บันทึกวิจัย · พฤศจิกายน 2026

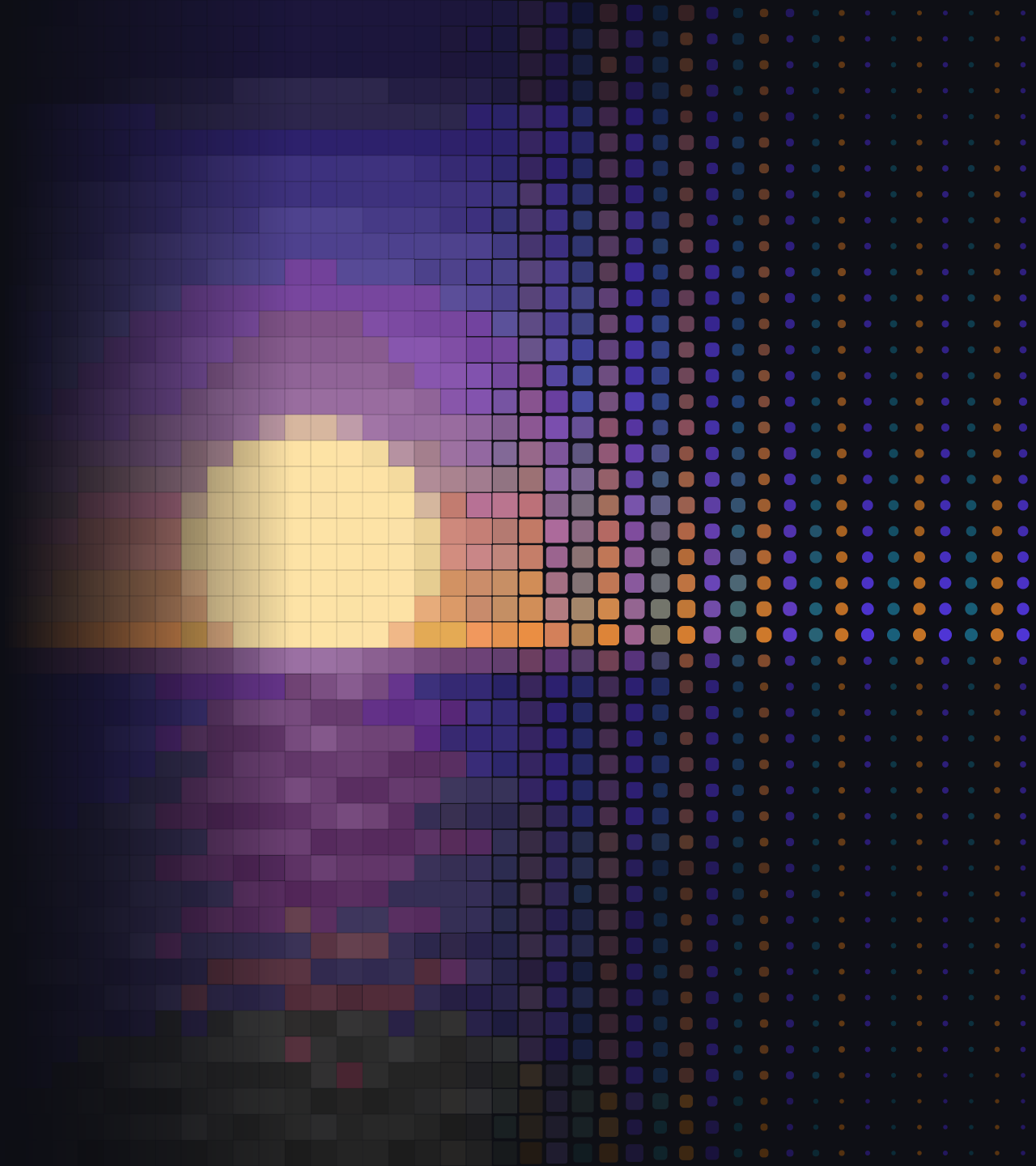
การเข้ารหัสแบบนิวรัล ในปี 2026

รอบโคเดก ไม่ใช่แทนโคเดก

การเล่นวิดีโอ · การถ่ายภาพ · การโทร · Generative AI · เกม · ซิลิคอน

ทัน ชว่าย

tanshuai.com · i@tanshuai.com



ข้อสรุป

เป็นกระแสหลักแล้ว แต่อยู่รอบโคเดก

01 โครงข่ายประสาทเทียมเป็นกระแสหลักแล้วที่ทั้งสองปลายของโคเดกมาตรฐาน

ISP ของกล้อง การลดขนาดภาพฝั่งเข้ารหัส และ super-resolution ฝั่งเล่นวิดีโอ บิตสตรีมตรงกลางยังเป็น AV1, HEVC, AVC หรือ VVC ที่ถอดรหัสด้วยฮาร์ดแวร์ฟังก์ชันตายตัว

02 โคเดกนิวรัลแบบ end-to-end ใช้งานจริงเฉพาะในวงปิด

การโทร tokenizer ของ Generative AI และเกม AAA หนึ่งเกม: บริษัทเดียวควบคุมทั้งสองฝั่ง และถอยกลับไปใช้โคเดกมาตรฐานได้

03 บิตสตรีมนิวรัลบนจอกระแสหลักต้องรอ H.267

เวอร์ชัน 1 มีกำหนดสิ้นปี 2029 ชิปรุ่นแรกราวปี 2031 และใช้ในกระแสหลักราวกลางทศวรรษ 2030 ประมาณการ

≈75%

ของทราฟฟิกข้อมูลมือถือเป็นวิดีโอ (สิ้นปี 2025)

1.26 พันล้าน

สมาร์ตโฟนที่จัดส่งในปี 2025 ทุกเครื่องถอดรหัสวิดีโอด้วยฮาร์ดแวร์ฟังก์ชันตายตัว

≈12×

พลังงานต่อพิกเซลของการออกแบบชิปถอดรหัสนิวรัลเพียงแบบเดียวที่มีการเผยแพร่ เทียบกับตัวถอดรหัส HEVC แบบฮาร์ดแวร์

สามทางที่โครงข่ายประสาทเทียมเข้าสู่โคเดก



แบบรอบโคเดกไม่ต้องขออนุญาตใคร แบบในมาตรฐานต้องมีมาตรฐาน แบบแทนโคเดกต้องควบคุมทั้งสองฝั่ง

