

The Effects of Partial Occlusion in Image-based CAPTCHAs on Pass Rates

อิทธิพลการทำให้แคปทชาแบบรูปภาพไม่สมบูรณ์ต่ออัตราการผ่านเข้าสู่ระบบ

รูปกรณ์ กฤษณารักษ์¹ และ ณัฐชนน หงส์วริทธิ์ธร²

ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยธรรมศาสตร์

Emails: tey_k@hotmail.com¹, chestu@gmail.com²

ABSTRACT - Several attempts in research of CAPTCHAs are to design and develop different kinds of CAPTCHAs in order to make them more deployable for computer system protection from internet bots. Text-based CAPTCHAs are commonly used in practice. They have been increasingly more distorted and camouflaged due to the increase of advanced developed bots. However, the increasing distortion of CAPTCHA text makes it difficult for human to decode and then take more time to get into systems. This research was to proposed new image-based CAPTCHAs which was developed base upon human visual perception so that the CAPTCHA is more usable to humans, but hard to decipher for auto-bots. The proposed image-based CAPTHCAs were generated based upon three factors to make the CAPTCHAs incomplete. Those factors are the two partial occlusion methods (constant and random occlusion of images), the two levels of image occlusion (30% and 45%), and the three different number of small frame as interference in an image (no frame, two frames and three frames). The proposed CAPTCHAs were evaluated by accuracy rate (pass rate) together with user satisfaction. A within-Subject design experiment was run with 40 participants. The results of the experiment indicated that the participants were satisfied with the image-based CAPTCHAs more than text-based CAPTCHAs at the 95% significant level. And importantly, the accuracy rate of image -based CAPTHCAs was significantly higher than text-based. The response time will be different depending on the completion of image CAPTCHAs. The most usable image-based CAPTCHAs should be made incomplete by constant occlusion of 30% of original picture, and can be interfered with zero or two small frames.

KEYWORDS – CAPTCHA, imaged-based CAPTCHA, Gestalt Theory

บทคัดย่อ - งานวิจัยเกี่ยวกับแคปทชาเป็นความพยายามที่จะออกแบบแคปทชาให้มีหลากหลายรูปแบบเพื่อป้องกันการโจมตีจากโปรแกรมอัตโนมัติแต่ต้องง่ายต่อผู้ใช้งาน แคปทชาแบบตัวอักษรเป็นแคปทชาที่พบบ่อยที่สุดและได้ถูกพัฒนาให้มีการบิดเบือนที่ซับซ้อนมากขึ้นเพื่อให้ทันต่อการพัฒนาของโปรแกรมอัตโนมัติ แต่ปัญหาที่เกิดขึ้นคือ ทำให้ผู้ใช้เกิดข้อผิดพลาดและความล่าช้าในการใช้งาน ดังนั้นงานวิจัยนี้จึงนำเสนอ แคปทชารูปภาพที่ออกแบบโดยตามหลักจิตวิทยาการรับรู้เพื่อให้ได้แคปทชารูปภาพที่ง่ายต่อผู้ใช้แต่ยากต่อการถูกโจมตีจากโปรแกรมอัตโนมัติ การออกแบบแคปทชารูปภาพจะใช้หลักการการทำให้รูปภาพไม่สมบูรณ์ด้วยปัจจัยย่อย 3 ปัจจัยได้แก่ รูปแบบการทำให้รูปภาพไม่สมบูรณ์ ประกอบด้วยแบบคงที่และแบบสุ่ม ระดับการทำให้รูปภาพไม่สมบูรณ์ ประกอบด้วยระดับที่ 1 ร้อยละ 30 ระดับที่ 2 ร้อยละ 45 และระดับการถูกสอดแทรกมี 3 ระดับ คือ ระดับที่ 1 ไม่ถูกสอดแทรก ระดับที่ 2 ถูกสอดแทรกด้วยภาพสี่เหลี่ยม 2 ภาพ ระดับที่ 3 ถูกสอดแทรกด้วยภาพสี่เหลี่ยม 3 ภาพ ประเมินผลด้วยอัตราความถูกต้องที่สามารถผ่านเข้าสู่ระบบ รวมถึงความพึงพอใจของผู้ใช้ด้วย ใช้รูปแบบการทดลองแบบ within-subject design ทดลองกับผู้เข้าร่วมการทดลอง 40 คน ผลการทดลองแสดงให้เห็นว่าผู้เข้าร่วมทดลองพึงพอใจแคปทชาแบบรูปภาพมากกว่าแคปทชาตัวอักษร อย่างมีนัยสำคัญทางสถิติที่ระดับความเชื่อมั่น 95% และที่สำคัญอัตราความถูกต้องของแคปทชาแบบรูปภาพสูงกว่าแบบตัวอักษร ความไม่สมบูรณ์ของรูปภาพจะส่งผลให้ภาพแคปทชาแต่ละแบบใช้เวลาในการตอบแตกต่างกันไป จากผลการทดลองแคปทชารูปภาพที่เหมาะสมคือ แคปทชารูปภาพที่ถูกทำให้ไม่สมบูรณ์ร้อยละ 30 ของภาพแบบคงที่ ซึ่งจะไม่มีการสอดแทรกด้วยภาพสี่เหลี่ยมหรือจะสอดแทรกด้วยภาพสี่เหลี่ยม 2 ภาพก็ได้

คำสำคัญ – แคปทชา, แคปทชาแบบรูปภาพ, ทฤษฎีเกสตัลต์

1. บทนำ

ปัจจุบันเทคโนโลยีได้มีบทบาทสำคัญในการแลกเปลี่ยนข้อมูลข่าวสารระหว่างกันและช่วยเพิ่มประสิทธิภาพการติดต่อสื่อสาร โปรแกรมอัตโนมัติ (bots) ถูกสร้างขึ้นเพื่อเพิ่มความสะดวกสบายให้กับผู้ใช้งานเครือข่ายอินเทอร์เน็ต และตอบสนองการแลกเปลี่ยนข้อมูลที่มีอัตราการขยายตัวที่เพิ่มมากขึ้น ขณะเดียวกันได้มีกลุ่มเขียนคอมพิวเตอร์ (hacker) นำโปรแกรมอัตโนมัติมาใช้ประโยชน์ในทางที่ผิด เช่น การส่งสแปม(Spam)หรือส่งอีเมลล์หลอกลวงโดยที่เจ้าของคอมพิวเตอร์ไม่รู้ตัว[1] ส่งผลกระทบและความเสียหายต่อผู้ใช้ทั่วไปขยายไปในวงกว้าง ในช่วงหลายปีที่ผ่านมาจึงมีงานวิจัยที่ศึกษาและออกแบบเครื่องมือตรวจจับโปรแกรมอัตโนมัติ เพื่อลดอัตราการถูกโจมตีระบบเครือข่ายอินเทอร์เน็ต โดยเครื่องมือหนึ่งที่ถูกออกแบบมาเพื่อแก้ปัญหานี้คือ “แคปต์ชา” (Captcha)

แคปต์ชาเป็นเครื่องมือที่ใช้เพื่อยืนยันตัวตนผู้ใช้ สามารถแยกแยะการใช้งานระหว่างมนุษย์กับโปรแกรมอัตโนมัติได้อย่างมีประสิทธิภาพ หลักการทำงานของแคปต์ชาจะสร้างแบบทดสอบเพื่อเข้าสู่ระบบคอมพิวเตอร์แบบโต้ตอบชนิดหนึ่งเพื่อทดสอบว่าผู้ใช้งานเป็นมนุษย์ไม่ใช่โปรแกรมอัตโนมัติ โดยแบบทดสอบนี้มนุษย์เท่านั้นที่จะสามารถเข้าใจและตอบได้ แคปต์ชาสามารถแบ่งออกเป็น 3 ดังนี้[2][3]

