

Project Proposal

Division of Applied Mathematics, Department of Mathematics

Faculty of Science, Silpakorn University

Date : 27 กันยายน 2561

Advisor : ผู้ช่วยศาสตราจารย์ ดร. นพดล ชุมชอบ

Student : นายภัคพล พงษ์ทวี รหัส 07580028

Project Title : ขั้นตอนวิธีเชิงตัวเลขชนิดใหม่สำหรับการต่อเติมภาพที่ใช้การแปรผันรวมกับการประยุกต์สำหรับซ่อมแซมภาพวาดศิลปะไทยและการลบบทบรรยายจากอนิเมะ
(A new numerical algorithm for TV-based image inpainting with its applications in restoring Thai painting images and removing subtitles from animes)

1 ที่มาและความสำคัญ

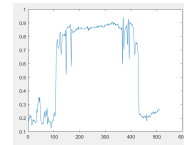
ภาพดิจิทัล (digital images) คือภาพที่นิยมใช้กันอย่างแพร่หลายในปัจจุบันอาจจะสร้างได้หลายวิธีทั้งการใช้กล้องถ่ายภาพเพื่อให้ได้ภาพ หรืออาจจะใช้อุปกรณ์ทางการแพทย์ต่างๆ จนไปถึงการใช้คลื่นที่มองไม่เห็นเพื่อถ่ายภาพดาราศาสตร์ต่างๆ ในอวกาศ ซึ่งภาพที่ได้ออกมานั้นมักจะผ่านการประมวลผลการประมวลผลอยู่เสมอ ตัวอย่างเช่น ภาพถ่ายพื้นผิวดวงจันทร์เมื่อส่งสัญญาณกลับมาจากดาวเทียมจะมีสัญญาณรบกวนเข้ามาแทรก จึงจำเป็นต้องผ่านการกำจัดสัญญาณรบกวนออกจากภาพ (images denoising) การติดตามอาการคนไข้ที่มีอาการเนื้องอกจะเป็นต้องทำการลงทะเบียนภาพ (image registration) เพื่อให้แพทย์สามารถติดตามการเปลี่ยนแปลงของเนื้องอกได้ การติดตามรถที่กระทำความผิดกฎหมายจราจร จำเป็นต้องแยกรถยนต์ออกจากพื้นหลังโดยใช้การแบ่งส่วนภาพ (image segmentation) และการลบวัตถุที่ไม่ต้องการออกไปจากภาพจะใช้การต่อเติมภาพ (image inpainting) เป็นต้น การต่อเติมภาพ เป็นหนึ่งในกระบวนการประมวลผลภาพที่จะเติมเต็มข้อมูลที่หายไปในพื้นที่ภาพที่กำหนด โดยมีจุดประสงค์เพื่อซ่อมแซมภาพที่เสียหาย โดยพื้นที่ภาพส่วนนั้นไม่สามารถพบได้จากการสังเกต โดยการกู้คืน สี, โครงสร้าง และพื้นผิว ที่เกิดการเสียหายเป็นวงกว้าง พิกเซลที่จะนำมาใช้ซ่อมแซมจะถูกคำนวณขึ้นมาใหม่จากข้อมูลที่พิกเซลที่อยู่โดยรอบที่ยังไม่เสียหาย [1] ซึ่งใช้สิ่งที่ไม่ต้องการออกจากภาพ ปัจจุบันมักเห็นได้ตามแอปพลิเคชันหน้าใส ที่ช่วยลบริ้วรอยที่ไม่ต้องการออกจากใบหน้า

1.1 การประยุกต์ใช้สำหรับซ่อมแซมภาพวาดศิลปะไทย

เนื่องจากการต่อเติมภาพด้วยการแปรผันจะสนใจที่ความต่อเนื่องของโครงสร้างทางเรขาคณิต ซึ่งวิธีการต่อเติมภาพรูปภาพด้วยการแปรผันมักจะให้ผลลัพธ์ได้ดีกับรูปภาพที่ราบเรียบเป็นช่วง (piecewise smooth image) [2] ซึ่งภาพวาดศิลปะไทยนั้น เป็นภาพซึ่งมีลักษณะราบเรียบเป็นช่วงจึงเหมาะสำหรับใช้แม่แบบวิธี ROF ซึ่งเป็นวิธีการเชิงแปรผัน สำหรับการต่อเติมภาพ เนื่องจากในปัจจุบันนี้ยังมีงานศิลปะไทยเก่าจำนวนมากที่ต้องการได้รับการซ่อมแซมจึงเป็นการช่วยให้จิตรกรผู้มีหน้าที่ซ่อมแซมภาพ สามารถเห็นตัวอย่างภาพที่ผ่านการต่อเติม เพื่อให้จิตรกรสามารถวางแผนในการซ่อมแซมได้ หรือการเข้าชมพิพิธภัณฑ์ ก็อาจพัฒนาเป็นแอปพลิเคชันสำหรับโทรศัพท์มือถือ ที่สามารถใช้ส่องภาพงานศิลปะและแสดงศิลปะที่สมบูรณ์ขึ้นมาทางหน้าจอได้



