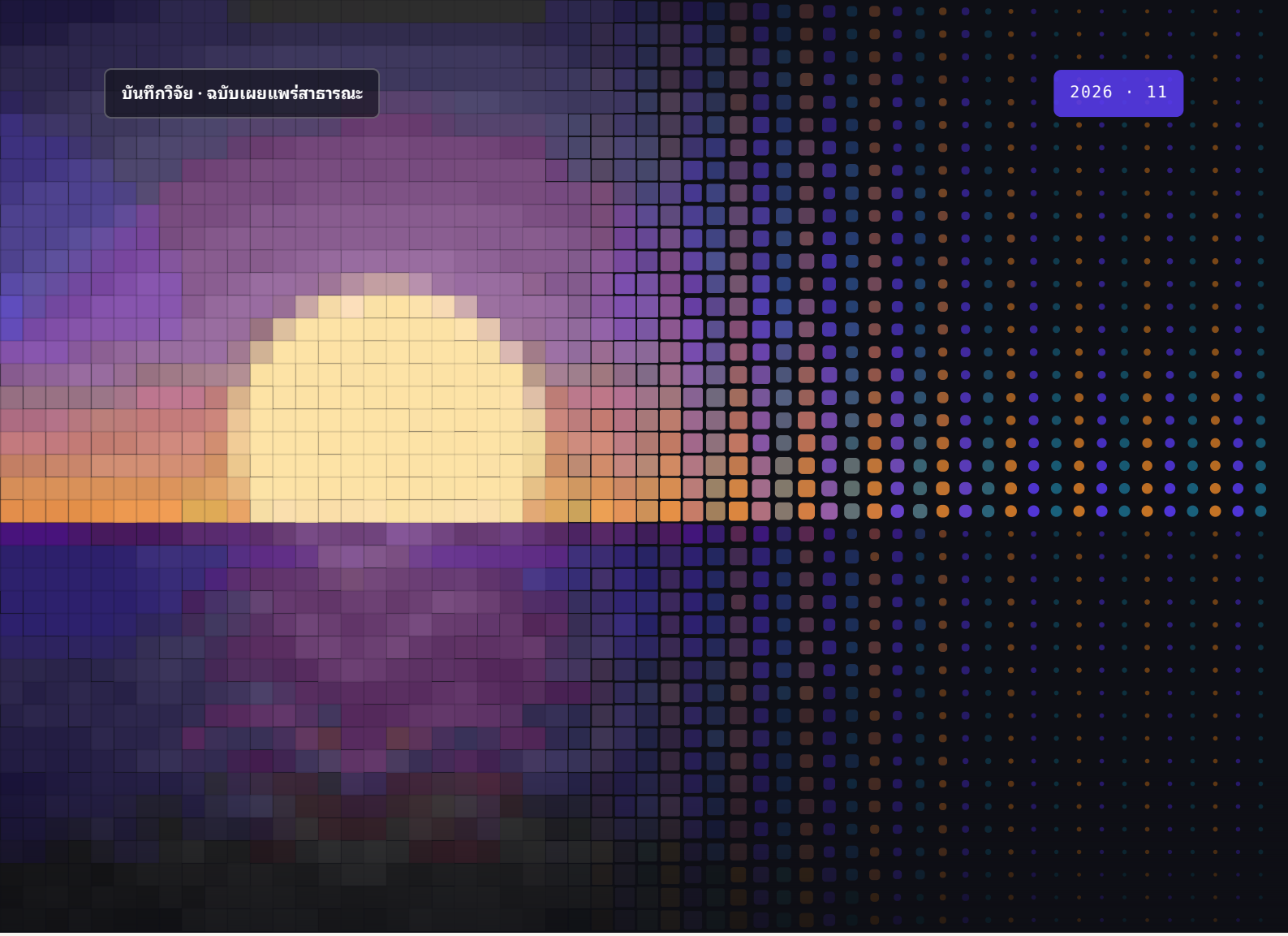


การบีบอัดแบบเรียนรู้ · การเล่นวิดีโอ × ซิลิคอน

การเข้ารหัสแบบนิวรัลในปี 2026

ใช้งานจริงรอบโคเดก ไม่ใช่แทนโคเดก

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

ข้อสรุป

โครงข่ายประสาทเทียมทำงานอยู่แล้วที่ทั้งสองปลายของโคเดกที่รับส่งวิดีโอของโลก โคเดกนิวรัลแบบ end-to-end ใช้งานจริงเฉพาะในวงปิด ส่วนจอกระแสหลักยังต้องรอ H.267

