

ขั้นตอนวิธีเชิงตัวเลขชนิดใหม่สำหรับการต่อเติมภาพที่ใช้การแปรผันรวมกับการประยุกต์
สำหรับซ่อมแซมภาพจิตรกรรมไทยโบราณและการลบบทบรรยายจากอนิเมะ

A new numerical algorithm for TV-based image inpainting with its applications
for restoring ancient Thai painting images and removing subtitles from animes

ภัคพล พงษ์ทวี

ภาควิชาคณิตศาสตร์
มหาวิทยาลัยศิลปากร

การนำเสนอความก้าวหน้าโครงการวิจัย
30 พฤศจิกายน 2561

ความก้าวหน้า

- การซ้อมแซมภาคศิลป์ไทย
- การลงทะเบียนนิเมะ

ตัวแบบการต่อเติมภาพเขตสีเทาที่ใช้การแปรผันรวม

$$\min_u \{ \mathcal{J}(u) = \frac{1}{2} \int_{\Omega} \lambda(u - z)^2 d\Omega + \int_{\Omega} |\nabla u| d\Omega \}$$

$$\lambda = \lambda(x) = \begin{cases} \lambda_0, & x \in \Omega \setminus D \\ 0, & x \in D \end{cases}$$

T.F. Chan and J. Shen , “Mathematical models of local non-texture inpaintings”, SIAM Journal on Applied Mathematics, vol. 62, no. 3, pp. 1019–1043, 2001.

ตัวแบบการต่อเติมภาพเอดสีเทาที่ใช้การแปรผันรวม (ต่อ)

$$\min_u \{ \mathcal{J}(u) = \frac{1}{2} \int_{\Omega} \lambda(u - z)^2 d\Omega + \int_{\Omega} |\nabla u| d\Omega \}$$



$$\begin{cases} -\nabla \cdot \left(\frac{\nabla u}{|\nabla u|} \right) + \lambda(u - z) = 0, & \mathbf{x} \in (1, n)^2 \\ \frac{\partial u}{\partial n} = 0, & x \in \partial\Omega \end{cases}$$

การเดินเวลาแบบชัดแจ้ง (explicit time marching)

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

$$u(\mathbf{x}, t_0) = z \quad t_k = t_0 + k\tau \ (\tau > 0) \quad t_0 = 0$$

$$u(\mathbf{x}, t_0), u(\mathbf{x}, t_1), u(\mathbf{x}, t_2), u(\mathbf{x}, t_3), \dots, u(\mathbf{x}, t^*)$$

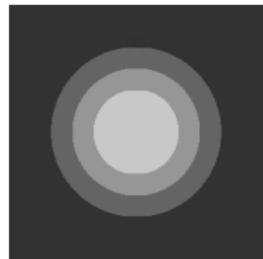
L. I. Rudin, S. Osher, E. Fatemi, "Nonlinear total variation based noise removal algorithms", Physica D: Nonlinear Phenomena, vol 60, issues 1–4, pp. 259–268, 1992.

การทำซ้ำแบบจุดตรึง (fixed-point iteration)

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

$u^{[0]}, u^{[1]}, u^{[2]}, u^{[3]}, \dots, u^*$

ปัญหาเชิงตัวเลข



รูปที่ 1: ตัวอย่างภาพที่เกิดปัญหาเชิงตัวเลข

$$\frac{1}{|\nabla u|} = \frac{1}{\sqrt{u_x^2 + u_y^2}} \rightarrow \infty$$

$$|\nabla u| \approx |\nabla u|_\beta = \sqrt{u_x^2 + u_y^2 + \beta}, \quad 0 < \beta \ll 1$$

วิธีการสปริทเบรกแม่น (Split Bregman method)

$$\min_{u,w} \{ \mathcal{J}(u, w) = \frac{1}{2} \int_{\Omega} \lambda(u - z)^2 d\Omega + \int_{\Omega} |\nabla w| d\Omega + \frac{\theta}{2} \int_{\Omega} (w - \nabla u + b) d\Omega \}$$

T. Goldstein and S. Osher, "The Split Bregman Method for L1-Regularized Problems", SIAM Journal on Imaging Sciences. vol. 2, issue 2, pp. 323-343, 2009.

