i เมื่อ $0 \leq i \leq d, 1 \leq j \leq |V_i|$. $|V_i|$ กำหนดให้จำนวนจุดบนระดับ i เซตของเส้นเชื่อม $E_{GT} := E_1 \cup E_2 \cup E_3 \cup E_4$ ของกลุ่มต้นไม้ที่ถูกกำหนดไว้แล้ว

รูปแบบของ E_1 เป็นเซตของเส้นเชื่อมที่อยู่ภายใต้รากของต้นไม้ T

E_2 เป็นเส้นเชื่อมความสัมพันธ์ของลำดับขั้นของจุดก่อนหน้า (up-edge) โดยที่รูปแบบความสัมพันธ์ดังนี้

$$E_2 := \{(v_{i+s,\eta^{i+s}}, v_{i,\eta^i}) \,\Big|\, v_{i+s,\eta^{i+s}}, v_{i,\eta^i} \in V, L\left(v_{i,\eta^i}\right) = L\left(v_{i+s,\eta^{i+s}}\right) - s,$$

$$1 \leq s \leq d \wedge \exists!\left(\left(v_{i,\eta^i}, v_{i+1,\eta^{i+1}}\right)\right), \dots, (v_{i+s-1,\eta^{i+s-1}}, v_{i+s,\eta^{i+s}})),$$

$$1 \leq \eta^i \leq |V_i|, \dots, 1 \leq \eta^{i+s-1} \leq |V_{i+s-1}|, 1 \leq \eta^{i+s} \leq |V_{i+s}|\},$$

E_3 เป็นเส้นเชื่อมที่แสดงความสัมพันธ์ของโครงสร้างบนต้นไม้ ทุกจุดที่อยู่ถัดไป (down-edge) โดยที่รูปแบบความสัมพันธ์ดังนี้

$$E_3 := \{(v_{i,\eta^i}, v_{i+s,\eta^{i+s}}) \,\Big|\, v_{i,\eta^i}, v_{i+s,\eta^{i+s}} \in V, L\left(v_{i+s,\eta^{i+s}}\right) = L\left(v_{i,\eta^i}\right) + s,$$

$$1 \leq s \leq d \wedge \exists!\left(\left(v_{i,\eta^i}, v_{i+1,\eta^{i+1}}\right)\right), \dots, (v_{i+s-1,\eta^{i+s-1}}, v_{i+s,\eta^{i+s}})),$$

$$1 \leq \eta^i \leq |V_i|, \dots, 1 \leq \eta^{i+s-1} \leq |V_{i+s-1}|, 1 \leq \eta^{i+s} \leq |V_{i+s}|\},$$

E_4 เป็นเส้นเชื่อมที่แสดงความสัมพันธ์ระหว่างจุดของโครงสร้างลำดับ โดยกำหนดให้ เซต

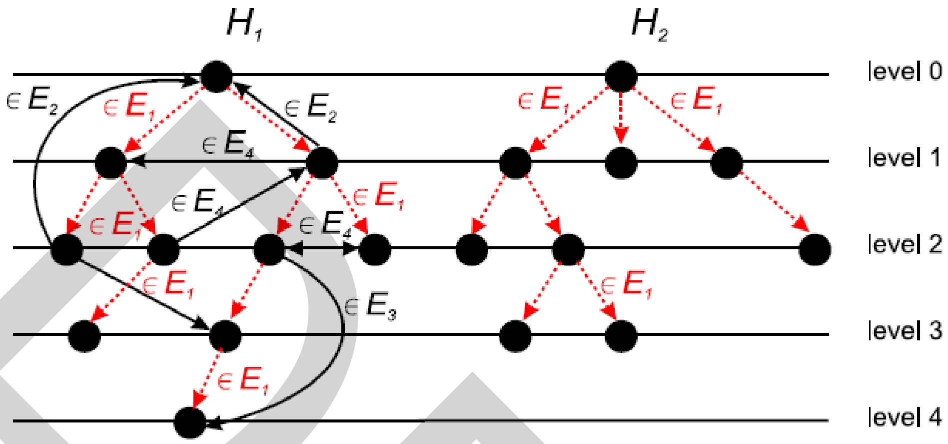
$E_4 := E_4^{i \to i} \cup E_4^{i+s \to i} \cup E_4^{i \to i+s}$ เป็นความสัมพันธ์ที่กำหนดการรวมกันระหว่างเซตย่อยของ E_4 เมื่อ

$$E_4^{i \to i} := \{(v_{i,\tilde{\eta}^i}, v_{i,\hat{\eta}^i}) \,\Big|\, v_{i,\tilde{\eta}^i}, v_{i,\hat{\eta}^i} \in V, 0 \leq i \leq d, L\left(v_{i,\tilde{\eta}^i}\right) = L\left(v_{i,\hat{\eta}^i}\right) \wedge (\tilde{\eta}^i < \hat{\eta}^i \vee \tilde{\eta}^i > \hat{\eta}^i)\},$$

$$E_4^{i+s \to i} := \{(v_{i+s,\eta^{i+s}}, v_{i,\eta}) \,\Big|\, v_{i+s,\eta^{i+s}}, v_{i,\eta^i} \in V, \left(v_{i+s,\eta^{i+s}}, v_{i,\eta^i}\right) \notin E_2, L\left(v_{i,\eta^i}\right) = L\left(v_{i+s,\eta^{i+s}}\right) - s, 1 \leq s \leq d\},$$

$$E_4^{i \to i+s} \coloneqq \{(v_{i,\eta^i}, v_{i+s,\eta^{i+s}}) \mid v_{i,\eta^i}, v_{i+s,\eta^{i+s}} \in V, (v_{i+s,\eta^i}, v_{i+s,\eta^{i+s}}) \notin E_1, E_3, L(v_{i+s,\eta^{i+s}}) = L(v_{i,\eta^i}) + s, 1 \le s \le d\}.$$

ถ้า E_2, E_3 ,หรือ E_4มีข้อมูลแล้ว $H = (V, E_{GT})$ แทนเป็นลักษณะของต้นไม้ทั่วไป (finite general tree) จากภาพที่ 3.17 แสดงความสัมพันธ์ของเส้นเชื่อมบนต้นไม้ทั่วไป H_1 และโครงสร้างของต้นไม้ ทั่วไป H_2 ที่สัมพันธ์กับระดับต่าง ๆ



ภาพที่ 3.17 ตัวอย่างการแทนความสัมพันธ์ตามแนวคิดกราฟ

กราฟลำดับขั้นแบบมีทิศทาง (Directed Universal Hierarchical Graphs) ในส่วนนี้ได้กล่าวถึงคลาสของกราฟทั่วไปในลักษณะของกราฟลำดับขั้นในรูปแบบโครงสร้างต้นไม้ทั่วไป เมื่อกำหนดให้โหนดรากอยู่ในระดับ 0 ซึ่งกราฟจะมีความซับซ้อนกว่าโครงสร้างต้นไม้ จะมีการกำหนดโครงสร้างพื้นฐานของกราฟแบบมีทิศทางไว้ดังนี้กำหนดเป็น

$$V \coloneqq \{v_{0,1}, v_{0,2}, \dots, v_{0,|V_0|}, v_{1,1}, v_{1,2}, \dots, v_{1,|V_1|} v_{2,1}, v_{2,2}, \dots, v_{2,|V_2|}, \dots, v_{d,1}, v_{d,2}, \dots, v_{d,|V_d|}\}$$

โดยที่กำหนดให้ $|V| < \infty$. และกำหนดให้ $L: V \to \kappa$ เป็นฟังก์ชันหลายระดับ จะถูกกำหนดให้ทุกจุดของอิลิเมนต์ในระดับเซต κ และ $d = |\kappa| - 1$ กำหนดให้ $v_{i,j}$ เป็นจุดของตำแหน่งที่ j บนระดับ i เมื่อ $0 \le i \le d, 1 \le j \le |V_i|$. $|V_i|$ กำหนดให้จำนวนจุดบนระดับ i เซตของเส้นเชื่อม $E_{DUHG} \coloneqq E_1 \cup E_2 \cup E_3$ ถูกกำหนดไว้แล้วดังนี้

E_1 เส้นเชื่อมลง (Down-edges) เป็นเส้นเชื่อมที่มีการเปลี่ยนแปลง 1 ระดับ

$$E_1 \coloneqq \{(v_{i,\eta^i}, v_{i+s,\eta_j^{i+s}}) \mid v_{i,\eta^i}, v_{i+s,\eta_j^{i+s}} \in V, 0 \le i \le d, 1 \le j \le |V_i|, L(v_{i+s,\eta_j^{i+s}}) = L(v_{i,\eta^i}) + s, 1 \le s \le d\}.$$

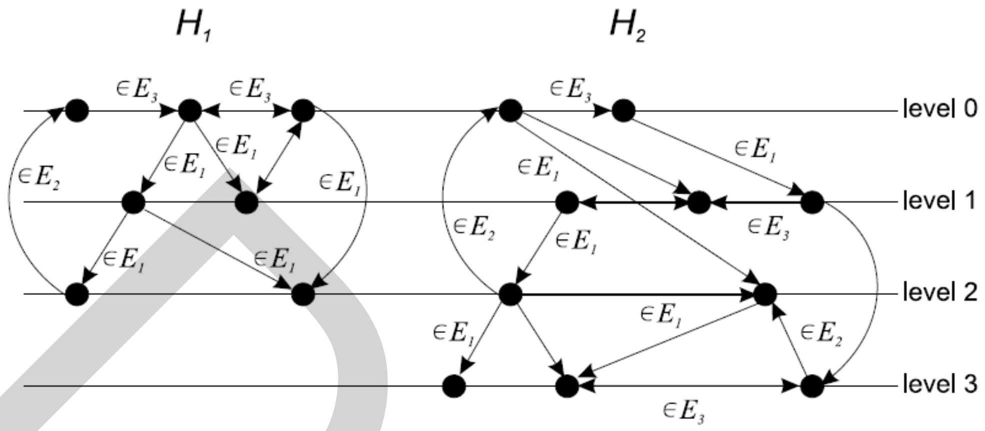
E_2 เส้นเชื่อมขึ้น (Up-edges) เป็นการเปลี่ยนแปลงอย่างน้อย 1 ระดับ

$$E_2 \coloneqq \{(v_{i+s,\eta_j^{i+s}}, v_{i,\eta^i}) \mid v_{i+s,\eta_j^{i+s}}, v_{i,\eta^i} \in V, 0 \le i \le d, 1 \le j \le |V_i|, L(v_{i,\eta^i}) = L(v_{i+s,\eta_j^{i+s}}) - s, 1 \le s \le d\},$$

