บทที่ 2
ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 การรับรู้เรื่องสีของมนุษย์ (Human Visual Perception) ด้วยภาษา

การสืบค้นข้อมูลจากอินเทอร์เน็ตหรือฐานข้อมูลในส่วนของภาพโดยใช้ภาษามนุษย์จำเป็นจะต้องแบ่งการวิเคราะห์เป็นสององค์ประกอบคือ องค์ประกอบที่ 1 ในเรื่องของการประมวลผลภาษาธรรมชาติ เพื่อทำความเข้าใจในคำองค์ประกอบของนามวลีที่ใช้ว่าแต่ละคำมีรูปแบบเป็นอย่างไร คำย่อยใดขยายส่วนใดในคำของนามวลีที่ใช้ ตลอดจนน้ำหนักของคำขยายรวมทั้งคำที่มาก่อนมาหลัง เช่น สีเหลืองอมส้ม ทั้งนี้การตีความต้องอาศัยจากการรับรู้ของมนุษย์เข้ามาเป็นส่วนประกอบ องค์ประกอบที่ 2 คือการวิเคราะห์สีจากภาพซึ่งมีหลายรูปแบบ โดยในส่วนนี้พัฒนามาจากพื้นฐานของทฤษฎี RGB เช่น CNS (Color Naming System) ซึ่งนำเสนอโดย Berk and Kaufman (Berk and Kaufman, 1982) โดย CNS เป็นโมเดลที่ประกอบด้วยสี แดง ส้ม น้ำตาล เหลือง เขียว น้ำเงินและ สีม่วง (สีม่วง : purple เป็นสีที่ผสมระหว่างสีแดงและน้ำเงิน โดยมีการเพิ่มส่วนของแสง (Light) เข้าไปประกอบด้วย) รวมทั้งมีการผนวก ส่วนของ สีดำ เทา และขาวเข้าไปด้วย (Conway , 1992) นอกจากนี้โมเดลสีที่นิยมใช้อีกประเภทคือ HSV (Hue , Saturation , Value) (Qingyong et al.,2006) ด้วยเหตุผลที่ว่าเป็นโมเดลที่มีลักษณะให้ผลใกล้เคียงกับการรับรู้ของมนุษย์มาก โดย Hue จะทำหน้าที่ในการแยกแยะสีต่างๆ(เช่น red, green, blue) ออกจากกัน รวมทั้งการกำหนดน้ำหนักของสีแต่ละประเภท ส่วน Saturation ทำหน้าที่กำหนดขนาดของแสงที่จะปรากฏในสีต่างๆเช่นสีแดงมีขนาดของ Saturation มากกว่าสีชมพู (Borghesani , 2009) ส่วน Value ใน HSV โมเดลหมายถึงความเข้มของแสง (Light Intensity) การทำ Color Quantization ในตัวแบบของสีต่างๆนั้น มีเจตนาที่จะลดเวลาในการคำนวณรวมทั้งยังสามารถแสดงความหมายของสีใกล้เคียงกับการรับรู้ของมนุษย์อีกด้วย การรับรู้เรื่องสีของมนุษย์จะเกี่ยวเนื่องกับ Hue มากกว่า Saturation และ Value การทำ Quantization แบ่งเป็นสองลักษณะคือแบบ Uniform และแบบ Non – uniform

การกำหนด Semantic Rule เซต M เพื่อทำการสร้างสัมพันธ์ภาพระหว่างมูลค่าของสีในทางคณิตศาสตร์ด้วยตัวแบบสีแต่ละประเภทนั้นทำโดยการสร้างความสัมพันธ์ระหว่าง Low Level feature มายัง Highly level fuzzy semantic (T (x) : $v \in V, t \in T(x)$) ในการประมวลผล

ในส่วนของสี โดยที่ค่าของ T (x) คือ [0,1] ซึ่งเรียกว่า degree of membership ของ v ที่สอดคล้องกับ linguistic value t ดังนั้นการศึกษาวิธีการสร้าง Semantic set M จึงเป็นเรื่องที่สำคัญในการแก้ปัญหา แนวทางการสร้างในประเด็นนี้มีงานศึกษาอยู่หลายวิธี เช่นใช้ ANN (Artificial Network) โดยงานวิจัยส่วนใหญ่มักจะทำโดยใช้เทคนิคของ ANN โดยมีความแตกต่างกันในส่วนของการออกแบบสถาปัตยกรรมในส่วนของ Hidden Layer และการกำหนดจำนวน Node รวมทั้งการเลือกใช้ activate function ของแต่ละ Node ในงานของ Qingyong (Qingyong et al.,2006)ใช้ hyperbolic tangent function ($\phi(x) = \tanh(\beta x)$: $\beta = 0.25$ นอกจากนี้ในงานวิจัยอื่นๆ มักจะนิยมใช้เทคนิค SVM ในการทำหน้าที่เป็น Classifier แต่เนื่องจาก SVM มีข้อจำกัดในส่วนที่เป็น Binary classifier ดังนั้น Qingyong (Qingyong et al.,2007) จึงนำหลักการของ genetic algorithm มาใช้ในการสร้าง Semantic rule เพื่อมาสนับสนุนการศึกษา

ปัจจัยที่สำคัญอีกส่วนหนึ่งในการศึกษาคือ การกำหนดไวยกรณ์ภาษาของสีในการสื่อถึงภาพ ในงานวิจัย (Qingyong et al.,2009) ได้กำหนด Syntax rule คือการกำหนด ไวยกรณ์ของภาษาดังนี้คือ

< Linguistic Expression >:= <Extended term> | <Basic term>

< Extended term> := <Extended> <Basic term>

< Extended > := many | somewhat | few

< Basic term> := red | orange | brown | yellow | green | blue | purple | black | grey | white

นอกเหนือจากการกำหนดไวยกรณ์เพื่อสนับสนุนการสร้าง Semantic rule แล้วยังต้องอาศัยกลไกทางคณิตศาสตร์ในการวิเคราะห์ค่าต่างๆในตัวแบบของสีเข้าร่วมด้วย เทคนิคเหล่านี้มีหลายวิธีด้วยกัน ดังที่กล่าวแล้ว โดยที่วิธีที่ค่อนข้างจะยืดหยุ่นกับลักษณะความเป็นภาษามนุษย์ที่นิยมใช้กันคือ ตรรกแบบคลุมเครือ (Fuzzy logic) (Bhoyer and Kakde , 2009 , Dhaval et al., 2003, Pal and Mitra, 2001 , Swiain and Ballard , 1999) ซึ่งเป็นรูปแบบหนึ่งของ multi value logic โดยสร้างมาจากพื้นฐานทฤษฎีของ Fuzzy set ซึ่งแตกต่างจาก “crisp logic” ซึ่งให้ค่าเพียง 2 ค่า (binary logic : Yes or No) โดย Fuzzy set ให้ค่าที่เรียกว่า degree of membership อยู่ระหว่าง 0 ถึง 1 โดยการวัดค่าสามารถสนองตอบต่อ การแทนความหมายในความรู้สึกในภาษามนุษย์ได้ค่อนข้างเหมาะสม (Lao and Savakis, 2001, Biacino and Geta,

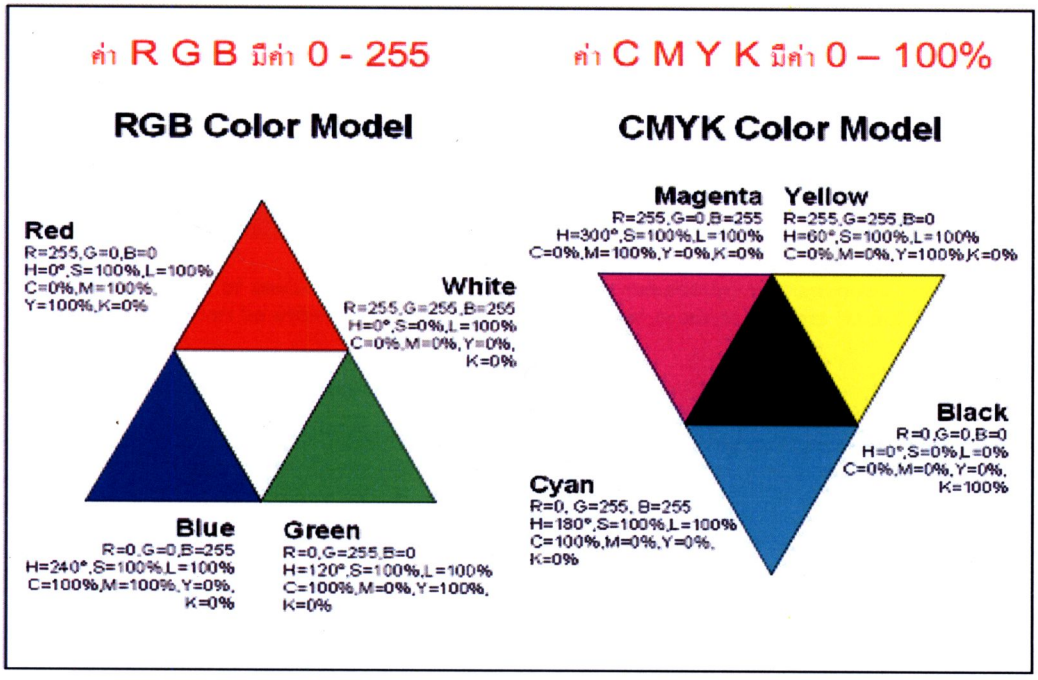
2002) ตัวอย่างเช่น ถ้าเป็นการกล่าวถึงอุณหภูมิโดยใช้ crisp set จะกล่าวว่า ไม่หนาว (ร้อน) กับหนาว แต่ถ้าเป็นการใช้ความเป็น membership จะกล่าวว่า degree of membership ของความร้อน คือ 0.7 ความร้อน ซึ่งถ้าใช้เป็นภาษาจะกล่าวว่า "ค่อนข้างร้อน" ในขณะที่ถ้า membership ของความร้อน คือ 0.95 ก็มีความหมายว่า "ร้อนมาก" (Very hot) เป็นต้น คำที่ใช้เพื่อสื่อความหมายในภาษานั้นมีความหมายแตกต่างกันตามที่มนุษย์ใช้ เช่น not hot, slightly warm, fairy cold ปรากฏการณ์การใช้ภาษาเช่นนี้มักเกิดขึ้นในหลายบริบท นอกเหนือจากเรื่องของสภาพอากาศ เช่นลักษณะของความหมายของสีก็เช่นกัน แต่เนื่องจากในส่วนของสีมีส่วนขององค์ประกอบที่แตกต่างกันและซับซ้อนมากกว่าในเรื่องของอากาศในประเด็นที่ว่าสีมีค่าของสีและมีรายละเอียดมากกว่า ตัวอย่าง เช่นในภาษาอังกฤษมีการใช้ภาษาสำหรับการสื่อถึงเรื่องของสีในหลายรูปแบบเช่นใช้คำมาขยายลักษณะสีเช่น "Many Red" , "Slightly Red" หรือไม่ก็ใช้ภาษาโดยเทียบเคียงกับสิ่งที่มนษุย์รู้จักเช่น "Green Grass" ("เขียวเหมือนหญ้า") ,"Blue Sky" หรือไม่ก็สร้างคำขึ้นมาแทนที่สีเช่น "น้ำตาลอมแดง (russet or reddish brown)" การอธิบายสีโดยใช้วัตถุหรือเหตุการณ์ที่มนุษย์รู้จักนั้นมีการใช้ในทุกภาษารวมทั้งภาษาไทย นอกจากนี้ในส่วนของภาษาไทยก็มีการสร้างรูปแบบที่พิเศษอีกแบบที่นิยมใช้เมื่อมีการกล่าวถึงวัตถุที่มีสองสี โดยการสร้างนามวลีขึ้นมาใช้ เช่น"สีส้มอมแดง", "สีแดงอมส้ม" ซึ่งลักษณะพิเศษนี้ยังมีความซับซ้อนเพิ่มขึ้นในส่วนที่ว่าการสลับที่ของคำที่บ่งสีในภาษาไทยก็มีผลทำให้ความหมายของสีของภาพในความเข้าใจของมนุษย์แตกต่างไปด้วย งานวิจัยนี้จะใช้รูปแบบของการสร้างนามวลี (Noun Phase) โดยการสร้างรูปแบบ Regular expression เป็น Syntax rule ของนามวลี โดยใช้ในกลุ่มคำนามของ สีแดง เหลือง ส้ม ประกอบกับคำกิริยาคือ คำว่า "อม" "แกม" และ "เจือ" โดยนามวลีที่ได้จะมีความหมายในความเข้าใจของมนุษย์ที่รับรู้