วิธีการสปริทเบรคแม่น (ต่อ)

$$\min_{u,w} \{ \mathcal{J}(u, w) = \frac{1}{2} \int_{\Omega} \lambda(u - z)^2 d\Omega + \int_{\Omega} |\nabla w| d\Omega + \frac{\theta}{2} \int_{\Omega} (w - \nabla u + b) d\Omega \}$$



$$u^{\text{New}} = \arg \min_u \{ \mathcal{J}_1(u) = \frac{1}{2} \int_{\Omega} \lambda(u - z)^2 d\Omega + \frac{\theta}{2} \int_{\Omega} (w^{\text{old}} - \nabla u + b^{\text{old}}) d\Omega \}$$

$$w^{\text{New}} = \arg \min_w \{ \mathcal{J}_2(w) = \int_{\Omega} |\nabla w| d\Omega + \frac{\theta}{2} \int_{\Omega} (w - \nabla u^{\text{New}} + b^{\text{old}}) d\Omega \}$$

$$b^{\text{New}} = b^{\text{old}} + \nabla u^{\text{New}} - w^{\text{New}}$$

Peak Signal Noise Ratio (PSNR)

$$\text{PSNR} = 10 \cdot \log_{10}\left(\frac{1}{\sqrt{\text{MSE}}}\right)$$

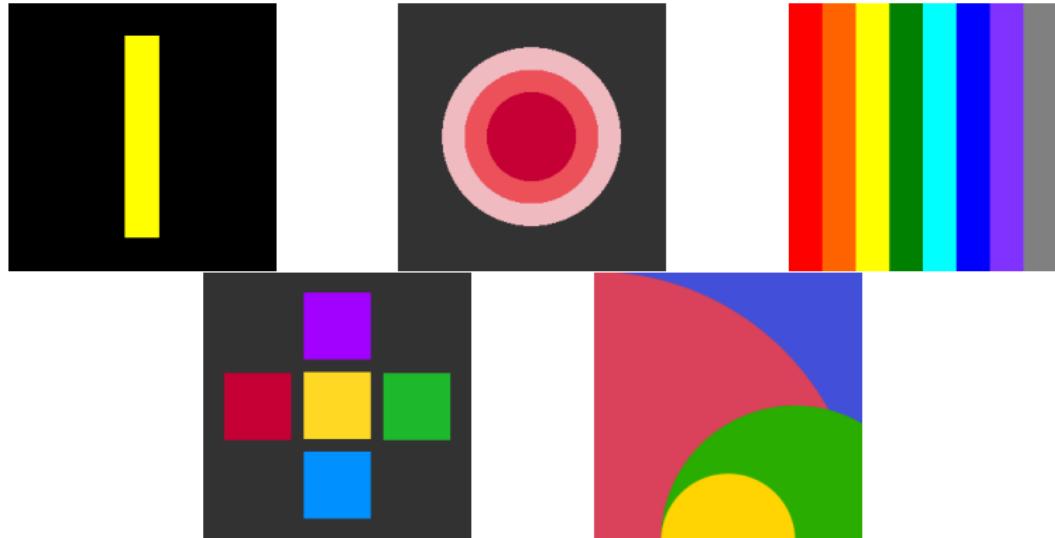
- MSE คือค่าคลาดเคลื่อนกำลังสองเฉลี่ยของภาพ โดยที่ $\text{MSE} = \frac{1}{nx \times ny} \sum (u - \tilde{u})^2$
- u แทนภาพต้นฉบับ
- \tilde{u} แทนภาพที่ได้จากการซ่อมแซมโดยวิธีเชิงตัวเลข

Structural Similarity (SSIM)

$$\text{SSIM}(u, \tilde{u}) = \frac{(2\mu_u\mu_{\tilde{u}} + 0.0001)(2\sigma_{u\tilde{u}} + 0.0009)}{(\mu_u^2 + \mu_{\tilde{u}}^2 + 0.0001)(\sigma_u^2 + \sigma_{\tilde{u}}^2 + 0.0009)}$$

