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 Introduction

ภาพดิจิตัล (digital images) คือภาพที่นิยมใช้กันอย่างแพร่หลายในปัจจุบันอาจจะสร้างได้หลายวิธีทั้งการใช้กล้องถ่ายภาพเพื่อให้ได้ภาพ หรืออาจจะใช้อุปกรณ์ทางการแพทย์ต่างๆ จนไปถึงการใช้คลื่นที่มองไม่เห็นเพื่อถ่ายภาพดาราจักรต่างๆ ในอวกาศ ซึ่งภาพที่ได้ออกมานั้นมักจะผ่านการประมวลการประมวลผลอยู่เสมอ ตัวอย่างเช่น ภาพถ่ายพื้นผิวดวงจันทร์เมื่อส่งสัญญาณกลับมาจากดาวเทียมจะมีสัญญาณรบกวนเข้ามาแทรก จึงจำเป็นที่จะต้องผ่านการกำจัดสัญญาณรบกวนออกจากภาพ (images denoising) การติดตามอาการคนไข้ที่มีอาการเนื้องอกจะเป็นต้องทำการลงทะเบียนภาพ (image registration) เพื่อให้แพทย์สามารถติดตามการเปลี่ยนแปลงของเนื้องอกได้ การติดตามรถที่กระทำผิดกฏจราจร จำเป็นต้องแยกรถยนตร์ออกจากพื้นหลังโดยใช้การแบ่งส่วนภาพ (image segmentation) และการลบวัตถุที่ไม่ต้องการออกไปจากภาพจะใช้การซ่อมแซมภาพ (image inpainting) เป็นต้น

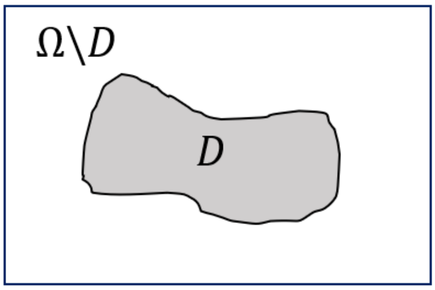
การต่อเติมภาพ คือเป็นหนึ่งในกระบวนการประมวลผลภาพที่จะเติมเต็มข้อมูลที่หายไปในพื้นที่ภาพที่กำหนด โดยมีจุดประสงค์เพื่อซ่อมแซมภาพที่เสียหาย โดยพื้นที่ภาพส่วนนั้นไม่สามารถพบได้จากการสังเกต โดยการกู้คืน สี, โครงสร้าง และพื้นผิว ที่เกิดการเสียหายเป็นวงกว้าง พิกเซลที่จะนำมาใช้ซ่อมแซมจะถูกคำนวณขึ้นมาใหม่จากข้อมูลที่พิกเซลที่อยู่โดยรอบที่ยังไม่เสียหาย [2] ซึ่งใช้ลบสิ่งที่ไม่ต้องการออกจากภาพ ปัจจุบันมักเห็นได้ตามแอปพลิเคชันหน้าใส ที่ช่วยลบริ้วรอยที่ไม่ต้องการออกจากใบหน้า

1.1 วิธีการทางคณิตศาสตร์สำหรับการต่อเติมภาพด้วยการแปรผัน

ในการต่อเติมภาพเฉดสีเทาด้วยวิธีการเชิงแปรผัน เราพิจารณาภาพ

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

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



รูปที่ 1: ตัวอย่าง โดเมนต่อเติม

ซึ่งสำหรับปัญหาการต่อเติมภาพโทนสีเทานั้น จะเรียกพื้นที่ซึ่งต้องการต่อเติมว่า โดเมนต่อเติม (inpaint domain) โดย D เป็นโดเมนซึ่ง $D \subset \Omega$

โดยการต่อเติมภาพเฉดเทานี้ จะใช้การแปรผันรวม (Total Variation) ซึ่งถูกคิดค้นโดย Chan และ Shen [1] ซึ่งประยุกต์มาจากการแปรผันรวมเพื่อกำจัดสัญญาณรบกวนบนตัวแบบ ROF [4]

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 คือเป็นพารามิเตอร์เร็กกิวลาร์ไรซ์เซชัน (regularization parameter) สำหรับกำหนดปริมาณของสัญญาณรบกวนที่ต้องการกำจัดออก

โดย Chan และ Shen ได้ทำการแก้ปัญหาการแปรผัน (Variational Problem) ได้ว่า

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

ซึ่งเมื่อทำให้ได้สมการออยเลอร์ลากรางซ์ สำหรับ fuctional $\mathcal{J}$ คือ

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

สำหรับ $\mathsf{x} \in \Omega$ ได้ว่า λ ถูกกำหนดโดย

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

จะเห็นได้ว่าเป็นสมการออยเลอร์ ลากรางซ์ดังกล่าวเป็น สมการอนุพันธ์ย่อยไม่เป็นเชิงเส้น จึงสามารถใช้วิธีเชิงตัวเลขด้วยวิธีต่างๆ ได้ดังนี้