(a) ภาพศิลปะไทยซึ่งเป็นเนคสี่เทา



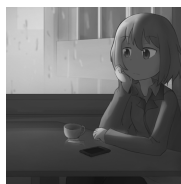
(b) ความเข้มของสีบริเวณคอลลัมน์ที่ 200

โดยจากตัวอย่างจะเป็นภาพวาด จิตรกรรม อุโบสถวัดคงคาราม ขนาด 512x512 ซึ่งถูกปรับให้เป็นเนคสี่เทา เมื่อนำข้อมูลความเข้มของสีบริเวณคอลลัมน์ที่ 200 มาเขียนกราฟจะเห็นว่าบริเวณจุดที่ 100 และจุดที่ 420 มีความแตกต่างของความเข้มอย่างชัดเจน โดยแบ่งเป็นช่วง (0,100), (100,420) และ (420,512) ซึ่งมีลักษณะราบเรียบเป็นช่วงนั่นเอง

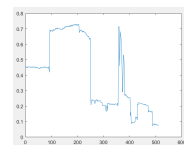
1.2 การประยุกต์ใช้สำหรับการลบทบรยายบนอนิเมะ

อนิเมะคือวิดีโอภาพวาดการ์ตูนสไตล์ญี่ปุ่น ซึ่งภาพวาดสไตล์ของอนิเมะเองก็เป็นภาพแบบราบเรียบเป็นช่วงจึงเหมาะกับการใช้แม่แบบวิธี ROF สำหรับการต่อเติมภาพ ซึ่งอนิเมะนั้นเนื่องจากผ่านการแปลภาษามาแล้ว ในบางครั้งจึงมีคำบรรยายแบบแข็งติดมาด้วย ซึ่งทบรยายแบบแข็งนั้นไม่สามารถเปิดและปิดได้ เมื่อผ่านการให้เสียงใหม่แล้วก็จะมืบทบรยายนั้นติดมาด้วย จึงเหมาะที่จะประยุกต์ใช้ในการลบทบรยายออก

สำหรับการต่อเติมภาพนั้น โดยปกติจะเป็นต้องหาโดเมนต่อเติมด้วยตัวเองซึ่งสามารถทำได้สำหรับภาพ 1 ภาพ แต่สำหรับไฟล์วิดีโออนิเมะนั้น ส่วนใหญ่จะมีภาพจำนวน 24 ภาพต่อวินาที ซึ่งนั่นหมายความว่าหาโดเมนต่อเติมด้วยตัวเองสำหรับ 24 ภาพต่อวินาทีเป็นเรื่องยาก การที่จะหาโดเมนต่อเติมซึ่งเป็นในส่วนคำบรรยายนั้นจึงเป็นเรื่องยาก จึงเป็นอีกความท้าทายที่จะต้องหาโดเมนต่อเติมแบบอัตโนมัติให้ได้



(a) ภาพอนิเมะซึ่งเป็นเนคสี่เทา



(b) ความเข้มของสีบริเวณคอลลัมน์ที่ 256

โดยจากตัวอย่างจะเป็นภาพอนิเมะ ขนาด 512x512 ซึ่งถูกปรับให้เป็นเนคสี่เทา เมื่อนำข้อมูลความเข้มของสีบริเวณคอลลัมน์ที่ 256 มาเขียนกราฟจะเห็นว่าบริเวณจุดที่ 100,250,350,400,420,480 มีความแตกต่างของความเข้มอย่างชัดเจน โดยแบ่งเป็นช่วง (0,100), (100,250), (250,350), (400,420) และ (480,512) ซึ่งมีลักษณะราบเรียบเป็นช่วงนั่นเอง

1.3 ความเร็วของขั้นตอนวิธีเชิงตัวเลข

ซึ่งวิธีการต่อเติมภาพด้วยวิธี Split Bergman ที่มีอยู่เดิมสำหรับภาพขาวดำขนาด 256x256 จะใช้เวลาประมวลผล 1.86 วินาที ดังที่จะเห็นได้จากรูปที่ 4 และรูปที่เป็นภาพสีขนาด 256x256 จะใช้เวลา 6.72 วินาที ดังที่จะเห็นได้จากรูปที่ 5 โดยสำหรับวิธีที่มีจำนวนเฟรม 24 เฟรมในหนึ่งวินาทีและมีขนาดเป็น 1920x1080 จะต้องใช้เวลาหลายวินาทีในการต่อเติมเฟรม 1 ภาพ จึงทำให้เวลาการต่อเติมวิดีโอ 1 วินาที จำเป็นต้องใช้เวลามากมายในการต่อเติม