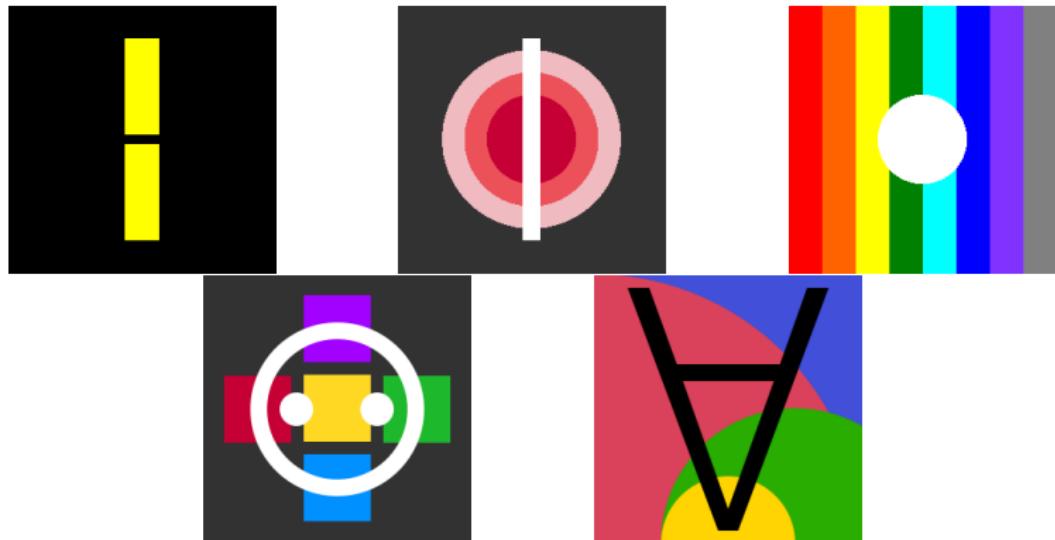
- u แทนภาพต้นฉบับ
- \tilde{u} แทนภาพต้นฉบับ และภาพที่ได้จากการซ่อมแซมโดยวิธีเชิงตัวเลข
- μ_u คือค่าเฉลี่ยของ u
- $\mu_{\tilde{u}}$ คือค่าเฉลี่ยของ \tilde{u}
- σ_u คือความแปรปรวนของ u
- $\sigma_{\tilde{u}}$ คือความแปรปรวนของ \tilde{u}