การประมวลผลด้วยภาพนั้น ส่วนสุดท้ายที่สำคัญที่สุดในระบบคือการรับรู้ภาพนั้นด้วยด้วยคอมพิวเตอร์และสายตาของมนุษย์ซึ่งต้องตรงกัน สืบเนื่องจากกระบวนการรับรู้ภาพของมนุษย์ใช้ทฤษฎีสีแบบ HSV (HVS: Human Visual System) ดังนั้นการประมวลผลภาพบนทฤษฎีสี แบบ HSV จึงเป็นส่วนหนึ่งที่จำเป็นจะต้องเรียนรู้เพื่อที่จะดำเนินการประมวลผลภาพด้วยคอมพิวเตอร์ (Han and Ma, 2010) เพื่อเทียบเคียงกับการรับรู้ด้วยสายตาของมนุษย์เพื่อการรับรู้สีของวัตถุในภาพได้อย่างถูกต้องและตรงกันระหว่างภาพในการวิเคราะห์ด้วยคอมพิวเตอร์กับการรับรู้ของมนุษย์ การศึกษาประเด็นนี้ประกอบด้วยพื้นฐานสนับสนุนหลายประการคือทฤษฎีสี Intensity, Luminance, Brightness, Simultaneous contrast, Mach bands, Transfer function , Color perception ตัวแบบคณิตศาสตร์ และกฎที่ใช้ในการตีความ

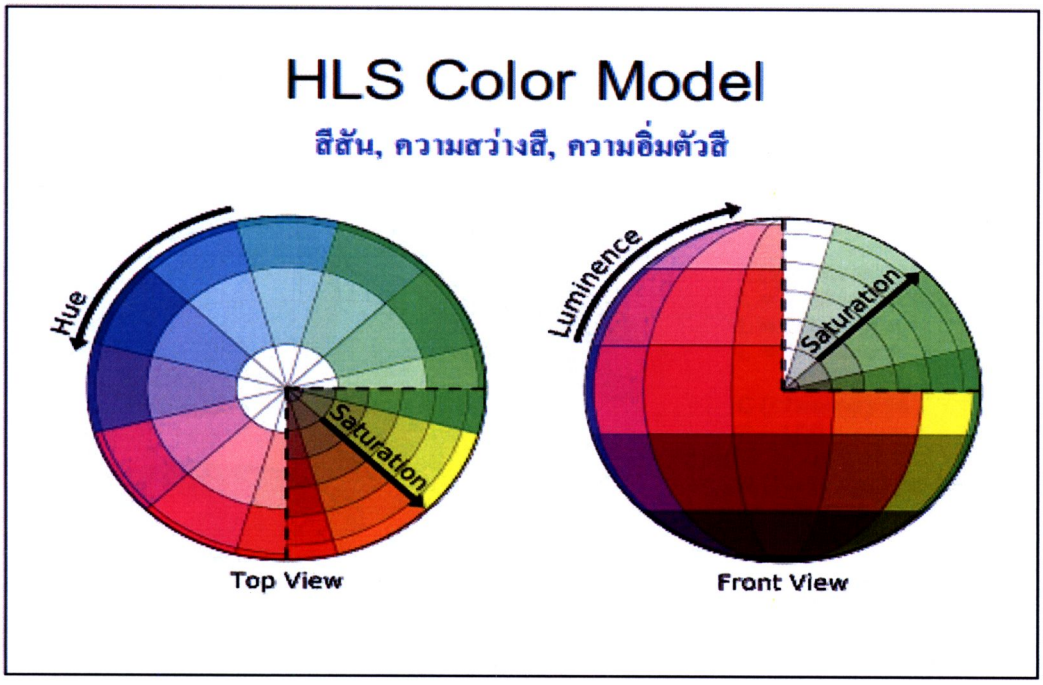
2.2 Color Perception

เนื่องจากภาพสีมีความซับซ้อนในการรับรู้ของมนุษย์มากกว่าภาพขาวดำ ดังนั้น การศึกษาลักษณะมูลค่าของภาพสีจึงเป็นความจำเป็น โดยอยู่บนพื้นฐานของ Color stimulus ซึ่งถูกกำหนดในเทอมของ SED (Spectral Energy System) คือ $C(\lambda)$ ซึ่งมีลักษณะเดียวกับค่าของ light intensity โดยที่ค่านี้จะส่งผลต่อการรับรู้เรื่องสี เช่นสีแดง สีเหลือง และสีเขียวเป็นต้น ความแตกต่างระหว่าง luminance กับ color คือสีไม่สามารถแทนด้วยค่าเพียงค่าเดียวได้ การแทนค่าสีประกอบด้วย brightness, hue และ saturation โดยที่ brightness หมายถึงการรับรู้ของมนุษย์ในส่วนที่เป็น luminance ของวัตถุนั้น hue หมายถึงค่า dominant ของ wavelength โดยการผสมระหว่าง light waves ซึ่งส่งผลให้เป็นสีตามการรับรู้เช่นสีแดง เหลือง หรือ ชมพู เป็นต้น ในขณะที่ ค่า saturation หมายถึงความสัมพันธ์ระหว่างปริมาณสีขาว ที่ผสมกับค่า hue ดังนั้นค่าทั้งสามนี้จึงเป็นคุณลักษณะที่ประกอบการเป็นสีในการรับรู้ของมนุษย์