วิธีการ Explicit Time Marching [5] โดยแนวคิดของวิธีการนี้คือการแนะนำตัวแปรเวลาสังเคราะห์ (time artificial variable) จากนั้นหาคำตอบแบบสภาวะคงตัว (steady-state solution) ของสมการเชิงอนุพันธ์ย่อยไม่เป็นเชิงเส้นที่ขึ้นอยู่กับเวลา และเพื่อจะแก้ความไม่เป็นเชิงเส้นของสมการเชิงอนุพันธ์ย่อย จะสามารถใช้รูปแบบที่ชัดแจ้งของออยเลอร์ (Euler’s explicit scheme) ที่กำหนดโดย

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

แต่เนื่องจากเมื่อ $|\nabla u(\mathsf{x}, t_k)| = 0$ จะทำให้ไม่สามารถคำนวณได้ เนื่องจากไม่สามารถหารด้วย 0 ได้ จึงกำหนดให้ $|\nabla u(\mathsf{x}, t_k)| = \sqrt{u_x^2 + u_y^2 + \beta}$ สำหรับ $\beta > 0$ และ β มีค่าใกล้ 0 เพื่อหลีกเลี่ยงเหตุการณ์ดังกล่าว

โดยเมื่อ $\tau > 0$ แทนขั้นเวลา (time step) ที่ได้จากการดิสครีตไทซ์โดเมนเวลา $[0, \infty)$ ซึ่งเห็นได้ว่าวิธีการเชิงตัวเลขดังกล่าวข้างต้นนั้นง่ายในการคำนวณ แต่การลู่เข้าสู่คำตอบที่เหมาะสมของปัญหา เชิงแปรผันค่อนข้างช้ามากเนื่องจากต้องใช้ τ ที่มีขนาดเล็กในการทำให้ลำดับของคำตอบลู่เข้า

วิธีการทำซ้ำแบบ Fixed Point [6] จะทำการแบ่งการทำซ้ำออกเป็น 2 ชั้น เรียกชั้นในกับชั้นนอก โดยการทำซ้ำชั้นในเป็นการทำซ้ำแบบ Gauss-seidel เพื่อให้หาค่า u เป็นลำดับถัดไป จากนั้นชั้นนอกจะเป็นการทำซ้ำแบบตรึงจุด (Fixed Point) เพื่อให้ค่า u ลู่เข้าสู่ค่าที่ต้องการ

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 $v = 0, 1, 2, ...$ ซึ่งคือค่าของจำนวนครั้งในการทำชั้นนอกที่ถูกทำไป

และเช่นเดียวกับวิธี Explicit Time Marching เนื่องจากวิธีนี้เมื่อ $|\nabla u(\mathsf{x}, t_k) = 0|$ จะทำให้ไม่สามารถคำนวณได้เนื่องจากไม่สามารถหารด้วย 0 ได้ จึงกำหนดให้ $|\nabla u(\mathsf{x}, t_k)| = \sqrt{u_x^2 + u_y^2 + \beta}$ สำหรับ $\beta > 0$ และ β มีค่าใกล้ 0 เพื่อหลีกเลี่ยงเหตุการณ์ดังกล่าว

ซึ่งทั้งวิธี Explicit Time Marching และวิธี Fixed Point จำเป็นต้องเพิ่ม β เพื่อหลีกเลี่ยงการหารด้วย 0 ซึ่งจะทำให้เกิดค่าคลาดเคลื่อน จึงได้มีการสร้างวิธี Split Bergman สำหรับแก้ปัญหานี้

วิธี Split Bergman [8] โดยวิธีการนี้คือการแยกส่วนการดำเนินการ (Spliting) และใช้การทำซ้ำ Bergman (Bergman Iteration) โดยวิธีนี้จะเป็นการแก้ปัญหาการแปรผัน โดยการเพิ่มตัวแปร w ในการแก้ปัญหา ซึ่งจะได้ปัญหาการแปรผันดังต่อไปนี้แทน

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

โดยเมื่อเพิ่มตัวแปร w สำหรับช่วยใจการคำนวณแล้ว จะมีพจน์ $\int_\Omega | \bigtriangledown w| d\Omega$ เพื่อบีบบังคับให้ w ไม่เปลี่ยนแปลงผลของคำตอบ, θ คือค่าบวกใดๆ ซึ่งเกี่ยวข้องกับความแรงของการต่อเติมซึ่งไม่ควรเล็กหรือใหญ่เกินไปเพื่อให้สามารถลู่เข้าได้ดี b คือตัวแปรช่วยสำหรับการทำซ้ำ Bergman ซึ่งปัญหาดังกล่าวสามารถแบ่งปัญหาได้เป็น 2 ส่วนคือ