การเข้ารหัสแบบนิวรัลใช้งานจริงที่ไหนในปี 2026

	รอบโคเดกมาตรฐาน ประมวลผลล่วงหน้า ปรับปรุงภาพ อัปสเกล	โคเดกนิวรัลแบบ end-to-end ตัวเข้ารหัสและตัวถอดรหัสแบบเรียนรู้	ในมาตรฐาน เครื่องมือนิวรัลในบิตสตรีมหรือ API มาตรฐาน
01 **การเล่นวิดีโอและสตรีมมิง**	● ใช้ในวงกว้าง ตัวลดขนาดภาพของ Netflix; super-resolution ของ YouTube และ Meta; ตัวอัปสเกล AI ในทีวี	● ยังไม่ใช้งาน ไม่มีบริการสตรีมมิงใดส่งวิดีโอในฟอร์แมตนิวรัล ทีวีและโทรศัพท์ถอดรหัสโคเดกมาตรฐานด้วยฮาร์ดแวร์	● กำลังจัดทำมาตรฐาน H.267 เวอร์ชัน 1 มีกำหนดสิ้นปี 2029 เครื่องมือนิวรัลอยู่ในกลุ่มตัวเลือก
02 **การถ่ายภาพและภาพนิ่ง**	● ใช้ในวงกว้าง AI ISP ในโทรศัพท์เรือธง; ISP นิวรัลเต็มรูปแบบเปิดตัวในรูป IP ที่ CES 2026	● สาธิต นำร่อง JPEG AI เป็นมาตรฐาน ISO: มีการสาธิตบนโทรศัพท์ ยังไม่มีผลิตภัณฑ์	● ยังไม่ใช้งาน ฟอร์แมตที่กำลังได้รับความนิยมเป็นแบบดั้งเดิม: Chrome 155 เปิด JPEG XL เป็นค่าเริ่มต้น
03 **การโทรและการประชุม**	● ใช้ในวงกว้าง super-resolution ฝั่งผู้รับ และการลดเสียงรบกวนแบบนิวรัล	● ใช้ในวงกว้าง เสียงแบบนิวรัลในการโทรสองฝ่ายบน Teams ตั้งแต่ปี 2021; วิดีโอ MLVC กำลังทยอยใช้ใน Teams	● ใช้งานจริงแบบจำกัด Opus 1.5 เพิ่มเครื่องมือ ML ที่เลือกใช้ได้และเข้ากันได้ย้อนหลัง
04 **Generative AI**	—	● ใช้ในวงกว้าง โมเดลวิดีโอและเสียงพูดทุกตัวใช้ tokenizer แบบเรียนรู้ เช่น VAE วิดีโอที่อัตรา 1:192	—
05 **เกมและกราฟิก**	● ใช้ในวงกว้าง การอัปสเกลและการสร้างเฟรมแบบนิวรัล: เฉพาะ DLSS 4 ก็อยู่ในเกมและแอปกว่า 250 รายการ	● ใช้งานจริงแบบจำกัด เท็กซ์เจอร์นิวรัลใช้จริงแล้วในเกม AAA หนึ่งเกม; SDK ของ NVIDIA ยังเป็นเบตา	● กำลังจัดทำมาตรฐาน DirectX กำลังเพิ่มการคำนวณเมทริกซ์ให้เชเดอร์ (พรีวิว)
06 **ลิงก์เฉพาะทาง: ดาวเทียม หุ่นยนต์**	● สาธิต นำร่อง การฟื้นฟูภาพจากฟีดบิตเรตต่ำฝั่งเซิร์ฟเวอร์	● สาธิต นำร่อง โคเดกนิวรัลที่ทนต่อการสูญหายของข้อมูล อยู่ในขั้นวิจัยและนำร่อง	—

เรียงแถวตามจำนวนคนที่แต่ละด้านเข้าถึง

นิวรัลที่สองปลาย มาตรฐานตรงกลาง

- **ที่ตัวเข้ารหัส**

  ตัวลดขนาดภาพแบบนิวรัลของ Netflix ใช้งานมาตั้งแต่ปี 2022 โดยไม่ต้องเปลี่ยนอะไรบนอุปกรณ์ใดเลย ผู้ชมราว 77% ชอบผลลัพธ์นี้มากกว่าในการทดสอบของ Netflix

  ผู้ขายอ้าง

- **ที่หน้าจอ**

  YouTube Super Resolution บนทีวีตั้งแต่ตุลาคม 2025 ส่วน Meta รัน super-resolution ทั้งบนเซิร์ฟเวอร์และบนโทรศัพท์

- **ในทีวี**

  ชิปประมวลผล NQ8 AI Gen3 ของ Samsung โฆษณาว่ามีโครงข่ายประสาทเทียม 768 ชุดสำหรับการอัปสเกล

บิตสตรีมยังเป็นมาตรฐาน

30%

ของการรับชม Netflix เป็น AV1 (ธันวาคม 2025)