1. แคปต์ชาแบบตัวอักษร (Text-based) ประเภทนี้จะให้ผู้ใช้พิมพ์ชุดตัวอักษรลักษณะบิดเบี้ยว สลับซับซ้อน โดยชุดอักษรอาจมีความหมายหรือไม่มีความหมาย

2. แคปต์ชาแบบรูปภาพ (Image-based) ประเภทนี้จะให้ผู้ใช้อธิบายรูปภาพ หมุนหรือย้ายภาพให้อยู่ตำแหน่งที่ต้องการ

3. แคปต์ชาแบบเสียง (Sound-based) ประเภทนี้จะแบ่งเป็น 2 กลุ่ม คือกลุ่มแรกจะให้ผู้ใช้งานทำแบบทดสอบด้วยการพิมพ์ตัวอักษรจากเสียงที่ได้ยิน และกลุ่มที่สองจะให้ผู้ใช้งานเลือกเสียงที่เกี่ยวข้องกับรูปภาพที่กำหนดให้

จากแคปต์ชาทั้ง 3 กลุ่มนี้ แคปต์ชาแบบตัวอักษร จะนิยมนำไปใช้มากในปัจจุบัน แต่การออกแบบแคปต์ชาแบบตัวอักษรที่ใช้การบิดเบือนตัวอักษรให้มีความซับซ้อน ทำให้ผู้ใช้งานอ่านตัวอักษรได้ยากและบางครั้งก็ไม่สามารถอ่านตัวอักษรได้เลย จำเป็นต้องใช้การคาดเดาจึงเกิดข้อผิดพลาดและความล่าช้าในการใช้งาน[4] จากปัญหาดังกล่าว หากมีการนำภาพมาใช้เป็นแคปต์ชาแทนการบิดเบือนหรือเพิ่มลวดลายให้ภาพ โดยออกแบบตามทฤษฎีเกสตัลต์ (Gestalt Theory) ก็อาจจะช่วยเพิ่ม

ประสิทธิภาพในการป้องกันการเข้าถึงข้อมูลจากโปรแกรมอัตโนมัติได้อีกทางหนึ่ง เนื่องจากภาพที่ออกแบบตามทฤษฎีเกสตัลต์นั้นจะทำให้โปรแกรมอัตโนมัติตรวจจับคุณลักษณะของภาพได้ยากขึ้น



รูปที่ 1.แสดงแคปต์ชารูปแบบตัวอักษรและแบบรูปภาพ ออกแบบโดยอาศัยทฤษฎีเกสตัลต์

รูปที่ 1. จะเป็นปัญหาสำคัญของโปรแกรมอัตโนมัติในการตรวจจับคุณลักษณะพิเศษ (Feature) ของภาพ ด้วยเทคนิคคอมพิวเตอร์วิทัศน์ (Computer Vision) แต่สำหรับมนุษย์แล้วจะสามารถรับรู้และเข้าใจแม้ว่าเค้าโครงภาพนั้นจะสมบูรณ์หรือไม่ก็ตาม เนื่องจากกระบวนการรับรู้ (Visual Perception) ของมนุษย์กับคอมพิวเตอร์นั้นมีความแตกต่างกัน โดยมนุษย์จะมีการจัดระเบียบการรับรู้ด้วยการเชื่อมโยงสิ่งต่างๆที่ตามองเห็นแล้วนำมาประมวลผลกับประสบการณ์ ต่างจากคอมพิวเตอร์ที่การรับรู้ภาพจะเกิดจากการคัดลอกลักษณะทางกายภาพของสิ่งที่เห็นอย่างตรงไปตรงมา [5][6] รายละเอียดของงานวิจัยในส่วนถัดไป จะกล่าวถึง งานวิจัยที่เกี่ยวข้อง ส่วนที่ 3 เป็นวิธีการที่นำเสนอ ส่วนที่ 4 เป็นผลลัพธ์ที่ได้จากการทดลอง และส่วนที่ 5 จะกล่าวถึง สรุป การอภิปรายและงานวิจัยในอนาคต

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ทฤษฎีการรับรู้

งานวิจัยนี้ต้องการศึกษาอิทธิพลการทำให้แคปต์ชาแบบรูปภาพไม่สมบูรณ์โดยที่ปริมาณความไม่สมบูรณ์นั้นอยู่ในปริมาณที่ผู้ใช้รับรู้หรือเข้าใจภาพได้ จึงเริ่มต้นจากแนวคิดด้านจิตวิทยาการรับรู้ของมนุษย์ (Human Perception) ซึ่งการรับรู้ของมนุษย์ไม่ใช่การลอกเลียนแบบตามความเป็นจริงเสมอไป แต่ยังมีประสบการณ์เข้ามาเกี่ยวข้อง การรับรู้ของมนุษย์มีหลายด้านตามประสาทสัมผัสทั้ง 5 ได้แก่ ตา-หู(เห็น),หู-ฟัง(ได้ยิน),ลิ้น-รู้รส,จมูก-ดม(ได้กลิ่น),ผิวหนัง-ถูกต้อง(รู้สึก)[13] แต่ในการศึกษาวิจัยครั้งนี้ได้เน้นศึกษาไปที่การออกแบบแคปต์ชาแบบรูปภาพ ทฤษฎีเกสตัลต์ (Gestalt Theory) เป็นทฤษฎีหนึ่งที่ทำ

ให้การรับรู้ในที่นี้จะเน้นการรับรู้ทางการรับรู้ภาพ (Visual Perception) การประมวลผลการรับรู้ของมนุษย์ตามหลักจิตวิทยาพุทธิปัญญา (Cognitive psychology) มี 2 ลักษณะ คือ 1.รับรู้ตรงไปตรงมา (bottom-up processing) ตามสภาพจริง และลักษณะที่ 2 การรับรู้โดยมีการใช้ประสบการณ์ หลักการและความรู้มาช่วย (top-down processing) [11] ทฤษฎีคอนสตรัคติวิสต์ เป็นทฤษฎีหนึ่งที่เกี่ยวข้องกับการรับรู้ของมนุษย์ที่ใช้ประสบการณ์ คือ มนุษย์สามารถรับรู้ทางการเห็นที่มีการเพิ่มเติม บิดเบือน เปลี่ยนแปลง เชื่อมโยงกับประสบการณ์เดิมเพื่อทำความเข้าใจกับสิ่งที่เห็น สิ่งที่รับรู้ [12] โดยทฤษฎีเกสตัลต์เป็นทฤษฎีที่ใช้อธิบายการรับรู้ทางภาพของมนุษย์ ที่มีการใช้การเพิ่มเติม เสริมแต่ง บิดเบือน ภาพที่เห็นให้เป็นไปตามมนุษย์คาดหวังหรือประสบการณ์ที่ผ่านมา โดยใช้ปัจจัยต่างๆ เช่น ปัจจัยจากความใกล้ชิด (The factor of Proximity) ปัจจัยจากความคล้ายคลึงกัน (The factor of Similarity) ปัจจัยจากการปิดส่วนที่ไม่สมบูรณ์ (The Factor of Closure) [13] เช่น



รูปที่ 2. แสดงภาพสุนัขที่ออกแบบด้วยปัจจัยจากการปิดส่วนที่ไม่สมบูรณ์

จากภาพข้างต้น แม้ว่าภาพจะมีองค์ประกอบไม่สมบูรณ์ แต่มนุษย์สามารถรับรู้และตีความได้ว่าเป็นภาพของสุนัข ด้วยการเติมส่วนที่ขาดหายไปได้ของภาพจากประสบการณ์ ดังนั้นในงานวิจัยนี้จะใช้คุณลักษณะการรับรู้ของมนุษย์ ในลักษณะการแต่งเติมหรือเพิ่มเติมภาพตามปัจจัยการปิดส่วนที่ไม่สมบูรณ์ของทฤษฎีเกสตัลต์ สร้างแคปต์ชาแบบรูปภาพให้เป็นภาพที่ไม่สมบูรณ์แต่มนุษย์ยังคงสามารถรับรู้หรือทำความเข้าใจได้