E_3 เส้นเชื่อมระหว่างกัน (Across-edges) ไม่มีการเปลี่ยนแปลงระดับ

$$E_3 \coloneqq \{(v_{i,\tilde{\eta}^i}, v_{i,\hat{\eta}^i}) \mid v_{i,\tilde{\eta}^i}, v_{i,\hat{\eta}^i} \in V, 0 \le i \le d, L(v_{i,\tilde{\eta}^i}) = L(v_{i,\hat{\eta}^i}) \wedge (\tilde{\eta}^i < \hat{\eta}^i \vee \tilde{\eta}^i > \hat{\eta}^i)\},$$

แล้ว $H = (V, E_{DUHG})$ ถูกเรียกว่ากราฟลำดับขั้นแบบมีทิศทาง จากภาพที่ 3.18 แสดงกราฟลำดับขั้นแบบมีทิศทาง 2 กราฟตามทฤษฎีที่กล่าวมาข้างต้น



ภาพที่ 3.18 ตัวอย่างความสัมพันธ์กราฟด้วยตามระดับและเส้นเชื่อม

3.2 ขั้นตอนการประมวลผล

สำหรับขั้นตอนการประมวลผลจะนำข้อมูลที่มีการจัดเก็บไว้แล้วมาประมวลผล ด้วยวิธีการที่นำเสนอคือการใช้แนวคิดกราฟ และความสัมพันธ์ภายในที่เป็นแบบลำดับชั้น รวมทั้งความสัมพันธ์ที่เกิดขึ้นระหว่างวัตถุภายในด้วยกันเองโดยจะแบ่งการประมวลผลเพื่อทดสอบค่าความถูกต้องและผลลัพธ์ที่ได้ออกเป็น 2 ส่วน คือการแทนค่าข้อมูลภาพลงบนอิลิเมนต์ในกราฟตามความสัมพันธ์ของวัตถุที่เกิดขึ้นบนกราฟด้วยตัวบุคคล และทำการเปรียบเทียบความสัมพันธ์ของกราฟทั้งหมดด้วยการหาค่าความเหมือนกันของกราฟลำดับแบบมีทิศทาง ซึ่งจะกล่าวต่อไปนี้

การเปรียบเทียบความเหมือน (Structural Similarity Measurement)

การเปรียบเทียบความเหมือนกันของโครงสร้างกราฟ (Structural Similarity of directed hierarchical graph) เป็นการหาความสัมพันธ์ของข้อมูลบนกราฟสายเชื่อมโยงข้อมูล ซึ่งสามารถเปรียบเทียบได้หลายรูปแบบซึ่งโครงสร้างกราฟที่นำมาใช้ในการเก็บจะเป็นกราฟลำดับขั้นแบบมีทิศทาง (Directed universal hierarchical graph) ซึ่งรูปแบบการเปรียบเทียบความเหมือนกันจะทำการวัดจาก กราฟย่อยภายในว่ามีความเหมือนกันหรือไม่ (Subgraph isomorphic relations) มีขั้นตอนดังนี้

1. เครื่องมือที่ใช้ในการวัดความเหมือนของโครงสร้างกราฟจะทำการกำหนด$H \in \square_{DUHG}$ และกำหนดให้ $S = \{s_1, s_2, \dots, s_d\}$,เมื่อ $s_i, 1 \leq i \leq d, d = |\kappa| - 1$เป็นสตริง และกำหนดให้โครงสร้างข้อมูลสตริงอยู่ภายในเงื่อนไขดังกล่าว
2. กำหนดฟังก์ชันการเปรียบเทียบภายในให้ขึ้นกับผลลัพธ์ของการกำหนดข้อมูลสตริงด้วย เพื่อที่ใช้เป็นตัวเปรียบเทียบความเหมือนกันของกราฟ
3. วนซ้ำข้อที่ 1 และ 2 ไปเรื่อย ๆ จนกระทั่งมีความคล้ายของข้อมูลสตริงภายในจนสุดท้ายจะถูกแปลงเป็นความเหมือนกันของกราฟ

จากขั้นตอนการวนซ้ำจะเป็นการเริ่มต้นของการกำหนดสตริงของกราฟ ซึ่งจะถูกกล่าวดังต่อไปนี้

การแทนข้อมูลสตริงลงในกราฟลำดับขั้นแบบมีทิศทาง

กำหนดให้ A แทนเซตข้อมูลตัวอักษรและเป็นเซตที่ไม่ว่าง สามารถเขียนเป็น $s \in A$ สตริงของ A^* เมื่อกำหนดให้ A^*เป็นเซตสตริงของทุก ๆ A

กำหนด$H \in \square_{DUHG}$ และกำหนดให้ $S^H = \{v_{0,1}^H, v_{1,1}^H \circ v_{1,1}^H \circ \dots \circ v_{1,|v_1|}^H, \dots, v_{d,1}^H, v_{d,2}^H \circ \dots \circ v_{d,|v_d|}^H\}$,เพื่อแทนค่าสตริง H

กำหนด$H \in \square_{DUHG}$ และกำหนดให้

$S_{out}^H = \{\delta_{out}(v_{0,1}^H) \circ \delta_{out}(v_{0,2}^H), \circ \dots \circ \delta_{out}(v_{0,|v_0|}^H), \delta_{out}(v_{1,1}^H) \circ \delta_{out}(v_{1,2}^H) \circ \dots \circ \delta_{out}(v_{1,|v_1|}^H), \dots, \delta_{out}(v_{d,1}^H) \circ \delta_{out}(v_{d,2}^H) \circ \dots \circ \delta_{out}(v_{d,|v_d|}^H)\}$,

เพื่อแทนค่าคุณสมบัติของดีกรีออก (S_{out}^H) ของสตริง H

และ $S_{in}^H = \{\delta_{in}(v_{0,1}^H) \circ \delta_{in}(v_{0,2}^H), \circ \dots \circ \delta_{in}(v_{0,|v_0|}^H), \delta_{in}(v_{1,1}^H) \circ \delta_{in}(v_{1,2}^H) \circ \dots \circ \delta_{in}(v_{1,|v_1|}^H), \dots, \delta_{in}(v_{d,1}^H) \circ \delta_{in}(v_{d,2}^H) \circ \dots \circ \delta_{in}(v_{d,|v_d|}^H)\}$, แทนคุณสมบัติของดีกรีเข้า (S_{in}^H) ของสตริง H

การเปรียบเทียบความเหมือนของกราฟจากลำดับดีกรี

สำหรับการเปรียบเทียบความเหมือนของกราฟในที่นี้จะมีการเพิ่มเติม จากการเปรียบเทียบทั่วไปที่มีเฉพาะโหนดหรือจุดเท่านั้น คือมีการใช้ลำดับดีกรีเข้าและออกด้วย ซึ่งเป็นคุณสมบัติของแต่ละโหนดที่อยู่บนโครงสร้างย่อยในกราฟ

กำหนดให้ X แทนเซต และฟังก์ชันจำนวนจริง แทนด้วย $s: X \times X \rightarrow [0,1]$ ถูกเรียกว่าตัววัดความเหมือนกัน ถ้า $\beta(x,y) = \beta(y,x), \forall x,y \in X$, $\beta(x,y) \leq \beta(x,x), \forall x,y \in X$

และ $\beta(x,x) = 1, \forall x \in X$

กำหนดให้ H_1 และ H_2 เป็นกราฟลำดับขั้นแบบมีทิศทางที่จะใช้ในการเปรียบเทียบความเหมือนกัน โดยเริ่มจากการเปรียบเทียบจากสตริงของคุณสมบัติของดีกรีออกก่อน (out degree)

$$S_i^{out} = \delta_{out}(v_{i,1}^H) \circ \delta_{out}(v_{i,2}^H) \circ \dots \circ \delta_{out}(v_{i,|v_i|}^H)$$

และคุณสมบัติของดีกรีเข้า (in degree) ของ $H \in \square_{DUHG}$

$$S_i^{in} = \delta_{in}(v_{i,1}^H) \circ \delta_{in}(v_{i,2}^H) \circ \ldots \circ \delta_{in}(v_{i,|v_i|}^H)$$

เมื่อมี $\delta_{out}(v_{i,j}^H) \geq \delta_{in}(v_{i,j+1}^H)$ และ $\delta_{in}(v_{i,j}^H) \geq \delta_{out}(v_{i,j+1}^H)$ กำหนดให้ n_i เป็น $n_i = \max(|v_i^{H_1}|, |v_i^{H_2}|)$ โดยการเพิ่ม 0 เข้าลงไปเพื่อให้ n_i ลำดับความยาวเท่ากัน สามารถเขียนสมการของดีกรีใหม่ได้ดังนี้

$$S_i^{out} = \delta_{out}(v_{i,1}^H) \circ \delta_{out}(v_{i,2}^H) \circ \ldots \circ \delta_{out}(v_{i,n_i}^H)$$

และ

$$S_i^{in} = \delta_{in}(v_{i,1}^H) \circ \delta_{in}(v_{i,2}^H) \circ \ldots \circ \delta_{in}(v_{i,n_i}^H)$$

ดังนั้น

$$S_i^{out}(H_1) = \delta_{out}(v_{i,1}^{H_1}) \circ \delta_{out}(v_{i,2}^{H_1}) \circ \ldots \circ \delta_{out}(v_{i,n_i}^{H_1})$$

$$S_i^{in}(H_2) = \delta_{in}(v_{i,1}^{H_2}) \circ \delta_{in}(v_{i,2}^{H_2}) \circ \ldots \circ \delta_{in}(v_{i,n_i}^{H_2})$$

โดยที่ i แทนตำแหน่งที่ของดีกรีเข้าและดีกรีออกของคุณสมบัติในกราฟ H_1 และ H_2

ดังนั้นกำหนดให้ margin i เป็นตำแหน่งของดีกรีเข้าและดีกรีออกดังนี้

$$\xi_{|i,1|}^{out}, \xi_{|i,2|}^{out}, \ldots, \xi_{|i,n_i|}^{out} \in \mathrm{N},$$

และ

$$\xi_{|i,1|}^{in}, \xi_{|i,2|}^{in}, \ldots, \xi_{|i,n_i|}^{in} \in \mathrm{N}$$

เมื่อ

$$\xi_{|i,1|}^{out} = \left\|\delta_{out}(v_{i,1}^{H_1}) - \delta_{out}(v_{i,1}^{H_2})\right\|,$$

$$\xi_{|i,2|}^{out} = \left\|\delta_{out}(v_{i,2}^{H_1}) - \delta_{out}(v_{i,2}^{H_2})\right\|,$$

…

$$\xi_{|i,n_i|}^{out} = \left\|\delta_{out}\left(v_{i,n_i}^{H_1}\right) - \delta_{out}(v_{i,n_i}^{H_2})\right\|,$$