ทัน ชว่าย

นักเทคโนโลยีและผู้ก่อตั้ง · นักวิจัยและนักพัฒนา AI

tanshuai.com · i@tanshuai.com

ข้อมูล ณ วันที่ 12 พฤศจิกายน 2026
จำนวนเงินเป็นดอลลาร์สหรัฐ

สารบัญ

สารบัญ



◆ **บทสรุปผู้บริหาร** 04
เป็นกระแสหลักแล้ว แต่อยู่รอบโคเดก

01 **ขอบเขตและวิธีการ** 07
“การเข้ารหัสแบบนิวรัล” ในรายงานนี้หมายถึงอะไร และผมศึกษาอย่างไร

02 **การเล่นวิดีโอและสตรีมมิง** 08
การเล่นวิดีโอและสตรีมมิง: นิวรัลที่สองปลาย มาตรฐานตรงกลาง

03 **การถ่ายภาพและภาพนิ่ง** 11
การถ่ายภาพและภาพนิ่ง: นิวรัลในกล้อง แบบดั้งเดิมบนเว็บ

04 **การโทรและการประชุม** 12
การโทรและการประชุม: ที่ที่โคเดกนิวรัลแบบ end-to-end ใช้งานจริง

05 **Generative AI** 14
Generative AI: การใช้งานที่ใหญ่ที่สุดซึ่งไม่มีใครเรียกว่าโคเดก

06 **เกมและกราฟิก** 15
เกมและกราฟิก: การสร้างภาพขึ้นใหม่แบบนิวรัลเป็นกระแสหลักแล้ว แอสเซตนิวรัลกำลังเริ่มต้น

07 **ลิงก์เฉพาะทาง** 17
ลิงก์เฉพาะทาง: ดาวเทียม โดรน และหุ่นยนต์

08 **ซิลิคอน** 18
มุมมองด้านซิลิคอน: พลังงาน การประมวลผล หน่วยความจำ และความตรงกันทุกบิต

09 **มาตรฐานและสิทธิบัตร** 22
มาตรฐานและสิทธิบัตร: จุดที่เครื่องมือนิวรัลกลายเป็นสิ่งที่เก็บค่าสิทธิได้

10 **ใครได้เงิน** 24
ใครได้เงิน และบริษัทจะเข้าไปเล่นตรงไหนได้

11 **คำกล่าวอ้างและหลักฐาน** 26
คำกล่าวอ้างที่คุณจะได้ยิน และหลักฐานว่าอย่างไร

12 **สัญญาณที่ต้องจับตา** 27
สิ่งที่ต้องจับตา ปี 2027 ถึง 2030

13 **คำถามที่ยังเปิดอยู่** 28

◆ **เกี่ยวกับผู้เขียน** 29
ทัน ชว่าย

◆ **แหล่งข้อมูล** 30

สารบัญภาพ

สารบัญภาพ



ภาพที่ 01	การเข้ารหัสแบบนิวรัลใช้งานจริงที่ไหนในปี 2026	06
ภาพที่ 02	สามทางที่โครงข่ายประสาทเทียมเข้าสู่โคเดก	07
ภาพที่ 03	จากมาตรฐานที่เสร็จสมบูรณ์ถึงจอกระแสหลักใช้เวลาเกือบทศวรรษ	10
ภาพที่ 04	ในการโทร โคเดกนิวรัลใช้บิตเรตน้อยกว่า HEVC แบบฮาร์ดแวร์ 65–76%	13
ภาพที่ 05	เมื่อบริษัทเดียวควบคุมทั้งสแตก การสร้างภาพขึ้นใหม่แบบนิวรัลก็แพร่หลายเร็ว	16
ภาพที่ 06	พลังงานต่อพิกเซลที่ถอดรหัส: ตัวถอดรหัสฟังก์ชันตายตัวนำอยู่ระดับสิบเท่า	19
ภาพที่ 07	เครื่องมือนิวรัลหนึ่งตัวใช้หลายหมื่น MAC ต่อพิกเซล ตัวถอดรหัสนิวรัลทั้งตัวใช้หลายแสน	20
ภาพที่ 08	การแบ่งงานที่ใช้จริง: ฮาร์ดแวร์ฟังก์ชันตายตัวสำหรับบิตสตรีม เอนจินนิวรัลสำหรับพิกเซล	21
ภาพที่ 09	ใครได้เงินจากการเข้ารหัสแบบนิวรัลในปี 2026	24