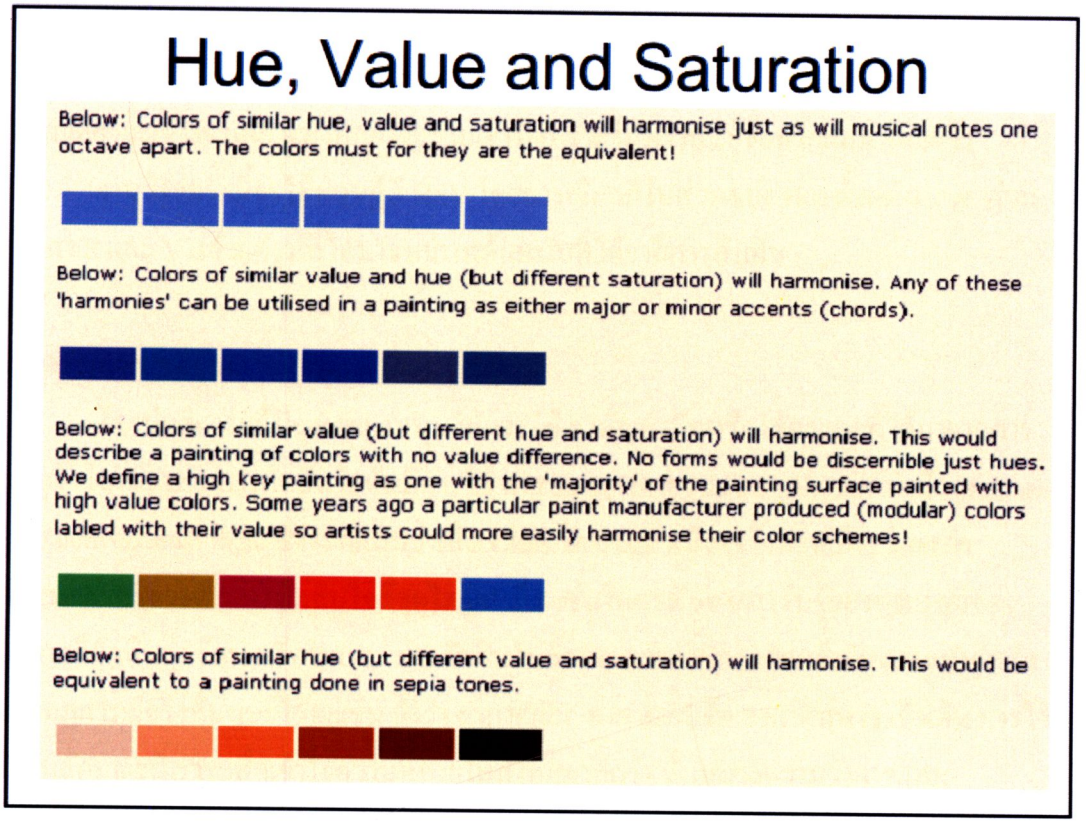
ค่า SED $C(\lambda)$ อาจเป็นค่าต่อเนื่องของ λ หรืออาจจะเป็นค่าในช่วงของ λ ก็ได้ เนื่องจาก $C(\lambda)$ ไม่ได้เป็นค่าเดียวในรูปแบบของการรับรู้ในส่วนของสี ในทางปฏิบัติเราสามารถกำหนดชุดของค่าพารามิเตอร์ในสีใดสีหนึ่งได้ตามการศึกษา เนื่องจากเรตินาในตาของมนุษย์ประกอบด้วย cone อยู่ 3 ประเภท โดยมีความทับซ้อนกันอยู่ในส่วนของ spectral sensitivities เช่นในกรณีมีค่าประมาณ 445 mm หมายถึงสีน้ำเงิน ในขณะที่ค่า 535 mm หมายถึงสีเขียวเป็นต้น การรับรู้ในส่วนที่เกี่ยวกับแสงนั้น cone จะทำการบูรณาการส่วนน้ำหนักของ intensity ในสัดส่วนที่เหมาะสมออกมาเป็นปริมาณในการรับรู้ของมนุษย์ กระบวนการดังกล่าวสามารถแทนได้ด้วยทฤษฎีการรับรู้สีที่เรียกว่า tri-stimulus theory ทฤษฎีนี้อธิบายว่า สีที่ปรากฏในการรับรู้คือ HSV ประกอบด้วยปัจจัย 3 ประเภทที่รวมกันแล้วจะแทนที่ด้วย spectral energy distribution โดย 3 ปัจจัยที่เป็นคุณลักษณะนี้ ประกอบด้วย ลักษณะของสีที่เรียกว่า primaries โดยที่ลักษณะที่เป็นสีหลัก (dominant color) คือแดง เขียว และน้ำเงิน ประกอบกับส่วนผสมที่เรียกว่า additive color คือ เหลือง สีแดงม่วง และสีอำพัน (สีน้ำเงินกับเขียว) ผลลัพธ์จากการผสมระหว่าง dominant กับ additive color เรียกว่า color subtraction สรุปก็คือสีแต่ละสีประกอบด้วยค่าเอกลักษณ์จากค่าพารามิเตอร์ 3 ค่าดังที่กล่าวมาแล้ว ทฤษฎีของสีมีหลายแบบด้วยกันเช่น RGB ใช้ในการแสดงภาพของจอภาพ ในขณะที่ CMYK เป็นทฤษฎีสีที่ใช้สำหรับการพิมพ์ ส่วนการรับรู้สีของมนุษย์ใช้ ทฤษฎี HSV (HSI)



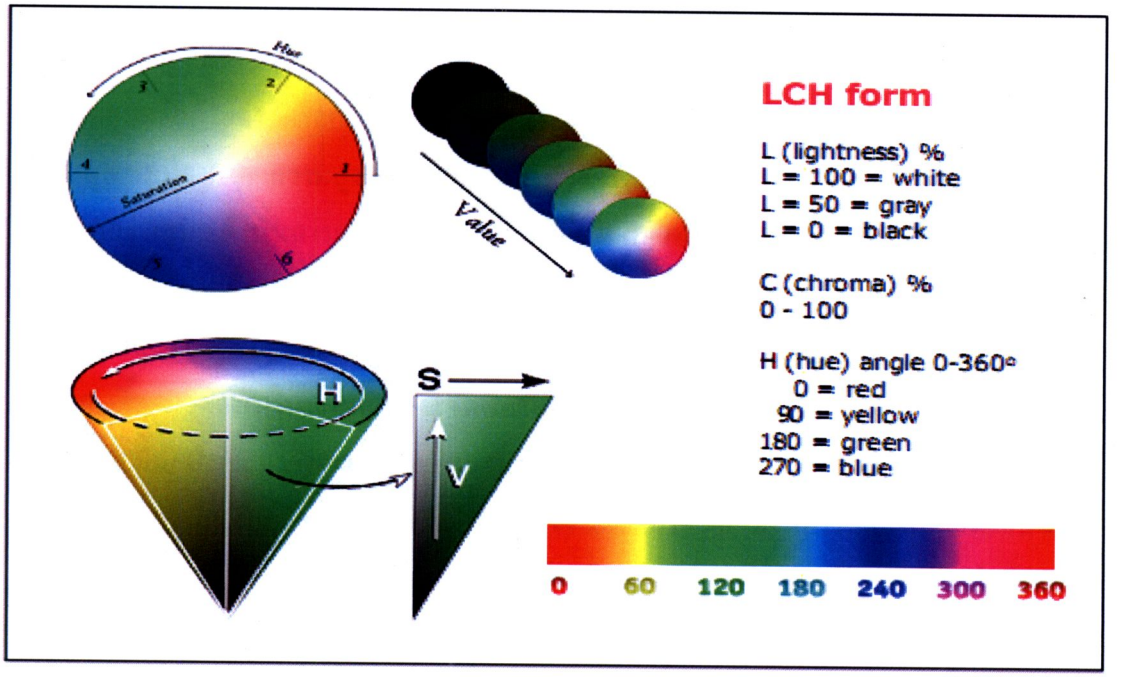
ภาพที่ 2. แสดงภาพสีบนจอภาพโดยใช้ทฤษฎีสีแบบ RGB ส่วนการแสดงสีของเครื่องพิมพ์ใช้ทฤษฎีสีแบบ CMYK



ภาพที่ 3. HLS (HSV) Hue คือค่าของสีแบบต่างๆ เช่นแดงเขียว ฟ้า Hue จะเกี่ยวข้องกับ Wavelength ค่าของ Value หรืออาจเรียกว่า Brightness หรือ Lightness (ความสว่าง) และ Saturation หรือที่เรียกว่าค่า Chroma คือค่าความอิ่มตัวของสี



ภาพที่ 4. เปรียบเทียบ Hue, Value และSaturation ในระดับต่างๆกัน



ภาพที่ 5. แสดงสีในระดับต่างของ HSV

ในส่วนการรับรู้เรื่องของสี (Kubat et al., 2007) ได้ทำการศึกษาการรับรู้ของมนุษย์ในบริบททางสังคมที่อยู่เช่น ใช้คำพูดสีน้ำตาลในความหมายที่เทียบเคียงกับ ช็อกโกแลต ซึ่งคนในสังคมเดียวจะเข้าใจ แต่ต่างสังคมอาจจะไม่เข้าใจเพราะไม่ทราบว่า ช็อคโกแลต คืออะไร

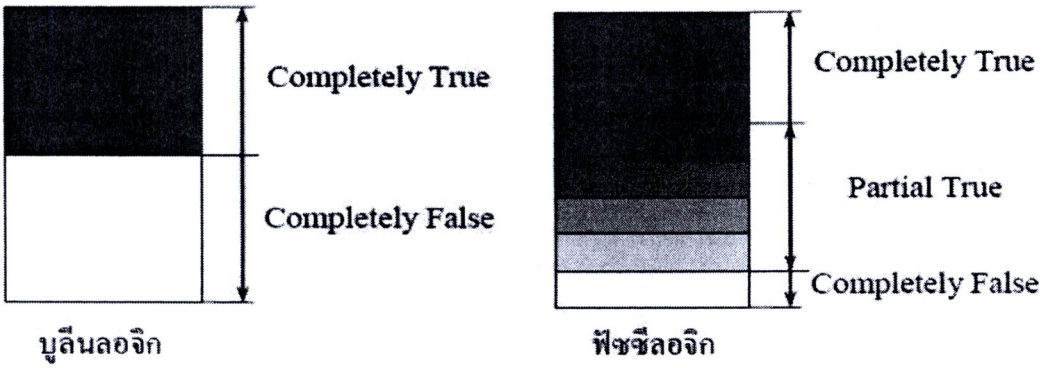
งานวิจัยครั้งนี้จะใช้ทฤษฎีสี HSV โดยการเลือกสีในส่วนของ Hue ตั้งแต่ 0 – π ส่วนค่าของ S และ V แล้วแต่รูปที่ได้มาจากภาพต้นฉบับที่ได้จากการสืบค้น

2.3 ตรรกะแบบคลุมเครือ (Fuzzy Logic)