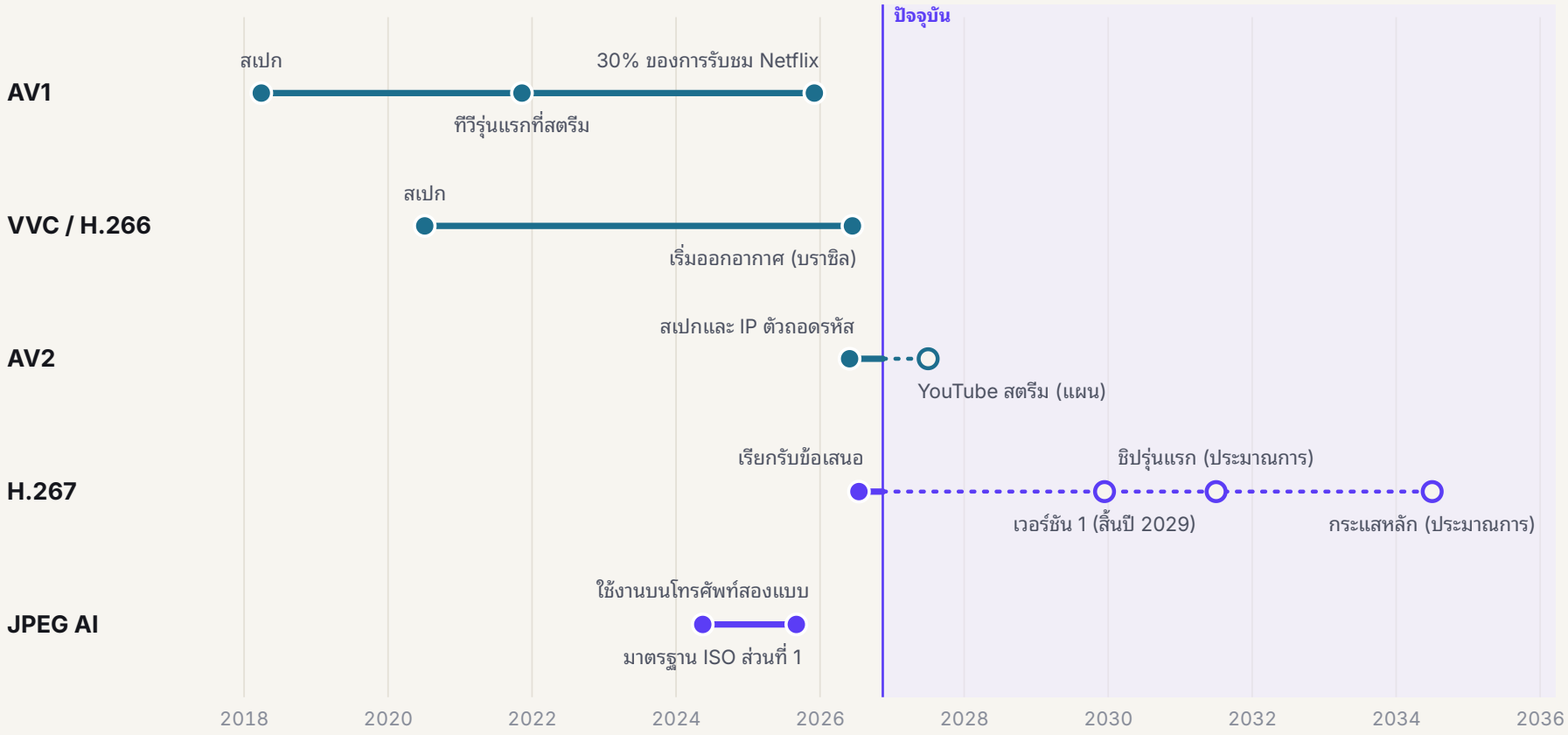
90%+

ของคลังวิดีโอ YouTube เมื่อถ่วงน้ำหนักด้วยเวลารับชม มีให้บริการใน AV1 (มีนาคม 2026)

3 สัปดาห์

จาก AV2 เวอร์ชัน 1.0.0 (28 พฤษภาคม 2026) ถึงการมีลูกค้า IP ตัวถอดรหัสของผู้ขายสามราย

จากมาตรฐานที่เสร็จสมบูรณ์ถึงจอกระแสหลักใช้เวลาเกือบทศวรรษ



AV1 ใช้เวลาเกือบแปดปีจากสเปก (2018) จนถึง 30% ของการรับชม Netflix (ธันวาคม 2025) เครื่องหมายแบบกลวงคือเป้าหมายหรือค่าประมาณการของผม

H.267 เป็นตัวตัดสินว่าเครื่องมือนิวรัลจะไปถึงทุกหน้าจอหรือไม่

ตุลาคม 2025
ประกาศเรียกรับหลักฐาน: ผลงานห้าชิ้นจาก 16 บริษัทและห้องแล็บประโยชน์ของนิวรัลชัดที่สุดกับ 8K HDR

กรกฎาคม 2026
ออกประกาศเรียกรับข้อเสนอที่เจนีวา

มกราคม 2027
ประเมินข้อเสนอที่บริสเบน

ครึ่งหลังของปี 2027
โมเดลทดสอบแรก

สิ้นปี 2029
มาตรฐานเวอร์ชัน 1

≈28%
ที่แนวทาง ECM แบบดั้งเดิมประหยัดได้เมื่อเทียบกับ VVC

6–14%
ที่ฟิลเตอร์ในลูปแบบนิวรัลประหยัดได้ ที่ 10,000 ถึงกว่า 22,000 MAC ต่อพิกเซล

SEI
สตรีม VVC ระบุฟิลเตอร์หลังถอดรหัสแบบนิวรัลให้ตัวถอดรหัสนำไปใช้ได้แล้ว

มุมมองของผม: แกนแบบดั้งเดิมที่มีเครื่องมือนิวรัลไม่กี่ตัว เฉพาะในจุดที่บิตที่ประหยัดได้คุ้มกับต้นทุนของตัวถอดรหัส

บทวิเคราะห์

นิวรัลในกล้อง แบบดั้งเดิมบนเว็บ

- **กล้อง**

  Cognitive ISP ของ Qualcomm ทำการแบ่งส่วนภาพแบบเรียลไทม์ในโทรศัพท์มาตั้งแต่ปี 2023 และมีการจัดแสดง AI ISP เต็มรูปแบบที่รันบน NPU ที่ CES 2026

- **JPEG AI**

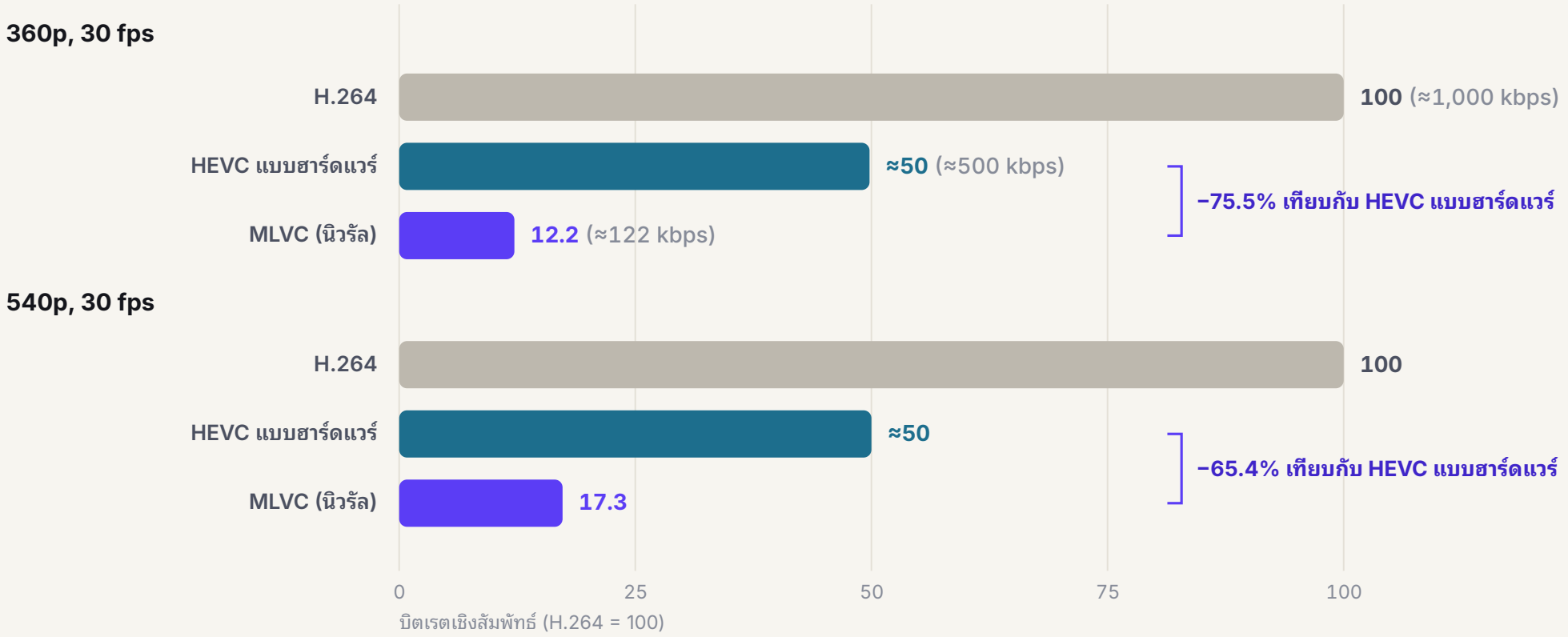
  เผยแพร่ในปี 2025: ดีกว่า VVC แบบ intra 12.5–27.9% ที่ 7,000 ถึง 216,000 MAC ต่อพิกเซล ผมไม่พบผลิตภัณฑ์ใดที่ใช้เป็นค่าเริ่มต้น