เพราะฉะนั้นแล้วผู้วิจัยจึงสนใจที่จะพัฒนาขั้นตอนวิธีเชิงตัวเลขให้สามารถต่อเติมภาพได้รวดเร็วยิ่งขึ้น ซึ่งวิธีการเชิงตัวเลขที่พัฒนาขึ้นสามารถประยุกต์ใช้ได้กับการซ่อมแซมภาพวาดศิลปะไทย และการลบทบรยายบนอนิเมะ นอกจากนี้ยังสามารถนำไปประยุกต์ใช้กับงานต่อเติมภาพอื่นๆ ได้อีกในอนาคต

2 บรรณกรรมและทฤษฎีบทที่เกี่ยวข้อง

2.1 การวัดคุณภาพของภาพหลังจากการต่อเติม

หลังจากการต่อเติมภาพแล้วจำเป็นต้องพิจารณาว่าการวัดคุณภาพของภาพที่ผ่านการต่อเติมดีมากน้อยเพียงใด โดยในวิจัยนี้จะสนใจคุณภาพของค่าในแต่ละพิกเซลที่ใกล้เคียงกับภาพต้นฉบับ และโครงสร้างโดยรวมที่ใกล้เคียงกับภาพต้นฉบับ โดยการวัดค่าดังต่อไปนี้

Peak signal-to-noise ratio (PSNR)[3] ใช้สำหรับวัดคุณภาพของภาพโดยเปรียบเทียบจากพิกเซลแต่ละพิกเซล โดยภาพที่มีความคล้ายต้นฉบับจะมีค่า PSNR เข้าใกล้อนันต์ หรือก็คือยิ่งมีค่ามากยิ่งคุณภาพดี ซึ่งสามารถคำนวณได้โดย

$$PSNR = 10 \cdot \log_{10} \left(\frac{peak^2}{\sqrt{MSE}} \right)$$

เมื่อ MSE คือ mean square error และ peak คือค่าสูงสุดโดยประเภทของภาพ ซึ่งสำหรับงานที่จะพูดถึงต่อไปนี้ จะพิจารณาภาพเป็นฟังก์ชันที่มีความเข้มของภาพอยู่ในช่วง $[0, 1]$ จึงได้ว่า peak มีค่าเป็น 1

Structural similarity (SSIM) [4] ใช้สำหรับว่าวัดคุณภาพของภาพจากโครงสร้างของภาพ โดยพิจารณาว่าภาพนั้นมีโครงสร้างแตกต่างหรือคล้ายคลึงกับภาพต้นฉบับมากน้อยเพียงใด โดยมีค่าอยู่ระหว่าง 0 ถึง 1 หากทั้งสองภาพมีความคล้ายคลึงกันมากค่า SSIM จะเข้าใกล้กับค่า 1 ซึ่ง SSIM นั้นสามารถคำนวณได้โดย

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}$$

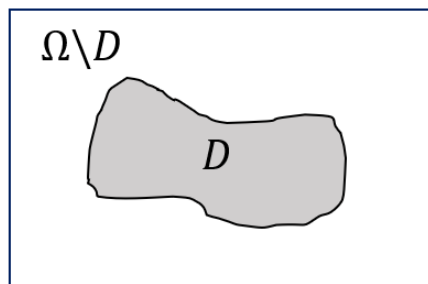
เมื่อ x, y คือภาพที่ต้องการเปรียบเทียบ μ คือค่าเฉลี่ยของภาพ σ^2 คือค่าความแปรปรวนของภาพ σ_{xy} คือความแปรปรวนร่วม $c_1 = (0.01L)^2$, $c_2 = (0.03L)^2$ และ L คือค่าสูงสุดโดยประเภทของภาพ ซึ่งสำหรับงานที่จะพูดถึงต่อไปนี้ จะพิจารณาภาพเป็นฟังก์ชันที่มีความเข้มของภาพอยู่ในช่วง $[0, 1]$ จึงได้ว่า L มีค่าเป็น 1

2.2 วิธีการทางคณิตศาสตร์สำหรับการต่อเติมภาพด้วยการแปรผัน

ในการต่อเติมภาพเฉดสีเทาด้วยวิธีการเชิงแปรผัน เราพิจารณาภาพ

$$u : \Omega \subset \mathbb{R}^2 \rightarrow V \subset [0, \infty)$$

เป็นฟังก์ชันต่อเนื่อง โดยที่ $\mathbf{x} = (x, y) \in \Omega$ แทนพิกัดทางกายภาพ (physical position) ของภาพ $u(\mathbf{x}) \in V$ แทนระดับความเข้มของภาพ (image intensity) ที่ \mathbf{x} และ Ω แทนโดเมนของภาพ ซึ่งในที่นี้เราสามารถสมมติได้โดย $\Omega = [0, n]^2$ และ $V = [0, 1]$ เมื่อ $n > 0$ เป็นจำนวนเต็มบวก ทั้งนี้ เราจะเรียกภาพ u ที่นิยามข้างต้นว่าภาพเฉดสีเทา (grayscale image)