2.2 งานวิจัยที่เกี่ยวข้องในการพัฒนาโปรแกรมแคปต์ชา

งานวิจัยที่เกี่ยวข้องในการพัฒนาโปรแกรมแคปต์ชา มีทั้งหมด 3 ด้าน [2] คือ

1. แคปต์ชาแบบตัวอักษร (Text-based) เป็นรูปแบบที่นิยมใช้ในปัจจุบัน การทดสอบโปรแกรมอัตโนมัติจะให้ผู้ใช้อัปโหลด

คำถามด้วยการกดเป็นตัวอักษรตามที่ปรากฏ [3] แต่ปัญหาของแคปต์ชาแบบนี้ คือ เทคนิคคอมพิวเตอร์วิทัศน์ (Computer vision) ในปัจจุบันสามารถแปลงไฟล์ภาพอักษรให้เป็นไฟล์ได้ถูกต้องมากขึ้น ดังนั้นการออกแบบแคปต์ชาแบบนี้จึงต้องเพิ่มความซับซ้อนของชุดตัวอักษรหรือลดขนาดพื้นที่ให้มีมีความซับซ้อนยิ่งขึ้นไป L.von Ahn [7] พัฒนาแคปต์ชาแบบตัวอักษรเรียกว่า “Recaptcha” โดยสร้างชุดอักษรด้วยการสแกนตัวอักษรจากหนังสือเก่า จึงสามารถป้องกันโปรแกรมอัตโนมัติได้เป็นอย่างดี แต่ตัวอักษรที่บิดเบี้ยวซ้อนกันก็ทำให้ผู้ใช้เกิดข้อผิดพลาดและความล่าช้าในการทำงานเพิ่มขึ้น

2. แคปต์ชาแบบเสียง (Sound-based) J.Holman [8] วิจัยและพัฒนาแคปต์ชาแบบเสียงเพื่อผู้ที่บกพร่องทางสายตาแคปต์ชาลักษณะนี้ทักษะของผู้ฟังถือว่าเป็นสิ่งสำคัญ เพราะเสียงที่อ่านชุดตัวอักษรหรือรูปภาพจะถูกแทรกด้วยเสียงรบกวนต่างๆ

3. แคปต์ชาแบบรูปภาพ (Image-based) อาศัยลักษณะการมองเห็นที่แตกต่างกันระหว่างมนุษย์กับคอมพิวเตอร์ แบ่งเป็น 2 รูปแบบคือ แบบแรกโดยส่วนใหญ่จะให้ผู้ใช้เคลื่อนย้ายภาพให้อยู่ตำแหน่งที่ถูกต้อง โดยที่ไม่ต้องพิมพ์ข้อความหรือชุดตัวอักษร W.Liao [5] วิจัยและพัฒนาแคปต์ชาแบบรูปภาพโดยใช้ใบหน้าคนและทำการสลับตำแหน่งส่วนของภาพสองตำแหน่งตามทฤษฎีเกสตัลต์ พบว่า การใช้ทฤษฎีเกสตัลต์ในแคปต์ชาจะลดความผิดพลาดของผู้ใช้ได้ ส่วนภาพที่นำมาใช้เป็นแคปต์ชาจะเป็นภาพใบหน้าคน เพราะมีพื้นผิวหรือรายละเอียดของภาพไม่มากจนเกินไป ในขณะที่ R.Gossweiler [4] ออกแบบแคปต์ชาแบบรูปภาพ พบว่า ภาพที่ใช้ไม่ควรเป็นภาพหน้าคน เพราะโปรแกรมอัตโนมัติจะสามารถตรวจจับภาพได้ง่าย และควรหลีกเลี่ยงภาพที่มีตัวอักษรแทรกอยู่ภายในรูปภาพหรือภาพที่มีพื้นเป็นท้องฟ้าหรือพื้นน้ำ ส่วน H.Gao [2] วิจัยและพัฒนาแคปต์ชาแบบรูปภาพจิ๊กซอว์ โดยจิ๊กซอว์แต่ละชิ้นจะถูกลบขอบของภาพเพื่อป้องกันการตรวจจับโดยใช้เทคนิค edge matching ของคอมพิวเตอร์วิทัศน์ การทดสอบพบว่า ภาพที่มีขนาดใหญ่จะมีประสิทธิภาพที่สูงทั้งในด้านเวลาที่ใช้งานและความถูกต้อง แต่แคปต์ชาจะมีปัญหาในการใช้งานเนื่องจากต้องใช้พื้นที่การแสดงผลภาพที่มีขนาดใหญ่ นอกจากนี้ต้องเลือกใช้ภาพที่ไม่กำกวมหลังจากที่หมุนหรือย้ายภาพแล้ว ซึ่งทำให้ผู้ใช้สับสน ส่วนแคปต์ชารูปภาพแบบที่สองจะให้ผู้ใช้อัปโหลดด้วยการพิมพ์อธิบายภาพซึ่งถือเป็นงานวิจัยส่วนน้อยใน

ปัจจุบัน เนื่องจากโปรแกรมอัตโนมัติสามารถทำแบบทดสอบนี้ได้ถูกต้องมากขึ้น การป้องกันโปรแกรมอัตโนมัติจึงต้องเพิ่มความซับซ้อนของภาพให้มากขึ้น จึงเกิดปัญหา คือ ถ้าใช้ภาพที่มีความซับซ้อนมากเกินไป จะทำให้ผู้ใช้ตอบแบบทดสอบผิดพลาดสูง แต่ถ้าภาพทดสอบง่ายเกินไปจะถูกโปรแกรมอัตโนมัติตรวจจับได้ การใช้ทฤษฎีเกสตัดต์ออกแบบแคปซูลจึงช่วยลดความซับซ้อนของภาพ และอาจทำให้ผู้ใช้เข้าใจภาพได้ง่าย [6]

3. วิธีการนำเสนอ

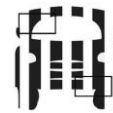
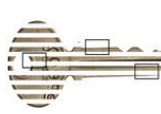
งานวิจัยนี้ศึกษาเปรียบเทียบการออกแบบแคปซูลรูปแบบตัวอักษรกับรูปภาพ โดยแบบรูปภาพจะแบ่งออกเป็น 3 ปัจจัยย่อย คือ ปัจจัยรูปแบบการทำให้รูปภาพไม่สมบูรณ์ (แบบคงที่และแบบสุ่ม) ระดับการทำให้รูปภาพไม่สมบูรณ์ (ระดับที่ 1 ร้อยละ 30 และระดับที่ 2 ร้อยละ 45) และระดับการถูกสอดแทรก (ไม่ถูกสอดแทรกภาพ ถูกสอดแทรกด้วยภาพสีเหลี่ยม 2 ภาพ และถูกสอดแทรกด้วยภาพสีเหลี่ยม 3 ภาพ)

การออกแบบแคปซูลแบบรูปภาพ มีหลักพื้นฐานที่ควรคำนึงถึง [4] คือ

1. ผู้ใช้ต้องสามารถเข้าใจภาพแคปซูลได้ง่าย สามารถทำแบบทดสอบได้ถูกต้อง และใช้เวลาที่เหมาะสม
2. แคปซูลสามารถสร้างหรือปรับปรุงได้ง่าย และนำไปใช้อย่างเหมาะสม