ป้ายกำกับในเนื้อหา

- ประมาณการ — ผู้เขียนคำนวณหรืออนุมานจากข้อมูลที่อ้างอิง
- ผู้ขายอ้าง — เผยแพร่โดยบริษัทที่ขายผลิตภัณฑ์นั้น ยังไม่ได้รับการตรวจสอบโดยอิสระ
- บทวิเคราะห์ — การตีความหลักฐานของผู้เขียน

ฉบับนี้แปลจากฉบับภาษาอังกฤษ หากเนื้อหาไม่ตรงกัน ให้ถือฉบับภาษาอังกฤษเป็นหลัก

ข้อมูล ณ วันที่ 12 พฤศจิกายน 2026 · จำนวนเงินเป็นดอลลาร์สหรัฐ

บทสรุปผู้บริหาร

เป็นกระแสหลักแล้ว แต่อยู่รอบโคเดก

ข้อสรุป

ในปี 2026 การเข้ารหัสแบบนิวรัลเป็นกระแสหลักแล้ว แต่ทำงานอยู่รอบโคเดกมาตรฐาน ไม่ได้มาแทนโคเดกเหล่านั้น

≈75%
ของทราฟฟิกข้อมูลมือถือเป็นวิดีโอ ณ สิ้นปี 2025 (Ericsson)

1.26 พันล้าน
สมาร์ตโฟนที่จัดส่งในปี 2025 ทุกเครื่องถอดรหัสวิดีโอด้วยฮาร์ดแวร์ฟังก์ชันตายตัว

≈12×
พลังงานต่อพิกเซลของการออกแบบชิปถอดรหัสนิวรัลเพียงแบบเดียวที่มีการเผยแพร่ เทียบกับชิปถอดรหัส HEVC ที่วัดจริงรวม DRAM

≈22 TOPS
กำลังประมวลผลต่อเนื่องที่ตัวถอดรหัสนิวรัลระดับ DCVC-RT ต้องใช้ที่ 1080p30 ประมาณการ

75.5%
บิตเรตที่ MLVC ใช้น้อยกว่า HEVC แบบฮาร์ดแวร์ที่ 360p ในการโทร บนชุดทดสอบของ Microsoft เอง ผู้ขายอ้าง

สิ้นปี 2029
H.267 เวอร์ชัน 1 มาตรฐานแรกที่น่าจะนำเครื่องมือนิวรัลไปสู่ทุกหน้าจอ

ใช้งานจริงที่ไหน

วันนี้โครงข่ายประสาทเทียมทำงานอยู่ที่ทั้งสองปลายของโคเดกที่รับส่งวิดีโอเกือบทั้งหมดของโลก ขาเข้าโครงข่ายเหล่านี้ปรับแต่งภาพ: หน่วยประมวลผลสัญญาณภาพ (ISP) แบบ AI ในกล้องโทรศัพท์ และตัวลดขนาดภาพแบบนิวรัลในไปป์ไลน์การเข้ารหัสของ Netflix ขาออก โครงข่ายเหล่านี้ซ่อมภาพ: Super Resolution ของ YouTube บนทีวี, super-resolution ที่ Meta และชิปประมวลผลภาพในทีวีที่โฆษณาว่ามีโครงข่ายประสาทเทียมหลายร้อยชุด ตรงกลางนั้น บิตสตรีมยังคงเป็น AV1, HEVC, AVC หรือ VVC ที่ถอดรหัสด้วยฮาร์ดแวร์ฟังก์ชันตายตัว ไม่มีบริการสตรีมมิงใดส่งวิดีโอในฟอร์แมตนิวรัลไปยังทีวีหรือโทรศัพท์ที่ใดที่มีตัวถอดรหัสวิดีโอนิวรัลทำงานบนอุปกรณ์ในวันนี้ ตัวถอดรหัสนั้นก็ทำงานอยู่ภายในแอปของบริษัทเดียว