รูปที่ 3: ตัวอย่าง โดเมนต่อเติม

ซึ่งสำหรับปัญหาการต่อเติมภาพโทนสีเทานั้น จะเรียกพื้นที่ที่ต้องการต่อเติมว่า โดเมนต่อเติม (inpainting domain) โดย D เป็นโดเมนซึ่ง $D \subset \Omega$

การต่อเติมภาพเฉดเทา นี้ จะใช้การแปรผันรวม (total variation) ซึ่งถูกคิดค้นโดย Chan และ Shen [5] ซึ่งประยุกต์มาจากการแปรผันรวมเพื่อกำจัดสัญญาณรบกวนบนตัวแบบ ROF [6] ได้ว่าจะสามารถหาภาพที่ถูกต่อเติมอย่างเหมาะสม u จะสามารถหาได้จาก

$$\min_u \{ \mathcal{J}(u) = \lambda \mathcal{D}(u, f) + \mathcal{R}(u) \}$$

ซึ่ง \mathcal{D} คือพจน์สำหรับวัดค่าเหมาะสม เพื่อไม่ให้ภาพก่อนต่อเติมและหลังจากต่อเติมมีความแตกต่างกันมากเกินไป \mathcal{R} คือพจน์สำหรับการต่อเติมภาพ และ λ คือเป็นพารามิเตอร์เร็กกิวลาริเซชัน (regularization parameter) สำหรับกำหนดปริมาณของสัญญาณรบกวนที่ต้องการกำจัดออก

Chan และ Shen ได้ทำการเสนอปัญหาการแปรผัน (variational problem) สำหรับการต่อเติมภาพไว้ว่า

$$\min_u \{ \mathcal{J}(u) = \frac{\lambda}{2} \int_{\Omega \setminus D} (u - z)^2 d\Omega + \int_{\Omega} |\nabla u| d\Omega \}$$

ซึ่งเมื่อทำให้ได้สมการออยเลอร์ลากรางจ์ สำหรับ functional \mathcal{J} คือ

$$\begin{cases} -\nabla \cdot \left(\frac{\nabla u}{|\nabla u|} \right) + \lambda(u - z) = 0 & x \in \Omega = (0, n)^2 \\ \frac{\partial u}{\partial n} = 0 & x \in \partial\Omega \end{cases}$$

สำหรับ $x \in \Omega$ ได้ว่า λ ถูกกำหนดโดย

$$\lambda(x) \begin{cases} \lambda & x \in \Omega \setminus D \\ 0 & x \in D \end{cases}$$

จะเห็นได้ว่าเป็นสมการออยเลอร์ ลากรางจ์ดังกล่าวเป็น สมการอนุพันธ์ย่อยไม่เป็นเชิงเส้น จึงสามารถใช้วิธีเชิงตัวเลขด้วยวิธีต่างๆ ได้ดังนี้

วิธีการ Explicit Time Marching [7] โดยแนวคิดของวิธีการนี้คือการแนะนำตัวแปรเวลาสังเคราะห์ (time artificial variable) จากนั้นหาคำตอบแบบสภาวะคงตัว (steady-state solution) ของสมการเชิงอนุพันธ์ย่อยไม่เป็นเชิงเส้นที่ขึ้นอยู่กับเวลา และเพื่อจะแก้ความไม่เป็นเชิงเส้นของสมการเชิงอนุพันธ์ย่อย จะสามารถใช้รูปแบบที่ชัดเจนของออยเลอร์ (Euler's explicit scheme) ที่กำหนดโดย

$$u(x, t_{k+1}) = u(x, t_k) + \tau \left(\nabla \cdot \left(\frac{\nabla u(x, t_k)}{|\nabla u(x, t_k)|} \right) + \lambda(x)(u(x, t_k) - z(x)) \right)$$

แต่เนื่องจากเมื่อ $|\nabla u(x, t_k)| = 0$ จะทำให้ไม่สามารถคำนวณได้ เนื่องจากหารด้วย 0 ไม่ได้ จึงกำหนดให้

$$|\nabla u(x, t_k)| = \sqrt{u_x^2 + u_y^2 + \beta} \text{ สำหรับ } \beta > 0 \text{ และ } \beta \text{ มีค่าใกล้ } 0 \text{ เพื่อหลีกเลี่ยงเหตุการณ์ดังกล่าว}$$