โดยปกติการแก้ปัญหาของมนุษย์มักจะใช้ สามัญสำนึกหรือใช้ตรรกะที่ได้มาจากการเรียนรู้ ในกระบวนการตัดสินใจนั้นมีการใช้ ตรรกะ อยู่หลายรูปแบบ รูปแบบหนึ่งคือ Boolean หรือ conventional logic ซึ่งจะตัดสินว่าเป็น Class ใด Class หนึ่ง (จริงหรือเท็จ) โดยการกำหนดค่าหรือระดับการตัดสินใจที่จุดใดจุดหนึ่ง เช่นในกรณีของการรับรู้ของอุณหภูมิของอากาศก็ตัดสินว่า ร้อน หรือหนาว การตัดสินในลักษณะเช่นนี้มักจะขัดแย้งกับธรรมชาติหรือปรากฏการณ์จริงที่มนุษย์ใช้อยู่เช่น คำว่า อากาศเย็น อากาศเย็นจัด อากาศหนาว เป็นต้น การใช้คำในลักษณะนี้เราจะสื่อเข้าใจความหมายในสังคมมากกว่า นั่นหมายความว่าการแทนความหมายหรือการตัดสินวัตถุ หรือปรากฏการณ์โดยใช้ค่า เพียง 2 ค่า (Binary Number) นั้นจะไม่ตรงกับความรู้สึกของมนุษย์อย่างแท้จริง ดังนั้นการใช้ fuzzy logic เข้ามาช่วยในการตัดสินโดยใช้ค่าที่มีความหมายในเชิงระดับ หรือที่เรียกว่า degree of membership จึงเหมาะสมกว่า ในทางทฤษฎี fuzzy logic หมายถึงแนวทางในการสร้างค่าที่แทนลักษณะหรือปรากฏการณ์ที่มีความคลุมครือด้วยค่า ที่เรียกว่า degree แทน ด้วยเหตุที่ว่าปรากฏการณ์ต่างๆรอบตัวมนุษย์ เช่น อุณหภูมิอากาศ ความสูง ความเร็ว ระยะทาง เป็นตัวอย่างของสิ่งที่มีลักษณะคลุมเครือปรากฏอยู่ เสมอ ดังนั้นการแทนความแตกต่างโดยใช้กฎเกณฑ์ในการตัดสินบนพื้นฐานของ binary logic เช่น ถ้า เราตัดสินว่า ขนาดความสูง 180 ซม.เป็นการแบ่งระหว่าง “สูง” กับ ‘เตี้ย” ดังนั้น ถ้านาย ก.สูง 190 ซม. จะตัดสินว่า นาย ก.สูง ในขณะที่ถ้านาย ข.สูง 179 ซม จะตัดสินว่า นาย ข. เตี้ย การตัดสินในลักษณะนี้จะขัดกับรู้สึกของมนุษย์เพราะความแตกต่างเพียงเล็กน้อยนี้ไม่ควรจะมีนัยสำคัญมากพอที่จะตัดสินว่า คนสองคนนี้อยู่คนละกลุ่ม ด้วยเหตุผลดังกล่าวจึงนำไปสู่แนวทางการแก้ปัญหาด้วยวิธีการใช้ fuzzy logic มาช่วยเพื่อให้การตัดสินมีความหมายเชิงมนุษย์เข้าไปกำกับด้วย ค่าที่ fuzzy logic สร้างเรียกว่า ค่า degree of fuzzy ซึ่งเป็นค่า multi value logic ทฤษฎีของ fuzzy logic พัฒนามาตั้งแต่ ปี คศ. 1930 ต่อมาได้มีการพัฒนาให้เหมาะสมกับการใช้งานในเชิงประยุกต์มากยิ่งขึ้นโดย L. A. Zadeh ในปี ค.ศ. 1965

สำนักงานคณะกรรมการวิจัยแห่งชาติ
ห้องสมุดงานวิจัย
วันที่ 17 ม.ค. 2555
เลขทะเบียน 248639
เลขเรียกหนังสือ

ตรรกะแบบฟัซซี(fuzzy logic) เป็นเครื่องมือที่ช่วยในการตัดสินใจภายใต้ความไม่แน่นอนของข้อมูลโดยยอมให้มีความยืดหยุ่นได้ตามธรรมชาติของมนุษย์ ซึ่งเป็นหลักการที่คล้ายกับการเลียนแบบวิธีความคิดที่ซับซ้อนของมนุษย์ ตรรกะแบบฟัซซีมีลักษณะที่พิเศษกว่าตรรกะแบบจริงเท็จ (Boolean logic) เนื่องจากเป็นแนวคิดที่มีการต่อขยายในส่วนของความจริง (partial true) โดยค่าความจริงจะอยู่ในช่วงระหว่างจริง (completely true) กับเท็จ (completely false) ในขณะที่ตรรกศาสตร์แบบเดิมจะมีแต่ค่าจริงกับเท็จเท่านั้น ดังแสดงด้วยภาพที่ 6.

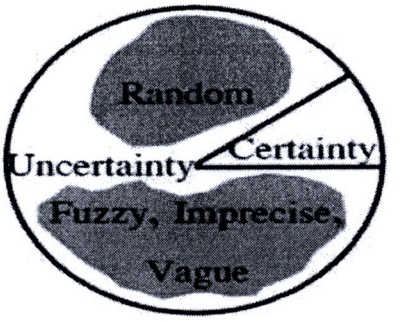


ภาพที่ 6. ตรรกะแบบจริงเท็จเทียบกับตรรกะแบบคลุมเครือ

ความเป็นฟัซซี (fuzziness) มีชื่อเรียกว่า มัลติวาลานซ์ (multivalance) เนื่องจากมีค่าความเป็นสมาชิกมากกว่า 2 ค่า ซึ่งแตกต่างกับไบวาลานซ์ (bivalance) ที่มีความเป็นสมาชิกเพียง 2 ค่า ฟัซซีเซต (Fuzzy set) เป็นเครื่องมือทางคณิตศาสตร์ที่สื่อถึง "ความไม่แน่นอน (uncertainty)" ซึ่งสามารถนำไปใช้ในกรณีที่เหตุการณ์มีหลายกรณี ไม่ใช่ เพียงแค่ 2 กรณี เช่น หากกำหนดว่า คนที่อ้วนคือคนที่มีน้ำหนักมากกว่า 75 กิโลกรัม ดังนั้นการไม่ใช้ฟัซซี จะให้ผลว่าคนที่มีน้ำหนัก 74.50 กิโลกรัม ไม่จัดเป็นคนที่อ้วน โดยการกำหนดรูปแบบ (modeling) ของลักษณะความไม่แน่นอนที่เป็นความคลุมเครือ ความไม่ตายตัว ในส่วนของทฤษฎีของฟัซซีเซตจะใช้ลักษณะความหมายตัวแปร (linguistic) มากกว่าปริมาณ (quantitative) ของตัวแปร เช่น การหาความหมายของ "คนที่อ้วน" เราไม่สามารถนิยามค่าความอ้วนที่ตรงกันและระบุเป็นหนึ่งเดียว (identical) ได้สำหรับคนที่อ้วน ตัวอย่างเช่นการให้ความหมายของคำว่า"อ้วน" ของ นายก.และนาย ข. ดังนี้ นาย ก. ให้ความหมายของ "คนอ้วน" ว่าหมายถึงคนที่มีน้ำหนักมากกว่า 70 กิโลกรัมในขณะที่ นาย ข. ให้ความหมายว่าเป็นคนที่มีน้ำหนัก

มากกว่า 75 กิโลกรัม ซึ่งทั้งสองคนต่างแสดงความหมายของคำว่าคนที่อ้วนโดยเปรียบเทียบและในมุมมองของตัวเองตามน้ำหนักของตน ในการทำงานในมุมมองแบบฐานสอง (Binary sense) จะได้ผลเป็น ใช่ หรือ ไม่ใช่ ทั้งที่บุคคลนั้นเป็นคนอ้วนน้ำหนักเกือบจะ 75 กิโลกรัม และในกรณีที่ถึงแม้ว่าบุคคลนี้จะมีน้ำหนัก 75 กิโลกรัมก็ตาม หากพิจารณาจากกลุ่มคนที่มีน้ำหนักเฉลี่ย 90 กิโลกรัม บุคคลที่มีน้ำหนัก 75 กิโลกรัม ก็จะไม่จัดอยู่ในกลุ่มคนที่อ้วน แสดงให้เห็นว่า ความอ้วนมีลักษณะความไม่แน่นอนแบบสุ่ม งานประยุกต์หลายด้าน รวมทั้งทางด้านการแพทย์ (Chang , 1998) ได้นำวิธีการของ Fuzzy logic มาประยุกต์ในการตรวจสอบ Micro calcification รวมทั้งการนำไปใช้ในการศึกษาปัญหาทั่ว ๆ ไป

ภาพที่ 7. เป็นการแสดงให้เห็นว่าแนวทางในการตัดสินใจของปัญหาทั้งหมดมีเพียงส่วนน้อยที่เป็นสิ่งที่แน่นอน (certainty) ที่เหลือคือสิ่งที่ไม่แน่นอน ซึ่งประกอบด้วยความไม่แน่นอนที่มีลักษณะแบบสุ่ม และความไม่แน่นอนที่มีลักษณะเป็นฟัซซี หรือคลุมเครือ ซึ่งมีมากกว่าร้อยละ 40 เพราะปัญหาส่วนมากเกี่ยวข้องกับการตัดสินใจของมนุษย์ซึ่งจะตัดสินใจตามพื้นฐานความคิดของคนเป็นหลัก



ภาพที่ 7. แสดงลักษณะปัญหาในโลกนี้ระหว่างความแน่นอนกับ ความไม่แน่นอน

ฟัซซีเป็นวิธีทางคณิตศาสตร์ที่แสดงถึงความคลุมเครือ ความไม่แน่นอนของระบบที่เกี่ยวข้องกับความคิดความรู้สึกของมนุษย์ เมื่อพิจารณาส่วนประกอบต่าง ๆ ในความไม่แน่นอนเพื่อกำหนดเงื่อนไขในการตัดสินใจ (Decision making) โดยอาศัยเซตของความเป็นสมาชิกในเซต (Set membership)

สำนักงานคณะกรรมการวิจัยแห่งชาติ ห้องสมุดงานวิจัย

2.3.1 เซตแบบฉบับ

เซตแบบฉบับ (classical set) หรือเซตทวินัย (crisp set) เป็นเซตที่มีค่าความเป็นสมาชิกเป็น 0 หรือ 1 {0, 1} เท่านั้น เซตในทฤษฎีเซตแบบฉบับจะมีขอบเขตแบบแข็ง (sharp boundary) ซึ่งเป็นขอบเขตที่ตัดขาดจากกันแบบทันทีทันใด เซตแบบฉบับมีการกำหนดค่าความเป็นสมาชิกตามแนวคิดเลขฐานสอง โดยที่ตัวแปรหนึ่ง ๆ จะมีค่าความเป็นสมาชิกเพียงสองค่า คือ 0 ไม่เป็นสมาชิก และ 1 เป็นสมาชิก ตัวอย่างเช่น เซตของคู่แต่งงาน จะสามารถบอกได้ว่าอย่างแน่ชัดว่าเป็นกลุ่มผู้แต่งงานหรือไม่แต่งงาน