วงปิด

โคเดกนิวรัลแบบ end-to-end ซึ่งทั้งตัวเข้ารหัสและตัวถอดรหัสเรียนรู้จากข้อมูล ใช้งานจริงเฉพาะในวงปิด คือในที่ที่บริษัทเดียวควบคุมทั้งสองฝั่งและถอยกลับไปใช้โคเดกมาตรฐานได้ Microsoft ใช้โคเดกเสียง AI ในการโทรสองฝ่ายทั้งหมดบน Teams และ Skype มาตั้งแต่ปี 2021 และในเดือนกรกฎาคม 2026 ก็เปิดซอร์ส MLVC โคเดกวิดีโอนิวรัลที่กำลังทยอยใช้ในการโทรแบบ peer-to-peer บน Teams: ที่ 360p MLVC ใช้บิตเรตน้อยกว่า HEVC แบบฮาร์ดแวร์ราว 75% ที่คุณภาพตามการให้คะแนนเท่ากัน โมเดลสร้างเสียงพูดและวิดีโอทุกตัวทำงานบนโคเดกแบบเรียนรู้ ตั้งแต่ tokenizer เสียงพูดที่ 12.5 Hz ไปจนถึง VAE วิดีโอที่มีอัตราการบีบอัด 1:192 การบีบอัดเท็กซ์เจอร์แบบนิวรัลใช้งานจริงแล้วในเกม AAA หนึ่งเกม ทั้งหมดนี้คือการใช้งานจริง และมีเงื่อนไขร่วมกันข้อหนึ่ง: ไม่มีใครอื่นต้องถอดรหัสสตรีมนั้น

จังหวะเวลา

การเล่นบิตสตรีมนิวรัลในวงกว้างต้องรอทั้งมาตรฐานและรอบการพัฒนาชิป มาตรฐานแรกที่น่าจะนำเครื่องมือนิวรัลเข้ามา คือมาตรฐานรุ่นถัดจาก VVC ที่ผมจะเรียกว่า H.267 ได้ออกประกาศเรียกรับข้อเสนอในเดือน

กรกฎาคม 2026 จะประเมินข้อเสนอในเดือนมกราคม 2027 และตั้งเป้าเวอร์ชันแรกไว้ที่สิ้นปี 2029 AV1 ใช้เวลาเกือบแปดปีจากสเปกที่เสร็จสมบูรณ์จนถึงสัดส่วน 30% ของการรับชม Netflix หากเป็นไปตามแบบแผนนั้น ชิป H.267 รุ่นแรกจะมาถึงราวปี 2031 และการใช้งานในกระแสหลักจะเกิดขึ้นในกลางทศวรรษ 2030 ประมาณการ ระหว่างนี้ โคเดกตัวถัดไปที่จะไปถึงหน้าจอคือ AV2 ซึ่งเสร็จในเดือนพฤษภาคม 2026 โดยไม่มีตัวถอดรหัสนิวรัลอยู่ข้างใน และผู้ขาย IP สามรายประกาศว่ามีลูกค้าตัวถอดรหัส AV2 ภายในสามสัปดาห์