โดยเมื่อ $\tau > 0$ แทนขั้นเวลา (time step) ที่ได้จากการดิสครีไทซ์โดเมนเวลา $[0, \infty)$ ซึ่งเห็นได้ว่าวิธีการเชิงตัวเลขดังกล่าวข้างต้นนั้นง่ายในการคำนวณ แต่การเข้าสู่คำตอบที่เหมาะสมของปัญหา เชิงแปรผันค่อนข้างช้ามากเนื่องจากต้องใช้ τ ที่มีขนาดเล็กในการทำให้ลำดับของคำตอบเข้าสู่

วิธีการทำซ้ำแบบ Fixed Point [8] จะทำการแบ่งการทำซ้ำออกเป็น 2 ชั้น เรียกชั้นในกับชั้นนอก โดยการทำซ้ำชั้นในเป็นการทำซ้ำแบบ Gauss-seidel เพื่อให้ค่า u เป็นลำดับถัดไป จากนั้นชั้นนอกจะเป็นการทำซ้ำแบบตรึงจุด (fixed point) เพื่อให้ค่า u เข้าสู่ค่าที่ต้องการ

ซึ่งการทำซ้ำชั้นในจะหาค่า u จากการใช้ Gauss-Seidel กับสมการ

$$-\nabla \cdot \left(\frac{\nabla u^{[v+1]}}{|\nabla u^{[v]}|} \right) + \lambda(u^{[v+1]} - z) = 0$$

โดยที่ $v = 0, 1, 2, \dots$ ซึ่งคือค่าของจำนวนครั้งในการทำซ้ำชั้นนอกที่ถูกทำไป และเช่นเดียวกับวิธี Explicit Time Marching

เนื่องจากวิธีนี้เมื่อ $|\nabla u| = 0$ จะทำให้ไม่สามารถคำนวณได้ เนื่องจากไม่สามารถหารด้วย 0 ได้ จึงกำหนดให้

$$|\nabla u| = \sqrt{u_x^2 + u_y^2 + \beta} \text{ สำหรับ } \beta > 0 \text{ และ } \beta \text{ มีค่าใกล้ } 0 \text{ เพื่อหลีกเลี่ยงเหตุการณ์ดังกล่าว}$$

ซึ่งทั้งวิธี Explicit Time Marching และวิธี Fixed Point จำเป็นต้องเพิ่ม β เพื่อหลีกเลี่ยงการหารด้วย 0 ซึ่งแม้จะเลือกค่า β ดังกล่าวให้ใกล้ 0 มากๆ ก็ยังทำให้เกิดค่าคลาดเคลื่อน จึงได้มีการสร้างวิธี Split Bergman สำหรับแก้ปัญหานี้

วิธี Split Bergman [9] โดยวิธีการนี้คือการแยกส่วนการดำเนินการ (splitting) และใช้การทำซ้ำ Bergman (bergman iteration) เพื่อที่จะกำจัดตัวแปร β ที่ทำให้เกิดการหารศูนย์ขึ้น วิธีนี้จะเป็นการแก้ปัญหาคงค่าแปรผัน โดยการเพิ่มตัวแปร w ในการแก้ปัญห ซึ่งจะได้ปัญหาคงค่าต่อไปนี้แทน

$$\min_u \{ \mathcal{J}(u) = \frac{\lambda}{2} \int_{\Omega \setminus D} (u - z)^2 d\Omega + \int_{\Omega} |\nabla w| d\Omega + \frac{\theta}{2} \int_{\Omega} (w - \nabla u + b) d\Omega \}$$

โดยเมื่อเพิ่มตัวแปร w สำหรับช่วยในการคำนวณแล้ว จะมีพจน์ $\int_{\Omega} |\nabla w| d\Omega$ เพื่อบีบบังคับให้ w ไม่เปลี่ยนแปลงผลของคำตอบ, θ คือค่าบวกใดๆ ซึ่งเกี่ยวข้องกับความเร็วของการต่อเติมซึ่งไม่ควรเล็กหรือใหญ่เกินไปเพื่อให้สามารถลู่เข้าได้ดี b คือตัวแปรช่วยสำหรับการทำซ้ำ Bergman ซึ่งปัญหาดังกล่าวสามารถแบ่งปัญหาได้เป็น 2 ส่วนคือ ปัญหาย่อย w โดยการคงค่า u ไว้จะได้ว่าปัญหาย่อยคือ

$$\min_u \{ \mathcal{J}(u) = \int_{\Omega} |\nabla w| d\Omega + \frac{\theta}{2} \int_{\Omega} (w - \nabla u + b) d\Omega \}$$