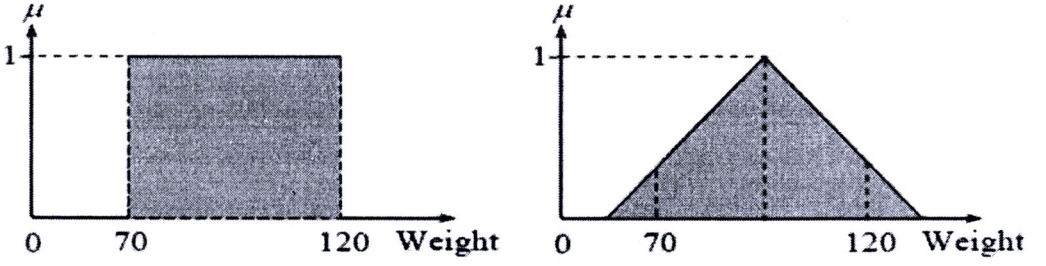
ผู้ที่แต่งงานแล้วจะมีค่าความเป็นสมาชิกในเซตของผู้ไม่แต่งงานเป็น 0 ส่วนผู้ที่ไม่แต่งงานมีค่าความเป็นสมาชิกภาพของเซตผู้ที่ไม่แต่งงานเป็น 1 ค่าความเป็นสมาชิกของทั้งสองเซตจะตัดขาดจากกันอย่างทันทีทันใด รูปแบบคณิตศาสตร์ของเซตแบบฉบับมีลักษณะดังนี้

$$\mu_A(x) = \begin{cases} 1, & x \in A \\ 0, & x \notin A \end{cases} \qquad \ldots(1)$$

เมื่อ A เป็นเซตแบบฉบับหรือเซตแบบทวินัยและ x เป็นสมาชิกในเซต $\mu_A(x)$ เป็นค่าความเป็นสมาชิกในเซต และ $\mu_A(x)$ เป็นฟังก์ชันความเป็นสมาชิกในเซต A

2.3.2 ฟัซซีเซต

ฟัซซีเซต (Fuzzy Set) เป็นเซตที่มีขอบเขตที่ราบเรียบ ทฤษฎีฟัซซีเซตจะครอบคลุมทฤษฎีเซตแบบฉบับ โดยฟัซซีเซตยอมให้มีค่าความเป็นสมาชิกของเซตมีค่าระหว่าง 0 และ 1 ในโลกแห่งความเป็นจริงเซตไม่ใช่มีเฉพาะเซตแบบฉบับเท่านั้น จะมีเซตแบบฟัซซีด้วย ฟัซซีเซตจะมีขอบเขตแบบฟัซซีไม่ใช่เปลี่ยนแปลงทันทีทันใดจากขาวเป็นดำ ตัวอย่างเช่น ความอ้วน นิยามคำว่าคนอ้วนในเซตทวินัยอาจกำหนดเป็นคนที่มีน้ำหนักตั้งแต่ 70 ถึง 120 กิโลกรัม โดยนิยามแบบฟัซซีเซตอาจกำหนดเป็นคนที่มีความอ้วนประมาณ 80 กิโลกรัม ซึ่งเป็นการให้นิยามที่ไม่แสดงถึงขอบเขตที่แน่นอน



ภาพที่ 8. การกำหนดค่าความเป็นสมาชิกของเซตทวินัยและเซตแบบฟัซซี

นิยามของฟัซซีเซต กำหนดให้ X เป็นเซตที่ไม่ว่าง ฟัซซีเซต A สามารถแสดงลักษณะเฉพาะได้จากฟังก์ชันความเป็นสมาชิก

$$\mu_{\tilde{A}}(x)"X \rightarrow [0,1] \qquad \ldots (2)$$

$$\tilde{A} = \left\{ \frac{\mu_{\tilde{A}}(x_1)}{x_1} + \frac{\mu_{\tilde{A}}(x_2)}{x_2} + \cdots + \frac{\mu_{\tilde{A}}(x_n)}{x_n} \right\} = \left\{ \sum_{i=1}^{n} \frac{\mu_{\tilde{A}}(x_i)}{x_i} \right\}$$

เมื่อพจน์ $\mu_{\tilde{A}}(x_i)/x_i,\ i = 1,2,\ldots,n$ หมายถึงค่าความเป็นสมาชิก $\mu_{\tilde{A}}(x_i)$ ของ x_i ในเซต $\tilde{A}$ และเครื่องหมายบวก "+" หมายถึงยูเนียน (union)

ทฤษฎีฟัซซีเซตสามารถแก้ปัญหาข้อจำกัดของเซตแบบดั้งเดิมได้ โดยฟัซซีเซตยอมให้มีค่าหรือดีกรีของความเป็นสมาชิก (degree of membership) ซึ่งแสดงด้วยค่าตัวเลขระหว่าง 0 และ 1 หรือเขียนเป็นสัญลักษณ์ [0, 1] โดย 0 หมายถึง ไม่เป็นสมาชิกในเซตและ 1 หมายถึงเป็นสมาชิกในเซต และค่าระหว่าง 0 กับ 1 เป็นสมาชิกบางส่วนในเซต การทำเช่นนี้ ทำให้เกิดความราบเรียบในการเปลี่ยนจากพื้นที่นอกเซตไปอยู่ในเซตของสมาชิกต่าง ๆ โดยมีฟังก์ชันสมาชิก (membership function) เป็นฟังก์ชันจัดเทียบ (mapping function) วัตถุในโดเมนใดๆสามารถ ให้เป็นค่าความเป็นสมาชิกในฟัซซีเซตได้ โดยความเป็นสมาชิกสำหรับฟัซซีเซต มีจำนวนระดับความเป็นสมาชิกเป็นอนันต์ คือค่าต่อเนื่องในช่วงตั้งแต่ 0 ถึง 1 ซึ่งครอบคลุมการกำหนดสมาชิกแบบฉบับ และเซตแบบฉบับหรือเซตทวินัย (crisp set) จะกำหนดตามดังสมการ

$$\mu_A(x) = \begin{cases} 1, & x \in A \\ 0, & x \notin A \end{cases} \qquad \ldots(3)$$

เมื่อ A เป็นเซตแบบฉบัยหรือเซตแบบทวินัย x เป็นสมาชิกในเซต μ_A เป็นค่าความเป็นสมาชิกในเซตและ $\mu_A(x)$ เป็นฟังก์ชันความเป้นสมาชิกในเซต A

2.3.3 ฟังก์ชันความเป็นสมาชิก

ฟังก์ชันความเป็นสมาชิก (membership function) เป็นฟังก์ชันที่มีการกำหนดระดับความเป็นสมาชิกของตัวแปรที่ต้องการใช้งาน โดยเริ่มจากการแทนที่กับตัวแทนที่มีความไม่ชัดเจน ไม่แน่นอน และคลุมเครือ ซึ่งเป็นส่วนที่สำคัญต่อคุณสมบัติหรือการดำเนินการของ

ฟัซซี เพราะรูปร่างของฟังก์ชันความเป็นสมาชิกมีผลต่อกระบวนการคิดและแก้ไขปัญหา โดยฟังก์ชันความเป็นสมาชิกจะไม่สมมาตรกันหรือสมมาตรกันทุกประการก็ได้

ชนิดของฟังก์ชันความเป็นสมาชิก

ชนิดของฟังก์ชันความเป็นสมาชิกที่ใช้งานทั่วไปมีหลายชนิด แต่ในที่นี้จะกล่าวถึงเพียงบาง 6 ชนิดที่ใช้นิยมใช้กันคือ

ก. ฟังก์ชันสามเหลี่ยม (triangular membership function)

ฟังก์ชันสามเหลี่ยมมีทั้งหมด 3 พารามิเตอร์คือ $\{a, b, c\}$

$$\text{Triangular (x: a, b, c)} = \begin{cases} 0 & ; \quad x < a \\ (x-a)/((b-a) & ; \quad a \le x \le b \\ (c-x)/(c-b) & ; \quad b \le x \le c \\ 0 & ; \quad x > c \end{cases} \qquad \ldots(4)$$

ข. ฟังก์ชันสี่เหลี่ยมคางหมู (trapezoidal membership function)

ฟังก์ชันสี่เหลี่ยมคางหมูมีทั้งหมด 4 พารามิเตอร์คือ $\{a, b, c, d\}$

$$\text{Trapezoidal (x: a, b, c, d)} = \begin{cases} 0 & ; \quad x < a \\ (x-a)/(b-a) & ; \quad a \le x < b \\ 1 & ; \quad b \le x < c \\ (d-x)/(d-c) & ; \quad c \le x < d \\ 0 & ; \quad x \ge d \end{cases} \qquad \ldots(5)$$

ค. ฟังก์ชันเกาส์เซียน (Gaussian membership function)

ฟังก์ชันเกาส์เซียนมีทั้งหมด 2 พารามิเตอร์คือ $\{m, \sigma\}$ ซึ่ง m หมายถึงค่าเฉลี่ย และ σ หมายถึง ค่าเบี่ยงเบนมาตรฐาน

$$\text{Gaussian}(x : m, \sigma) = \exp(-\frac{(x-m^2)}{\sigma^2}) \qquad \ldots(6)$$

ง. ฟังก์ชันระฆังคว่ำ (Bell-shaped membership function)

ฟังก์ชันรูประฆังคว่ำมีพารามิเตอร์ทั้งหมด 3 ค่าคือ {a, b, c}

$$\text{Bell-Shaped (x: a, b, c)} = \frac{1}{1+\left|\frac{x-c}{a}\right|^{2b}} \qquad \ldots(7)$$

จ. ฟังก์ชันตัวเอส (Smooth Membership Function)

ฟังก์ชันรูปตัวเอสมีพารามิเตอร์ทั้งหมด 2 ค่าคือ $\{a, b\}$

$$\text{S(x: a, b)} = \begin{cases} 0 & ; \quad x < a \\ 2\left(\frac{x-b}{b-a}\right)^2 & ; \quad a \le x \le \frac{a+b}{2} \\ 1-2\left(\frac{x-b}{b-a}\right)^2 & ; \quad \frac{a+b}{2} \le x < b \\ 1 & ; \quad x \ge b \end{cases} \qquad \ldots(8)$$

ฉ. ฟังก์ชันตัวแซด (Z-membership function)