- **เว็บ**

  Chrome 155 เปิดใช้ JPEG XL เป็นค่าเริ่มต้นเมื่อ 6 ตุลาคม 2026 ด้วยตัวถอดรหัสที่ปลอดภัยด้านหน่วยความจำซึ่งเขียนด้วย Rust

> **ฟอร์แมตที่ต้องเปิดได้ทุกที่ยังเป็นแบบดั้งเดิม การเข้ารหัสแบบนิวรัลชนะภายในไปป์ไลน์ของบริษัทเดียว**

การโทรคือที่ที่โคเดกนิวรัลแบบ end-to-end ใช้งานจริง



Satin ใช้ในการโทรสองฝ่ายทั้งหมดบน Teams และ Skype มาตั้งแต่ปี 2021 MLVC ซึ่งเป็นโอเพนซอร์สภายใต้ MIT ตั้งแต่ 24 กรกฎาคม 2026 กำลังทยอยใช้ในการโทรแบบ peer-to-peer บน Teams ที่สูงสุด 540p30 โดยใช้ NPU ไม่เกินครึ่งหนึ่ง ผู้ขายอ้าง

ไม่ใช่ทุกแพลตฟอร์มที่หันมาใช้นิวรัล และเหตุผลก็ให้บทเรียนที่ดี

ในมาตรฐาน แบบเลือกใช้ได้

Opus 1.5 และ 1.6

การปรับปรุงเสียงแบบนิวรัล การปกปิดการสูญหายของแพ็กเก็ต ข้อมูลซ้ำซ้อนเชิงลึก และการขยายแบนด์วิดท์ ทั้งหมดปิดไว้เป็นค่าเริ่มต้น และเข้ากันได้เต็มที่กับมาตรฐานปี 2012

แบบดั้งเดิม โดยตั้งใจ

Meta MLow (2024)

การโทรของ Meta กว่า 20% ทำบนโทรศัพท์ ARMv7 และการโทรผ่าน WhatsApp หลายสิบล้านครั้งต่อวันทำบนโทรศัพท์อายุเกินสิบปี MLow ให้คุณภาพราวสองเท่าของ Opus ที่ 6 kbps โดยใช้การประมวลผลน้อยกว่า 10%

งบความซับซ้อนของฐานอุปกรณ์ที่ใช้งานอยู่คือตัวตัดสิน โคเดกนิวรัลใช้งานจริงก่อนในที่ที่อุปกรณ์เป็นรุ่นใหม่และผู้ขายรายเดียวควบคุมทั้งสองฝั่ง

การใช้งานที่ใหญ่ที่สุดซึ่งไม่มีใครเรียกว่าโคเดก

- **โคเดกแบบเรียนรู้อยู่ทุกที่**

  โมเดลเสียงพูดและวิดีโอทุกตัวอ่านและเขียนโทเค็นหรือค่า latent ของโคเดกแบบเรียนรู้

- **ไม่ใช่ฟอร์แมตสำหรับเผยแพร่**

  แต่ละโมเดลมี latent space ของตัวเอง: ไม่มีการควบคุมบิตเรต ไม่มีการทดสอบความสอดคล้อง ไม่มีตัวถอดรหัสตัวที่สอง

- **รายการต้นทุน**

  สำหรับผู้ให้บริการวิดีโอจาก Generative AI ค่า latent ที่บีบอัดได้มากขึ้นช่วยลดต้นทุนการสร้างได้มากกว่าสัดส่วน บทวิเคราะห์

สองตัวอย่าง

12.5 Hz · 1.1 kbps

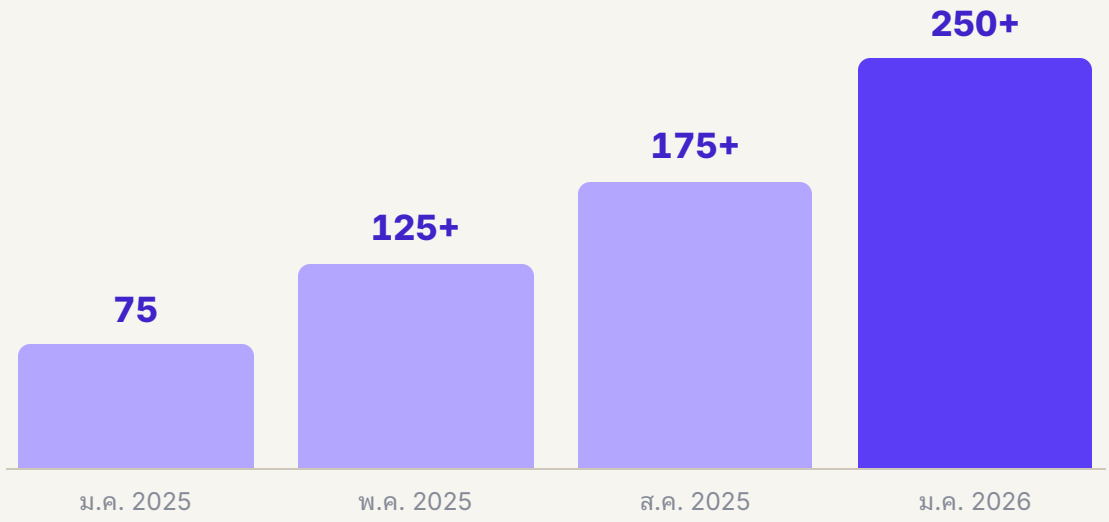
โทเค็นเสียงพูด Mimi ของ Kyutai ต่ำกว่าโคเดกโทรศัพท์ใด ๆ

1:192

VAE วิดีโอของ LTX-Video: โทเค็น latent หนึ่งตัวแทน 32×32 พิกเซลตลอด 8 เฟรม

การสร้างภาพขึ้นใหม่แบบนิวรัลเป็นกระแสหลัก แอสเซตนิวรัลกำลังเริ่มต้น

DLSS 4 Multi Frame Generation: เกมและแอป



การบีบอัดเท็กซ์เจอร์แบบนิวรัล

ฉากสาธิตของ NVIDIA, GTC 2026

หน่วยความจำเท็กซ์เจอร์ 6.5 GB → 0.97 GB

หน่วยความจำลดลง 85%

Assassin's Creed Mirage วางจำหน่ายแล้ว

เฉพาะแอสเซตที่เลือก; เทียบกับฟอร์แมตแบบบล็อก

บีบอัดได้ดีกว่า ≈30%

วัดต่างกัน: เทียบกันโดยตรงไม่ได้

DirectX Linear Algebra เข้าสู่ public preview ในเดือนเมษายน 2026: เส้นทางโค้ดเดียวสำหรับการถอดรหัสแบบนิวรัลภายในเชเดอร์บน GPU ของผู้ผลิตรายใดก็ได้