ภาพสังเคราะห์



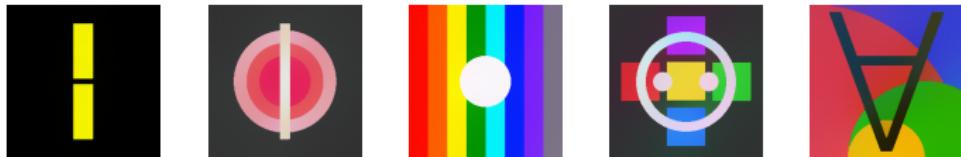
รูปที่ 2: ภาพต้นฉบับ

ภาพสังเคราะห์ที่จะซ่อมแซม

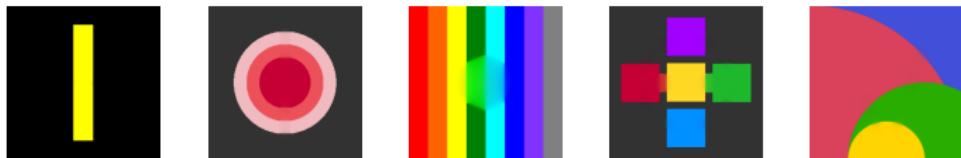


รูปที่ 3: ภาพที่จะทำการซ่อมแซม

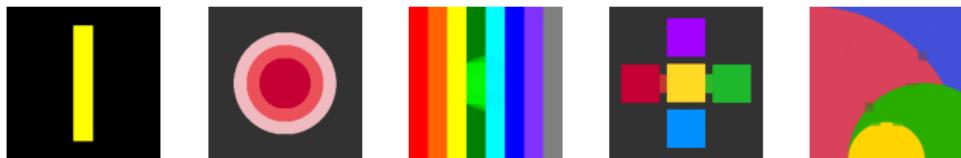
ผลการทดสอบ



รูปที่ 4: ผลการซ่อมแซมจากการวิจัยการเดินเวลา



รูปที่ 5: ผลการซ่อมแซมจากการวิจัยการทำเส้นทางแบบจุดต้อง



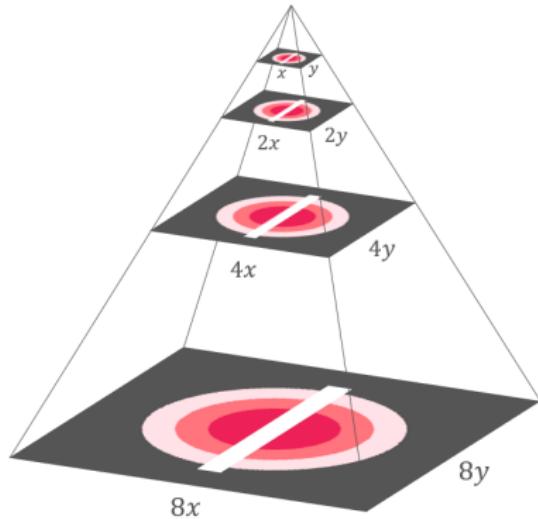
รูปที่ 6: ผลการซ่อมแซมจากการวิจัยการสปริทเบรกแม่น

ประสิทธิภาพของวิธีการเชิงตัวเลขทั้ง 3 วิธี

วิธีการ	เวลาประมาณ (วินาที)	PSNR (dB)	SSIM
การเดินเวลา	120.68	16.72	0.9960
การทำซ้ำจุดตرجิ	74.81	38.67	0.9999
การสปริทเบรกแม่น	14.06	39.42	0.9999

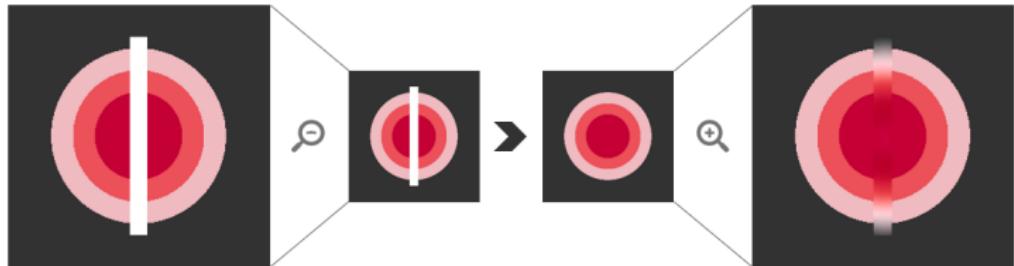
ตารางที่ 1: แสดงการซ่อมแซมเฉลี่ยของวิธีการเชิงตัวเลข

พีระมิดรูปภาพ



รูปที่ 7: วิธีการพีระมิดรูปภาพ

พีระมิดรูปภาพ (ต่อ)



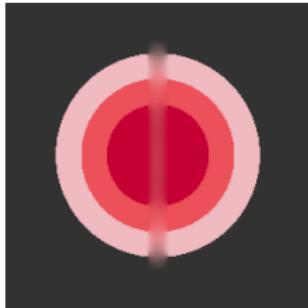
รูปที่ 8: วิธีการพีระมิดรูปภาพ

ผลการซ่อมแซมเมื่อใช้พิระมิดรูปภาพ

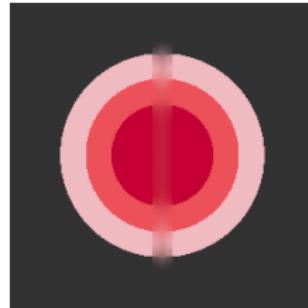
รูปแบบการทำซ้ำ	เวลาประมาณ (วินาที)	PSNR (dB)	SSIM
ไม่ใช้พิระมิดรูปภาพ	17.38	39.42	0.9999
10/1/1/10000	13.52	39.38	0.9999
10/3/3/10000	11.86	39.54	0.9999
10/10/10/10000	9.26	40.17	0.9999
100/1/1/10000	10.28	39.04	0.9999
100/3/3/10000	10.28	39.80	0.9999
100/10/10/10000	9.27	40.12	0.9999