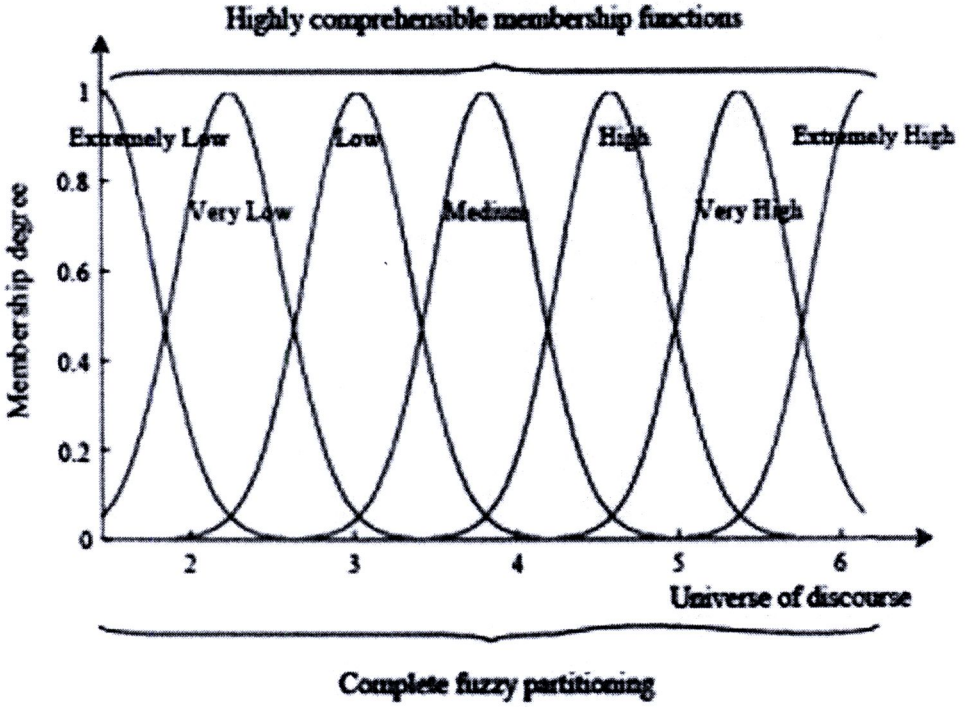
ฟังก์ชันรูปตัวเซดมีพารามิเตอร์ทั้งหมด 2 ค่าคือ $\{a, b\}$

$$\text{Z(x: a, b)} = \begin{cases} 1 & ; \quad x < a \\ 1-2\left(\frac{x-b}{b-a}\right)^2 & ; \quad a \le x \le \frac{a+b}{2} \\ 2\left(\frac{x-b}{b-a}\right)^2 & ; \quad \frac{a+b}{2} \le x < b \\ 0 & ; \quad x \ge b \end{cases} \qquad \ldots(9)$$

การเลือกฟังก์ชันของความเป็นสมาชิก จะต้องเลือกตามความเหมาะสมความครอบคลุมของข้อมูลที่จะรับเข้ามา โดยสามารถที่ทับซ้อนกันเพื่อให้การดำเนินงานราบเรียบ ซึ่งมีความเป็นสมาชิกหลายค่าได้ และฟังก์ชันความเป็นสมาชิกเปลี่ยนแปลงแก้ไขให้เหมาะกับงานที่กำลังปฏิบัติงานหรือตามความต้องการ การวิจัยครั้งนี้จะทดลองโดยใช้ฟังก์ชันทั้ง 6 แบบเพื่อดูความเหมาะสมโดยเทียบกับค่าความถูกต้องที่ได้กับข้อมูลชุดที่ใช้ในการทดลอง (Supervised training data)

2.4 ตัวแปรภาษา (Linguistic variable)

เซตแบบฟัซซีสามารถประยุกต์ใช้ในการอธิบายค่าของตัวแปรเช่นเดียวกับเซตแบบดั้งเดิม เช่น ประโยค "อุณหภูมิในห้องเย็น" คำว่า "เย็น" เป็นคำที่ใช้แสดงปริมาณอุณหภูมิ ในทางรูปนัย สามารถเขียนได้เป็น ปริมาณอุณหภูมิ ในห้อง เย็น หรือ "Temperature_ Quantity is Cold" ตัวแปร Temperature_ Quantity เป็นตัวแปรภาษา (linguistic variable) ซึ่งเป็นแนวคิดที่สำคัญมากในตรรกะแบบฟัซซี ตัวแปรภาษาช่วยกำหนดค่าของสิ่งที่จะอธิบายทั้งในรูปคุณภาพ โดยใช้พจน์ภาษา (linguistic term) และในรูปปริมาณ โดยใช้ฟังก์ชันความเป็นสมาชิก (membership function) ซึ่งแสดงความของของเซตแบบฟัซซี พจน์ภาษานั้นใช้สำหรับการแสดงแนวคิดและองค์ความรู้ในการสื่อสารของมนุษย์ ส่วนฟังก์ชันความเป็นสมาชิกมีประโยชน์ในการจัดการกับอินพุตที่เป็นข้อมูลเชิงตัวเลข

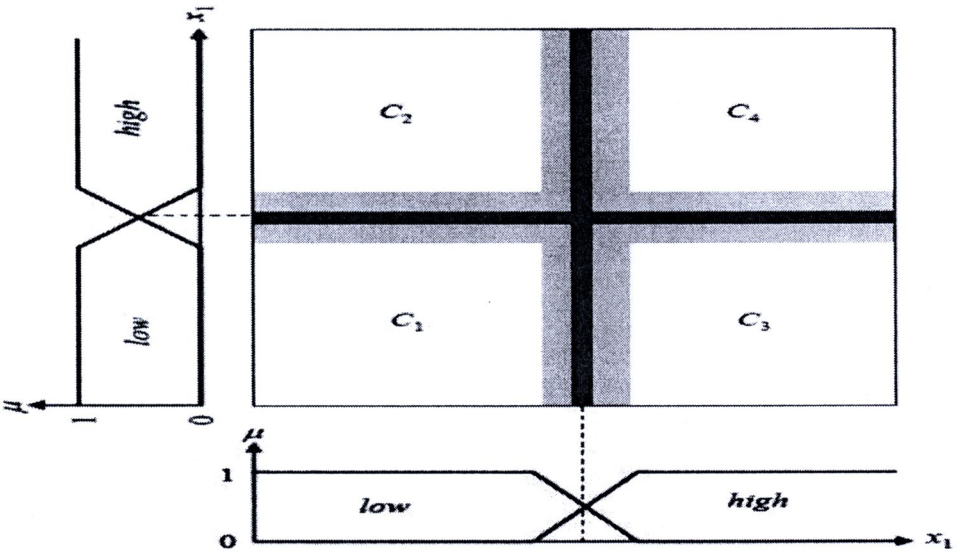


ภาพที่ 9. ตัวอย่างตัวแปรภาษา

ตัวแปรภาษาเป็นการประกอบกันของตัวแปรสัญลักษณ์ (symbolic variable) และตัวแปรเชิงเลข (numerical variable) ตัวอย่างตัวแปรสัญลักษณ์ เช่น "รูปร่าง เป็นทรงกระบอก" (Shape = Cylinder) คำว่า "รูปร่าง" เป็นตัวแปรที่บอกถึงรูปร่างของวัตถุ ตัวอย่างตัวแปรเชิงเลข เช่น "ความสูงเท่ากับ 4 ฟุต" (Height = 4') ตัวแปรเชิงตัวเลขจะมีใช้กันใน

สาขางานด้านวิทยาศาสตร์ วิศวกรรมศาสตร์ คณิตศาสตร์ การแพทย์ และอื่น ๆ ส่วนตัวแปรสัญลักษณ์มีความสำคัญในวิทยาการเกี่ยวกับปัญญาประดิษฐ์และการตัดสินใจ การใช้ตัวแปรภาษาเป็นการรวมตัวแปรเชิงตัวเลขกับตัวแปรสัญลักษณ์เข้าด้วยกัน ภาพที่ 9 แสดงตัวอย่างเซตตัวแปรภาษาของเซตฟัซซี ได้แก่ Extremely Low, Very Low, Low, Medium, High, Very High และ Extremely High

กฎฟัซซี (Fuzzy rules)



ภาพที่ 10. ตัวอย่างรูปแบบการจัดกลุ่มด้วยฟัซซี

จากภาพที่ 10. สามารถเขียนเป็นกฎในรูปประโยคภาษาได้ดังนี้

กฎข้อ 1: ถ้า x_1 มีค่า *low* และ x_2 มีค่า *low* แล้ว ข้อมูล (x_1, x_2) เป็นกลุ่ม C_1

กฎข้อ 2: ถ้า x_1 มีค่า *low* และ x_2 มีค่า *high* แล้ว ข้อมูล (x_1, x_2) เป็นกลุ่ม C_2

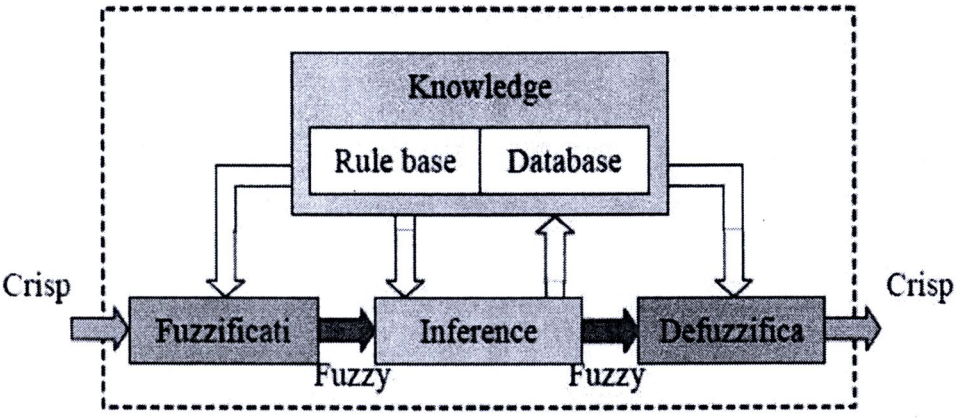
กฎข้อ 3: ถ้า x_1 มีค่า *high* และ x_2 มีค่า *low* แล้ว ข้อมูล (x_1, x_2) เป็นกลุ่ม C_3

กฎข้อ 4: ถ้า x_1 มีค่า *high* และ x_2 มีค่า *high* แล้ว ข้อมูล (x_1, x_2) เป็นกลุ่ม C_4

เมื่อ x_1 เป็นตัวแปรภาษาในมิติที่ 1, x_2 เป็นตัวแปรภาษาในมิติที่ 2, *low* และ *high* เป็นพจน์ภาษา (linguistic terms), ข้อมูล (x_1, x_2) เป็นคู่ลำดับของวัตถุที่ต้องการจัดกลุ่ม และ C_1, C_2, C_3 และ C_4 เป็นกลุ่มข้อมูล 1, 2, 3 และ 4

โครงสร้างพื้นฐานของการประมวลผลแบบฟัซซีลอจิก

โครงสร้างพื้นฐานของการประมวลผลแบบฟัซซี ซึ่งประกอบด้วยส่วนที่สำคัญ 4 ส่วน ดังภาพที่ 11.