ดาวเทียม โดรน และหุ่นยนต์: เหมาะที่สุด แต่ตลาดเล็ก

- **ทำไมจึงเหมาะ**

  เจ้าของรายเดียวควบคุมทั้งสองปลาย แบนด์วิดท์มีจำกัด และเนื้อหาแคบพอที่จะฝึกโมเดลได้

- **จีน**

  ประกาศฉบับที่ 49 ปี 2026 ซึ่งมีผล 18 พฤศจิกายน ปรับปรุงการควบคุมการส่งออกโดรน โดยมีรายการหนึ่งใช้การบีบอัดภาพเชิงความหมายเป็นเกณฑ์

- **สหรัฐอเมริกา**

  FCC เพิ่มโดรนที่ผลิตในต่างประเทศและชิ้นส่วนสำคัญเข้าใน Covered List เมื่อ 22 ธันวาคม 2025

> **ตลาดเฉพาะที่ดีสำหรับบริษัทเฉพาะทางที่มีความสามารถด้านการปฏิบัติตามกฎควบคุมการส่งออก ไม่ใช่ตลาดกระแสหลัก**

พลังงานต่อพิกเซลมากกว่าในระดับสิบเท่า

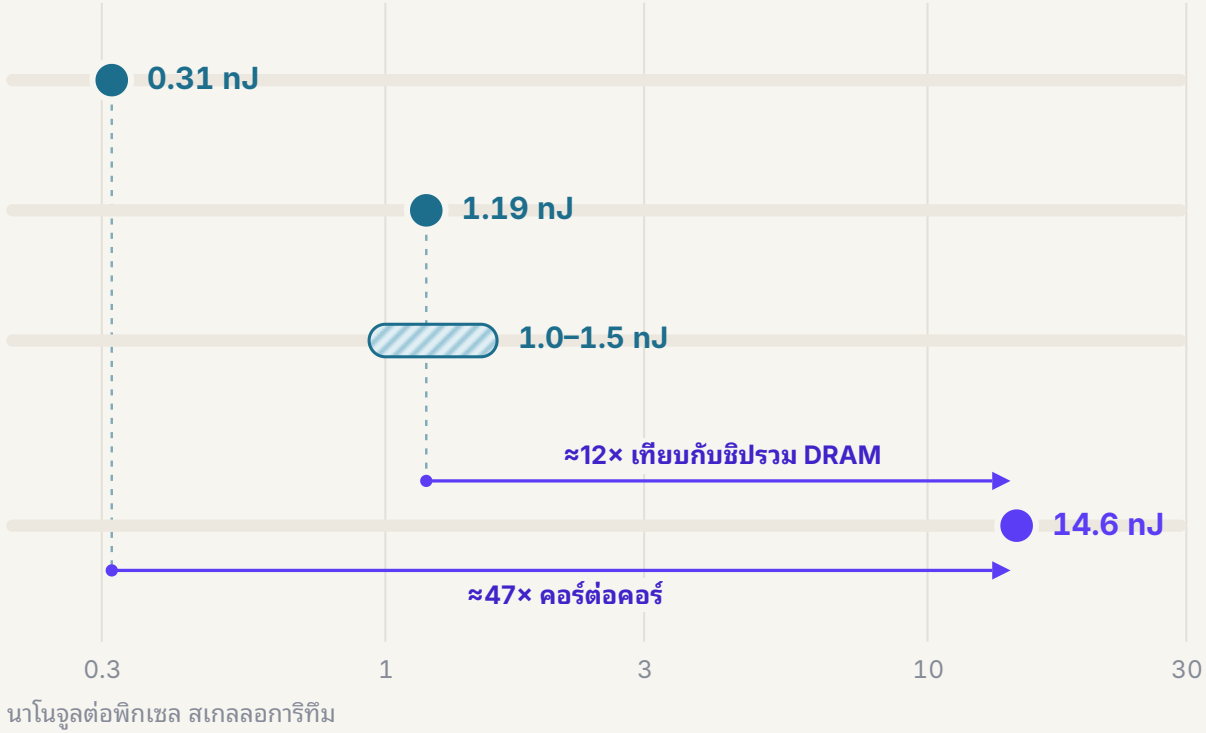
ชิปถอดรหัส HEVC เฉพาะคอร์
40 nm, 4K30, วัดจริง

ชิปถอดรหัส HEVC รวม DRAM
40 nm, 4K30, คอร์วัดจริงบวก DRAM จากการจำลอง

ตัวถอดรหัส VVC
28 nm, 4K60, ค่าคาดการณ์

ตัวถอดรหัสวิดีโอนิวรัล (NVCA)
28 nm, 1080p25, ผลสังเคราะห์วงจรเท่านั้น ไม่นับค่าน้ำหนักและ DRAM

ที่ 1080p30 คิดเป็นราว 74 mW เทียบกับ 908 mW



ที่ 1080p30: ราว 74 mW สำหรับชิป HEVC รวม DRAM เทียบกับราว 900 mW สำหรับ NVCA ชิป HEVC อยู่บนโหนดที่เก่ากว่า และ NVCA ไม่นับค่าน้ำหนักและ DRAM: การเปรียบเทียบนี้จึงเข้าข้างการออกแบบนิวรัล