ปัญหาย่อย w โดยการคงค่า u ไว้จะได้ว่าปัญหาย่อยคือ

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

ซึ่งเมื่อทำการแก้ปัญหาย่นอนี้แล้วจะได้ว่า

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

ปัญหาย่อย u โดยการคงค่า w ไว้จะได้ว่าปัญหาย่อยคือ

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

ซึ่งเมื่อทำการแก้ปัญหาย่อยนี้แล้วจะได้ว่า

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

เราจะประมาณคำตอบนี้โดยการใช้ หนึ่งรอบ Gauss-seidel ต่อหนึ่งรอบการทำซ้ำของ Bergman ซึ่ง subproblem จะถูกแก้หนึ่งครั้ง ต่อหนึ่งรอบ bergman iteration แต่ทั้งนี้การทำซ้ำ Gauss-seidel หลายครั้ง จะทำให้การแก้ปัญหาย่อยมีความแม่นยำขึ้น ส่วนตัวแปรช่วย b มีค่าเริ่มต้นเป็น 0 จากนั้นทำการปรับค่าโดย

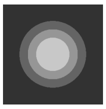
$$b^{k+1} = b^k + \bigtriangledown u - w$$

จึงได้ว่าวิธีการ Split Bergman มีการทำงานในภาพรวมเป็นดังนี้

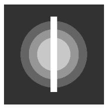
initialization $u = 0, d = 0, b = 0$
while $||u_{cur} - u_{prev}||_2 > Tol$ **do**
  Solve the w subproblem
  Solve the u subproblem
  $b = b + \bigtriangledown u - w$
end

โดยการทำซ้ำนี้จะกระทำจนกระทั่ง นอร์ม L2 ระหว่างรอบปัจจุบันต่างกับรอบก่อนหน้าไม่เกินค่า Tol ที่กำหนดไว้หรือจำนวนรอบการทำซ้ำมากจนถึงจุดสิ้นสุดที่เพียงพอที่จะให้ลู่เข้าซึ่งไม่ควรใหญ่เกินไปเพื่อไม่ให้เสียเวลาประมวลผลจนนานเกินไป

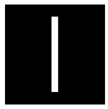
โดยผลการต่อเติมภาพจากทั้ง 3 วิธีข้างต้น สำหรับการกำหนดรอบการทำซ้ำไม่เกิน 1000 รอบและค่านอร์ม L2 ภาพปัจจุบันและภาพก่อนหน้าต่างกันไม่เกิน 0.0001 ได้ผลลัพธ์ดังนี้



(a) ภาพต้นฉบับ

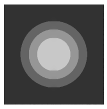


(b) ภาพที่ต้องการทำการต่อเติม

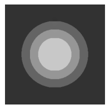


(c) โดเมนต่อเติม

(d) ภาพจากวิธี Explicit Time Marching ใช้เวลา 3.10 วินาที PSNR 19.6733



(e) ภาพจากวิธี Fixed Point ใช้เวลา 6.93 วินาที PSNR 42.6631



(f) ภาพจากวิธี Split Bergman ใช้เวลา 1.86 วินาที PSNR 44.4275

จากการทดลองจะเห็นว่าวิธีการ Split Bergman ได้ผลลัพธ์ที่ดีที่สุด จึงได้มีความสนใจที่จะศึกษาวิธี Split bergman เป็นลำดับถัดไป

1.2 วิธีการทางคณิตศาสตร์สำหรับการต่อเติมภาพด้วยการแปรผันบนภาพสี

สำหรับการต่อเติมภาพเชิงแปรผันซึ่งเป็นภาพสีในระบบ RGB จะพิจารณา

$$u : \Omega \to V^3$$

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

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

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

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

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

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

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

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

ปัญหาย่อย w โดยการคงค่า u ไว้จะได้ว่าปัญหาย่อยคือ

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

ซึ่งเมื่อทำการแก้ปัญหาย่อยนี้จะได้คำตอบที่แม่นตรงคือ

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

ปัญหาย่อย u โดยการคงค่า w ไว้จะได้ว่าปัญหาย่อยคือ

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

ซึ่งเมื่อทำการแก้ปัญหาย่อยนี้แล้วจะได้ระบบสมการออยเลอร์ลากรางจ์คือ

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

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

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

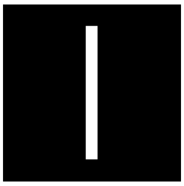
ซึ่งเมื่อแก้ระบบสมการดังกล่าวด้วยวิธี Gauss-seidel แล้วจะนำค่าที่ได้มาเปลี่ยนค่าตัวแปร b โดยที่