ซึ่งเมื่อทำการแก้ปัญหาย่อยนี้แล้วจะได้ว่า

$$w_{i,j} = \frac{\nabla u_{i,j} + b_{i,j}}{|\nabla u_{i,j} + b_{i,j}|} \max\{|\nabla u_{i,j} + b_{i,j}| - \frac{1}{\theta}, 0\}$$

ปัญหาย่อย u โดยการคงค่า w ไว้จะได้ว่าปัญหาย่อยคือ

$$\min_u \{ \mathcal{J}(u) = \frac{\lambda}{2} \int_{\Omega \setminus D} (u - z)^2 d\Omega + \frac{\theta}{2} \int_{\Omega} (w - \nabla u + b) d\Omega \}$$

ซึ่งเมื่อทำการแก้ปัญหาย่อยนี้แล้วจะได้ว่า

$$\frac{1}{\theta} \lambda u - \Delta u = \frac{1}{\theta} \lambda z - \nabla \cdot (w - b)$$

ส่วนตัวแปรช่วย b มีค่าเริ่มต้นเป็น 0 จากนั้นทำการปรับค่าโดย

$$b^{k+1} = b^k + \nabla u - w$$

จึงได้ว่าวิธีการ Split Bergman มีการทำงานในภาพรวมเป็นดังนี้

```

initialization  $u = 0, d = 0, b = 0$ 
while  $\|u_{cur} - u_{prev}\|_2 > Tol$  do
    Solve the  $w$  subproblem
    Solve the  $u$  subproblem
     $b = b + \nabla u - w$ 
end

```

การทำซ้ำนี้จะทำงานกระทั่ง นอร์ม L2 ระหว่างรอบปัจจุบันต่างกับรอบก่อนหน้าไม่เกินค่า Tol ที่กำหนดไว้หรือจำนวนรอบการทำซ้ำมากจนถึงจุดสิ้นสุดที่เพียงพอที่จะให้ลู่เข้าซึ่งไม่ควรใหญ่เกินไปเพื่อไม่ให้เสียเวลาประมวลผลจนนานเกินไป

ผลการต่อเติมภาพจากทั้ง 3 วิธีข้างต้น สำหรับการกำหนดรอบการทำซ้ำไม่เกิน 1000 รอบและค่านอร์ม L2 ภาพปัจจุบันและภาพก่อนหน้าต่างกันไม่เกิน 0.0001 ได้ผลลัพธ์ดังนี้



รูปที่ 4: การต่อเติมรูปภาพเฉดเทา

จากการทดลองจะเห็นว่าวิธีการ Split Bergman ได้ผลลัพธ์ที่ดีที่สุด จึงได้มีความสนใจที่จะศึกษาวิธี Split Bergman เป็นลำดับถัดไป

2.3 วิธีการทางคณิตศาสตร์สำหรับการต่อเติมภาพด้วยการแปรผันบนภาพสี

สำหรับการต่อเติมภาพเชิงแปรผันซึ่งเป็นภาพสีในระบบ RGB จะพิจารณา

$$u = (u_1, u_2, u_3)^T : \Omega \rightarrow V^3$$

เมื่อ (u_1, u_2, u_3) โดยที่ $u_1, u_2, u_3 : \Omega \rightarrow V$ แทนภาพเฉดสีแดง สีเขียว และสีน้ำเงินของ u ตามลำดับ จะสามารถปรับปรุงตัวแบบ ROF ได้ดังนี้

$$\min_u \{ \mathcal{J}(u) = \lambda \bar{\mathcal{D}}(u, z) + \bar{\mathcal{R}}(u) \}$$

โดยที่ $\bar{\mathcal{D}}$ เป็นพจน์วัดค่าเหมาะสมเพื่อให้ภาพก่อนการต่อและหลังต่อเติมต่างกันมาเกินไปซึ่ง

$$\begin{aligned} \bar{\mathcal{D}}(u, z) &= \sum_{l=1}^3 \mathcal{D}(u_l, z_l) \\ &= \frac{1}{2} \int_{\Omega} (u_1 - z_1)^2 d\Omega + \frac{1}{2} \int_{\Omega} (u_2 - z_2)^2 d\Omega + \frac{1}{2} \int_{\Omega} (u_3 - z_3)^2 d\Omega \end{aligned}$$

และ $\bar{\mathcal{R}}$ คือพจน์สำหรับการต่อเติมภาพ

$$\bar{\mathcal{R}}(u) = \sum_{l=1}^3 \mathcal{R}(u_l) = \int_{\Omega} |\nabla u_1| d\Omega + \int_{\Omega} |\nabla u_2| d\Omega + \int_{\Omega} |\nabla u_3| d\Omega$$