ซิลิคอน

เหตุผลอยู่ที่ซิลิคอน การออกแบบชิปถอดรหัสวิดีโอนิวรัลเพียงแบบเดียวที่มีการเผยแพร่ใช้พลังงาน 14.6 nJ ต่อพิกเซล ราว 12 เท่าของชิปถอดรหัส HEVC ที่วัดจริงเมื่อรวม DRAM และราว 47 เท่าของคอร์ของชิปนั้นอย่างเดียว DCVC-RT โคเดกเชิงวิจัยแบบเรียลไทม์ของ Microsoft ต้องใช้การคูณสะสม (multiply-accumulate, MAC) ราว 175,000 ครั้งต่อพิกเซลที่ถอดรหัส ตามที่ผมนับจากตารางรายโมดูลของผู้พัฒนาเอง หรือกำลังประมวลผลต่อเนื่องราว 22 TOPS สำหรับ 1080p30 ประมาณการ : ราวครึ่งหนึ่งของ 40 TOPS ที่ใช้นิยาม AI PC แต่ใช้ไปกับสตรีมวิดีโอเพียงสตรีมเดียว ตัวถอดรหัสนิวรัลยังต้องให้ผลลัพธ์ที่ตรงกันทุกบิตบนทุกอุปกรณ์ มิฉะนั้นการถอดรหัสเอนโทรปีจะพังทลาย นี่คือเหตุที่ MLVC ยอมสละประสิทธิภาพการบีบอัดเพื่อให้ถอดรหัสได้ผลเหมือนกันบน NPU ทุกตัว และทำงานที่ 540p ไม่ใช่ 4K

ธุรกิจ

ในเชิงธุรกิจ เงินไหลตามมาตรฐานและแพลตฟอร์ม ไม่ใช่ตามโคเดกนิวรัลที่ขายแยกต่างหาก แพลตฟอร์มเก็บส่วนที่ประหยัดได้จากการประมวลผลล่วงหน้าและภายหลังแบบนิวรัลไว้เอง ผู้ขาย IP ขายไลเซนส์ตัวถอดรหัสฟังก์ชันตายตัวสำหรับมาตรฐานใหม่แต่ละตัว พร้อม NPU สำหรับงานภาพ กลุ่มสิทธิบัตรเก็บค่าสิทธิต่อเครื่อง: สำหรับ VVC คือ 0.50 ดอลลาร์สหรัฐต่อโทรศัพท์ และ 1.00–1.50 ดอลลาร์สหรัฐต่อทีวี 4K ในสหรัฐฯ และตลาดอื่นใน Region 1 สตาร์ตอัปโคเดกนิวรัลที่มีชื่อเสียงที่สุดสามราย คือ WaveOne, iSIZE และ Deep Render ถูก Apple, Sony และ InterDigital ซื้อไป โอกาสที่ยังเปิดอยู่คือการประมวลผลล่วงหน้าและภายหลังแบบนิวรัล ผลิตภัณฑ์แบบวงปิด ซิลิคอนที่จับคู่โคเดกฟังก์ชันตายตัวกับ NPU และงานประเมินผลกับงานมาตรฐานที่อยู่ระหว่างกลาง