และ

$$\xi_{|i,1|}^{in} = \left\|\delta_{in}(v_{i,1}^{H_1}) - \delta_{in}(v_{i,1}^{H_2})\right\|,$$

$$\xi_{|i,2|}^{in} = \left\|\delta_{in}(v_{i,2}^{H_1}) - \delta_{in}(v_{i,2}^{H_2})\right\|,$$

…

$$\xi_{|i,n_i|}^{in} = \left\|\delta_{in}\left(v_{i,n_i}^{H_1}\right) - \delta_{in}(v_{i,n_i}^{H_2})\right\|,$$

ตามลำดับ

3.3 ขั้นตอนการวัดประสิทธิภาพ

ในงานวิจัยนี้ได้นำเสนอขั้นตอนการวัดประสิทธิภาพ (evaluation of similarity measurement) ที่เหมาะสมกับวิธีการจัดเก็บแบบโครงสร้างกราฟที่มีทั้งความสัมพันธ์ของอนุกรมวิธานและความสัมพันธ์ของวัตถุภายในภาพ ดังนั้นจึงเป็นการวัดความคล้ายกันของภาพด้วยโครงสร้างความสัมพันธ์ลำดับขั้น (semantic similarity with relational hierarchy structure: RHS) ดังนั้นเป็นขั้นตอนสุดท้ายของกระบวนการจำแนกความหมายภาพ โดยจะมีการแยกเปรียบเทียบสองส่วนคือ ส่วนที่เกิดจากวัตถุภายในภายซึ่งจะเปรียบเทียบความต่างของกลุ่มวัตถุและเปรียบเทียบความต่างของภาพซึ่งถูกแทนด้วยกราฟดังนั้นจึงได้แบ่งการเปรียบเทียบเป็น 2 ส่วนดังนี้

3.3.1 การเปรียบเทียบความเหมือนของวัตถุ

กำหนดให้ $\xi_1, \xi_1, \dots, \xi_n$ เป็นจำนวนเต็มและให้ O_i, O_j เป็นวัตถุ ดังนั้นเขียนเป็นสมการเพื่อแสดงความสัมพันธ์ของวัตถุดังนี้ $\Upsilon\left(O_i, O_j, \xi_1, \xi_1, \dots, \xi_n\right) \coloneqq \frac{\beta(\xi_1), \beta(\xi_1), \dots, \beta(\xi_n)}{n}$ เมื่อ $\beta(\xi_i) \leq 1, 1 \leq i \leq n,$ เมื่อ $\Upsilon(\xi_1, \xi_1, \dots, \xi_n)$ คือการวัดความเหมือนของวัตถุ และได้มีการกำหนดเงื่อนไขเพิ่มเติมของการวัดค่าความเหมือน เมื่อ $\beta(\xi_i) \leq 1, 1 \leq i \leq n,$ และ $\Upsilon(\xi_1, \xi_1, \dots, \xi_n)$ เป็น ภายใต้สมมุติฐานเมื่อ $O_i = O_j, \beta(\xi_i) = 0, 1 \leq i \leq n$ นอกเหนือจากนี้ยังมีการวัดความเหมือนของวัตถุจากจำนวนดีกรีเข้าและออกในระดับ i จากโหนดดังนี้

$$\Upsilon_i^{out} = \Upsilon_i^{out}(\xi_{|i,1|}^{out}, \xi_{|i,2|}^{out}, \dots, \xi_{|i,n|}^{out}) \in [0,1]$$

และ

$$\Upsilon_i^{in} = \Upsilon_i^{in}(\xi_{|i,1|}^{in}, \xi_{|i,2|}^{in}, \dots, \xi_{|i,n|}^{in}) \in [0,1]$$

3.3.2 การเปรียบเทียบความเหมือนของภาพ

กำหนดให้ H_1 และ H_2 เป็นกราฟลำดับขั้นที่แทนภาพที่ 1 และภาพที่ 2 ตามลำดับและ $v \coloneqq \max(d^{H_1}, d^{H_2})$ โดยที่ $d^{H_i} = |\mathfrak{I}^{H_i}| - 1$ ดังนั้นความเหมือนของกราฟสามารถเขียนเป็นสมการได้ดังนี้

$$s(H_1, H_2) \coloneqq \frac{\prod_{i=0}^{v} \overline{\Upsilon}_i}{\sum_{i=0}^{v} \overline{\Upsilon}_i / v + 1}$$

$$s(H_1, H_2) = s(H_2, H_1),$$

$$s(H_1, H_1) = 1,$$

$$0 < s(H_1, H_2) \leq s(H_1, H_1)$$

เมื่อกำหนดให้

$$\overline{\Upsilon}_i = \overline{\Upsilon}\left(\xi_{|i,1|}^{out}, \xi_{|i,2|}^{out}, \dots, \xi_{|i,n|}^{out}, \xi_{|i,1|}^{in}, \xi_{|i,2|}^{in}, \dots, \xi_{|i,n|}^{in}\right);$$

$$=\alpha.\Upsilon_i^{out}\left(\xi_{|i,1|}^{out}, \xi_{|i,2|}^{out}, \dots, \xi_{|i,n|}^{out}\right) + (1-\alpha).\Upsilon_i^{in}\left(\xi_{|i,1|}^{in}, \xi_{|i,2|}^{in}, \dots, \xi_{|i,n|}^{in}\right), \alpha \in [0,1]$$

เมื่อมีการกำหนด $\bar{s}(H_1, H_2) \coloneqq \sum_{i=0}^{v} \overline{\Upsilon}_i/(v+1)$ แล้วจากสมการข้างต้นททที่กล่าวถึงความสมมาตรกันของกราฟนั้นสามารถเขียนได้ว่า $\bar{s}(H_1, H_2) = \bar{s}(H_2, H_1)$ เมื่อ $\overline{\Upsilon}_i \leq \alpha.1 + (1-\alpha).1 = 1, \bar{s}(H_1, H_2) = \frac{1+1+\cdots+1_{(v+1-times)}}{v+1} = 1$

เมื่อนำผลการเปรียบเทียบความคล้ายกันของภาพที่จัดได้มาทำการวัดประสิทธิภาพของสิ่งที่นำเสนอข้างต้น โดยจะตรวจสอบกลุ่มของภาพที่จัดได้ว่ามีค่าเป็นอย่างไร เมื่อเทียบกับกลุ่มของภาพที่ถูกต้อง ซึ่งต้องมีการวัดค่าความระลึก (recall) และค่าความแม่นยำ (precision) จะเป็นค่าที่แสดงว่า การค้นคืนข้อมูลได้ตรงกับความต้องการเพียงใด ส่วนค่าความระลึกจะเป็นค่าที่แสดงถึงความครอบคลุมในการจัดกลุ่มภาพ หลังจากนั้นจะนำค่ามาคำนวณในรูปของค่าความถูกต้อง (accuracy) และ F-measure ต่อไปและนำค่าทั้งหมดมาแปลผลและประเมินผลลัพธ์ที่ได้ว่ามีความเหมาะสม หรือตรงกับวัตถุประสงค์ที่ต้องการหรือไม่ในรูปที่สามารถเข้าใจได้ง่าย สามารถอ่านเพิ่มเติมได้ในบทที่ 2

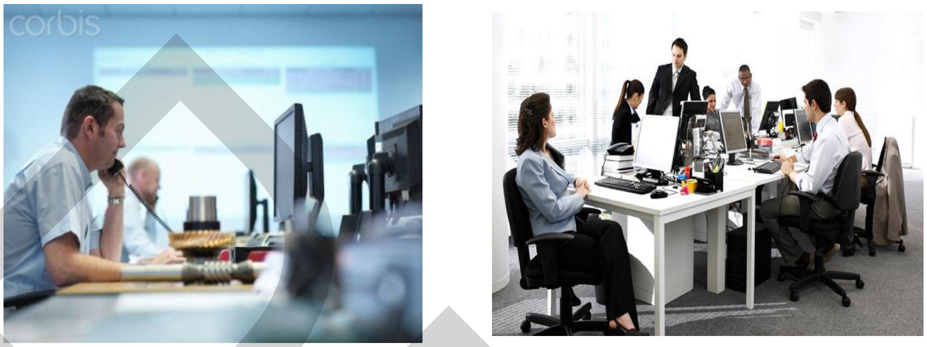
บทที่ 4
ผลการทดลอง

ในการทำวิจัยครั้งนี้เพื่อค้นหาภาพที่มีความเหมือนกันทางด้านความหมายโดยนำเสนอวิธีการวัดความคล้ายของภาพด้วยอนุกรมวิธาน WordNet ที่มีการจำลองโครงสร้างในรูปแบบกราฟลำดับขั้น ดังนั้นการวัดความหมายหรือความคล้ายกันของภาพจากรูปแบบความสัมพันธ์ของคำศัพท์ Synset ตามอนุกรมวิธาน และขั้นตอนสุดท้ายขั้นตอนการประมวลผล จึงพยายามเปรียบเทียบตามลักษณะของการแทนด้วยโครงสร้างของกราฟ ซึ่งจะมีส่วนช่วยให้การค้นหาภาพมีความหมายที่ใกล้เคียงกันมากที่สุด ดังนั้นขั้นตอนสุดท้ายของการค้นหาความหมายภาพ ด้วยการเปรียบเทียบความคล้ายข้อมูล ซึ่งในงานวิจัยนี้ได้ใช้วิธีการเปรียบเทียบกันทั้งหมด 4 วิธี การวัดความคล้ายด้วยอนุกรมวิธาน (WordNet Similarity: WS) การวัดความคล้ายด้วยคิวรี่คำศัพท์ (Keyword-based Query Similarity: KQS) การวัดความคล้ายด้วยคิวรี่ย้อนกลับ (Relevance Feedback based on Query Similarity: RFQS) และนำเสนอวิธีการวัดความคล้ายกันของภาพด้วยโครงสร้างกราฟตามความสัมพันธ์ลำดับขั้น (Semantic Graph Similarity with Relational Hierarchy Structure: RHS)