3. โปรแกรมอัตโนมัติต้องไม่สามารถตรวจจับภาพได้ง่าย งานวิจัยต่างๆ พยายามทำให้คอมพิวเตอร์ “มองเห็น” ภาพให้ใกล้เคียงกับมนุษย์ โดยทั่วไปการเห็นภาพของคอมพิวเตอร์จะใช้ “คุณลักษณะ (Feature)” ของภาพมาใช้วิเคราะห์เทคนิค Bag of features (BOF)[9] เป็นเทคนิคหนึ่งที่ใช้คุณลักษณะของภาพมาวิเคราะห์เพื่อตรวจจับภาพ โดยแยกลักษณะเด่นของภาพออกมา เช่น แยกพื้นผิวและรูปร่างของภาพออกเป็นส่วนย่อยๆ เพื่อนำคุณลักษณะของภาพมาเปรียบเทียบกับฐานข้อมูล และสร้างคำอธิบายให้ภาพ อย่างไรก็ตามความถูกต้องในการตรวจจับหรืออธิบายภาพจะขึ้นอยู่กับที่ความสมบูรณ์ของโครงสร้าง (contour) และความชัดเจนของคุณลักษณะภาพ (Feature)[10] และจากงานวิจัยที่เกี่ยวข้องแสดงให้เห็นว่า การออกแบบแคปซูลทั้งแบบตัวอักษรและรูปภาพโดยใช้ทฤษฎีเกสตัดต์ช่วยเพิ่มประสิทธิภาพการป้องกันโปรแกรมอัตโนมัติได้เนื่องจากภาพที่ได้จะมีเค้าโครงรูปร่างที่ไม่สมบูรณ์ซึ่งเป็น

ปัญหาของโปรแกรมอัตโนมัติ งานวิจัยนี้จึงทำการพัฒนาโปรแกรมแคปซูลโดยใช้ทฤษฎีเกสตัดต์มาเป็นหลักการออกแบบ ออกแบบได้ 12 รูปแบบ (ดังรูปที่ 3)

	
1) ภาพโทรศัพท์ที่ถูกทำให้ไม่สมบูรณ์ร้อยละ 30 ของภาพแบบคงที่และไม่ถูกสอดแทรก	2) ภาพโทรศัพท์ที่ถูกทำให้ไม่สมบูรณ์ร้อยละ 45 ของภาพแบบคงที่และไม่ถูกสอดแทรก
	
3) ภาพไม้บรรทัดที่ถูกทำให้ไม่สมบูรณ์ร้อยละ 30 ของภาพแบบคงที่และถูกสอดแทรกด้วยภาพสีเหลี่ยม 2 ภาพ	4) ภาพรถยนต์ที่ถูกทำให้ไม่สมบูรณ์ร้อยละ 45 ของภาพแบบคงที่และถูกสอดแทรกด้วยภาพสีเหลี่ยม 2 ภาพ
	
5) ภาพลูกบาศก์ที่ถูกทำให้ไม่สมบูรณ์ร้อยละ 30 แบบคงที่และถูกสอดแทรกด้วยภาพสีเหลี่ยม 3 ภาพ	6) ภาพกุญแจที่ถูกทำให้ไม่สมบูรณ์ร้อยละ 45 ของภาพแบบคงที่ และถูกสอดแทรกด้วยภาพสีเหลี่ยม 3 ภาพ
	
7) ภาพแอปเปิ้ลที่ถูกทำให้ไม่สมบูรณ์ร้อยละ 30 ของภาพแบบสุ่มและไม่ถูกสอดแทรก	8) ภาพถุงมือที่ถูกทำให้ไม่สมบูรณ์ร้อยละ 45 ของภาพแบบสุ่มและไม่ถูกสอดแทรก
	
9) ภาพแตงโมที่ถูกทำให้ไม่สมบูรณ์ร้อยละ 30 ของภาพแบบสุ่มและถูกสอดแทรกด้วยภาพสีเหลี่ยม 2 ภาพ	10) ภาพหมวกที่ถูกทำให้ไม่สมบูรณ์ร้อยละ 45 ของภาพแบบสุ่ม และถูกสอดแทรกด้วยภาพสีเหลี่ยม 2 ภาพ
	
11) ภาพกีตาร์ที่ถูกทำให้ไม่สมบูรณ์ร้อยละ 30 ของภาพแบบสุ่มและถูกสอดแทรกด้วยภาพสีเหลี่ยม 3 ภาพ	12) ภาพไมโครโฟนที่ถูกทำให้ไม่สมบูรณ์ร้อยละ 45 แบบสุ่ม และถูกสอดแทรกด้วยภาพสีเหลี่ยม 3 ภาพ

รูปที่ 3. แสดงแคปซูลแบบรูปภาพโดยใช้ทฤษฎีเกสตัดต์ทั้ง 12 รูปแบบ

การเลือกภาพ

การเลือกภาพที่ใช้สร้างแคปต์ชาจะใช้ภาพที่เกี่ยวข้องกับสิ่งของสัตว์ ที่ผู้ทดลองเคยมีประสบการณ์ในการรับรู้และเห็นภาพเหล่านี้ นอกจากนี้การเลือกภาพที่เป็นกลุ่มใหญ่จะช่วยป้องกันโปรแกรมอัตโนมัติได้ยิ่งขึ้น [4]

การตรวจสอบการป้องกันโปรแกรมอัตโนมัติ

งานวิจัยนี้ทดสอบการป้องกันโปรแกรมจากเทคนิค Bag of Features (BoF) ด้วยการนำแคปต์ชาที่ออกแบบด้วยทฤษฎีเกสตัดต์ทั้ง 12 รูปแบบ ไปทดสอบการค้นหาคำจากเว็บไซต์กูเกิ้ล (search by image) ซึ่งรูปแบบแคปต์ชาที่สามารถป้องกันการตรวจจับจากโปรแกรมอัตโนมัติจะต้องเป็นภาพที่เว็บไซต์กูเกิ้ลไม่สามารถสร้างคำอธิบายได้หรือให้คำอธิบายไม่ถูกต้อง

3.1 การออกแบบการทดลอง

งานวิจัยนี้ใช้รูปแบบการทดลองเป็นแบบ 2x2x3 Within-Subjects Design ประกอบด้วยกลุ่มควบคุม 1 กลุ่ม คือ แคปต์ชาตัวอักษร และกลุ่มทดลอง 12 กลุ่ม รวมทั้งหมด 13 กลุ่มทดลอง เพื่อวัดผลความถูกต้องในการตอบคำถาม และความพึงพอใจของผู้ใช้

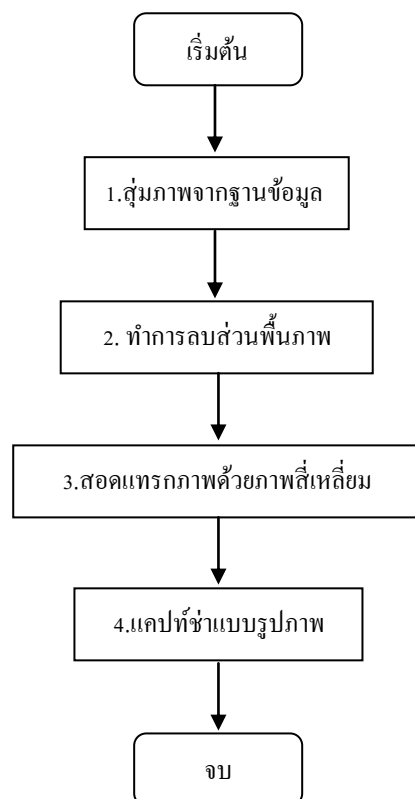
ตารางที่ 1. แสดงรายละเอียดรูปการออกแบบการทดลองเปรียบเทียบแคปต์ชาตัวอักษรกับรูปภาพ

กลุ่มควบคุม	กลุ่มทดลอง					
แคปต์ชาแบบตัวอักษร	แคปต์ชาแบบรูปภาพ					
	การทำให้รูปภาพไม่สมบูรณ์แบบคงที่			การทำให้รูปภาพไม่สมบูรณ์แบบสุ่ม		
	ระดับการทำให้รูปภาพไม่สมบูรณ์					
	ร้อยละ30	ร้อยละ45		ร้อยละ30	ร้อยละ45	
	ระดับการถูกสอดแทรกด้วยภาพสี่เหลี่ยม					
	ไม่ถูกสอดแทรก	2 ภาพ	3 ภาพ	ไม่ถูกสอดแทรก	2 ภาพ	3 ภาพ