ตารางที่ 2: ผลการซ่อมแซมภาพโดยวิธีการเชิงตัวเลขที่นำเสนอในรูปของค่าเฉลี่ย

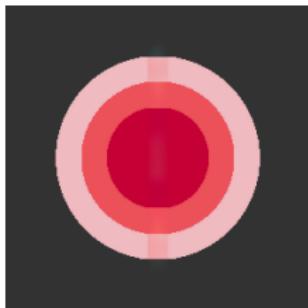
การทำซ้ำความละเอียดคอมพิวเตอร์สุด



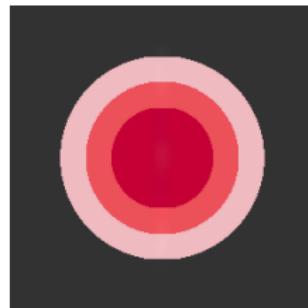
(a) 5 ครั้ง



(b) 10 ครั้ง



(c) 50 ครั้ง



(d) 100 ครั้ง

รูปที่ 9: พีระมิดที่ลำดับการทำซ้ำเป็น 10/10/10 และที่ระดับความคมชัดละเอียดสุดใช้จำนวนการทำซ้ำที่ต่างกัน

ผลการซ่อมแซมภาพในระดับความคมชัดละเอียดสุด 10 ครั้ง

รูปแบบการทำซ้ำ	เวลาประมาณ (วินาที)	PSNR (dB)	SSIM
ไม่ใช้พรมมิตรรูปภาพ	0.37	17.26	0.9963
10/1/1/10	0.40	28.54	0.9993
10/3/3/10	0.33	29.83	0.9994
10/10/10/10	0.38	32.56	0.9995
100/1/1/10	0.34	31.50	0.9999
100/3/3/10	0.36	31.99	0.9999
100/10/10/10	0.38	33.39	0.9998

ตารางที่ 3: ผลการซ่อมแซมภาพเฉลี่ยเมื่อใช้การทำซ้ำในระดับความคมชัดละเอียดสุด 10 ครั้ง

การซ่อมแซมภาพจิตรกรรมไทยโบราณ



(a) วัดแก้วไพทุรย์



(b) วัดแก้วไพทุรย์



(c) วัดพระยืนพุทธบาทยุคล



(d) วัดคงคaram



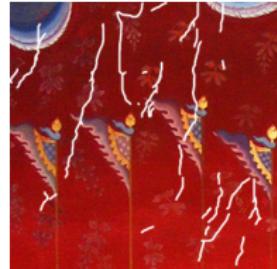
(e) วัดท่าถนน



(f) รอยความเสียหาย

รูปที่ 10: ภาพต้นฉบับสำหรับใช้ในการทดสอบ

ภาพที่ทำให้เสียหาย



รูปที่ 11: ภาพที่ทำให้เสียหาย

ผลการซ่อมแซมโดยวิธีการสปริทเบรกแม่น



รูปที่ 12: ผลการซ่อมแซมโดยวิธีการสปริทเบรกแม่น

ผลการซ่อมแซมโดยวิธีการที่พัฒนาขึ้น



รูปที่ 13: ผลการซ่อมแซมโดยวิธีการที่พัฒนาขึ้น

เปรียบเทียบระหว่าง 2 วิธี

ผลการซ่อมแซมภาพศิลปะไทย

วิธีการ	เวลาประมาณ (วินาที)	PSNR (dB)	SSIM
สเปรย์เบรกเมน	2.72	34.89	1.0000
วิธีการที่พัฒนาขึ้น	0.39	35.30	1.0000

ตารางที่ 4: แสดงผลการซ่อมแซมภาพศิลปะไทยในรูปค่าเฉลี่ย

การลับบทบรรยายบนอนิเมะ



รูปที่ 14: Festival Asia Special Video - feat. Inori Aizawa

Lorem Ipsum

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat.