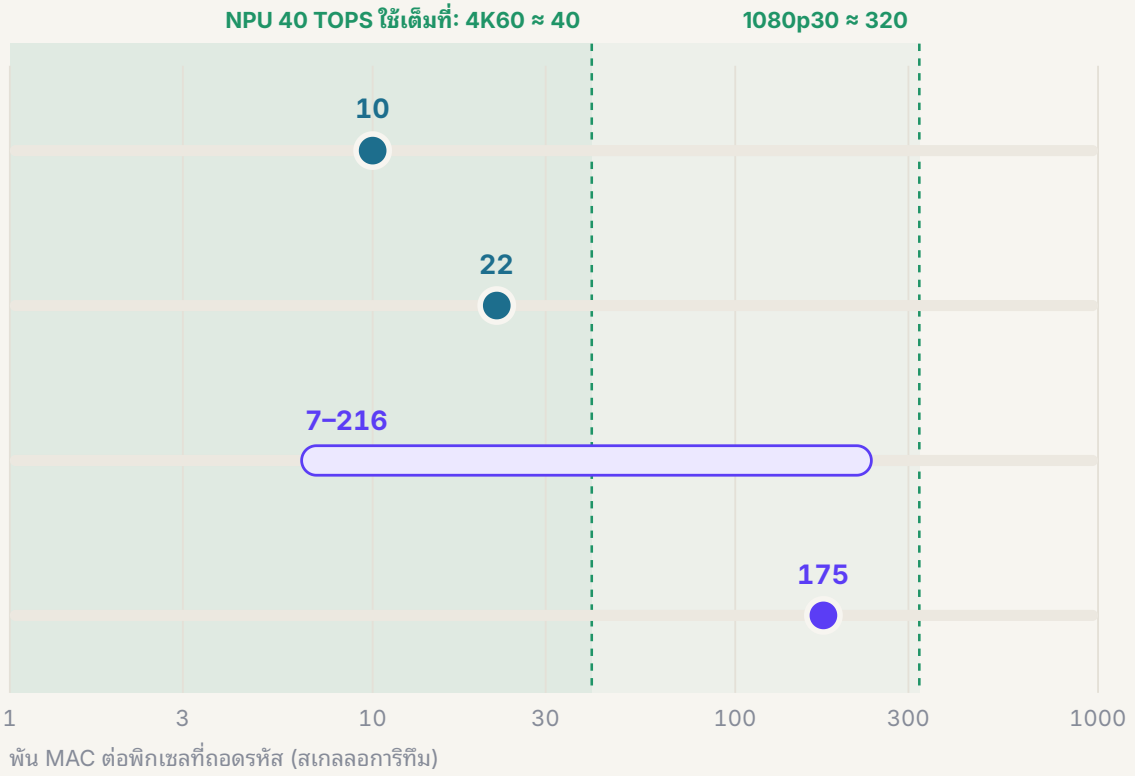
สิ่งที่ตัวถอดรหัสนิวรัลเรียกร้องจาก NPU

ฟิลเตอร์ในลูปแบบนิวรัล ความซับซ้อนต่ำมาก
งานสำรวจของ JVET · ประหยัดบิต ≈6%
1.2 TOPS ที่ 1080p30 · 10 ที่ 4K60

ฟิลเตอร์ในลูปแบบนิวรัล ความซับซ้อนต่ำ
งานสำรวจของ JVET · ≈8%
2.7 TOPS ที่ 1080p30 · 22 ที่ 4K60

ตัวถอดรหัส JPEG AI ระดับพื้นฐานถึงสูง
12.5–27.9% เทียบกับ VVC แบบ intra

DCVC-RT ตัวถอดรหัสวิดีโอนิวรัลทั้งตัว
21% เทียบกับตัวอ้างอิงของ VVC ตามการตั้งค่าในบทความ
22 TOPS ที่ 1080p30 · 174 ที่ 4K60



DCVC-RT ต้องใช้ราว 175,000 MAC ต่อพิกเซล: ราว 22 TOPS ที่ 1080p30 และ 174 ที่ 4K60 เมื่อใช้งานเต็มที่ ประมาณการ NPU 40 TOPS ที่ใช้งานเต็มที่รองรับได้ราว 40,000 ที่ 4K60

ทราฟฟิกหน่วยความจำและความตรงกันทุกบิตคือขีดจำกัดที่แท้จริง

DRAM เป็นตัวหลัก

หน่วยความจำมาก่อน MAC

ในชิป HEVC ของ MIT นั้น DRAM คิดเป็นราวสามในสี่ของพลังงานทั้งระบบ: 0.88 จาก 1.19 nJ ต่อพิกเซล ตัวถอดรหัสนิวรัลแบก feature map ขนาดใหญ่ไว้ระหว่างเฟรม และผู้เขียน DCVC-RT พบว่าสิ่งที่จำกัดความเร็วคือ I/O ของหน่วยความจำ ไม่ใช่การคำนวณ

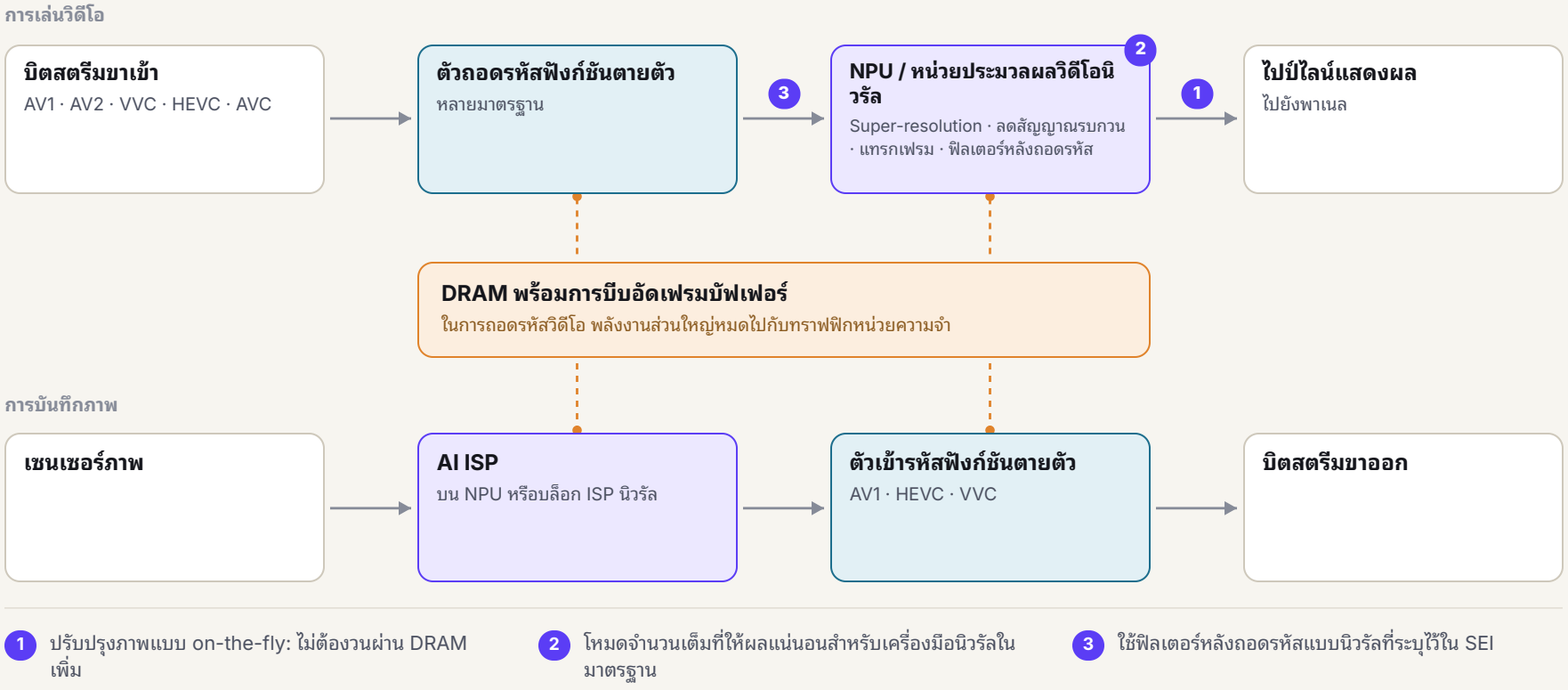
ความตรงกันทุกบิต

ตัวถอดรหัสทุกตัวต้องคำนวณได้บิตเดียวกัน

การถอดรหัสเอนโทรปีจะล้มเหลวหากอุปกรณ์สองเครื่องคำนวณความน่าจะเป็นต่างกัน ทางแก้คือ โมเดลจำนวนเต็ม (DCVC-RT) hyperprior แบบ shared-scale (MLVC) การอนุมานจำนวนเต็ม 16 บิต (SADL ของ JVET) และจุดตรวจความสอดคล้อง (JPEG AI)