3.2 เครื่องมือที่ใช้ในการทดลอง

โปรแกรมแคปต์ชาที่ใช้ในงานวิจัยนี้ถูกพัฒนาด้วยภาษา HTML เมื่อโปรแกรมได้พัฒนาเสร็จสิ้นแล้ว ได้ถูกนำมาติดตั้งการทำงานบนเว็บเบราว์เซอร์ Microsoft Internet Explorer 8 ในคอมพิวเตอร์แบบพกพา จอภาพขนาด 14.1 นิ้ว ที่มีความละเอียดจอภาพอยู่ที่ 1280*800 pixels และใช้โปรแกรม Camtasia Studio version 8 ถูกใช้เพื่อบันทึกการทำงานของผู้ใช้งาน นอกจากนี้ยังใช้แบบสอบถามเพื่อเก็บข้อมูลทั่วไปของผู้เข้าร่วมการทดลองก่อนทำการทดลองและใช้แบบประเมินความพึงพอใจหลังการทดลองด้วย

หลักการทำงานของแคปต์ชาที่นำเสนอ จะมีการทำงานดังต่อไปนี้



รูปที่ 4. Flow Chart การทำงานของแคปต์ชาแบบรูปภาพโดยใช้ทฤษฎีเกสตัดต์

ขั้นตอนที่ 1 ทำการสุ่มรูปภาพประเภทสิ่งของหรือสัตว์จากฐานข้อมูลที่ได้จัดเตรียมไว้

ขั้นตอนที่ 2 นำภาพที่ได้จากการสุ่ม ทำการลบส่วนที่เป็นพื้นภาพ เพื่อให้ภาพไม่สมบูรณ์ แบ่งเป็น 2 ระดับคือ ร้อยละ 30 และร้อยละ 45 ของภาพ โดยมีรูปแบบการลบภาพ 2 รูปแบบ คือ แบบคงที่ และแบบสุ่ม

ขั้นตอนที่ 3 จากนั้นจะทำการสอดแทรกด้วยภาพสี่เหลี่ยม ซึ่งแบ่งเป็น 3 ระดับ คือ ระดับแรก จะไม่สอดแทรกด้วยภาพสี่เหลี่ยม ระดับสอง จะสอดแทรกด้วยภาพสี่เหลี่ยม 2 ภาพ และระดับสาม จะสอดแทรกด้วยภาพสี่เหลี่ยม 3 ภาพ โดยใช้การสุ่มตำแหน่งที่จะสอดแทรกภาพสี่เหลี่ยมในภาพแคปหน้าจอ

ขั้นตอนที่ 4 ภายหลังจากการทำให้รูปภาพไม่สมบูรณ์แล้ว จะให้ผู้ใช้งานทำการทดสอบ ซึ่งคำตอบจะถูกส่งกลับเข้ามาแล้วนำไปตรวจสอบด้วยการเปรียบเทียบจากฐานข้อมูล



รูปที่ 5. หน้าจอการทำงานของแคปหน้าจอแบบรูปภาพที่ออกแบบโดยทฤษฎีเกสตัลต์

3.3 การวัดผลการทดลอง

3.3.1 ประสิทธิภาพการใช้งาน

ด้านความถูกต้อง

การวัดผลการทดลองด้านความถูกต้องจะใช้ "อัตราความถูกต้อง" คำนวณได้จากสัดส่วนของเวลาต่อคำตอบที่ถูกต้อง 1 ข้อ เช่น 1 ข้อ ใช้เวลา 10 วินาที โดยรูปแบบที่มีอัตราความถูกต้องดีที่สุด จะใช้เวลาตอบคำถามน้อยที่สุด ซึ่งผู้ใช้งานต้องอธิบายรายละเอียดของภาพแคปหน้าจอในแต่ละข้อให้ถูกต้องตรงกับที่โจทย์กำหนด ในส่วนเวลาที่ใช้งานเพื่อความแม่นยำ การทดลองครั้งนี้ได้เลือกใช้การบันทึกภาพการทำงานของผู้ใช้ด้วยโปรแกรม Camtasia Studio Version 8 โดยการคำนวณเวลานั้นจะเริ่มนับเมื่อโจทย์คำถามถูกแสดงจนถึงเวลาที่ผู้ใช้งานกดขึ้นต้นการตอบคำถามขึ้นนั้น

ด้านความพึงพอใจ

จะทำการวัดผลด้วยแบบสอบถามที่พัฒนาขึ้น จำนวน 7 ข้อ โดยกำหนดวัดระดับความพึงพอใจออกเป็น 5 ระดับ คือ ระดับ 1 หมายถึง พึงพอใจน้อยที่สุด และระดับที่ 5 หมายถึง พึงพอใจมากที่สุด โดยเก็บข้อมูลจากแบบสอบถามหลังจากการตอบแบบทดสอบโปรแกรมแคปหน้าจอทั้งแบบตัวอักษรและแบบรูปภาพ เครื่องมือการวัดความพึงพอใจแบ่งเป็น 7 หัวข้อ คือ ความเข้าใจง่าย ความเหมาะสมในการใช้เป็นแคปหน้าจอ ความสวยงามในการใช้เป็นแคปหน้าจอ ความเหมาะสมของประเภทรูปภาพหรือรูปแบบตัวอักษร ความเหมาะสมของขนาดแคปหน้าจอ ความชัดเจนของแคปหน้าจอ และความง่ายในการตอบคำถามแคปหน้าจอ

3.4 วิธีทดลอง

ขั้นตอนแรก ผู้เข้าร่วมการทดลองจะได้รับแบบสอบถามเพื่อเก็บข้อมูลทั่วไปก่อนผู้ทดลองจะเริ่มทำการทดลอง จากนั้นผู้เข้าร่วมการทดลองจะได้รับฟังคำแนะนำเบื้องต้นในการทดลอง และฝึกทดลองใช้งานด้วยการตอบแบบทดสอบโปรแกรมแคปหน้าจอทั้งแบบตัวอักษรและแบบรูปภาพอย่างละ 1 ข้อ เพื่อสร้างความคุ้นเคยให้กับผู้ใช้เป็นเวลา 5 นาที จากนั้นจึงเริ่มการทดลอง ผู้เข้าร่วมการทดลองจะต้องทำแบบทดสอบแคปหน้าจอด้วยการพิมพ์อธิบายภาพแคปหน้าจอทั้งแบบอักษรและแบบรูปภาพ ที่กำหนดให้ด้วยตัวเองอีก 48 ข้อ แบ่งเป็นแบบทดสอบ แคปหน้าจอตัวอักษร 24 ข้อ และแบบทดสอบแคปหน้าจอแบบรูปภาพอีก 24 ข้อ ซึ่งแบบทดสอบจะถูกเรียงสลับข้อระหว่างแคปหน้าจอตัวอักษร และแคปหน้าจอแบบรูปภาพไปเรื่อยๆ จนครบทั้งหมด 48 ข้อ ลักษณะของโปรแกรมที่ใช้ทดลองจะเหมือนกับเกมที่คุณเล่นจะได้รับคะแนนมากหากอธิบายภาพแคปหน้าจอได้ถูกต้องตรงตามที่โจทย์กำหนดให้ในแต่ละข้อ หลังจากการทดลองผู้เข้าร่วมการทดลองจะขอความช่วยเหลือแบบสอบถามเกี่ยวกับทัศนคติความพึงพอใจและความคิดเห็นต่อโปรแกรมแคปหน้าจอที่ใช้ในการทดลอง

4. ผลการทดลอง

งานวิจัยนี้เป็นการศึกษาการออกแบบแคปหน้าจอโดยอาศัยทฤษฎีเกสตัลต์ ที่มีรูปแบบแตกต่างกัน 12 รูปแบบ โดยวิเคราะห์ความแตกต่างของคะแนนความถูกต้องและเวลาที่ใช้งาน ด้วย Repeated Measures ANOVA และวิเคราะห์ความแตกต่างของ