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

ซึ่งจากวิธีการ Split Bergman สำหรับภาพสี จะได้ภาพผลลัพธ์ดังนี้

(a) ภาพต้นฉบับ

(b) ภาพที่ต้องการทำการต่อเติม



(c) โดเมนต่อเติม

(d) ภาพจากวิธี Explicit Time Marching ใช้เวลา 6.72 วินาที PSNR 37.2374

1.3 การประยุกต์ใช้สำหรับต่อเติมภาพวาดศิลปะไทย

เนื่องจากการต่อเติมภาพด้วยการแปรผันจะสนใจที่ความต่อเนื่องของโครงสร้างทางเรขาคณิต ซึ่งวิธีการต่อเติมภาพรูปภาพด้วยการแปรผันมักจะให้ผลลัพธ์ได้ดีกับรูปภาพที่ราบเรียบเป็นช่วง (piecewise smooth image) [3] ซึ่งภาพวาดศิลปะไทยนั้น เป็นภาพซึ่งมีลักษณะราบเรียบเป็นช่วงจึงเหมาะสำหรับใช้แม่แบบวิธี ROF สำหรับการต่อเติมภาพ เพื่อช่วยให้จิตรกรผู้มีหน้าที่ซ่อมแซมภาพสามารถเห็นตัวอย่างภาพที่ผ่านการต่อเติม เพื่อให้จิตรกรสามารถวางแผนในการซ่อมแซมได้ หรือการเข้าชมพิพิธภัณฑ์ ก็อาจพัฒนาเป็นแอปพลิเคชันสำหรับโทรศัพท์มือถือ ที่สามารถใช้ส่องภาพงานศิลปะและแสดงศิลปะที่สมบูรณ์ขึ้นมาทางหน้าจอได้

1.4 กาประยุกต์ใช้สำหรับการลบคำบรรยายบนอนิเมะ

อนิเมะคือวิดีโอภาพวาดการ์ตูนสไตล์ญี่ปุ่น ซึ่งภาพวาดสไตล์ของอนิเมะเองก็

SAVE POINT

ซึ่งการซ่อมแซมรูปภาพมีวิธีการทางคณิตศาสตร์ที่ใช้แตกต่างกันไปจำนวนมาก แต่การซ่อมแซมด้วยการแปรผันจะสนใจที่ความต่อเนื่องของโครงสร้างทางเรขาคณิต ซึ่งวิธีการซ่อมแซมรูปภาพด้วยการแปรผันมักจะให้ผลลัพธ์ได้ดีกับพื้นที่แคบและเล็ก ในรูปภาพที่ราบเรียบเป็นช่วง (piecewise smooth image) หรือที่เรียกกันว่าภาพการ์ตูน เนื่องจากวิธีการนี้ไม่สามารถทำการสร้างพื้นผิว (Texture) ขึ้นมาได้ [3] โดยวิธีการแปรผันที่สนใจ จะใช้ตัวแบบ Rudin-Osher-Fatemi (ROF) [4]ซึ่งถูกนำเสนอในรูปเชิงแปรผัน (variational formulation) ไว้ดังนี้

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

เมื่อ $\mathcal{D}$ คือ พจน์สำหรับวัดค่าความเหมาะสมข้อมูล (Data fitting Term) λ คือ ตัวแปรจัดระเบียบ (Regularization parameter) $\mathcal{R}$ พจน์จัดระเบียบ (Regularization Term)

ซึ่งตัวแบบ ROF พจน์สำหรับวัดค่าความเหมาะสมข้อมูลได้หลายวิธี ซึ่งผู้ศึกษาสนใจที่จะใช้วิธีแปรผันรวม (Total Variation) [7] ในการแก้ตัวแบบนี้ โดยเป็นการแก้ปัญหาการแปรผันมีขอบเขต (bounded variation หรือ BV) ทั้งหมดโดยที่ภาพ u อยู่ใน $BV(\Omega)$ เมื่อสามารถหาปริพันธ์ได้และจะมี Radon measure Du ซึ่ง

$$\int_\Omega u(x) div \vec{g}(x) dx = \int_\Omega \langle \vec{g}, Du(x) \rangle \qquad \forall \vec{g} \in C_c^1(\Omega, \mathbb{R}^2)^2$$

และจาก Du เป็น distributional gradient ของ u เมื่อ u ราบเรียบแล้ว $Du(x) = \bigtriangledown u(x) dx$ โดย total variation seminorm ของ u คือ

$$||u||_{TV(\Omega)} := \int_{\Omega} |Du| := sup\Bigg\{ \int_{\Omega} u \; div \; \vec{g} \; dx \;\; : \vec{g} \; \in C^1_c(\Omega, \mathbb{R}^2)^2 \; , \; \sqrt{g_1^2 + g_2^2} \leq 1 \Bigg\}$$

จาก u ราบเรียบแล้ว การแปรผันรวมสมมูลกับอินทิกรัลของขนาดเกรเดียนท์

$$||u||_{TV(\Omega)} = \int_{\Omega} | \bigtriangledown u| dx$$

จึงได้ว่าจะหาฟังก์ชันแปรผันมีขอบเขต u หาได้จาก minimization problem

$$\underset{u \in BV(\Omega)}{arg \; min} ||u||_{TV(\Omega)} + \frac{\lambda}{2} \int_{\Omega \setminus D} (f(x) - u(x))^2 dx$$