ภาพที่ 11. การประมวลผลแบบฟัซซี

ส่วนที่แปลงอินพุตทั่วไปเพื่อเปลี่ยนเป็นการอินพุตแบบตัวแปรฟัซซี (Fuzzification) หรือในรูปแบบเซตฟัซซีหรือเรียกว่าเป็นตัวแปรภาษา (Linguistic variable)

ฐานความรู้ (Knowledge base) เป็นส่วนที่จัดเก็บรวบรวมความรู้ ในการควบคุม ประกอบ 2 ส่วนคือ ฐานกฎ (Rule base) และฐานข้อมูล (Database)

ฐานกฎ (Rule base) ส่วนของการกำหนดวิธีการควบคุม ซึ่งได้จากผู้เชี่ยวชาญใน รูปแบบของชุดข้อมูลแบบกฎของภาษา (Linguistic rule)

ฐานข้อมูล (Database) เป็นการจัดเตรียมส่วนที่จำเป็นเพื่อที่จะใช้ในการกำหนดกฎการ ควบคุม และการจัดการข้อมูลของตรรกศาสตร์ฟัซซี

เครื่องอนุมานหรือการตีความ (Inference Engine) เป็นส่วนที่ทำหน้าที่ตรวจสอบ ข้อเท็จจริงและกฎ เพื่อใช้ในการตีความหาเหตุผล เหมือนกลไกสำหรับควบคุมการใช้ความรู้ใน การแก้ไขปัญหา รวมทั้งการกำหนดวิธีการของการตีความเพื่อหาคำตอบ

ส่วนที่ทำการแปลงเอาต์พุตให้อยู่ในช่วงที่เหมาะสม (Defuzzification) โดยทำการแปลงข้อมูลที่อยู่ในรูปแบบฟัซซีให้เป็นค่าที่สรุปผลหรือค่าการควบคุมระบบ

วิธีการหาจุดศูนย์ถ่วง (Central of Gravity: COG) เป็นวิธีการเฉลี่ยผลที่ได้จากการตีความหาเหตุ ที่นิยมใช้ในปัจจุบัน ค่าที่ได้จะคำนวณจุดศูนย์ถ่วงโดยรวมจะหาได้จากการประมาณค่าจากสมการดังนี้

$$\text{COG} = \frac{\sum_{i=1}^{N} \alpha_i w_i}{\sum_{i=1}^{N} \alpha_i} \qquad \ldots(10)$$

โดยค่าของตัวแปรภายในสมการมีความหมายดังนี้

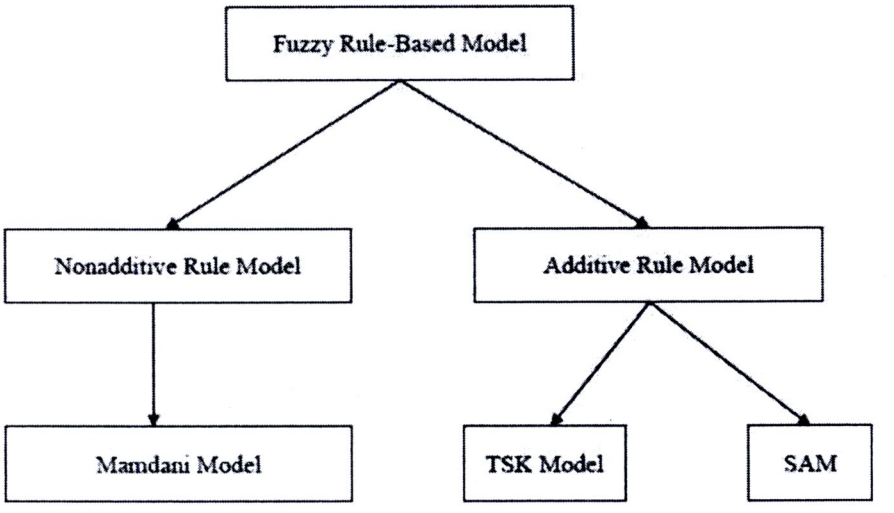
COG แทนค่าของจุดศูนย์ถ่วง (Central of Gravity)

N แทนค่าตั้งแต่ตำแหน่งที่ 1 ถึง N

α_i แทนค่าฟัซซีของเอาต์พุตในฟัซซีตำแหน่งที่ i

w_i แทนค่าพื้นที่ใต้โค้งของฟัซซีตำแหน่งที่ i

ชนิดของระบบกฎฟัซซี



ภาพที่ 12. ประเภทของกฎฟัซซี

ในการประมาณค่าฟังก์ชัน (function approximation) ระบบกฎฟัซซีที่ใช้มี 3 ชนิด ใหญ่ ๆ ได้แก่ (1) รูปแบบ Madani (2) รูปแบบ Takagi-Sugeno-Kang (TSK) และ (3) รูปแบบ Standard Additive Model (SAM) โดยรูปแบบ Madani รวมผลการอนุมาน (inference) ของกฎโดยวิธีการซ้อนทับ (superimposition) จากกฎหลาย ๆ ข้อ ซึ่งไม่เป็นแบบบวกกัน จึงเรียกระบบแบบนี้ว่าเป็น non-additive rule model แต่สำหรับ TSK และ SAM มีการอนุมาณแบบรวมค่าน้ำหนัก (weighted sum) จากหลาย ๆ กฎ เพื่อรวมเป็นข้อสรุปสุดท้าย จึงเรียกระบบแบบนี้ว่า additive rule model การจัดกลุ่มของระบบกฎแบบฟัซซีแสดงในภาพ 12 งานวิจัยครั้งนี้จะใช้รูปแบบของ TSK

2.5 Image Retrieval Based on Fuzzy Color Semantics

การสืบค้นรูปโดยอาศัยสาระสำคัญในรูปที่เรียกกันทั่วๆไปว่า (CBIS : Content –Based Image Retrieval นั้น โดยทั่วไปจะใช้ visual feature เช่น color, texture, shape และ spatial feature แต่ด้วยเหตุที่ว่าความต้องการของการสืบค้นนั้นจะตั้งอยู่บนความต้องการในเชิงความหมายในการรับรู้ของมนุษย์ มากกว่าการมองในลักษณะองค์ประกอบของ feature vector คำว่าความต้องการในเชิงการรับรู้ของมนุษย์ที่ใช้กัน ใช้คำว่า semantic concept ซึ่งโดยปกติมักจะใช้ภาษาธรรมชาติ (ภาษามนุษย์ : NLP) ในการสื่อถึงภาพที่ต้องการ การสืบค้นภาพจากฐานข้อมูลหรือจากอินเทอร์เน็ต (Berk et al., 1982) จำเป็นต้องอาศัยการบูรณาการเทคนิคต่างๆทางด้าน IR (Information Retrieval) เข้ามาช่วยดำเนินการกับข้อมูลที่เป็น low level ร่วมกับความรู้ในส่วนของการรับรู้ของมนุษย์รวมทั้งในส่วนของคุณสมบัติของข้อมูลที่เป็นลักษณะนามธรรมคือ semantic feature ซึ่งจัดว่าเป็น high level feature นั่นหมายความสิ่งที่ผู้สืบค้นต้องการนั้นเป็นลักษณะนามธรรมซึ่งสามารถสื่อให้เข้าใจด้วยภาษามนุษย์ เพื่อให้ระบบสืบค้นนั้นสามารถสืบค้นได้จากภาพที่มีองค์ประกอบเป็น low level feature ปัญหาที่สำคัญในการศึกษาเรื่องนี้ก็คือการหาเทคนิคเพื่ออุดช่องว่างระหว่างภาพในมิติการมอง (visual data) กับภาพในมิติที่เป็นจริงในสถานะของข้อมูลในคอมพิวเตอร์ที่เรียกว่า low level feature (Convey , 1992) Zhao และ Grosky (Zhao and Grosky,2007) ได้ทำการศึกษาเรื่องนี้โดยใช้เทคนิคทางสถิติ Latent correlation ในการสร้างสัมพันธภาพระหว่าง low level feature กับ high level feature เพื่อให้เกิดเป็น semantic concept ซึ่งทำให้ผู้ใช้สามารถสืบค้นภาพได้โดยอาศัย semantic based แนวทางในการศึกษานี้ ทำการสร้างสัมพันธภาพระหว่าง visual features และ semantic features แต่ถึงกระนั้นก็ตาม semantic ที่ได้ก็ยังไม่สามารถไปสนับสนุนในส่วนของ

การใช้ภาษาในการสืบค้นได้ด้วยเหตุผลที่ว่ายังมีความกำกวมเกิดขึ้น ในการจัดการกับปัญหานี้มีการศึกษารายละเอียดในส่วนของ high level semantics เพื่อที่จะนำมาสนับสนุนในการแก้ปัญหาดังกล่าว รวมทั้งการหาเทคนิคอื่นๆเข้ามาใช้เช่น fuzzy semantic เพื่อที่จะทำให้ระบบสามารถสืบค้นด้วยภาษากับสิ่งที่เป็น semantic concept (Haslam et al., 2007) งานของ (Qingyong et al., 2009) ได้ทำการศึกษาในส่วนของ global fuzzy color semantic description รวมทั้งการผนวกเอาเทคนิคของ Neural Network เข้ามาช่วยในการสกัดเอา color semantic feature vector ออกมาใช้งานโดยอาศัยการสนับสนุนของวิธีการ Fuzzy งานวิจัยนี้ใช้ข้อมูลจาก Corel Image Set ในส่วนของภาษา (Regier and Kay, 2009) ที่ใช้ในการสืบค้นคือภาษาอังกฤษ ก็มีการกำหนดเป็น linguistic variable ประกอบกับ basic concepts โดยการนิยามว่า Basic Term หมายถึง Linguistic Value Set ที่มีสมาชิกในเซตที่ไม่สามารถแตกย่อยได้อีกในส่วนของความหมาย ในขณะที่ส่วน Extended Terms หมายถึงเซตของ Linguistic Value Set ซึ่งสมาชิกภายในถูกสร้างขึ้นจาก basic term ที่มีส่วนขยายตามกฎเกณฑ์ทางภาษา (Syntax rule) ดังปรากฏในตารางที่ 1