คะแนนความพึงพอใจด้วย t-test แบบจับคู่ (Paired Samples T-test) ได้ผลการทดลองดังแสดงรายละเอียดดังนี้

4.1 อัตราความถูกต้อง

ตารางที่ 2 แสดงให้เห็นว่า รูปแบบการทำให้รูปภาพไม่สมบูรณ์ต่างกันส่งผลให้อัตราความถูกต้องต่างกัน คือ รูปแบบการทำให้รูปภาพไม่สมบูรณ์แบบคงที่ (Mean = 9.035; S.D.=.517) จะใช้เวลาน้อยกว่าแบบสุ่ม (Mean = 11.892; S.D.=.450) ในส่วนระดับการทำให้รูปภาพไม่สมบูรณ์ที่ต่างกันส่งผลให้อัตราความถูกต้องต่างกัน คือ ระดับการทำให้รูปภาพไม่สมบูรณ์ร้อยละ 30 (Mean = 8.850; S.D.=.435) จะใช้เวลาน้อยกว่าร้อยละ 45 (Mean=12.077; S.D.=.506) และระดับการถูกสอดแทรกด้วยภาพสี่เหลี่ยมที่ต่างกันก็ส่งผลให้อัตราความถูกต้องแตกต่างกันด้วย คือ สอดแทรกด้วยภาพสี่เหลี่ยม 2 ภาพ จะใช้เวลาน้อยสุด (Mean=10.059; S.D.= .419) และสอดแทรกด้วยภาพสี่เหลี่ยม 3 ภาพ (Mean=11.084; S.D.=.824) จะใช้เวลามากที่สุด โดยจากการวิเคราะห์เปรียบเทียบรายคู่ของแต่ละระดับการถูกสอดแทรกด้วยภาพสี่เหลี่ยม (ตารางที่ 3) ซึ่งเห็นว่า แคปซันที่ไม่มีภาพสี่เหลี่ยมสอดแทรกกับแคปซันที่มีภาพสี่เหลี่ยมสองภาพ มีอัตราความถูกต้องไม่แตกต่างกัน (Sig. =0.280)

ตารางที่ 2. การทดสอบ With-in Subjects

Source	F	Sig.
รูปแบบการทำให้รูปภาพไม่สมบูรณ์	972.757	.000
ระดับการทำให้รูปภาพไม่สมบูรณ์	1.457	.000
ระดับการถูกสอดแทรกด้วยภาพสี่เหลี่ยม	33.248	.000
รูปแบบการทำให้รูปภาพไม่สมบูรณ์ X ระดับการทำให้รูปภาพไม่สมบูรณ์	141.976	.000
รูปแบบการทำให้รูปภาพไม่สมบูรณ์ X ระดับการถูกสอดแทรกด้วยภาพสี่เหลี่ยม	4.375	.016
ระดับการทำให้รูปภาพไม่สมบูรณ์ X ระดับการถูกสอดแทรกด้วยภาพสี่เหลี่ยม	35.780	.000
รูปแบบการทำให้รูปภาพไม่สมบูรณ์ X ระดับการทำให้รูปภาพไม่สมบูรณ์ X ระดับการถูกสอดแทรกด้วยสี่เหลี่ยม	.420	0.65

ตารางที่ 3. การทดสอบ Pairwise Comparisons ของระดับการถูกสอดแทรก

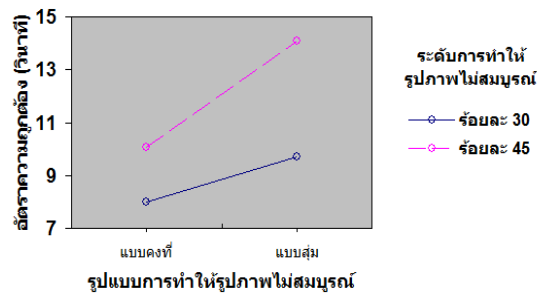
(I)	(J)	Mean different (I-J)	Std. Error	Sig. ^a
ไม่ถูกสอดแทรก	ภาพสี่เหลี่ยม 2 ภาพ	.188	.109	.280
	ภาพสี่เหลี่ยม 3 ภาพ	-.838 [*]	.141	.000
ภาพสี่เหลี่ยม 3 ภาพ	ภาพสี่เหลี่ยม 2 ภาพ	1.025 [*]	.148	.000

รูปที่ 6-8 เป็นการแสดงการเกิดปฏิสัมพันธ์ตัวแปรทั้งหมดระหว่างสองตัวแปร

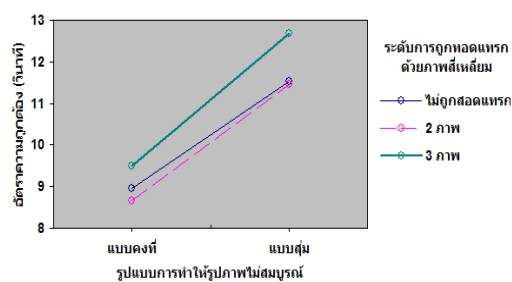
โดยที่รูปที่ 6 เป็นการเกิดอิทธิพลของตัวแปรร่วมระหว่างรูปแบบการทำให้รูปภาพไม่สมบูรณ์กับระดับการทำให้รูปภาพไม่สมบูรณ์ ซึ่งเห็นว่า อัตราความถูกต้องในภาพที่ถูกทำให้ไม่สมบูรณ์แบบสุ่มร้อยละ 45 จะใช้เวลามากที่สุด ส่วนอัตราความถูกต้องไม่ค่อยแตกต่างกันมากนักในรูปภาพที่ถูกทำให้ไม่สมบูรณ์ร้อยละ 30 ไม่ว่าจะเป็นรูปแบบคงที่หรือแบบสุ่ม ในขณะที่ตัวแปรรูปภาพที่ถูกทำให้ไม่สมบูรณ์ร้อยละ 45 จะแตกต่างกันมาก

ในรูปที่ 7 เป็นการเกิดอิทธิพลของตัวแปรร่วมระหว่างรูปแบบการทำให้รูปภาพไม่สมบูรณ์กับระดับการถูกสอดแทรกด้วยภาพสี่เหลี่ยม ซึ่งเห็นว่า อัตราความถูกต้องของภาพที่ถูกทำให้ไม่สมบูรณ์แบบคงที่และแบบสุ่ม ที่ไม่ถูกสอดแทรกด้วยภาพสี่เหลี่ยมหรือถูกสอดแทรกด้วยภาพสี่เหลี่ยม 2 ภาพ มีอัตราความถูกต้องไม่ค่อยแตกต่างกันมากนัก ส่วนอัตราความถูกต้องของภาพที่ถูกทำให้ไม่สมบูรณ์แบบคงที่และแบบสุ่มที่มีการถูกสอดแทรกด้วยภาพสี่เหลี่ยม 3 ภาพจะแตกต่างกันมาก

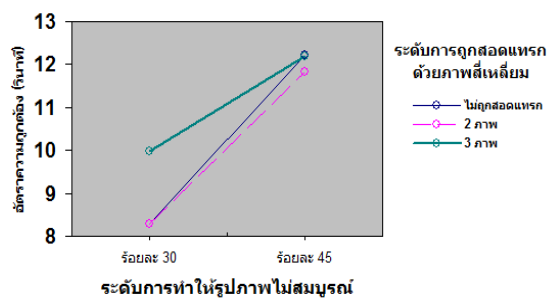
และรูปที่ 8 เป็นการเกิดอิทธิพลของตัวแปรร่วมระหว่างระดับการทำให้รูปภาพไม่สมบูรณ์กับระดับการถูกสอดแทรกด้วยภาพสี่เหลี่ยม ซึ่งเห็นว่า อัตราความถูกต้องในภาพที่มีระดับการทำให้รูปภาพไม่สมบูรณ์ร้อยละ 45 ทั้งที่ไม่ถูกสอดแทรกด้วยภาพสี่เหลี่ยมถูกสอดแทรกด้วยภาพสี่เหลี่ยม 2 ภาพและ 3 ภาพ จะมีอัตราความถูกต้องไม่ค่อยแตกต่างกันมากนัก ในขณะที่ตัวแปรระดับการทำให้รูปภาพไม่สมบูรณ์ร้อยละ 30 จะแตกต่างกันมาก