ซึ่งงานนี้ได้สนใจที่จะนำวิธี Split Bergman มาประยุกต์ใช้กับปัญหาเชิงแปรผัน จึงแนะนำเวกเตอร์เสริม w_1, w_2, w_3 พร้อมทั้งตัวแปรสำหรับการทำซ้ำ Bregman b_1, b_2, b_3 และตัวแปรช่วย $\theta_1, \theta_2, \theta_3 > 0$ เพื่อแก้ปัญหาเชิงแปรผัน ซึ่งจะได้ปัญหาเชิงแปรผันเป็น

$$\min_u \{ \bar{\mathcal{J}} = \lambda \bar{\mathcal{D}}(u, z) + \sum_{l=1}^3 \int_{\Omega} |w_l| d\Omega + \frac{\theta_l}{2} \sum_{l=1}^3 \int_{\Omega} (w_l - \nabla u_l - b_l)^2 d\Omega \}$$

ปัญหาย่อย w โดยการคงค่า u ไว้จะได้ว่าปัญหาย่อยคือ

$$\min_u \{ \bar{\mathcal{J}} = \sum_{l=1}^3 \int_{\Omega} |w_l| d\Omega + \frac{\theta_l}{2} \sum_{l=1}^3 \int_{\Omega} (w_l - \nabla u_l - b_l)^2 d\Omega \}$$

ซึ่งเมื่อทำการแก้ปัญหาย่อยนี้จะได้คำตอบที่แน่นอนตรงคือ

$$w_{l,i,j} = \frac{\nabla u_{l,i,j} + b_{l,i,j}}{|\nabla u_{l,i,j} + b_{l,i,j}|} \max\{|\nabla u_{l,i,j} + b_{l,i,j}| - \frac{1}{\theta}, 0\} \quad l = 1, 2, 3$$

ปัญหาย่อย u โดยการคงค่า w ไว้จะได้ว่าปัญหาย่อยคือ

$$\min_u \{ \bar{\mathcal{J}} = \lambda \bar{\mathcal{D}}(u, z) + \frac{\theta_l}{2} \sum_{l=1}^3 \int_{\Omega} (w_l - \nabla u_l - b_l)^2 d\Omega \}$$

ซึ่งเมื่อทำการแก้ปัญหาย่อยนี้แล้วจะได้ระบบสมการย่อยเลอร์ลากรางจ์คือ

$$\frac{1}{\theta} \lambda u_1 - \Delta u_1 = \frac{1}{\theta} \lambda z_1 - \nabla \cdot (w_1 - b_1)$$

$$\frac{1}{\theta} \lambda u_2 - \Delta u_2 = \frac{1}{\theta} \lambda z_2 - \nabla \cdot (w_2 - b_2)$$

$$\frac{1}{\theta} \lambda u_3 - \Delta u_3 = \frac{1}{\theta} \lambda z_3 - \nabla \cdot (w_3 - b_3)$$

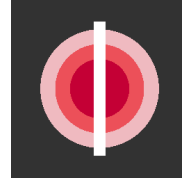
ซึ่งเมื่อแก้ระบบสมการดังกล่าวแล้วจะนำค่าที่ได้มาเปลี่ยนค่าตัวแปร b โดยที่

$$b_l^{k+1} = b_l^k + \nabla u_l - w_l \quad l = 1, 2, 3$$

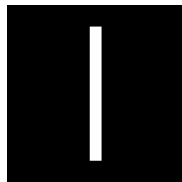
ซึ่งจากวิธีการ Split Bergman สำหรับภาพสี จะได้ภาพผลลัพธ์ดังนี้



(a) ภาพต้นฉบับ



(b) ภาพที่ต้องการทำการต่อเติม



(c) โดเมนต่อเติม



(d) วิธี Split Bergman ใช้เวลา 6.72 วินาที PSNR 37.2374 SSIM 0.9976

รูปที่ 5: การต่อเติมรูปภาพสี

3 วัตถุประสงค์โครงการวิจัย

วัตถุประสงค์ของโครงการวิจัยมีดังต่อไปนี้

- (1) ศึกษาวิธีการแปรผันและวิธีการเชิงตัวเลขที่มีประสิทธิภาพเพื่อเติมข้อมูลที่ขาดหายในภาพหรือวิดีโอ
- (2) สร้างวิธีการเชิงตัวเลขใหม่สำหรับซ่อมแซมภาพศิลปะไทยและลบบทบรรยายออกจากอนิเมะ
- (3) นำวิธีการที่สร้างขึ้นเพื่อซ่อมแซมภาพไทย และลบบทบรรยายในอนิเมะ