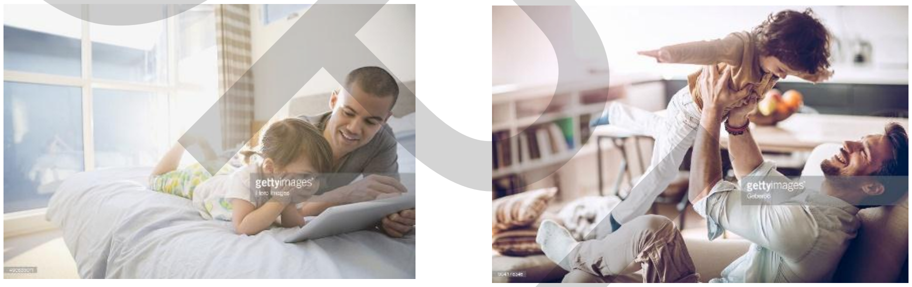
4.1 การกำหนดข้อมูลภาพ

ข้อมูลภาพสำหรับการทดลองได้ใช้ฐานข้อมูลที่เก็บรวบรวมและมาจากแอพพลิเคชัน LabelMe [B. C. Russell,2008];[A. Torralba,2010] โดยแอพพลิเคชันนี้สามารถทำงานได้อย่างเต็มรูปแบบบนเว็บในลักษณะของเครื่องมือให้ความหมาย และได้มีการคัดเลือกภาพสำหรับการทดลองให้ครอบคลุมกับการทดลองโดยเป็นภาพส่วนบุคคลที่อยู่ในหมวดหมู่ของภาพภายใน (indoor) และภาพภายนอก (outdoor) มาจากฐานข้อมูล Corbis [Corbis] และคำศัพท์ได้อ้างอิงจากฐานข้อมูล NIST TRECVID 2015 [NIST] และคำศัพท์ได้อ้างอิงจากฐานข้อมูลอนุกรมวิธาน WordNet [Miller G.A., 1990];[Zinger S., 2005] คัดเลือกคำศัพท์ภาษาอังกฤษมาทั้งหมด 200 คำหลัก โดยคำศัพท์ที่นำมาใช้จะเป็นคำหลักที่มาจากวัตถุทั่วไปที่มักพบในกลุ่มภาพตามหมวดหมู่ที่สอดคล้องกับภาพที่คัดเลือกมาเป็นหลัก โดยที่คำศัพท์ที่ถูกบันทึกลงบนภาพจะไม่มีคำคุณศัพท์ เช่น *long, short, lengthy* หรือ *very, decidedly, intensely, rather, quite, somewhat, pretty,*

extremely เป็นต้น และคำหลักที่ถูกสร้างขึ้นนั้นได้พยายามหลีกเลี่ยงคำหลักที่มีความหมายกำกวม และคำหลักที่เป็นทั้งคำเหมือน (Synonym) เช่น *“abbey”, “church”, “cathedral”* เป็นต้น เพราะอาจเป็นส่วนหนึ่งที่ทำให้ผลการทดลองนั้นเกิดความผิดพลาดได้ จะใช้ชื่อวัตถุที่อยู่บนภาพเท่านั้นเป็นคำศัพท์ที่นำมาใช้ในการเก็บข้อมูล ซึ่งจะสอดคล้องกับคลาสบนอนุกรมวิธาน WordNet



ก. ภาพทำงานในสำนักงาน



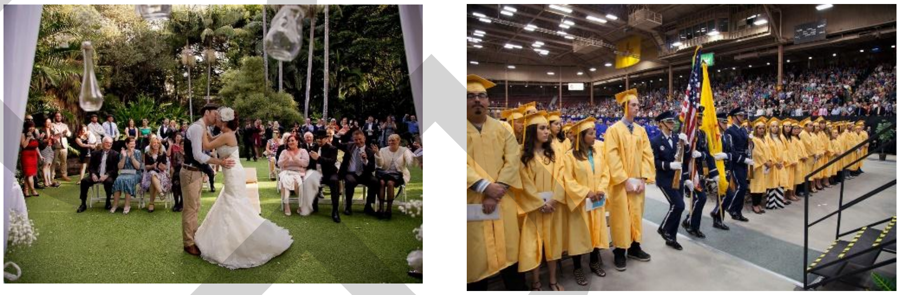
ข. ภาพครอบครัว

ภาพที่ 4.1 ตัวอย่างกลุ่มภาพทำงานในสำนักงาน และภาพครอบครัว

ดังนั้นในการทดลองนี้จะมีการแท็กคำหลักตามที่กำหนดไว้บนภาพที่ถูกคัดเลือกมาตามความหมายพื้นฐานจากการสุ่มและตัดสินใจจากการจำแนกความหมายภาพด้วยคนเป็นหลัก (human scenes classification) [Jianxiong Xiao, 2010] ข้อมูลภาพทั้งหมดจะถูกแท็กไว้บนภาพจากผู้ใช้ในแอพพลิเคชัน LabelMe โดยกำหนดให้จำกัดขอบเขตของคำหลักที่ใช้ในการทดลองนี้รวมทั้งหมด ภาพรวมทั้งหมดที่ใช้ในการทดลอง 2,250 ภาพ โดยแต่ละชุดจะทำการสุ่มข้อมูลภาพและจำกัดขอบเขตของคำศัพท์ที่ใช้เพื่อให้ครอบคลุมโครงสร้างการทำงาน โดยทำการเปรียบเทียบความคล้ายกันของความหมายภาพด้วย 4 วิธีการดังนี้ WS, KQS, RFQS และ RHS เปรียบเทียบความหมายลงใน 5 กลุ่มเหตุการณ์ประกอบด้วย ภาพพักผ่อน (outdoor leisure), ภาพพิธีการ

(ceremony), ภาพทำงานในสำนักงาน (office working), ภาพเล่นกีฬา (sport game) และ ภาพครอบครัว (family time) ดังแสดงภาพตัวอย่างในภาพที่ 4.1 และ ภาพที่ 4.2

ก. ภาพพักผ่อน



ข. ภาพพิธีการ

ภาพที่ 4.2 ตัวอย่างกลุ่มภาพพักผ่อน และ ภาพพิธีการ

4.2 ผลการทดลอง

การทดลองจะแบ่งชุดข้อมูลภาพเพื่อใช้ทำการทดลองออกเป็น 3 ชุด ด้วยการสุ่มและจัดคำหลักที่มีความสอดคล้องกันตามชุดข้อมูล แบ่งชุดการทดลองออกเป็น 3 ชุด ชุดที่ 1 ใช้ภาพสุ่ม 500 ภาพ คำศัพท์เกี่ยวข้อง 80 คำหลัก, ชุดที่ 2 ใช้ภาพสุ่ม 750ภาพ คำศัพท์เกี่ยวข้อง 120 คำหลัก ชุดที่ 3 ใช้ภาพสุ่ม 1,200 ภาพ คำศัพท์เกี่ยวข้อง 180 คำหลัก ข้อมูลของคำหลักในกลุ่มใดมีผลต่อการเปรียบเทียบความคล้ายกันของความหมายภาพลงใน 5 กลุ่มเหตุการณ์ประกอบด้วย outdoor leisure, ceremony, office working, sport game และ family time โดยที่จะทำการเปรียบเทียบทั้งหมด 4 วิธีการดังนี้ WS, KQS, RFQS และ RHS จากตารางที่ 4.1

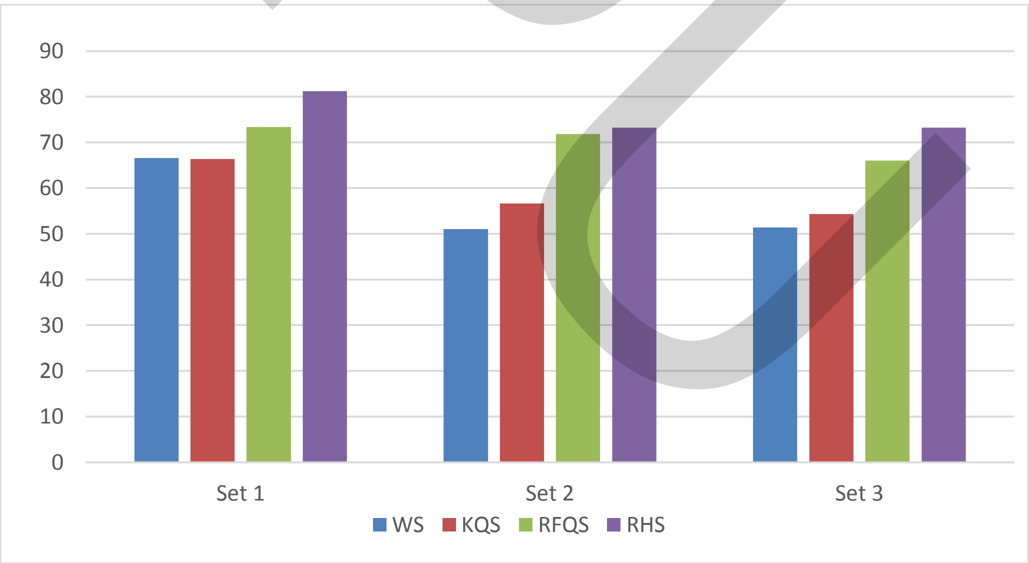
Measure	Class	Performance (%)								
		Set 1			Set 2			Set 3		
		Prec.	Recall	F_1	Prec.	Recall	F_1	Prec.	Recall	F_1
WS	Outdoor leisure	69.9	65.0	67.4	54.6	53.0	53.8	55.3	52.0	53.6
	ceremony	69.1	67.0	68.0	50.0	51.0	50.5	54.0	54.0	54.0
	Office indoor	61.5	64.0	62.7	50.9	55.0	52.9	45.3	48.0	46.6
	sport game	67.3	68.0	67.7	47.6	49.0	48.3	50.5	51.5	51.0
	family time	65.7	69.0	67.3	52.2	47.0	49.5	52.5	51.5	52.0
	Accuracy	66.6			51			51.4		
KQS	Outdoor leisure	66.7	64.0	65.3	58.1	54.0	56.0	53.3	57.0	55.1
	ceremony	67.3	68.0	67.7	58.7	54.0	56.3	53.2	58.0	55.5
	Office indoor	66.0	64.0	65.0	54.2	52.0	53.1	54.2	52.0	53.1
	sport game	66.7	68.0	67.3	56.4	62.0	59.0	56.3	53.5	54.8
	family time	65.1	67.6	66.3	56.0	61.0	58.4	54.8	51.0	52.8
	Accuracy	66.3			56.6			54.3		
RFQS	Outdoor leisure	71.2	74.0	72.5	68.6	72.0	70.2	67.0	65.0	66.0
	ceremony	74.5	73.0	73.7	72.1	75.0	73.5	68.3	69.0	68.7
	Office indoor	81.0	68.0	73.9	77.2	71.0	74.0	66.0	64.0	65.0
	sport game	72.6	77.0	74.8	70.6	72.0	71.3	65.4	68.0	66.7
	family time	69.4	75.0	72.1	71.1	69.0	70.1	63.4	64.0	63.7
	Accuracy	73.4			71.8			66.0		
RHS	Outdoor leisure	83.0	78.0	80.4	70.6	72.0	71.3	74.5	76.0	75.2
	ceremony	81.7	76.0	78.8	72.3	73.0	72.6	75.5	76.2	75.9
	Office indoor	80.4	86.0	83.1	81.4	70.0	75.3	72.6	69.0	70.8
	sport game	81.6	84.0	82.8	72.6	77.0	74.8	73.7	73.0	73.4
	family time	79.6	82.0	80.8	70.5	74.0	72.2	69.9	72.0	70.9
	Accuracy	81.2			73.2			73.3		