รูปที่ 6. กราฟ Interaction effect ของรูปแบบการทำให้รูปภาพไม่สมบูรณ์กับระดับการทำให้รูปภาพไม่สมบูรณ์



รูปที่ 7. กราฟ Interaction effect ของรูปแบบการทำให้รูปภาพไม่สมบูรณ์กับระดับการถูกสอดแทรกด้วยภาพสี่เหลี่ยม



รูปที่ 8. กราฟ Interaction effect ของระดับการทำให้รูปภาพไม่สมบูรณ์กับระดับการถูกสอดแทรกด้วยภาพสี่เหลี่ยม

ตารางที่ 4 พบว่า อัตราความถูกต้องของแคปซูลแบบรูปภาพรูปแบบ ที่ถูกทำให้รูปภาพไม่สมบูรณ์ร้อยละ 30 ของภาพแบบคงที่และไม่ถูกสอดแทรก (Mean=7.48; S.D.=.702) กับรูปแบบ 3 ที่ถูกทำให้รูปภาพไม่สมบูรณ์ร้อยละ 30 ของภาพแบบคงที่และไม่ถูกสอดแทรกด้วยภาพสี่เหลี่ยม 2 ภาพ (Mean=7.50; S.D.=.898) จะดีกว่าแคปซูลแบบตัวอักษร (Mean=12.07; S.D.= 1.430)

ตารางที่ 4. อัตราความถูกต้องของแคปซูลแบบตัวอักษรและแบบรูปภาพ 12 แบบ

รูปแบบ	อัตราความถูกต้อง (วินาที)
แคปซูลแบบตัวอักษร	12.07
1) แคปซูลรูปภาพที่ถูกทำให้ไม่สมบูรณ์ร้อยละ 30 แบบคงที่และไม่ถูกสอดแทรก	7.48
2) แคปซูลรูปภาพที่ถูกทำให้ไม่สมบูรณ์ร้อยละ 45 แบบคงที่และไม่ถูกสอดแทรก	10.42
3) แคปซูลรูปภาพที่ถูกทำให้ไม่สมบูรณ์ร้อยละ 30 แบบคงที่และถูกสอดแทรกด้วยภาพสี่เหลี่ยม 2 ภาพ	7.5
4) แคปซูลรูปภาพที่ถูกทำให้ไม่สมบูรณ์ร้อยละ 45 แบบคงที่และถูกสอดแทรกด้วยภาพสี่เหลี่ยม 2 ภาพ	9.82
5) แคปซูลรูปภาพที่ถูกทำให้ไม่สมบูรณ์ร้อยละ 30 แบบคงที่ และถูกสอดแทรกด้วยภาพสี่เหลี่ยม 3 ภาพ	8.98
6) แคปซูลรูปภาพที่ถูกทำให้ไม่สมบูรณ์ร้อยละ 45 แบบคงที่ และถูกสอดแทรกด้วยภาพสี่เหลี่ยม 3 ภาพ	9.98
7) แคปซูลรูปภาพที่ถูกทำให้ไม่สมบูรณ์ร้อยละ 30 แบบสั่นและไม่ถูกสอดแทรก	9.06
8) แคปซูลรูปภาพที่ถูกทำให้ไม่สมบูรณ์ร้อยละ 45 แบบสั่นและไม่ถูกสอดแทรก	14.01
9) แคปซูลรูปภาพที่ถูกทำให้ไม่สมบูรณ์ร้อยละ 30 แบบสั่น และถูกสอดแทรกด้วยภาพสี่เหลี่ยม 2 ภาพ	9.08
10) แคปซูลรูปภาพที่ถูกทำให้ไม่สมบูรณ์ร้อยละ 45 แบบสั่น และถูกสอดแทรกด้วยภาพสี่เหลี่ยม 2 ภาพ	13.82
11) แคปซูลรูปภาพที่ถูกทำให้ไม่สมบูรณ์ร้อยละ 30 แบบสั่น และถูกสอดแทรกด้วยภาพสี่เหลี่ยม 3 ภาพ	10.97
12) แคปซูลรูปภาพที่ถูกทำให้ไม่สมบูรณ์ร้อยละ 45 แบบสั่น และถูกสอดแทรกด้วยภาพสี่เหลี่ยม 3 ภาพ	14.38

4.2 การป้องกันโปรแกรมอัตโนมัติ

ผลการทดสอบการป้องกันโปรแกรมอัตโนมัติด้วยวิธีการค้นหาด้วยภาพจากเว็บไซต์กูเกิ้ล พบว่า เว็บไซต์กูเกิ้ลไม่สามารถสร้างคำอธิบายหรือตรวจจับภาพแคปชั่นที่ออกแบบด้วยทฤษฎีเกสตัลท์ทั้ง 12 รูปแบบที่ใช้ในงานวิจัยนี้



รูปที่ 9. แสดงตัวอย่างการทดสอบการป้องกันโปรแกรมอัตโนมัติ

4.3 ความพึงพอใจ

ตารางที่ 5. แสดงความพึงพอใจด้านต่างๆโดยเฉลี่ยของแคปชั่นแบบรูปภาพ

รายการประเมิน	N	Mean	Std. Deviation
ภาพที่ใช้เข้าใจง่าย	40	4.02	.530
ความเหมาะสมในการใช้ภาพเป็นแคปชั่น	40	4.10	.744
ความสวยงามในการใช้ภาพเป็นแคปชั่น	40	3.95	.749
ความเหมาะสมของประเภทภาพ	40	3.82	.675
ความเหมาะสมของขนาดภาพ	40	3.93	.764
ความชัดเจนของภาพ	40	3.70	.791
ตอบคำถามของแคปชั่นได้ง่าย	40	3.95	.846

จากตารางที่ 5 พบว่า กลุ่มผู้ทดลองมีความพอใจเฉลี่ยแคปชั่นแบบรูปภาพเกี่ยวกับความเหมาะสมในการใช้ภาพเป็นแคปชั่นมากที่สุด (Mean=4.10; S.D.=.744) และพอใจเฉลี่ยด้านความชัดเจนของภาพต่ำสุด (Mean=3.70; S.D.=.791)

ตารางที่ 6. แสดงความพึงพอใจด้านต่างๆโดยเฉลี่ยของแคปชั่นแบบอักษร

รายการประเมิน	N	Mean	Std. Deviation
แบบตัวอักษรที่ใช้เข้าใจง่าย	40	3.30	.939
ความเหมาะสมในการใช้ตัวอักษร	40	3.53	.506
ความสวยงามในการใช้ตัวอักษร	40	3.00	.506
ความเหมาะสมของรูปแบบตัวอักษร	40	3.40	.632
ความเหมาะสมของขนาดตัวอักษร	40	3.75	.543
ความชัดเจนของตัวอักษร	40	3.28	.506
ตอบคำถามของแคปชั่นได้ง่าย	40	3.50	.716

จากตารางที่ 6 พบว่า กลุ่มผู้ทดลองมีความพอใจเฉลี่ยแคปชั่นแบบตัวอักษรเกี่ยวกับความเหมาะสมของขนาดตัวอักษรมากที่สุด (Mean=3.75; S.D.=.543) และพอใจเฉลี่ยความสวยงามในการใช้ตัวอักษรเป็นแคปชั่นต่ำสุด (Mean=3.00; S.D.=.506)