คำถาม	การประเมิน	ตัวเลขสำคัญ (แหล่งที่มาอยู่ในแต่ละบท)
1 วันนี้การเข้ารหัสแบบนิวรัลใช้งานจริงในวงกว้างที่ไหน?	รอบโคเดก: ISP ของกล้อง การลดขนาดภาพฝั่งเข้ารหัส และ super-resolution ฝั่งเล่นวิดีโอ ส่วนแบบ end-to-end มีเฉพาะในวงปิด: การโทร แอสเซตเกม และ tokenizer ของ Generative AI	วิดีโอ ≈75% ของข้อมูลมือถือ; โทรศัพท์ 1.26 พันล้านเครื่องและทีวี ≈200 ล้านเครื่องที่จัดส่งในปี 2025 ล้วนถอดรหัสด้วยฮาร์ดแวร์ฟังก์ชันตายตัว; DLSS 4 อยู่ในเกมและแอปกว่า 250 รายการ
2 โคเดกนิวรัลจะมาแทน AV1, HEVC และ VVC ในการเล่นวิดีโอหรือไม่? เมื่อไร?	ไม่ใช่ก่อนจะมีมาตรฐานและรอบการเปลี่ยนอุปกรณ์ H.267 คือช่องทางแรกที่เป็นไปได้จริง และน่าจะผสมเครื่องมือแบบดั้งเดิมกับแบบนิวรัลเข้าด้วยกัน	H.267 ประกาศเรียกรับข้อเสนอกรกฎาคม 2026 ประเมินมกราคม 2027 เวอร์ชัน 1 สิ้นปี 2029; เข้าสู่กระแสหลักกลางทศวรรษ 2030 ประมาณการ
3 ต้นทุนบนซิลิคอนเป็นเท่าไร?	พลังงานต่อพิกเซลสูงกว่าการถอดรหัสด้วยฮาร์ดแวร์ฟังก์ชันตายตัวในปัจจุบันระดับสิบเท่า บวกทราฟฟิกหน่วยความจำและข้อกำหนดเรื่องความตรงกันทุกบิต	14.6 nJ/พิกเซล (นิวรัล) เทียบกับ 1.19 (HEVC รวม DRAM); ≈175 kMAC ต่อพิกเซลที่ถอดรหัส ≈ 22 TOPS ที่ 1080p30 ประมาณการ
4 มาตรฐานและสิทธิบัตรเปิดทางให้แค่ไหน?	สตรีม VVC ระบุฟิลเตอร์หลังถอดรหัสแบบนิวรัลได้แล้ว; JPEG AI เป็นมาตรฐาน ISO; H.267 จะเป็นตัวตัดสินว่าเครื่องมือนิวรัลใดจะกลายเป็นสิ่งที่เก็บค่าสิทธิได้	ECM (แบบดั้งเดิม) ดีกว่า VVC ≈28%; ฟิลเตอร์ในลูปแบบนิวรัล 6–14% ที่ 10–22+ kMAC/พิกเซล; ByteDance มี 645 ตระกูลสิทธิบัตรในงาน beyond-VVC
5 ใครได้เงิน และบริษัทจะเข้าไปเล่นตรงไหนได้?	แพลตฟอร์ม ผู้ขาย IP และผู้ถือสิทธิบัตร สตาร์ตอัปที่ทำโคเดกนิวรัลอย่างเดียวถูกซื้อกิจการไปแล้ว ช่องทางอยู่ที่การประมวลผลล่วงหน้า/ภายหลัง วงปิด ซิลิคอนโคเดกบวก NPU และงานประเมินผล	กลุ่มสิทธิบัตร VVC 0.50 ดอลลาร์สหรัฐต่อโทรศัพท์; IP ตัวถอดรหัส AV2 ขายไลเซนส์ได้ภายในไม่กี่สัปดาห์หลังสเปกออก; การซื้อกิจการสามครั้ง ไม่เปิดเผยเงื่อนไข

คำถาม	การประเมิน	ตัวเลขสำคัญ (แหล่งที่มาอยู่ในแต่ละบท)
⑥ **คำกล่าวอ้างแบบใดควรฟังหูไว้หู?**	ค่าที่ประหยัดได้ซึ่งเทียบกับโคเดกรุ่นเก่า วัดที่ความละเอียดต่ำ ไม่บอกต้นทุนการถอดรหัส หรือมาจากการสาธิตแทนผลิตภัณฑ์ที่วางจำหน่ายแล้ว	MLVC: น้อยกว่า H.264 88% แต่น้อยกว่า HEVC แบบฮาร์ดแวร์ 65–76%; เท็กซ์เจอร์: ประหยัด VRAM 85% ในการสาธิต และราว 30% ในเกมที่วางจำหน่ายแล้ว

ภาพที่ 01

การเข้ารหัสแบบนิวรัลใช้งานจริงที่ไหนในปี 2026

โครงข่ายประสาทเทียมทำงานอยู่แล้วที่ทั้งสองปลายของโคเดกกระแสหลัก โคเดกนิวรัลแบบ end-to-end ใช้งานจริงเฉพาะในที่ที่บริษัทเดียวควบคุมทั้งตัวเข้ารหัสและตัวถอดรหัส

	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** การฟื้นฟูภาพจากฟิดบิตเรตต่ำฝั่งเซิร์ฟเวอร์	● **สาธิต นำร่อง** โคเดกนิวรัลที่ทนต่อการสูญหายของข้อมูล อยู่ในขั้นวิจัยและนำร่อง	—

ที่มา: เอกสารของบริษัท องค์กรมาตรฐาน และนักพัฒนาที่อ้างถึงในรายงาน สถานะ ณ วันที่ 12 พฤศจิกายน 2026 เรียงแถวตามจำนวนคนที่แต่ละด้านเข้าถึง