4 ขอบเขตการศึกษา

ขอบเขตของโครงการมีดังต่อไปนี้

- (4.1) ภาพศิลปะที่ใช้ศึกษา เป็นภาพจิตรกรรมฝาผนังไทย ที่อยู่ภายใต้เว็บไซต์ Wikipedia.org ซึ่งได้รับการอนุญาตให้ใช้งานแบบ Creative Commons หรือแบบ Public Domain
- (4.2) วิดีโอที่ใช้ศึกษาเป็นวิดีโอประเภทอนิเมะ โดยศึกษากับไฟล์อนิเมะที่ใช้ Color space แบบ RGB เท่านั้น
- (4.3) บทบรรยายที่ใช้ทดสอบ จะถูกล้อมรอบไว้ด้วยสีดำ ขนาดความหนาขนาดไม่น้อยกว่า 5 พิกเซล
- (4.4) วิดีโอที่ใช้ศึกษาขนาดไม่เกิน 1920x1080
- (4.5) คอมพิวเตอร์ที่ใช้ทดลองใช้หน่วยประมวลผล I7-6700HQ ใช้การ์ดจอ Nvidia GTX 960M แรม 16GB ฮาร์ดดิสก์แบบ SSD
- (4.6) จะศึกษาคุณภาพของภาพหลังจากผ่านการต่อเติมโดยใช้ค่า PSNR และ SSIM โดยสำหรับไฟล์วิดีโอนั้นจะทำการเฉลี่ยจากจำนวนเฟรมทั้งหมดในวิดีโอ หากเฟรมใดได้ค่า PSNR เป็นอนันต์ จะทำการข้ามเฟรมนั้นไป

5 ระเบียบวิธีวิจัย

วิธีการมีดังต่อไปนี้

- (5.1) ศึกษาการคณิตศาสตร์ต่อเติมข้อมูลที่ขาดหายบนรูปภาพ
- (5.2) พัฒนาวิธีการเชิงตัวเลขสำหรับการซ่อมแซมรูปภาพ
- (5.3) ทดสอบวิธีการเชิงตัวเลขที่พัฒนาขึ้นโดยโปรแกรมคอมพิวเตอร์บนภาพสังเคราะห์
- (5.4) อภิปรายผลที่ได้จากการทดลองเชิงตัวเลข
- (5.5) สรุปผลการดำเนินงานวิจัยและจัดทำรูปเล่มฉบับสมบูรณ์

6 แผนการดำเนินงานวิจัย

แผนการดำเนินงานตลอดทั้งโครงการสามารถสรุปได้โดยย่อจากตารางต่อไปนี้

แผนการดำเนินงาน	เดือนที่											
	1	2	3	4	5	6	7	8	9	10	11	12
ศึกษาการคณิตศาสตร์ต่อเติมข้อมูลที่ขาดหายบนรูปภาพ	x	x										
พัฒนาวิธีการเชิงตัวเลขสำหรับการซ่อมแซมรูปภาพ			x	x								
ทดสอบวิธีการเชิงตัวเลขที่พัฒนาขึ้นโดยโปรแกรม-คอมพิวเตอร์บนภาพสังเคราะห์					x	x						
อภิปรายผลที่ได้จากการทดลองเชิงตัวเลข							x	x				
สรุปผลการดำเนินงานวิจัยและจัดทำรูปเล่มฉบับสมบูรณ์									x	x	x	x

7 บรรณานุกรม

- [1] Borko Furht., Encyclopedia of Multimedia. Springer. pp. 315-316, 2006.
- [2] Işık Barış Fidaner. "A Survey on Variational Image Inpainting , Texture Synthesis and Image Completion." 2007. **https://www.semanticscholar.org/paper/_/36f4d32ce45f72091510ab4d4d1cc3bf81ffe879**
- [3] David Salomon. Data Compression: The Complete Reference (4 ed.). Springer. pp. 281. 2007.
- [4] Zhou Wang, Alan Conrad Bovik, Hamid Rahim Sheikh and Eero P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," in IEEE Transactions on Image Processing, vol. 13, no. 4, pp. 600-612, 2004.
- [5] Tony F. Chan and Jianhong Shen , "Mathematical models of local non-texture inpaintings", SIAM Journal on Applied Mathematics, vol. 62, no. 3, pp. 1019-1043, 2001.
- [6] Leonid I. Rudin, Stanley Osher, Emad Fatemi, "Nonlinear total variation based noise removal algorithms", Physica D: Nonlinear Phenomena, Volume 60, Issues 1-4, pp. 259-268, 1992.
- [7] Antonio Marquina and Stanley Osher, "Explicit algorithms for a new time dependent model based on level set motion for nonlinear deblurring and noise removal", SIAM Journal on Scientific Computing, vol. 22 issue 2, 387-405. 2006.
- [8] Curtis R. Vogel and M.E. Oman, "Iterative methods for total variation denoising", SIAM Journal on Scientific Computing. vol. 17, pp. 227-238, 1996.
- [9] Pascal Getreuer, Total Variation Inpainting using Split Bregman, Image Processing On Line, 2 (2012), pp. 147-157.