สำหรับทีมชิป นี่คือวินัยแบบการแปลงจำนวนเต็มของ HEVC อีกครั้ง: มาตรฐานต้องกำหนดการคำนวณไว้อย่างแม่นยำ

การแบ่งงานที่ใช้จริง: โคเดกฟังก์ชันตายตัวบวก NPU



1. ปรับปรุงภาพแบบ on-the-fly: ไม่ต้องวนผ่าน DRAM เพิ่ม
2. โหมดจำนวนเต็มที่ให้ผลแน่นอนสำหรับเครื่องมือนิวรัลในมาตรฐาน
3. ใช้ฟิลเตอร์หลังถอดรหัสแบบนิวรัลที่ระบุไว้ใน SEI

> Canghai ของ Tencent: ใช้งานชิปรุ่นแรกแล้วกว่า 100,000 ชิ้น รุ่นที่สองเพิ่ม H.266 และ super-resolution บนชิป Chips&Media และ Allegro DVT ขาย IP โคเดกและ IP วิดีโอนิวรัลควบคู่กัน

สิ่งที่ผมจะสร้าง แยกตามบทบาท

ทีม SoC และชิปทีวี

- ตัวถอดรหัสฟังก์ชันตายตัวสำหรับทุกมาตรฐานที่ต้องเล่นได้ AV2 ตั้งแต่ตอนนี้
- NPU ขนาดพอสำหรับการปรับปรุงภาพที่ความละเอียดของจอ โดยไม่ต้องวนผ่าน DRAM เพิ่ม
- SEI สำหรับฟิลเตอร์หลังถอดรหัสแบบนิวรัล และโหมดจำนวนเต็มที่ให้ผลแน่นอน
- H.267 เป็นจุดเริ่มออกแบบในปี 2029 หรือหลังจากนั้น

ผู้ขาย IP

- ตัวถอดรหัส AV2 และ VVC ตั้งแต่ตอนนี้
- ตัวเร่งฟิลเตอร์หลังถอดรหัส และการบีบอัดเฟรมบัฟเฟอร์
- เครื่องมือทดสอบความสอดคล้องสำหรับการประมวลผลแบบนิวรัล

แพลตฟอร์มวิดีโอ

- การประมวลผลล่วงหน้าแบบนิวรัลที่ตัวเข้ารหัสคืนทุนก่อน
- super-resolution ฝั่งไคลเอนต์ในที่ที่อุปกรณ์รองรับ
- โคเดกนิวรัลแบบ end-to-end เฉพาะในวงปิดที่คุณควบคุม

จุดที่เครื่องมือนิวรัลกลายเป็นสิ่งที่เก็บค่าสิทธิได้

มาตรฐาน	เนื้อหาด้านนิวรัล	สถานะ พฤศจิกายน 2026
H.267 (ชื่อชั่วคราว)	ยังเปิดกว้าง: ทั้งเครื่องมือแบบดั้งเดิม (ECM) และเครื่องมือนิวรัลเป็นตัวเลือก	ประเมินข้อเสนอมกราคม 2027; เวอร์ชัน 1 สิ้นปี 2029
VVC และ HEVC พร้อม VSEI	ฟิลเตอร์หลังถอดรหัสแบบนิวรัลที่ระบุใน SEI; ตัวถอดรหัสเลือกใช้ได้	มีผลบังคับใช้; ขยายในเดือนตุลาคม 2025
AV2	แบบดั้งเดิม ปลอดค่าสิทธิ	เวอร์ชัน 1.0.0 เมื่อ 28 พฤษภาคม 2026; YouTube วางแผนสตรีมบนเว็บต้นปี 2027
JPEG AI	โคเดกภาพแบบเรียนรู้ end-to-end	ส่วนที่ 1 เผยแพร่ปี 2025; ไม่พบผลิตภัณฑ์ที่รองรับเป็นค่าเริ่มต้น
IEEE 1857.11	การเข้ารหัสภาพแบบนิวรัลที่ออกแบบสำหรับ NPU	เผยแพร่ธันวาคม 2024
Video Coding for Machines	การเข้ารหัสเพื่อการวิเคราะห์โดยเครื่องจักรแทนการรับชม	ร่างมาตรฐานสากลต้นปี 2026
ส่วนขยาย ML ของ Opus	เลือกใช้ได้และเข้ากันได้ย้อนหลัง	อยู่ใน libopus 1.5 และ 1.6

645

ตระกูลสิทธิบัตร beyond-VVC ที่ ByteDance ถือครอง (GreyB, 2025)

0.50 · 1.00–1.50 ดอลลาร์

อัตราของกลุ่มสิทธิบัตร VVC ต่อโทรศัพท์และต่อทีวี 4K, Region 1

0 ดอลลาร์

ค่าสิทธิของ AV1 และ AV2 จากสมาชิก AOMedia

เงินไหลตามมาตรฐานและแพลตฟอร์ม

มีรายได้แล้ว

แพลตฟอร์ม

เก็บส่วนที่ประหยัดได้

ตัวลดขนาดภาพแบบนิวรัลของ Netflix; super-resolution ของ YouTube และ Meta; Microsoft เปิดซอร์ส MLVC ภายใต้ MIT

มีรายได้แล้ว

ผู้ผลิตอุปกรณ์และชิป

ขายความแตกต่าง

ISP ของกล้อง; ชิปอัปสเกลในทีวี; ชิปโคเดก Canghai ที่ Tencent พัฒนาเอง

มีรายได้แล้ว

ผู้ขาย IP

ขายไลเซนส์ตามมาตรฐานใหม่แต่ละตัว

Chips&Media, Allegro DVT และ VeriSilicon มีลูกค้าตัวถอดรหัส AV2 ภายในสามสัปดาห์หลังสเปกออก

มีรายได้แล้ว

ผู้ผลิต GPU

ผูกผู้ใช้ไว้กับแพลตฟอร์ม

DLSS 4 ในเกมและแอปกว่า 250 รายการ; SDK เท็กซ์เจอร์นิวรัลยังเป็นเบตา

ต่อเครื่อง

ผู้ถือสิทธิบัตรและกลุ่มสิทธิบัตร

เก็บค่าสิทธิต่อเครื่อง

VVC: 0.50 ดอลลาร์ต่อโทรศัพท์ 1.00–1.50 ดอลลาร์ต่อทีวี 4K; เครื่องมือของ H.267 หลังปี 2029