เมื่อ λ มีค่าบวก ปัญหา minimalization นี้จะเหมือนกับปัญหาการลบสิ่งรบกวนของ Rudin, Osher และ Fatemi เพียงแต่ปริพันธ์ลำดับอยู่บน $\Omega - D$ แทนที่จะเป็น Ω ถ้าผลลัพธ์ที่แม่นตรงอยู่ใน BV และมีค่าอยู่ในช่วง $[0, 1]$ แล้วจะมี minimizer u แต่มักจะไม่มีเพียงหนึ่งเดียว

การซ่อมแซมรูปภาพอาจมองเป็นลักษณะการลบสิ่งรบกวนที่มี spatially-varying regularization strength เป็น $\lambda(x)$ ทำให้ได้ว่า

$$\underset{u}{arg \; min} ||u||_{TV(\Omega)} + \frac{1}{2} \int_{\Omega} \lambda(x) (f(x) - u(x))^2 dx$$

โดยที่ $\lambda(x)$ จะมีค่าเป็น 0 เมื่ออยู่ใน D และ $\lambda(x) > 0$ เมื่ออยู่นอก D ทำให้เมื่อ $x \in D$ ที่ $\lambda(x) = 0$ ค่า $f(x)$ จะไม่ถูกใช้ ทำให้ $u(x)$ ได้รับผลจาก $||u||_{tv}$ เท่านั้น ส่วนที่ด้านนอก D จะเป็น TV-regularize denoising พฤติกรรมลดสิ่งลบกวนนี้อาจเป็นที่น่าพอใจเมื่อยากที่จะระบุโดเมนที่ต้องซ่อมแซมได้อย่างถูกต้อง และเมื่อใช้ ขนาดใหญ่จะทำให้การลดสิ่งรบกวนมีผลน้อยมากจนทำให้พื้นที่นอก D แทบไม่เปลี่ยนแปลง

จากโมเดลทางคณิตศาสตร์ที่ได้กล่าวมาข้างต้น จะสามารถใช้วิธีการทางเชิงตัวเลขสำหรับการซ่อมแซมรูปภาพโดยใช้ความแปรปรวนทั้งหมดได้หลายวิธีการ จึงขอยกตัวอย่างวิธีการไทม์มาร์ชชิ่ง (time marching method) [5] ปีนวิธีการที่ง่ายและสะดวกในการแก้สมการเชิงอนุพันธ์ย่อยไม่เป็นเชิงเส้น แนวคิดของวิธีการนี้คือการแนะนำตัวแปรเวลาสังเคราะห์ (time artificial variable) จากนั้นหาคำตอบแบบสภาวะคงตัว (steady-state solution) ของสมการเชิงอนุพันธ์ย่อยไม่เป็นเชิงเส้นที่ขึ้นอยู่กับเวลา และเพื่อจะแก้ความไม่เป็นเชิงเส้นของสมการเชิงอนุพันธ์ย่อย จะสามารถใช้รูปแบบที่ชัดแจ้งของออยเลอร์ (Euler's explicit scheme) ที่กำหนดโดย

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

เมื่อ $\tau > 0$ แทนขั้นเวลา (time step) ที่ได้จากการดิสครีตไทซ์โดเมนเวลา $[0, \infty)$

หลังจากใช้การประมาณแบบไฟไนต์ดิฟเฟอเรนซ์ จะได้รูปแบบการทำซ้ำเป็น

$$(u^{[k+1]})_{i,j} = (u^{[k]})_{i,j} + \tau \left(\mathcal{K}(u^{[k]})_{i,j} + \lambda((u^{[k]})_{i,j} + (f)_{i,j}) \right)$$

เห็นได้ว่าวิธีการเชิงตัวเลขดังกล่าวข้างต้นนั้นง่ายในการคำนวณ แต่การลู่เข้าสู่คำตอบที่เหมาะสมของปัญหา เชิงแปรผันค่อนข้างช้ามากเนื่องจากต้องใช้ τ ที่มีขนาดเล็กในการทำให้ลำดับของคำตอบลู่เข้า

เนื่องจากวิธีไทม์มาร์ชชิ่งการลู่เข้าของคำตอบค่อนข้างช้า จึงมีอีกวิธีที่สามารถลู่เข้าสู่คำตอบได้ไวขึ้น นั่นคือวิธี Split Bergman [8] ซึ่งคือการแยกส่วนการดำเนินการ (splitting) และการทำซ้ำ bergman (bergman iteration)

จากความแปรปรวนทั้งหมดสามารถประมาณได้โดย $| \bigtriangledown u_{i,j}|$ บนทุกพิกเซลนั่นคือ

$$||u||_{TV(\Omega)} \approx \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} | \bigtriangledown u_{i,j}|$$

เมื่อ $\bigtriangledown u_{i,j}$ คือ discrete gradient วิธี split bergman จะนำมาใช้เพื่อแก้ minimization problem

$$\begin{cases} \underset{d,u}{arg\,min} \sum_{i,j} |d_{i,j}| + \frac{1}{2}\sum_{i,j} \lambda_{i,j}(f_{i,j} - u_{i,j})^2 \\ subject\ to\ d = \bigtriangledown u \end{cases}$$