ตารางที่ 7. เปรียบเทียบความพึงพอใจของผู้ใช้ระหว่างแคปชั่นรูปภาพกับตัวอักษร

	รูปแบบแคปชั่น	N	Mean	Std. Deviation	t	Sig. (2-tailed)
ความพึงพอใจ	รูปภาพ	40	3.92	.48092	6.18	.000
	ตัวอักษร	40	3.39	.42025		

จากตารางที่ 7 ผลการวิเคราะห์เปรียบเทียบคะแนนความพึงพอใจระหว่างแคปชั่นแบบรูปภาพกับตัวอักษร โดยใช้แบบสอบถามทั้งหมด 7 ข้อ พบว่า ผู้ทดลองมีความพอใจเฉลี่ยแคปชั่นแบบรูปภาพ (Mean =3.92; S.D.=.480) มากกว่าแคปชั่นแบบตัวอักษร (Mean =3.39; S.D.=.420) ที่ระดับนัยสำคัญ .05 ($t=6.18 > 0$, $df=39$; $p=0.000 < .05$)

5. บทสรุป

รูปแบบงานวิจัยที่เกี่ยวข้องกับการออกแบบแคปชั่นในช่วงหลายปีที่ผ่านมา มีแนวโน้มที่จะถูกโปรแกรมอัตโนมัติ

ตรวจจับได้ง่ายยิ่งขึ้น ดังนั้น การออกแบบแคปต์ชาเพื่อป้องกันโปรแกรมอัตโนมัติในปัจจุบันจึงต้องเพิ่มลักษณะการบิดเบือนแคปต์ชาให้ซับซ้อนยิ่งขึ้น แต่ก็ทำให้ผู้ใช้งานเกิดข้อผิดพลาดและความล่าช้าในการใช้งานเช่นเดียวกัน งานวิจัยนี้จึงนำเสนอการศึกษาเปรียบเทียบรูปแบบการออกแบบแคปต์ชารูปแบบตัวอักษรกับรูปภาพ โดยรูปภาพจะแบ่งออกเป็น 3 ปัจจัยย่อยคือปัจจัยรูปแบบการทำให้รูปภาพไม่สมบูรณ์ (แบบคงที่และแบบสุ่ม) ระดับการทำให้รูปภาพไม่สมบูรณ์ (ระดับที่ 1 ร้อยละ 30 และระดับที่ 2 ร้อยละ 45) และระดับการทดสอบแทรกในภาพ (ไม่สอดแทรกภาพ สอดแทรกด้วยภาพสี่เหลี่ยม 2 ภาพ และสอดแทรกด้วยภาพสี่เหลี่ยม 3 ภาพ) เพื่อทดสอบประสิทธิภาพในการใช้ ผลการทดลองแสดงให้เห็นว่าแคปต์ชาแบบรูปภาพจะมีอัตราความถูกต้องดีกว่าแคปต์ชาแบบตัวอักษร แต่ก็ขึ้นอยู่กับความไม่สมบูรณ์ของรูปภาพ โดยรูปแบบการทำให้รูปภาพไม่สมบูรณ์ ระดับการทำให้รูปภาพไม่สมบูรณ์ และการสอดแทรกด้วยภาพสี่เหลี่ยม จะทำให้แคปต์ชาแต่ละแบบใช้เวลาในการตอบแตกต่างกันไป ซึ่งรูปแบบ 1 ที่ถูกทำให้รูปภาพไม่สมบูรณ์ร้อยละ 30 ของภาพแบบคงที่และไม่ได้ถูกสอดแทรก กับรูปแบบ 3 ที่ถูกทำให้รูปภาพไม่สมบูรณ์ร้อยละ 30 ของภาพแบบคงที่และถูกสอดแทรกด้วยภาพสี่เหลี่ยม 2 ภาพ จะมีอัตราความถูกต้องดีที่สุด คือใช้เวลาตอบคำถามได้ถูกต้องต่อหนึ่งข้อน้อยที่สุด ดังนั้นรูปแบบที่เหมาะสมในการออกแบบแคปต์ชารูปภาพ คือ แคปต์ชารูปภาพที่ถูกทำให้ไม่สมบูรณ์ร้อยละ 30 ของภาพแบบคงที่ ซึ่งจะไม่มีการสอดแทรกด้วยภาพสี่เหลี่ยมหรือจะสอดแทรกด้วยภาพสี่เหลี่ยม 2 ภาพก็ได้ แต่ถ้าเพิ่มลักษณะการสอดแทรกจะส่งผลให้ใช้เวลาที่ตอบมากขึ้นทันที เช่น เพิ่มการสอดแทรกด้วยภาพสี่เหลี่ยม 3 ภาพ ในส่วนความพึงพอใจ ผู้ใช้มีความพึงพอใจแคปต์ชาแบบรูปภาพมากกว่าแบบตัวอักษรอย่างมีนัยสำคัญ

อย่างไรก็ตามการทดลองในครั้งนี้ยังไม่ได้พิจารณาถึงปัจจัยในการเลือกประเภทของภาพที่นำมาใช้เป็นแคปต์ชา ซึ่ง อาจจะเป็นทางเลือกที่เหมาะสมกว่าที่ช่วยส่งผลด้านเวลาที่ใช้ และปริมาณความถูกต้องในการใช้งาน ดังนั้นการท้าววิจัยในครั้งต่อไป ควรทำการศึกษาเพิ่มเติม โดยนำปัจจัยในการเลือกประเภทของภาพมาเป็นปัจจัยในการทดลองด้วย

เอกสารอ้างอิง

- [1] Klatch, "The Value of Captcha-Humans vs Robots", 2010, from <http://www.scsccbkk.org/CAPTCHA%20Humans%20vs%20Robots.pdf>
- [2] H.Gao,D.Yao,H.Liu,X.Liu,L.Wang,and L.Wang "A novel imaged based captcha using jigsaw puzzle," in IEEE 2010 International Conference on Computational Science and Engineering, pp.351-356, DEC.2010.
- [3] J.Yan,and Ahmad, " Usability of CAPTCHAs or usability issues in CAPTCHA design ," in SOUPS '08 Proceedings of the 4th symposium on Usable privacy and security,pp. 44-52,2008.
- [4] R.Gossweiler,M.Kamvar, and S.Baluja, "A captcha based on image orientation,"in International World Wide Web Conference Committee (IW3C2), 2009.
- [5] W.Liao,"A captcha mechanism by exchanging image blocks," in Pattern Recognition, 2006. ICPR 2006. 18th International Conference,pp.1179-1183,2006.
- [6] W.Liao and C.Chang," Embedding information within dynamic visual patterns,"in Multimedia and Expo, 2004. ICME '04. 2004 IEEE International Conference,pp.895-898, Jun.2004.
- [7] L.von Ahn,"CAPTCHA:telling humans and computers apart Automatically" from <http://www.captcha.net/J.Holman,J.Lazer,J.Feng.andJ.D'Arcy>,"Developing usable captchas for blind users,"in Assets '07 Proceedings of the 9th international ACM SIGACCESS conference on Computers and accessibility, pp.245-246,2007
- [9] A.Vetro, "Mobile visual search:Architectures, technologies, and the emerging mpeg standard",in IEEE MultiMedia, pp.86-94, Mar.2011.
- [10]J.Wright,"Search by text, voice, or image" from <http://insidesearch.blogspot.com/2011/06/search-by-text-voice-or-image.html>
- [11] Solso, Robert L, Cognitive psychology, 2d ed. Boston.Allyn and Bacon, 1988.
- [12] สุมาลี ชัยเจริญ. เทคโนโลยีการศึกษา: หลักการ ทฤษฎี ผู้การปฏิบัติ.ขอนแก่น: คลังนานา,2551.
- [13] รังจิ นพเกตุ.จิตวิทยาการรับรู้.พิมพ์ครั้งที่ 1.กรุงเทพฯ : ประกายพริ้ง, 2540.