ตารางที่ 4.1 ผลลัพธ์การเปรียบเทียบความคล้ายกันของความหมายภาพ

จากตารางที่ 4.1 แสดงให้เห็นว่ากลุ่มภาพ set 1 ที่ถูกเปรียบเทียบความคล้ายของภาพ ด้วย WS จะมีค่าความถูกต้องสูงถึง 66.6% จะเห็นว่าสามารถบอกความหมายของภาพได้ดีกว่า KQS ถึง 0.03% แต่อย่างไรก็ตามเมื่อเปรียบเทียบกับ set 2 และ 3 และ KQS กลับมีความค่าความถูกต้องที่สูงกว่า 4.6% และ 2.9% ตามลำดับ RFQS สามารถเปรียบเทียบความคล้ายของภาพใน set 1 ได้อย่างดีค่าความถูกสูงถึง 73.4% แต่เมื่อ set 2 และ 3 กลับมีค่าความถูกต้องที่ลดลงเมื่อมีจำนวนภาพ

และจำนวนของคำศัพท์ที่มากขึ้นเช่นเดียวกันกับวิธีการ RHS แต่อย่างไรก็ตามค่าความถูกต้องใน set 1 มีค่าความถูกต้องที่สูงถึง 81.2% และเมื่อใช้กลุ่มข้อมูลของ set 2 และ 3 กลับมีค่าความถูกต้องที่ต่ำลง 73.2% และ 73.3% แต่มีความแตกต่างไม่มากทั้ง ๆที่จำนวนของภาพใน set 2 และ set 3 มีจำนวนภาพเพิ่มขึ้นเท่าตัวทั้งนี้ขึ้นกับจำนวนของลำดับความสัมพันธ์ที่เกิดขึ้นบนกราฟมีความซับซ้อน แต่ยังสามารถเปรียบเทียบเพื่อค้นหาความหมายของภาพได้อย่างสมบูรณ์จนทำให้มีค่าใกล้เคียงกัน

จากการเปรียบเทียบวิธี KQS จะเห็นว่าสามารถเปรียบเทียบความหมายภาพใน set 1 ลงในกลุ่ม ceremony มีค่า Prec. 67.3% Recall 68% และ sport game มีค่า Prec. 66.7% Recall 68% แต่ใน set 1 ที่มีการเปรียบเทียบด้วย RFQS ลงในกลุ่ม Office indoor มีค่า Prec. 81% Recall 68% สำหรับการเปรียบเทียบความหมายใน set 3 ด้วย RHS จะเห็นว่าสามารถวัดความคล้ายของภาพได้สูงถึง 73.3% และลงในกลุ่ม ceremony ได้ค่า F_1 ถึง 75.9% เช่นเดียวกันกับวิธีอื่น WS, KQS และ RFQS จะสามารถเปรียบเทียบความคล้ายของภาพในกลุ่ม Outdoor leisure ได้ค่า F_1 สูงที่สุด 75.2% ใน set 3 ด้วยวิธีการ RHS แต่ในทางกลับกัน กลุ่มที่มีการเปรียบเทียบได้ค่า F_1 ที่น้อยนั้นส่วนใหญ่เป็นกลุ่ม family time ด้วยวิธี RHS จะได้เพียง 70.9%



ภาพที่ 4.3 กราฟแสดงการเปรียบเทียบความคล้ายกันด้วยค่าความถูกต้องระหว่างกลุ่มภาพ

ก. ภาพกลุ่ม outdoor leisure

ข. ภาพกลุ่ม ceremony

ค. ภาพกลุ่ม office indoor

ง. ภาพกลุ่ม sport game

ฉ. ภาพกลุ่ม family time

ภาพที่ 4.4 ตัวอย่างผลลัพธ์ของการเปรียบเทียบความคล้ายกันของความหมายภาพ

จากภาพที่ 4.3 แสดงการเปรียบเทียบความคล้ายกันด้วยค่าความถูกต้องระหว่างกลุ่มภาพ set 1 set 2 และ set 3 ด้วยวิธีการ WS, KQS, RFQS และ RHS จะเห็นว่าการเปรียบเทียบความคล้ายด้วยกลุ่มข้อมูลภาพ set 1 นั้นมีค่าความถูกต้องสูงที่สุดเนื่องมาจากจำนวนภาพที่ใช้ในการทดลองและจำนวนคำศัพท์ที่นำมาใช้มีจำนวนน้อย ความสัมพันธ์ที่เกิดขึ้นในโครงสร้างของกราฟน้อยการเปรียบเทียบจำนวนขั้นของลำดับภายในไม่มากนัก จึงทำให้มีความสามารถในการเปรียบเทียบได้มากกว่า กลุ่มข้อมูลภาพอื่น ๆ แต่อย่างไรก็ตามใน set 2 และ 3 จะเห็นว่ามีค่าความถูกต้องที่น้อยลงจริงแต่วิธี RHS นั้นมีค่าความถูกต้องที่ใกล้เคียงกันทั้ง ๆ ที่มีจำนวนภาพและคำศัพท์เพิ่มขึ้นเท่าตัว ดังนั้นจะเห็นว่าประสิทธิภาพของการเปรียบเทียบภาพด้วยการใช้กราฟแบบลำดับขั้นที่เปรียบเทียบตามกลุ่มคำของ WordNet ที่กำหนดไว้แล้วจะมีผลช่วยในการจัดกลุ่มความหมายได้มากกว่าวิธีอื่น ๆ แสดงภาพตัวอย่างในภาพที่ 4.4

DPU

บทที่ 5
สรุปผลการทดลอง

ในบทนี้ได้ทำการสรุปผลการทดลอง และข้อเสนอแนะ รวมทั้งข้อผิดพลาดที่เกิดขึ้นจากการทดลอง รวมไปถึงสิ่งที่ควรจะปรับปรุงเพิ่มเติม เพื่อแก้ไขข้อผิดพลาดและแนวทางการทำวิจัยต่อในเรื่องของการแปลความหมายภาพซึ่งเป็นงานวิจัยที่ปัจจุบันได้มีนักวิจัยให้ความสนใจอย่างแพร่หลาย สามารถนำไปประยุกต์ใช้ได้หลายด้าน เช่น การแพทย์ การขนส่งคมนาคม การรักษาความปลอดภัย หรือทางภาคอุตสาหกรรม เป็นต้น

5.1 สรุปและอภิปรายผลการวิจัย

งานวิจัยนี้ได้เสนอหัวข้อวิจัยทางด้านการประมวลผลภาพ ในส่วนของการเปรียบเทียบความคล้ายกันของความหมายภาพเพื่อให้ได้ความหมายของภาพที่อยู่ในหมวดหมู่เดียวกัน โดยปกติทั่วไปนั้นการใช้อัลกอริทึมที่มาสกัดข้อมูลภาพนั้นมักจะใช้สกัดเพียงข้อมูลที่เกิดขึ้นภายในภาพ แล้วนำมาประมวลผลเพื่อใช้ในการสืบค้นข้อมูลภาพ แต่ปัจจุบันได้มีการนำคำหลักที่ได้จากการให้ความหมายของการแท็กวัตถุบนภาพมาหาความสัมพันธ์ภายใน โดยพยายามหาความสัมพันธ์ที่คล้ายกันของวัตถุในหมวดหมู่เดียวกัน ในงานวิจัยนี้ก็เช่นเดียวกันแต่จะเฉพาะเจาะจงลงในกลุ่มของภาพส่วนบุคคลเฉพาะเหตุการณ์กิจกรรม โดยใช้แนวคิดกราฟที่แทนข้อมูลภาพในภาพทั้งหมดและข้อมูลคำศัพท์ภาพจะมีความสัมพันธ์สอดคล้องกับโครงสร้างของ WordNet [Miller, George A.,1990] เพื่อแสดงถึงกลุ่มความหมายของคำศัพท์นั้น ๆซึ่งจะนำพาความสัมพันธ์ให้เกิดขึ้น โดยทำการทดลองเพื่อเปรียบเทียบการจัดกลุ่มทั้งหมด 4 วิธีการคือ ภาพ โดยแต่ละชุดจะทำการสุ่มข้อมูลภาพและจำกัดขอบเขตของคำศัพท์ที่ใช้เพื่อให้ครอบคลุมโครงสร้างการทำงาน โดยทำการเปรียบเทียบความคล้ายกันของความหมายภาพด้วย 4 วิธีการดังนี้ WS, KQS, RFQS และ RHS เปรียบเทียบความหมายลงใน 5 กลุ่มเหตุการณ์ประกอบด้วย ภาพพักผ่อน (outdoor leisure), ภาพพิธีการ (ceremony), ภาพทำงานในสำนักงาน (office indoor), ภาพเล่นกีฬา (sport game) และ ภาพครอบครัว (family time)

จากการทดลองในบทที่ 4 สามารถสรุปได้ว่าจากตารางที่ 4.1 แสดงให้เห็นว่ากลุ่มภาพ set 1 ที่ถูกเปรียบเทียบความคล้ายของภาพ ด้วย WS จะมีค่าความถูกต้องสูงถึง 66.6% RFQS สามารถเปรียบเทียบความคล้ายของภาพใน set 1 ได้อย่างดีค่าความถูกสูงถึง 73.4% แต่เมื่อ set 2 และ 3 กลับมีค่าความถูกต้องที่ลดลงเมื่อมีจำนวนภาพที่มากขึ้นที่ แต่อย่างไรก็ตามค่าความถูกต้องใน set 1 มีค่าความถูกต้องที่สูงถึง 81.2% และเช่นเดียวกันเมื่อใช้กลุ่มข้อมูลของ set 2 และ 3 กลับมีค่าความถูกต้องที่ต่ำลง 73.2% และ 73.3% ลักษณะเหตุผลของค่าความถูกต้องที่ไม่มีการเปลี่ยนแปลงอาจจะเกิดจากความหลากหลายของคำศัพท์ทำให้การเปรียบเทียบตามเส้นการเชื่อมโยงนั้น ยังมีความสัมพันธ์ที่น้อยเกินไป แต่อย่างไรก็ตามค่าความถูกต้องที่ได้จากการจัดกลุ่มด้วย RFQS ได้มากกว่าวิธีการอื่น ๆ เพราะฉะนั้นจากการทดลองสามารถสรุปได้ว่า การที่นำทฤษฎีของกราฟที่มีโครงสร้างลำดับและใช้ WordNet ที่มีการจัดกลุ่มของคำศัพท์เข้ามาช่วยในการเปรียบเทียบความหมายของภาพได้