โดยตัวแปรช่วย d คือเวคเตอร์ที่บีบบังคับ $\bigtriangledown u$ และใช้วิธีการทำซ้ำ bergman เพื่อแก้ปัญหาค่าเหมาะสมแบบมีข้อจำกัด ซึ่งในแต่ละการทำซ้ำ bergman จะเป็นการแก้

$$\underset{d,u}{arg\,min} \sum_{i,j} |d_{i,j}| + \frac{\lambda}{2}\sum_{i,j} \lambda_{i,j}(f_{i,j} - u_{i,j})^2 + \frac{\gamma}{2}\sum_{i,j} |d_{i,j} - \bigtriangledown u_{i,j} - b_{i,j}|^2$$

เมื่อ b เป็นตัวแปรของวิธีการทำซ้ำ bergman และ γ เป็นค่าคงที่บวกใดๆ โดยการ minimization บน d และ u จะแก้โดย alternative direction method โดยแต่ละขั้นของการหาค่าต่ำสุด ตัวแปร d และ u จะให้ตัวแปรอื่นคงค่าไว้

d subproblem เมื่อเราคงค่า u ไว้ จะได้ว่า d subproblem คือ

$$\underset{d}{arg\,min} \sum_{i,j} |d_{i,j}| + \frac{\gamma}{2}\sum_{i,j} |d_{i,j} - \bigtriangledown u_{i,j} - b_{i,j}|^2$$

โดยปัญหานี้เมื่อทำการแก้แล้วจะได้ว่า

$$d_{i,j} = \frac{\bigtriangledown u_{i,j} + b_{i,j}}{|\bigtriangledown u_{i,j} + b_{i,j}|} max\{|\bigtriangledown u_{i,j} + b_{i,j}| - \frac{1}{\gamma}, 0\}$$

u subproblem เมื่อเราคงค่า d ไว้ จะได้ว่า u subproblem คือ

$$\underset{u}{arg\,min} \frac{1}{2}\sum_{i,j} \lambda_{i,j}(f_{i,j} - u_{i,j})^2 + \frac{\gamma}{2}\sum_{i,j} |d_{i,j} - \bigtriangledown u_{i,j} - b_{i,j}|^2$$

เมื่อแก้แล้วจะได้ว่า

$$\frac{1}{\gamma} - \triangle u = \frac{1}{\gamma}\lambda f - div(d - b)$$

โดยที่ div คือ discrete divergence และ $\bigtriangledown u$ คือ discrete lapacian เราจะประมาณคำตอบนี้โดยการใช้ หนึ่งรอบ Gauss-seidel ต่อหนึ่งรอบการทำซ้ำของ Bergman ซึ่ง subproblem จะถูกแก้หนึ่งครั้ง ต่อหนึ่งรอบ bergman iteration แต่ทั้งนี้การทำซ้ำ Gauss-seidel หลายครั้ง จะทำให้การแก้ subproblem มีความแม่นยำขึ้น ส่วนตัวแปรช่วย b มีค่าเริ่มต้นเป็น 0 จากนั้นทำการปรับค่าโดย

$$b^{k+1} = b^k + \bigtriangledown u - d$$

โดยที่ความเกี่ยวข้องกันของแต่ละพื้นที่จะแรงขึ้นเมื่อ γ ใหญ่ขึ้น ดังนั้น γ ไม่ควรเล็กหรือใหญ่จนเกินไป จะทำให้ทั้งสอง subproblem ลู่เข้าได้ดี จึงได้ว่าวิธีการในภาพรวมเป็นดังนี้

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

โดยการทำซ้ำนี้จะกระทำจนกระทั่ง นอร์ม L2 ระหว่างรอบปัจจุบันต่างกับรอบก่อนหน้าไม่เกินค่า Tol ที่กำหนดไว้หรือจำนวนรอบการทำซ้ำมากจนถึงจุดสิ้นสุดที่เพียงพอที่จะให้ลู่เข้าซึ่งไม่ควรใหญ่เกินไปเพื่อไม่ให้เสียเวลาประมวลผลจนนานเกินไป

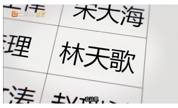
ซึ่งวิธีเชิงตัวเลขข้างต้นเป็นวิธีการสำหรับภาพเฉดเทา (Gray-scale) สำหรับการประยุกต์ใช้กับภาพสีนั้น ภาพสี จะประกอบขึ้นด้วยสี 3 สี คือ แดง เขียว น้ำเงิน ซึ่งเราสามารถใช้วิธีเชิงตัวเลขข้างต้น แยกสำหรับแต่ละสี เพื่อทำการซ่อมแซมภาพก่อนจะนำมารวมกลับเป็นภาพสีอีกครั้ง และสำหรับวิดีโอนั้นประกอบด้วยภาพจำนวนหลายภาพต่อหนึ่งหน่วยเวลา เราจะเรียกภาพหนึ่งภาพในวิดีโอว่า เฟรม (frame) ซึ่งเฟรมนี้เป็นภาพสี เราจึงสามารถแบ่งใช้โมเดลกับแต่ละสีและรวมกันกับมาเป็นวิดีโออีกครั้งได้ จึงสามารถซ่อมแซมวิดีโอได้ด้วย