ออกด้วยการถูกซื้อกิจการ

สตาร์ตอัปโคเดกนิวรัล

ถูกซื้อกิจการ

WaveOne → Apple (2023); iSIZE → Sony (2023); Deep Render → InterDigital (2025)

สตาร์ตอัปโคเดกนิวรัลที่มีชื่อเสียงที่สุดสามราย คือ WaveOne, iSIZE และ Deep Render ถูก Apple, Sony และ InterDigital ซื้อไป

หกช่องทาง และความเสี่ยงของแต่ละทาง

ช่องทาง	ใครซื้อ	ทำไมต้องตอนนี้	ความเสี่ยงหลัก
การประมวลผลล่วงหน้าแบบนิวรัลที่ตัวเข้ารหัส	แพลตฟอร์มสตรีมมิงและวิดีโอจากผู้ใช้ ผู้แพร่ภาพ	ไม่ต้องเปลี่ยนอุปกรณ์; พิสูจน์แล้วที่ Netflix	แพลตฟอร์มใหญ่สร้างเอง
IP และงานปรับจูนการปรับปรุงภาพฝั่งแสดงผล	ผู้ผลิตชิปทีวี กล่องรับสัญญาณ และโทรศัพท์	NPU อยู่ในอุปกรณ์แล้ว	กลายเป็นสินค้าโภคภัณฑ์; ทีมภายในของผู้ผลิตชิป
โคเดกนิวรัลแบบวงปิด	ผู้ขายระบบสื่อสารองค์กร ระบบอุตสาหกรรม และเทเลเพรเซนซ์	MLVC พิสูจน์ว่าใช้ได้ และโค้ดเปิดเผย	คุณต้องควบคุมทั้งสองฝั่ง; NPU กระจัดกระจายหลายแบบ
ชิลิคอนหรือ IP โคเดกบวก NPU	ผู้ผลิต SoC	รอบของ AV2 กำลังเริ่ม	ผู้เล่นรายเดิมแข็งแกร่ง
การประเมินผล ความสอดคล้อง และการอนุมานที่ให้ผลแน่นอน	ทุกรายที่นำเครื่องมือนิวรัลไปใช้จริง	ความตรงกันทุกบิตเป็นข้อกำหนดที่เลี่ยงไม่ได้	ตลาดเล็ก เว้นแต่ขายรวมกับอย่างอื่น
การส่งข้อเสนอเข้ามาตรฐาน H.267	ผู้ให้ไลเซนส์ แพลตฟอร์มใหญ่	ข้อเสนอจะถูกประเมินในเดือนมกราคม 2027	ใช้เวลานาน แพง ไม่แน่นอน

คำกล่าวอ้างที่คุณจะได้ยิน

ฟังหูไว้หูจนกว่าคำกล่าวอ้างจะระบุครบสี่เรื่อง

คำกล่าวอ้าง	หลักฐานว่าอย่างไร
“โคเดกนิวรัลชนะ VVC แล้ว”	ในห้องแล็บ DCVC-RT ประหยัด 21% เมื่อเทียบกับตัวเข้ารหัสอ้างอิงของ VVC ส่วนตัวถอดรหัสใช้ราว 175,000 MAC ต่อพิกเซล
“โคเดก AI จะมาแทน H.26x และ AV1 ภายในไม่กี่ปี”	ไม่ใช่ในการเล่นวิดีโอ: H.267 เวอร์ชัน 1 สิ้นปี 2029 ชิปราวปี 2031 ประมาณการ โคเดกที่กำลังมาถึงตอนนี้คือ AV2 ซึ่งเป็นแบบดั้งเดิม
“MLVC ลดแบนด์วิดท์ได้ 88%”	เมื่อเทียบกับ H.264 ที่ 360p ส่วนเมื่อเทียบกับ HEVC แบบฮาร์ดแวร์: 75.5% ที่ 360p และ 65.4% ที่ 540p บนชุดทดสอบของ Microsoft เอง
“เท็กซ์เจอร์นิวรัลลดหน่วยความจำวิดีโอได้ 85–96%”	ในฉากสาธิต เกมที่วางจำหน่ายแล้วได้ราว 30% และใช้เฉพาะกับแอสเซตที่เลือกไว้
“JPEG AI เป็นมาตรฐาน ISO กล้องและเบราว์เซอร์จึงจะนำไปใช้”	ไม่พบการรองรับเป็นค่าเริ่มต้นในผลิตภัณฑ์ เว็บเลือกใช้ JPEG XL ซึ่งเป็นฟอร์แมตแบบดั้งเดิม

ขอให้ระบุจุดอ้างอิง ตัวชี้วัด ความละเอียด และต้นทุนการถอดรหัส คำกล่าวอ้างที่ระบุทั้งสี่ข้อไม่ได้คือผลการวิจัย ไม่ใช่ผลิตภัณฑ์

สิ่งที่ต้องจับตา และคำถามที่ยังเปิดอยู่

สัญญาณที่ต้องจับตา

มกราคม 2027	ประเมินข้อเสนอ H.267: กี่ฉบับที่พึ่งเครื่องมือนิวรัล และด้วยต้นทุนต่อพิกเซลเท่าไร
ต้นปี 2027	สตรีม AV2 บนเว็บครั้งแรกของ YouTube; ตัวถอดรหัส AV2 ใน SoC ที่ประกาศออกมา
ครึ่งหลังของปี 2027	โมเดลทดสอบแรกของ H.267
2027	MLVC ขยายเกินการโทรแบบ peer-to-peer บน Teams
2027–2028	ผลิตภัณฑ์แรกที่ใช้ JPEG AI เป็นค่าเริ่มต้น
สิ้นปี 2029	H.267 เวอร์ชัน 1 เริ่มรอบการพัฒนาชิป

คำถามที่ยังเปิดอยู่

01 ความตรงกันทุกบิตในวงกว้าง: ถอดรหัสจำนวนเต็มให้ตรงกันทุกบิตบน NPU ของผู้ผลิตต่างราย แลกกับการเสียประสิทธิภาพการบีบอัดเท่าไร?

02 หน่วยความจำมาก่อน MAC: ทราฟฟิก DRAM ต่อพิกเซลของการประมวลผลภายหลังแบบนิวรัล

03 ต้นทุนต่อพิกเซลที่ทำให้เครื่องมือนิวรัลไม่คุ้มอีกต่อไปใน H.267

04 วงปิดใดนอกเหนือจากการโทรจะนำโคเดกนิวรัลแบบ end-to-end ไปใช้ก่อน

05 จะวัดความเที่ยงตรงอย่างไรเมื่อตัวถอดรหัสสร้างรายละเอียดขึ้นเอง

เหล่านี้คือเรื่องที่เราสำรวจร่วมกันได้ ผ่านการแลกเปลี่ยนและความร่วมมือทางเทคนิค

i@tanshuai.com · tanshuai.com