5.2 ข้อเสนอแนะ

ในงานวิจัยส่วนใหญ่เน้นไปที่การสกัดข้อมูลภาพในรูปแบบของการสกัดด้วยอัลกอริทึมที่แตกต่างกัน และค้นหาภาพเพียงความเหมือนกันของรูปทรงหรือลักษณะเฉพาะ หรือเพียงแต่วัตถุบนภาพ เท่านั้น ทั้งที่มีความหมายและไม่มีความหมาย แต่อย่างไรก็ตาม เบื้องต้นของผลลัพธ์ที่ได้จากการค้นคืน จำแนกภาพ และการจัดกลุ่มภาพ คือความเหมือนกันทางกายภาพ เช่น รูปทรง สี หรือชนิดของวัตถุ แต่ลักษณะการวิเคราะห์และพิจารณาของการเหมือนกันทางความหมายภาพ นั้นจะมีลักษณะการวิเคราะห์ที่แตกต่างกันออกไป ในอีกรูปแบบหนึ่งซึ่งเป็นรูปแบบที่เกิดจากความคิดของมนุษย์ที่มีการแปลงความจากภาพ สิ่งควรจะมีการปรับปรุงเพิ่มเติมเพื่อให้การแปลความหมายของภาพได้ดียิ่งขึ้น ก็คือการพิจารณารายละเอียดของวัตถุของภาพ ซึ่งคำหลักที่ถูกเลือกมาจากกลุ่มภาพใน LabelMe [A. Torralba,2010];[B. C. Russell,2008] เพื่อเพิ่มความหลากหลายของกลุ่มคำศัพท์ควรจะมีการจัดกลุ่มของวัตถุที่เข้ามาทำการทดลองให้มีความหมายที่รัดกุมมากขึ้น และเมื่อภาพที่มีส่วนของสภาพแวดล้อมที่คล้ายคลึงกันมาก จะถูกจัดอยู่ในกลุ่มเดียวกันทั้งที่ วัตถุภายในภาพมีความแตกต่างกันของรายละเอียดภายในซึ่งบางครั้งการจัดคำหลักจะมีผลต่อการแปลความหมายโดยตรง แต่อย่างไรก็ตามการให้ความสัมพันธ์ของวัตถุภายในภาพนั้นสามารถช่วยทำการแปลความหมายที่ได้มีความสมบูรณ์ขึ้นอย่างเห็นได้ชัดเจน เพราะฉะนั้นในงานวิจัยที่นำเสนอนี้เป็นอีก

แนวทางหนึ่งที่พยายามจะคิดค้นวิธีการที่จะหาความหมายที่เกิดขึ้นจากภาพ ในอีกมุมมองหนึ่งซึ่งยังคงต้องมีการพิจารณาและวิเคราะห์ปรับปรุงการทดลองต่อไป

สำหรับในกระบวนการจำแนกข้อมูลภาพในงานวิจัยนี้ได้ ทำการพิจารณาภาพ ประกอบด้วยวัตถุของภาพเป็นหลัก และมีการใช้ความสัมพันธ์ที่เกิดจากวัตถุเพื่อให้สามารถสื่อความหมายของภาพได้มากยิ่งขึ้น ผลที่ได้จากการจำแนกภาพโดยทั่วไปสามารถที่จะจำแนกได้อย่างไม่มีปัญหา แต่เมื่อวัตถุของภาพมีจำนวนมากขึ้นทำให้เกิดการแสดงความสัมพันธ์ที่ซับซ้อนมากขึ้นตามลำดับ จนในบางครั้งไม่สามารถทำการจำแนกได้อย่างถูกต้อง รวมทั้งจำนวนของวัตถุที่มีการจัดเก็บเพื่อเป็นข้อมูล อาจจะมีข้อผิดพลาดในการให้ข้อมูลของภาพและความสัมพันธ์ หรือ ควรจะมีการเพิ่มวิธีการที่นำมาใช้ในการช่วยการจำแนกเพิ่มขึ้น เช่น การแสดงออกด้วยท่าทางของมนุษย์นั้นสามารถสื่อความหมายของภาพได้อย่างสมบูรณ์ เช่น การกระโดด การกอด การกุมขมับ การเอนหลัง เป็นต้น และสิ่งที่ควรจะมีการเพิ่มขึ้นคือการใช้หลักการของวัตถุที่โดดเด่น รวมทั้งขนาดและตำแหน่งของวัตถุที่เด่นบนภาพ เพื่อแสดงให้เห็นถึงความหมายหลักที่ควรจะนำไปพิจารณา

****

DPU

บรรณานุกรม

Albert Gordo and Diane Larlus, (2017). Beyond instance-level image retrieval: Leveraging captions to learn a global visual representation for semantic retrieval. Conference on Computer Vision and Pattern Recognition (CVPR17), 6589-6598.

Allan Hanbury (2008). A survey of methods for image annotation. Journal of Visual Languages and Computing archive. Academic Press, Inc. 19(5), 617-627.

Benitez, A.B., and S.-F. Chang. (2002). Semantic Knowledge Construction From Annotated Image Collections. International Conference On Multimedia & Expo (ICME-2002). Lausanne. 26-29.

Benjamin Yao, Xiong Yang and Tianfu Wu. (2009) Image Parsing with Stochastic Grammar: The Lotus Hill Dataset and Inference Scheme. First Workshop SIG-09: First International Workshop on Stochastic Image Grammars.

Everingham, M., Van Gool, L., Williams, C. K. I., Winn, J. and Zisserman, A. (2010). The PASCAL Visual Object Classes (VOC) Challenge. International Journal of Computer Vision. 88(2). 303-338.

F. Giunchiglia, P. Shvaiko, M. Yatskevich. (2004). S-Match: an Algorithm and an Implementation of Semantic Matching. In Proceedings of ESWS.

Feichao Wang. (2011) A Survey on Automatic Image Annotation and Trends of the New Age. Procedia Engineering. 23. 434-438.

G. Varelas, E. Voutsakis, P. Raftopoulou, E. Petrakis, E. Milios. (2005). Semantic similarity methods in wordNet and their application to information retrieval on the web. Proceedings of the 7th annual ACM international workshop on Web information and data management (WIDM’05).

Galleguillos C., Belongie S. (2010). Context Based Object Categorization: A Critical Survey. Computer Vision and Image Understanding (CVIU). 114. 712-722.

Hollink, L., Nguyen, G., Schreiber, G., Wielemaker, J., Wielinga,B., and Worring, M.. (2004). Adding spatial semantics to image annotations. In Proceedings International Workshop on Knowledge Markup and Semantic Annotation.

http://host.robots.ox.ac.uk/pascal/VOC/voc2012/examples/index.html.

https://www.flickr.com/ ค้นคืนเมื่อวันที่ 23 เม.ย. 2560.

DPU

I. Simon, N. Snavely, and S. Seitz. (2007). Scene summarization for online image collections. in IEEE Int. Conf. Computer Vision. 1–8.

J. Euzenat, P. Shvaiko. (2013) Ontology Matching. Springer. 2nd ed. 511. 103.

Jain, A. K. and Vilaya, A. (1996). Image Retrieval Using Color and Shape", Pattern Recognition.29(8). 1233-1244.

Joo-Hwee Lim, Jun Li, Philippe Mulhem, Qi Tian. (2003). Content-based summarization for personal image library. Proceedings of the 3rd ACM/IEEE-CS Joint Conference on Digital Libraries(JCDL03). 393.

JS Hare, PAS Sinclair, PH Lewis, K Martinez, PGB Enser, CJ Sandom. (2006).Bridging the semantic gap in multimedia information retrieval: Top-down and bottom-up approaches. 3rd European Semantic Web Conference.

Junshi Huang, Rogerio S. Feris, Qiang Chen, Shuicheng Yan. (2015). Cross-domain Image Retrieval with a Dual Attribute-aware Ranking Network. Computer Vision and Pattern Recognition, 1062-1070.

K. Barnard, P. Duygulu, D. Forsyth, N. de Freitas, D. M. Blei, and M.I. Jordan. (2003). Matching words and pictures. 3. 1107–1135.

Kevin Lin, Huei-Fang Yang, Kuan-Hsien Liu, Jen-Hao Hsiao, Chu-Song Chen. (2015). Rapid Clothing Retrieval via Deep Learning of Binary Codes and Hierarchical Search. ACM International Conference on Multimedia Retrieval (ICMR).

Kow Kuroda, Francis Bond and Kentaro Torisawa. (2010). Why Wikipedia needs to make Friends with WordNet. The 5th International Conference of the Global WordNet Association (GWC-2010). 9-16.

L. Von Ahn, Liu, and M. Blum. (2006) Peekaboom: a game for locating objects in images. Proceedings SIGCHI conference on Human Factors in Computing Systems. 55–64.

M. Johnson and R. Cipolla. (2005). Improved image annotation and labeling through multi-label boosting. Brit. Machine Vision Conf.

Mathias Lux , Jutta Becker , Harald Krottmaier. (2003). Semantic Annotation and Retrieval of Digital Photos. In Proc. Of CAiSE Proceedings Information Systems for a Connected Society.

Mathias Lux. (2009). Caliph & Emir: MPEG-7 Photo Annotation and Retrieval. Proceedings of the Seventeen ACM International Conferences on Multimedia, Beijing, China. 925-926.

Miller, George A. (1990). WordNet: An on-line lexical database. International Journal of Lexicography.3. 235–312.

Mohammed Alkhawlani and Mohammed Elmogy. (2015). Text-based, Content-based, and Semantic-based Image Retrievals: A Survey. International Journal of Computer and Information Technology. 4(1).

Patwardhan and Pedersen. (2006). Using WordNet Based Context Vectors to Estimate the Semantic Relatedness of Concepts. Proceedings of the EACL 2006 Workshop Making Sense of Sense- Bringing Computational Linguistics and Psycholinguistics Together. 1-8.

Pedro F. Felzenszwalb, Ross B. Girshick, David McAllester and Deva Ramanan. (2010). Object Detection with Discriminatively Trained Part Based Models. IEEE Transactions on Pattern Analysis and Machine Intelligence. 32(9).

Philippe Mulhem, and Joo Hwee Lim. (2002) Symbolic photograph content-based retrieval. In Proceedings of the eleventh international conference on Information and knowledge management. McLean. 94 – 101.

R. Zhao and W. I. Grosky. (2002). Narrowing the Semantic Gap–Improved Text-Based Web Document Retrieval Using Visual Features. IEEE Transactions on Multimedia. 4(2) 189-200.