ผู้พัฒนาจึงสนใจที่จะพัฒนาวิธีการเชิงตัวเลขเพื่อให้สามารถซ่อมแซมได้รวดเร็วยิ่งขึ้น เนื่องจากหากต้องการซ่อมแซมวิดีโอแบบเรียลไทม์ จะเป็นจะต้องสามารถซ่อมแซมได้เร็วถึง 25 ภาพต่อวินาที ซึ่งแต่ละภาพเป็นภาพสี ซึ่งวิธีการ Split Bergman หากทำให้ได้คุณภาพดีจะใช้เวลาที่นานขึ้น และหากทำให้เวลาทันสำหรับ 25 ภาพต่อวินาที ในแต่ละภาพจะยังไม่ถูกซ่อมแซม

ซึ่งวิธีการซ่อมแซมภาพด้วยการแปรผันรวมนั้น ใช้ในการซ่อมแซมภาพที่ราบเรียบเป็นช่วง โดยภาพจิตรกรรมฝาผนังนั้น เป็นหนึ่งในภาพที่ราบเรียบเป็นช่วง จึงเหมาะสมที่จะใช้วิธีการนี้ในการซ่อมแซม และเมื่อพัฒนาให้สามารถซ่อมแซมวิดีโอแบบเรียลไทม์ได้ อาจพัฒนาต่อยอดเป็นเทคโนโลยี augmented reality ที่สามารถยกโทรศัพท์มือถือขึ้นมาส่องยังภาพวาดฝาผนังที่ได้รับความเสียหายและแสดงภาพที่ถูกซ่อมแซมแล้วบนหน้าจอได้อย่างทันที

การจะซ่อมแซมภาพได้นั้น จำเป็นจะต้องมีการหาโดเมนซ่อมแซม (Inpaint Domain) ซึ่งสำหรับในการศึกษาครั้งนี้ ซึ่งโดยทั่วไป มักจะต้องให้ผู้ที่ใช้งานหาโดเมนซ่อมแซมเอง แต่โครงการวิจัยชิ้นนี้นอกจากจะสร้างวิธีการเชิงตัวเลขในการซ่อมแซมภาพแล้ว ยังได้เสนอวิธีการหาวิธีการหาบทบรรยายในวิดีโอแบบอนิเมะไว้อีกด้วย เพื่อทำให้การค้นหาโดเมนซ่อมแซมสำหรับไฟล์วิดีโอแบบอนิเมะเป็นไปได้อย่างอัตโนมัติ

โดยบทบรรยายของอนิเมะนั้น มักจะขึ้นบริเวณด้านล่างของหน้าจอ และนอกจากนี้ บทบรรยายอนิเมะมักจะใช้ขอบของตัวอักษรเป็นสีดำอีกด้วย ด้วยสมบัตินี้เองทำให้เราสามารถหาบริเวณบนเฟรมที่เป็นบทบรรยายได้โดยจะมีวิธีหาพื้นที่ซึ่งเป็นบทบรรยายดังนี้



(a) ภาพเฟรมอนิเมะที่มีบทบรรยาย

(b) ภาพหลังทำการตัดส่วนล่างและ thresholding

ตัดเฟรมมาเฉพาะส่วนล่างของเฟรมที่น่าจะมีบทบรรยายปรากฏอยู่ จากนั้นทำการ thresholding เพื่อหาบริเวณที่เป็นสีดำ เนื่องจากบทบรรยายจะถูกล้อมรอบด้วยสีดำเสมอ

(a) ภาพหลังทำการสลับสี

(b) ภาพหลังทำการเปลี่ยนพื้นที่สีขาว

ทำการสลับสีระหว่างสีดำกับสีขาวของภาพที่ทำการ thresholding หลังจากนั้นทำการเปลี่ยนพื้นที่สีขาวซึ่งติดกับขอบของเฟรมทั้งหมดให้เป็นสีดำ เพราะว่า บทบรรยายไม่อยู่ติดกับหน้าจอ เราจะถือว่าสิ่งที่อยู่ติดกับหน้าจอไม่ใช่บทบรรยาย

(a) ภาพหลังการ erode และ opening

(b) ภาพหลังการ dilate

จากนั้นนำวัตถุที่มีขนาดเล็กเกินไป หรือใหญ่เกินไปออกจากภาพด้วยวิธีการ erode และ opening จะได้ว่าส่วนที่เหลือเป็นสีขาวในภาพคือบทบรรยาย แต่ว่าขอบของบทบรรยายก็ต้องถูกลบออกไปด้วย จึงทำการ dilate เพื่อขยายขอบของบทบรรยายให้เท่ากับบทบรรยายที่อยู่ในเฟรมวิดีโอ และสิ่งที่เหลืออยู่คือโดเมนซ่อมแซม ที่จะนำไปใช้ในการซ่อมแซมภาพต่อไป