ตารางที่ 1. แสดง Linguistic value for color semantic

Basic term	Extended term
Red	Many red, Somewhat red, Few red
Orange	Many orange, Somewhat orange, Few orange
Brown	Many brown, Somewhat brown, Few brown
Yellow	Many yellow, Somewhat yellow, Few yellow
Green	Many green, Somewhat green, Few green
Blue	Many blue, Somewhat blue, Few blue
Purple	Many purple, Somewhat purple, Few purple
Black	Many black, Somewhat black, Few black
Gray	Many gray, Somewhat gray, Few gray
White	Many white, Somewhat white, Few white

ถึงแม้ว่าการจัดการของคอมพิวเตอร์ในส่วนของสีนั้นจะมีได้เป็นล้านสี แต่โดยพื้นฐานแล้วมนุษย์เราสามารถกำหนดชื่อสีได้ประมาณ 10 -20 สีเท่านั้น (Bhoyar and Kakde, 2009) ใน

ส่วนของการศึกษาเรื่อง Color Naming Models ที่มีส่วนของการใช้ Numerical Color Space ที่มีความสามารถในการแสดงมิติของความหมายด้วยภาษามนุษย์นั้น Berk et all.(Berk et all., 1982) ได้เสนอ Color Naming System (CNS) ขึ้นมาใช้โดยที่ระบบนี้ประกอบด้วยเซตสีที่ประกอบด้วยสี red, orange, brown, yellow, green, blue, และ purple นอกจากนี้ยังมีการผนวก achromatic terms คือ black, gray, white เข้าไปกับสีต่างๆในเซตของสีดังกล่าวด้วย ซึ่งในที่นี้ก็คือ basic term set ในงานของ (Qingyong et al., 2009) นำ CNS มาใช้ โดยขยายการศึกษาในส่วนของการสร้างกฎเพิ่มเติมกับกฎของไวยกรณ์ภาษาของสี งานวิจัยนี้ใช้ HSV เป็น color model และสร้าง semantic rule ในการ mapping จาก low-level feature ไปยัง high –level fuzzy semantics T โดย T (x) คือค่า membership ที่มีความหมายของ สีในเชิง linguistic value (t) การศึกษาในงานนี้ใช้ ANN แบบมาตรฐานคือ 3 layer perceptron แนวทางการศึกษาคือการสร้าง operators บนพื้นฐานของ fuzzy set บน basic term แล้วจึงทำการกำหนด operators ให้กับ extended terms ความหมายโดยทั่วไปของผลที่สร้างคือ t หมายถึง linguistic value และ M คือ semantic rule ของ linguistic value t นั่นเอง ดังตัวอย่างเช่น คำว่า 'many' กำหนดให้เป็น m operator ได้เป็น $m(x) = \{M^2(x) \mid x \in U\}$ โดยที่ $M(x) \in [0,1]$ นั่นหมายความว่า m จะลดขนาดของค่า membership ยกเว้นกรณีที่ membership มีค่าเท่ากับ 0 หรือ 1 ในกรณีของคำว่า somewhat การกำหนด s operator ปรากฏดังนี้ $s(x) = \{SIN(M(x) * \pi \mid x \in U\}$ เนื่องจาก $0 \le M(x) \le 1$ ดังนั้น $s(x) \in [0,1]$ ดังนั้น s จะทำให้ค่า ขนาดของ membership เดิมที่มีขนาดต่ำ กับสูง มีค่าลดลงในขณะที่ ถ้า membership เดิมมีระดับปานกลางจะเพิ่มขึ้น ซึ่งทำให้สอดคล้องกับความหมายในภาษาที่ใช้ งานวิจัย (Bhoyar and Kakde ,2009) ทำการศึกษาในส่วนของ Fuzzy Color Semantic โดยใช้ RGB color histogram กับ Neuro –Fuzzy เพื่อทำการแยกประเภทของสีของมนุษย์ สถาปัตยกรรมของ ANN ที่ใช้ คือ three layer feed forward โดยมี 3 neuron in put layer , 45 neuron in hidden layer และ 11 neuron in output layer ส่วนฟังก์ชันที่ใช้ในการหาค่า membership คือ Bell shape เพื่อแบ่งเทอม dark , somewhat dark และ bright

$$f(i, a, b, c) = \frac{1}{1 + \left| \frac{x - c}{a} \right|^{2b}}$$; where i is the intensity of RGB component of the color pixel

คำว่า Dark มีค่า a, b, c = (46 , 2.5 ,- 7.5)

คำว่า Somewhat Dark มีค่า a, b, c = (27 , 2.2 ,70)

คำว่า Bright มีค่า a, b, c = (45 , 2.75 , 245)

สีที่งานนี้ทำการจำแนกคือ red, green, blue, yellow, purple, orange, white, brown, pink, black, gray กับข้อมูล 25 ตัวอย่าง ผลการทดลองได้กฎทั้งสิ้น 64 กฎ ผลลัพธ์ของการทดลอง โดยการสืบค้นด้วยคำว่า mostly green , AND somewhat red จากสามโดเมนของรูปได้ค่าความถูกต้องคือ ร้อยละ 48.4675 , 20.5572, 19.8203 ตามลำดับ การศึกษาเรื่องชื่อของสีในงานของ Weijer และพวก (Weijer et al., 2009) ทำการเปรียบเทียบระหว่าง Chip based color naming กับ วิธี color naming from real world ซึ่งได้รูปสีจากสภาวะแวดล้อมต่างๆกัน เช่นลักษณะกล้อง มุมกล้อง ปริมาณแสงเป็นต้น ซึ่งส่งผลให้การให้ชื่อสีแตกต่างกันไป งานวิจัยนี้ใช้รูปที่สืบค้นจาก Google จำนวน 250 รูปโดยมีชื่อของสีจำแนกเป็น 11 ชื่อและใช้ Probabilistic Latent Semantic Analysis ในการศึกษาโดยอาศัยพื้นฐานของ Yanei และ Barnnard (Yanei and Barnnard,2005) ซึ่งทำการวิเคราะห์ค่าของสีโดยอาศัยคำคุณศัพท์กับคุณลักษณะ (feature) ของรูปโดยคำนวณค่า Entropy ผลการทดลองกับรูปภาพ 4 ประเภท(car ,shoes, dresses, poetry) ได้ค่าเฉลี่ยความถูกต้อง ร้อยละ 95.4% งานวิจัยต่างๆที่ทำในลักณะของการจำแนกด้วยสีด้วยภาษามนุษย์ในตระกูลต่างๆเช่นภาษาอังกฤษ ภาษารัสเซีย (Shamir , 2007) ยังไม่มีการสร้างนามวลีในลักษณะเช่นภาษาไทย ดังนั้นการศึกษาเรื่องนี้จะทำให้เราได้แนวทางการกำกับภาพสีแบบอัตโนมัติโดยใช้ความหมายของสีในรูปแบบที่เป็นนามวลี ในภาษาไทย

2.6 องค์ประกอบของคำในภาษาไทย

คำนามที่เป็นสีในภาษาไทยเช่นสีแดง สีส้ม ได้มีการนำมาสร้างนามวลีโดยใช้คำกริยา เช่น “อม” “แกม” “เจือ” ซึ่งมีความหมายเมื่อเทียบกับภาษาอังกฤษดังนี้

[LEXITRON : http://lexitron.nectec.or.th/2009_1/]

อม (1) - V – embezzle

อม (1) - V – hide

แกม - V - attach

Syn. :: แนม#3,แกม#1

Def :: บริเวณหน้าทั้ง 2 ข้างถัดตาลงมา

เหลือบ (2) - ADJ - glassy

Def :: สีเลื่อมพรายเวลากระทบแสงเป็นหลายสีคล้ายสีตัวเหลือบ

Example in Thai :: หางบ่วงบางชนิดขนเป็นมันสีเหลือบ ดูสวยมาก

เจือ (1) - ADJ - impure

Syn. :: เจือ#1,ไม่แท้#1,ไม่บริสุทธิ์#1

Def :: เอาส่วนน้อยประสมลงไปในส่วนมากให้ระคนปนกัน, ตามปกติใช้กับของเหลว

เจือ (2) - V - mix

Def :: เอาส่วนน้อยประสมลงไปในส่วนมากให้ระคนปนกัน, ตามปกติใช้กับของเหลว

การนำคำนามและคำกริยาเหล่านี้ไปสร้างนามวลีเพื่อนำมาใช้งานในภาษาไทย โดยการสืบค้นคำเหล่านี้จาก Google พบคำเหล่านี้ในการกล่าวถึงการแสดงสีของสิ่งของดังตัวอย่างเช่น "ซัฟไฟร์สีส้มอมแดง (Reddish Orange Sapphire) " , "ทับทิมสีแดงอมส้ม " , "ผ้าไหมสีแดงอมส้ม " ,"โลหะเก็บไว้นานจนเป็นสีแดงเจือเหลือง" ,"เนื้อไม้สีแดงแกมเหลือง", "แคทลียาสีเหลืองแกมแดง" ซึ่งการสืบค้นต้องอาศัยการกำกับภาษากับแต่ละรูปด้วยมนุษย์ ทำให้กระบวนการนี้เสียเวลามาก รวมทั้งบางครั้งการกำกับภาพนั้นด้วยข้อความว่าเป็นรูปสีร่วมกับข้อความอื่นด้วยมนุษย์เพียงคนเดียวก็อาจจะเกิดความผิดพลาด(noise) ได้

ความแตกต่างของงานวิจัยนี้เทียบกับงานวิจัยในงานอื่น ก็คือการใช้ธรรมชาติของภาษาไทยที่แตกต่างจากภาษาอังกฤษเพื่อนำมาสร้างคำนามวลีที่มีความหมายของและภาพ เพื่อนำไปสืบค้น รวมทั้งการที่อาจจะนำความรู้ที่ได้ไปช่วยในการสืบค้นภาพที่ไมมีการกำกับข้อความได้ ซึ่งในส่วนนี้ยังเป็นข้อจำกัดของ Search engine ต่างๆเช่น Google