But I must explain to you how all this mistaken idea of denouncing of a pleasure and praising pain was born and I will not give you a complete account of the system, and expound the actual teachings of the great explorer of the truth, the master-builder of human happiness. No one rejects, dislikes, or avoids pleasure itself, because it is pleasure

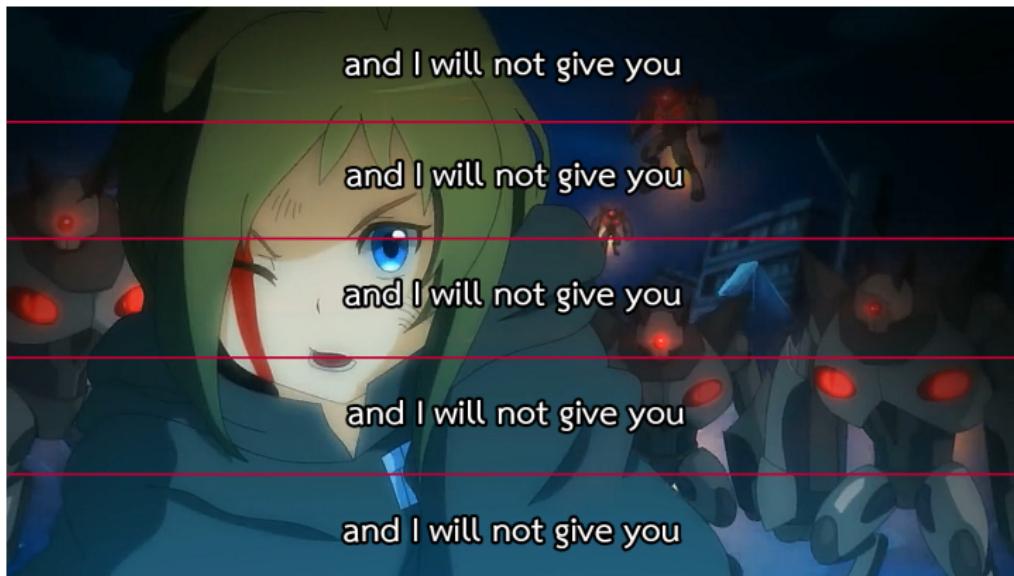
しかし私は、喜びを非難して苦痛を賞賛するという誤ったこの考えがすべてどのようにして誕生したかをあなたに説明しなければならないから、私はあなたにその体系を完璧に説明し、真実を求める偉大な探究家、人間の喜びを築く建築家の実践的な教えを詳しく説明しよう

หากแต่ข้าพเจ้าต้องขออภัยต่อท่านในเหตุที่มิโนคิดอันความชิงชักอันหลงผิดทั้งหลาย ในการไฟห้าแฉะสุดีบัดແພດ
ว่าบ่มเพาะขึ้นได้อย่างไรแล้ว ข้าพเจ้าจักให้สาระแห่งระบบอันครอบถ้วนแด่ท่าน และประสิทธิ์คำสอนที่แท้แห่งการแสวงหาอัน
ยิ่งใหญ่ในความจริง ซึ่งเป็นบูรพคณาจารย์ผู้สร้างร่างความมาสุขแห่งมวลมนุษย์ ทุกจะมีผู้ได้มิรับ มิชอบ หรือเลี้ยงหลีก
ความสุข ด้วยเพราเวณเป็นสุขก็หาไม่

รูปที่ 15: Lorem Ipsum ทั้ง 4 ภาษา¹

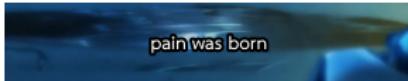
¹ ส่วนหนึ่งของ Lorem ipsum ในภาษาต่างๆ จาก https://en.wikipedia.org/wiki/Lorem_ipsum, https://ja.wikipedia.org/wiki/Lorem_ipsum และ <https://th.wikipedia.org/wiki/ລອມເມື່ອສັນເກີດ> สืบค้นเมื่อ 25 พฤษภาคม 2561