2 Objective

วัตถุประสงค์ของโครงการวิจัยมีดังต่อไปนี้

(1) ศึกษาวิธีการแปรผันและวิธีการเชิงตัวเลขที่มีประสิทธิภาพเพื่อเติมข้อมูลที่ขาดหายในภาพหรือวิดีโอ

(2) สร้างวิธีการเชิงตัวเลขใหม่สำหรับซ่อมแซมภาพศิลปะไทยและลบบทบรรยายออกจากอนิเมะ

(3) นำวิธีการที่สร้างขึ้นเพื่อซ่อมแซมภาพไทย และลบบทบรรยายอนิเมะ

3 Scope of Study

ขอบเขตของโครงงานมีดังต่อไปนี้

(3.1) ภาพศิลปะที่ใช้ศึกษา เป็นภาพจิตรกรรมฝาผนังไทย ที่อยู่ภายใต้เว็บ Wikipedia.org ซึ่งได้รับการอนุญาตให้ใช้งานแบบ Creative Commons หรือแบบ Public Domain

(3.2) วิดีโอที่ใช้ศึกษาเป็นวิดีโอประเภทอนิเมะ โดยศึกษากับไฟล์อนิเมะที่ใช้ Color space แบบ RGB เท่านั้น

(3.3) บทบรรยายที่ใช้ทดสอบ จะถูกล้อมรอบไว้ด้วยสีดำ ขนาดความหนาขนาดไม่น้อยกว่า 5 พิกเซล

(3.4) วิดีโอที่ใช้ศึกษาขนาดไม่เกิน 1920x1080

(3.5) คอมพิวเตอร์ที่ใช้ทดลองใช้หน่วยประมวลผล I7-6700HQ ใช้การ์ดจอ Nvidia GTX 960M แรม 16GB ฮาร์ดดิกส์แบบ SSD

4 Methodology

วิธีการมีดังต่อไปนี้

(4.1) ศึกษาการคณิตศาสตร์ต่อเติมข้อมูลที่ขาดหายบนรูปภาพ

(4.2) พัฒนาวิธีการเชิงตัวเลขสำหรับการซ่อมแซมรูปภาพ

(4.3) ทดสอบวิธีการเชิงตัวเลขที่พัฒนาขึ้นโดยโปรแกรมคอมพิวเตอร์บนภาพสังเคราะห์

(4.4) อภิปรายผลที่ได้จากการทดลองเชิงตัวเลข

(4.5) สรุปผลการดำเนินงานวิจัยและจัดทำรูปเล่มฉบับสมบูรณ์

5 Time Periods

แผนการดำเนินงานตลอดทั้งโครงการสามารถสรุปได้โดยย่อจากตารางต่อไปนี้

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

6 References

[1] T.F. Chan, J. Shen, “Mathematical models of local non-texture inpaintings”, SIAM Journal on Applied Mathematics, vol. 62, no. 3, pp. 1019–1043, 2001. http://www.jstor.org/stable/ 3061798

[2] Furht B., อ้างอิง 2561: Encyclopedia of Multimedia [จาก `https://doi.org/10.1007/0-387-30038-4_98`] สืบค้นเมื่อ 5 สิงหาคม 2561

[3] Işık Barış Fidaner, อ้างอิง 2561: A Survey on Variational Image Inpainting , Texture Synthesis and Image Completion [จาก `https://www.semanticscholar.org/paper/_/36f4d32ce45f72091510ab4d4d1cc3bf81ffe879`] สืบค้นเมื่อ 5 สิงหาคม 2561

[4] 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. `https://doi.org/10.1016/0167-2789(92)90242-F`

[5] A. Marquina และ S. Osher, อ้างอิง 2561: Explicit algorithms for a new time dependent model based on level set motion for nonlinear deblurring and noise removal [จาก `https://doi.org/10.1137/S1064827599351751`] สืบค้นเมื่อ 10 กันยายน 2561

[6] C.R. Vogel and M.E. Oman,”Iterative methods for total variation denoising”, SIAM J. Sci. Comput. 17, pp. 227-238, 1996.

[7] Pascal Getreuer, อ้างอิง 2561: Rudin-Osher-Fatemi Total Variation Denoising using Split Bregman , [จาก `https://doi.org/10.5201/ipol.2012.g-tvd`] สืบค้นเมื่อ 5 สิงหาคม 2561

[8] Pascal Getreuer, อ้างอิง 2561: Total Variation Inpainting using Split Bregman [จาก `https://doi.org/10.5201/ipol.2012.g-tvi`] สืบค้นเมื่อ 5 สิงหาคม 2561