Russell, B.C., Torralba, A., Murphy, K.P., and Freeman, W.T. (2008). LabelMe: a database and web-based tool for image annotation. International Journal Computer Vision. 77.

S. Shatford Layne. (1994). Some Issues in the Indexing of Imahes. Journal of The America Society for Information Science. 45(8). 583-588.

Smeulders, A. W. M., Worring, M., Santini, S., Gupta, A., and Jain, R. (2000). Content-Based Image Retrieval at the End of the Early Years. IEEE Transactions on Pattern Analysis and Machine Intelligence. 22(12).

Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun, (2016). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. IEEE Transactions on Pattern Analysis and Machine Intelligence,39(6), 1137 – 1149.

Sun Ting and Geng Guohua. (2016). Image Retrieval Method for Deep Neural Network. international Journal of Signal Processing, Image Processing and Pattern Recognition, 9(7),33-42.

Tie Hua Zhou, Ling Wang , and Keun Ho Ryu. (2015). Supporting Keyword Search for Image Retrieval with Integration of Probabilistic Annotation. 7. 6303-6320.

Wan, Ji; Wang, Dayong; HOI, Steven C. H.; Wu, Pengcheng; Zhu, Jianke; Zhang, Yongdong; and Li, Jintao. (2014). Deep learning for content based image retrieval: A comprehensive study. MM '14: Proceedings of the 22nd ACM International Conference on Multimedia, 157-166.

Xin-Jing Wang, Lei Zhang, Xirong Li, Wei-Ying Ma. (2008). Annotating images by mining image search results. IEEE Pattern Analysis and Machine Intelligence.

Y. Wu, J.-Y. Bouguet, A. Nefian, and I. Kozintsev. (2007). Learning concept templates from web images to query personal image databases. IEEE Int. Conf. Multimedia and Expo. 1986–1989.

Ying Liua, Dengsheng Zhanga, Guojun Lua, and Wei-Ying Mab. (2007). A survey of content-based image retrieval with high-level semantics. Pattern Recognition. 40, (1). 262–282.

Zhou, T.H.; Wang, L.; Ryu, K.H. (2015). Supporting Keyword Search for Image Retrieval with Integration of Probabilistic Annotation. Sustainability. 7. 6303-6320.

Zijia Lin, Guiguang Ding, Mingqing Hu, Jianmin Wang and Jiaguang Sun. (2012).Automatic image annotation using tag-related random search over visual neighbors. CIKM '12 Proceedings of the 21st ACM international conference on Information and knowledge management. 1784-1788.

ประวัติผู้วิจัย

ผู้ช่วยศาสตราจารย์นัศพ์ชาณัณ ชินปัญช์ธนะ

สถิติประยุกต์ (คอมพิวเตอร์) สถาบันบัณฑิตพัฒนบริหารศาสตร์

วิทยาศาสตร์บัณฑิต (วิทยาการคอมพิวเตอร์) มหาวิทยาลัยหอการค้าไทย

ปัจจุบัน

อาจารย์ประจำภาควิชาเทคโนโลยีสารสนเทศธุรกิจ คณะเทคโนโลยีสารสนเทศ

มหาวิทยาลัยธุรกิจบัณฑิตย์

คณะกรรมการปรับปรุงกลุ่มผลิตชุดวิชาคอมพิวเตอร์เบื้องต้น

คณะกรรมการกลุ่มผลิตชุดวิชาการระบบสำนักงานอัติโนมัติและพาณิชย์อิเล็กทรอนิกส์

คณะกรรมการกลุ่มผลิตชุดวิชาการจัดการเว็บเซอร์วิส

คณะกรรมการกลุ่มผลิตชุดวิชาการสื่อสารข้อมูลและระบบเครือข่ายคอมพิวเตอร์

สาขาวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยสุโขทัยธรรมาธิราช

รางวัลและเกียรติประวัติ

- รับรางวัล บทความดีเด่น กลุ่มวิทยาศาสตร์เรื่อง ,"การแปลความหมายภาพดิวยวิธีการวัดความคลิายกันของกราฟแบบจับคู่, จาก ประชุมทางวิชาการระดับชาติพะเยาวิจัย ครั้งที่ 6 “วิทยาศาสตร์ เทคโนโลยี และนวัตกรรม เพื่อความเขิมแข็งของชุมชน”, มหาวิทยาลัยพะเยา, จังหวัดพะเยา, มกราคม 2560.
- รับรางวัล The best paper เรื่อง ,"การคัดเลือกข้อมูลเพื่อใช้จำแนกโครงสร้างท่าทางมนุษย์ด้วยซัพพอร์ตเวคเตอร์แมชชีน" จาก The 8th National Conference on Applied Computer Technology and Information Systems 2015.
- รับรองคุณภาพการสอน ปีการศึกษา 2557 จาก มหาวิทยาลัยธุรกิจบัณฑิตย์
- รับรางวัลนักวิจัยดีเด่น สาขาเทคโนโลยีสารสนเทศ 2558 จาก มหาวิทยาลัยธุรกิจบัณฑิตย์
- Microsoft Office Specialist (MOS) Certification (PowerPoint and Word) 2015
- Certificate of Achievement for iOS Development June 2013
- ได้รับทุนการศึกษาประเภทเรียนดี มหาวิทยาลัยหอการค้าไทย พ.ศ. 2534 – 2537
- ได้รับปริญญาตรีเกียรตินิยมอันดับ 1 จาก มหาวิทยาลัยหอการค้าไทย

งานวิจัยที่สนใจ

สนใจงานวิจัยทางด้านการค้นคืนความหมายภาพ (Image retrieval) ทั้งภาพธรรมชาติทั่วไป (natural images) และ ภาพที่เป็นส่วนบุคคล (personal images) ในหัวข้อเกี่ยวกับ การค้นคืน human activity หรือในหัวข้อการแปลความหมายภาพ (semantic human image) โดยเจาะจงทางด้านการใช้ท่าทางของมนุษย์เพื่อแสดงถึงความหมายของภาพ รวมทั้งการจำแนกภาพต่าง ๆ เป็นกลุ่ม (Image classification) และการประยุกต์หลักการประมวลผลภาพ เพื่อนำมาใช้ในทางอุตสาหกรรมได้จริง

หนังสือ

นัศพ์ชาณัฌ ชินปัญช์ธนะ. หน่วยที่ 13 การจัดแฟ้มข้อมูลและการป้องกันระบบคอมพิวเตอร์, *สถาปัตยกรรมคอมพิวเตอร์และระบบปฏิบัติการ (Computer Architecture and Operation Systems) (99315)*, มหาวิทยาลัยสุโขทัยธรรมาธิราช, สาขาวิทยาศาสตร์และเทคโนโลยี 2557.

นัศพ์ชาณัฌ ชินปัญช์ธนะ. หน่วยที่ 1 ความรู้เบื้องต้นเกี่ยวกับเครือข่ายคอมพิวเตอร์ และ หน่วยที่ 3 ชุมชนเครือข่าย, *หลักการบริหารและจัดการเครือข่าย (Network Management) (99412)* ,มหาวิทยาลัยสุโขทัยธรรมาธิราช, สาขาวิทยาศาสตร์และเทคโนโลยี 2556.

นัศพ์ชาณัฌ ชินปัญช์ธนะ. หน่วยที่ 10 โครงสร้างแบบต้นไม้ และ หน่วยที่ 11 โครงสร้างแบบกราฟ, *โครงสร้างข้อมูลและขั้นตอนวิธี (Data Structure and Algorithms) (993314)*, มหาวิทยาลัยสุโขทัยธรรมาธิราช, สาขาวิทยาศาสตร์และเทคโนโลยี 2556.

นัศพ์ชาณัฌ ชินปัญช์ธนะ. หน่วยที่ 12 โลจิสติกส์และโซ่อุปทานในระบบพาณิชย์อิเล็กทรอนิกส์ และหน่วยที่ 13 การชำระเงินในระบบพาณิชย์อิเล็กทรอนิกส์, *ระบบสำนักงานอัตโนมัติและพาณิชย์อิเล็กทรอนิกส์ (Office Automation System and Electronic Commerce) (99311)*, มหาวิทยาลัยสุโขทัยธรรมาธิราช สาขาวิทยาศาสตร์และเทคโนโลยี 2555.

นัศพ์ชาณัฌ ชินปัญช์ธนะ. หน่วยที่ 3 แบบจำลองเครือข่ายและโพรโทคอลกับการบริการผ่านเว็บ หน่วยที่ 5 ความรู้พื้นฐานของภาษาเอ็กซ์เอ็มแอล, *เทคโนโลยีการบริการผ่านเว็บและการประยุกต์ (Web Service) (99301)*, มหาวิทยาลัยสุโขทัยธรรมาธิราช สาขาวิทยาศาสตร์และเทคโนโลยี 2556.

นัศพ์ชาณัฌ ชินปัญช์ธนะ. หน่วยที่ 3 ตัวกลางในการสื่อสารข้อมูล และ หน่วยที่ 10 โพรโทคอล, *การสื่อสารข้อมูล*

DPU

*และระบบเครือข่ายคอมพิวเตอร์ (Data Communications and Networking),*มหาวิทยาลัยสุโขทัยธรรมาธิราช สาขา วิทยาศาสตร์และเทคโนโลยี, 2553.

งานวิจัย

เรื่องการจัดกลุ่มความหมายภาพส่วนบุคคลดิวยขิอความกิจกรรมเหตุการณ์ (Semantic Clustering for Personal Image with Context of Activity Events) มหาวิทยาลัยธุรกิจบัณฑิตย์ 2557.

นัศพ์ชาณัณ ชินปัญช์ธนะ, การแปลความหมายภาพด้วยแนวคิดพื้นฐานความสัมพันธ์ของกราฟแบบลำดับชั้น (Semantic Annotation Model of Hierarchical Relationships based on Conceptual Graph Representation), มหาวิทยาลัยธุรกิจบัณฑิตย์, 2556.

นัศพ์ชาณัณ ชินปัญช์ธนะ, การรู้จำท่าทางมนุษย์จากภาพส่วนบุคคลผ่านการวิเคราะห์การใช้พลังงานร่างกาย (Human Action Recognition from Personal Photos via Analysis of Body Energy Expenditure), มหาวิทยาลัยธุรกิจบัณฑิตย์, 2555.

นัศพ์ชาณัณ ชินปัญช์ธนะ. การแปลความหมายภาพด้วยวิธีการวัดความคล้ายกันของกราฟแบบจับคู่, มหาวิทยาลัยธุรกิจบัณฑิตย์, 2554.