แบ่งเป็น 5 ส่วน

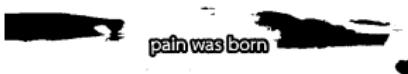


รูปที่ 16: แบ่งวิดีโอออกเป็น 5 ส่วน

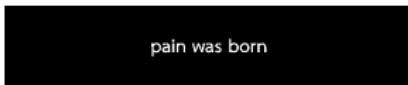
การหาบทบรรยาย



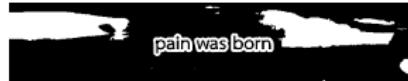
(a) ภาพเฟรมอนิเมะบริเวณที่มีบทบรรยาย



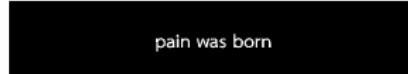
(c) สลับสีขาวเป็นสีดำ สีดำเป็นสีขาว



(e) ลบสิ่งที่เล็กเกินไปหรือใหญ่เกินไป



(b) ให้สีดำในภาพเป็นสีขาว สีอื่นเป็นสีดำ



(d) ภาพหลังทำการเปลี่ยนพื้นที่สีขาว



(f) ขยายโดเมนต่อเติมให้ครอบคลุม

รูปที่ 17: ขั้นตอนวิธีการหาบทบรรยาย

ผลลัพธ์การ hababrery

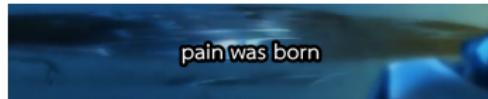
ภาษา	จำนวนพิกเซลในโดเมน	จำนวนพิกเซลที่ตัวตรวจพบ	จำนวนพิกเซลที่ผิดพลาด	ร้อยละการผิดพลาด
ไทย	23,222,220	24,083,125	2,141,201	9.22
อังกฤษ	27,278,745	28,598,424	3,714,321	13.62
ญี่ปุ่น	28,544,173	30,103,466	3,740,971	13.11

ตารางที่ 5: ความคลาดเคลื่อนเฉลี่ยของการหาโดเมนต่อitem ในบทบรรยายภาษาต่างๆ

การข้ามเฟรมและยึดเฟรม



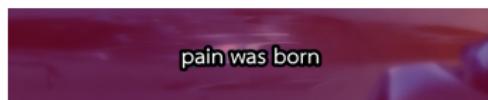
(a) เฟรมก่อนหน้า



(b) เฟรมปัจจุบัน



(c) บริเวณคำนวณ SSIM ของเฟรมก่อนหน้า



(d) บริเวณคำนวณ SSIM ของเฟรมปัจจุบัน

รูปที่ 18: บริเวณที่คำนวณ SSIM สำหรับการยึดเฟรมและข้ามเฟรม

ผลการลบโดยใช้วิธีข้ามเฟรมและยึดเฟรม

วิธีการ	เวลาประมาณ (วินาที)	PSNR (dB)	SSIM
สปริทเบรคแม่นและพิรัมมิครูปภาพ	141.29	31.39	0.9510
ยึดเฟรม	132.78	32.20	0.9655
ข้ามเฟรม	89.29	29.07	0.9408
ยึดเฟรมและข้ามเฟรม	75.76	29.33	0.9454

ตารางที่ 6: การลบคำบรรยายเฉลี่ยด้วยวิธีต่างๆ

ผลการลบคำบรรยาย



(a) ก่อนลบคำบรรยาย



(b) หลังลบคำบรรยาย

รูปที่ 19: ก่อนลบคำบรรยายและหลังลบคำบรรยาย

ผลการลบคำบรรยาย (ต่อ)

วิธีการ	เวลาประมาณ (วินาที)	PSNR (dB)	SSIM
สปริทเบรกแม่น	*	*	*
วิธีการที่พัฒนาขึ้น	75.76	29.33	0.9454

ตารางที่ 7: ผลการลบบทบรรยายออกจากอนิเมะโดยวิธีการสปริทเบรกแม่นและวิธีการที่พัฒนาขึ้น

วิธีการดำเนินงาน

แผนการดำเนินงาน	เดือนที่											
	1	2	3	4	5	6	7	8	9	10	11	12
ศึกษาตัวแบบและขั้นตอนวิธีการต่อเติมภาพที่ใช้การแปรผันรวมในเชิงลึก	x	x										
พัฒนาขั้นตอนวิธีสำหรับการต่อเติมภาพที่ใช้การแปรผันรวมชนิดใหม่			x	x	x	x						
ทดสอบขั้นตอนวิธีการต่อเติมภาพที่พัฒนาขึ้นโดยโปรแกรม-คอมพิวเตอร์บนภาพสังเคราะห์และภาพจริง				x	x	x						
อภิปรายผลที่ได้จากการทดลองเชิงตัวเลข					x	x	x					
สรุปผลการดำเนินงานวิจัยและจัดทำรูปเล่มฉบับสมบูรณ์								x	x	x	x	

ขอขอบคุณ