นัศพ์ชาณัณ ชินปัญช์ธนะ, ระบบตรวจนับวัตถุอัตโนมัติด้วยเทมเพลตแมชชิ่งแบบนอร์มอลไลซ์คอรรีเลชั่น (Automatic Counting Objects System by using Template Matching with Normalized Correlation) 2553.

นัศพ์ชาณัณ ชินปัญช์ธนะ. การจำแนกความหมายของภาพจากวัตถุโดยใช้หลักการโครงสร้างสเกตรีตรอน, มหาวิทยาลัยธุรกิจบัณฑิตย์, 2552.

ผลงานทางวิชาการ

นัศพ์ชาณัณ ชินปัญช์ธนะ, “ศึกษาการสกัดคุณลักษณะข้อมูลสำหรับระบบค้นคืนข้อมูลภาพ,” วารสารมหาวิทยาลัยคริสเตียน, 2560.

N. Chinpanthana, “Similarity Images with Hierarchical Graph based on WordNet Thesaurus,” The Journal of Information Science and Technology (JIST) (NCIT2017 selected paper), 2017.

N. Chinpanthana, “Personal Image Retrieval with Hierarchical Similarity Measure Based on WordNet Framework” 2017 International Conference on Materials Engineering and Functional Materials (ICMFM 2017) Hanoi, Vietnam, 2017.

อนวัช พริบไหว,นัศพ์ชาณัณ ชินปัญช์ธนะ และนันทิกา ปริญญาพล, “การพัฒนาแอปพลิเคชันบนมือถือเพื่อค้นหาพื้นที่เพาะปลูกข้าวและพื้นที่ป่าไม้ในประเทศไทยด้วยภาพถ่ายดาวเทียมและระบบสารสนเทศภูมิศาสตร์,” งานประชุมวิชาการระดับชาติ วิทยาศาสตร์และเทคโนโลยีระหว่างสถาบัน ครั้งที่5, โรงแรมมิราเคิลแกรนด์ คอนเวนชั่น, กรุงเทพ 2560.

นัศพ์ชาณัณ ชินปัญช์ธนะ และ สาราญ ไผ่นวล, การแปลความหมายภาพดิวยวิธีการวัดความคลิายกันของกราฟแบบจับคู่, งานประชุมทางวิชาการระดับชาติพะเยาวิจัย ครั้งที่ 6 “วิทยาศาสตร์ เทคโนโลยี และนวัตกรรม เพื่อความเขิมแข็งของชุมชน”, มหาวิทยาลัยพะเยา, จังหวัดพะเยา, มกราคม 2560. (บทความดีเด่น กลุ่มวิทยาศาสตร์)

นัศพ์ชาณัณ ชินปัญช์ธนะ และเตชค์ฐสิณป์ เพียซิาย, ระบบตรวจนับวัตถุอัตโนมัติดิวยเทมเพลตแมชชิ่งแบบนอร์มัลไลซ์คอรีเลชัน” การประชุมวิชาการระดับชาติและนานาชาติ ราชภัฏสุราษฎร์ธานีวิจัย ครั้งที่ 12, มหาวิทยาลัยราชภัฏสุราษฎร์ธานี, ตุลาคม,2559.

นัศพ์ชาณัณ ชินปัญช์ธนะ, การวัดความคลิายกันของภาพดิวยการเชื่อมโยงกราฟแบบลาดับขั้นบนพื้นฐานของอนุกรมวิธาน WordNet,การประชุมวิชาการระดับประเทศทางดิานเทคโนโลยีสารสนเทศ (NCIT),2559.

นัศพ์ชาณัณ ชินปัญช์ธนะ และ สำราญ ไผ่นวล, “การแปลความหมายภาพด้วยวิธีการวัดความคล้ายกันของกราฟแบบจับคู่,” งานประชุมทางวิชาการระดับชาติพะเยาวิจัย ครั้งที่ 6 วิทยาศาสตร์ เทคโนโลยี และนวัตกรรม เพื่อความเข้มแข็งของชุมชน , มหาวิทยาลัยพะเยา, จังหวัดพะเยา, มกราคม 2560. (The best paper award)

N. Chinpanthana, “Semantic Clustering Based on Context of Activity Events in Personal Photo Collections,” International Workshop on Smart Info-Media Systems in Asia, Chiba Institute of Technology Narashina, Japan, 2015.

นัศพ์ชาณัณ ชินปัญช์ธนะ, สำราญ ไผ่นวล และ ริญญรัตน์ โชติสุริยสินสุข, “การศึกษางานวิจัยการประมวลผลภาพดิจิทัลและการประยุกต์ใช้งานในแอพพลิเคชัน,” การประชุมวิชาการระดับชาติ มหาวิทยาลัยราชภัฏเพชรบูรณ์ ครั้งที่ 3, 2559.

วิธารัตน์ ประดับ, นัศพ์ชาณัณ ชินปัญช์ธนะ และ ริญญรัตน์ โชติสุริยสินสุข, “ทิศทางงานวิจัยของการแปลความหมายภาพด้วยการแบ่งส่วนข้อมูลภาพ,” การประชุมวิชาการระดับชาติ มหาวิทยาลัยราชภัฏนครปฐม ครั้งที่ 8, จังหวัดนครปฐม, 31 มีนาคม - 1 เมษายน 2559.

นัศพ์ชาณัณ ชินปัญช์ธนะ, “การจัดกลุ่มความหมายภาพเหตุการณ์กิจกรรมด้วยฟัซซี่ซีมีน” งานประชุมทางวิชาการระดับชาติพะเยาวิจัย ครั้งที่ 5 “วิทยาศาสตร์ เทคโนโลยี และนวัตกรรม เพื่อความเข้มแข็งของชุมชน”, มหาวิทยาลัยพะเยา, จังหวัดพะเยา, มกราคม 2559.

นัศพ์ชาณัณ ชินปัญช์ธนะ, “ทิศทางอนาคตของงานวิจัยทางด้านการประมวลผลภาพดิจิทัลในการแปลความหมายภาพ, ในงานประชุมวิชาการระดับชาติและนานาชาติมหาวิทยาลัยราชภัฏอุตรดิตถ์ ครั้งที่ 11, มหาวิทยาลัยราชภัฏอุตรดิตถ์, จังหวัดอุตรดิตถ์, ธันวาคม 2558.

นัศพ์ชาณัณ ชินปัญช์ธนะ, “การจำแนกความหมายภาพด้วยการสกัดฟีเจอร์จากโครงสร้างสเกตรีตรอนบนพื้นฐานแนวคิดกราฟลำดับชั้น” , วารสารแม่โจ้เทคโนโลยีสารสนเทศและนวัตกรรม,1, 2558.

นัศพ์ชาณัณ ชินปัญช์ธนะ, “การคัดเลือกข้อมูลเพื่อใช้จำแนกโครงสร้างท่าทางมนุษย์ด้วยซัพพอร์ตเวคเตอร์แมชชีน”, “The 8th National Conference on Applied Computer Technology and Information Systems”, วันที่ ๓๐ – ๓๑ มกราคม ๒๕๕๘, มหาวิทยาลัยนครพนม,จังหวัดนครพนม. (The best paper award)

นัศพ์ชาณัณ ชินปัญช์ธนะ, “การจำแนกความหมายภาพด้วยเรเดียลเบสิสฟังก์ชันบนพื้นฐานของแนวคิดกราฟลำดับชั้น” ,นานาวิทยาการวิจัยและนวัตกรรมเพื่อการพัฒนาที่ยั่งยืนสู่เอเซีย” วันศุกร์ที่ 10 ตุลาคม 2557 , วิทยาลัยบัณฑิตเอเซีย จังหวัดขอนแก่น

N. Chinpanthana and T. Phiasai., “Kernel-based on Data Fusion for Image Classification with Body Energy Action Model”, International Journal of Signal Processing Systems, vol. 1, no. 2, December, 2013.

N. Chinpanthana and T. Phiasai., “Kernel-based on Data Fusion for Image Classification with Body Energy Action Model”, The 2013 5th International Conference on Signal Processing Systems (ICSPS 2013), Sydney, Australia, 2013.

N. Chinpanthana, “Semantic Similarity Measure with Conceptual Graph-Based Image Annotation”, International Conference on Advanced Computer Science Application and Technologies (ASCAT 2012), Kualalumpur,Malaysia, Nov., 2012.

N. Chinpanthana and T. Phiasai., “Automatic Counting System With Normalized Correlation Coefficient Template Matching”, International Conference on Computer and Information Technology, Amsterdam, Netherlands, July 13-15, 2011.

N. Chinpanthana, “Integrating Qualitative Features with Feature Selection for Semantic Image Classification”, International Conference on Management technology and applications (ICMTA2010), Singapore,10-12 Sept.,2010.

N. Chinpanthana and T. Phiasai., “Multi-Layer Perception Networks for Semantic Image Classification with Structural Skeleton Framework”, International Technical Conference on Circuit/Systems Computers and Communications, Pattaya, Thailand,July., 2010.

N. Chinpanthana, “Extracting Features with Structural Skeleton Framework for Semantic Image Classification by using Supporting Vector Machine”, The 3rd International Conference on Advanced Computer Theory and Engineering, Chengdu, China, 20-22 Aug.,2010.

N. Chinpanthana, “Semantic Semantic Salient Images Based on Similarity Matching with Conceptual Graph”, International Technical Conference on Circuits/Systems, Computers and Communications , Jeju, Korea 2009.

S. Chinpanchana, S. Maneewongvatana, and B. Thipakorn, “Semantic Human Image Classification Based on Energy Action Model with Essential Reference points”, International Symposium on Communications and Information Technologies, 16-19 Oct, Sydney, AUS, 2007.

S. Chinpanchana, S. Maneewongvatana, and B. Thipakorn, “High-Level Semantic Image Classification by Using Energy Expenditure”, International Workshop on Smart Info-Media Systems, Thailand, 2007.

S. Chinpanchana, “Semantic Human Action Classification Based on Energy-Action Model”, Tencon 2006 IEEE Region 10 , Hongkong, China, 2006.

S. Chinpanchana, S. Maneewongvatana, and B. Thipakorn, “Semantic Personal Image Classification by Energy Expenditure,” International Symposium on Communications and Information Technologies, Beijing, China, 2005.

S. Chinpanchana, S. Maneewongvatana and B. Thipakorn, “Semantic Personal Image Pattern Classification Based on Human Body,” Asia Information Retrieval Symposium, Beijing, China, Oct. 2004.

S. Chinpanchana and B. Thipakorn, “Semantic Classification of Personal Images Based on Human Action and Associate Bayesian Rule,” International Technical Conference on Circuits/Systems, Computers and Communications, Sendai, Japan, 2004.

****