ขอบเขตและวิธีการ

01 “การเข้ารหัสแบบนิวรัล” ในรายงานนี้หมายถึงอะไร และผมศึกษาอย่างไร

สามทางที่โครงข่ายประสาทเทียมเข้าสู่โคเดก

“การเข้ารหัสแบบนิวรัล” ครอบคลุมสามเรื่องที่ต่างกัน และข้อถกเถียงส่วนใหญ่เกิดจากการเอาสามเรื่องนี้มาปนกัน

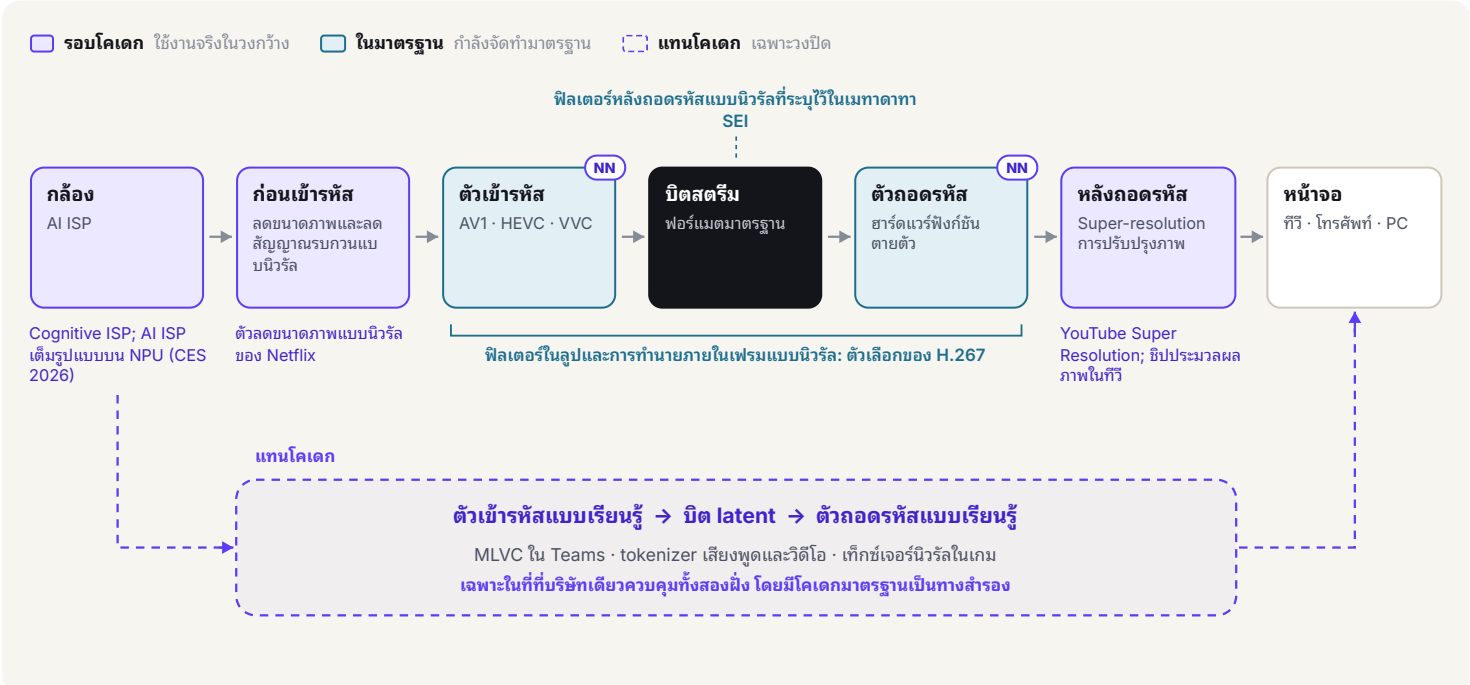
01 **รอบโคเดก** โครงข่ายประมวลผลพิกเซลก่อนการเข้ารหัสหรือหลังการถอดรหัส โดยบิตสตรีมยังเป็นแบบมาตรฐาน ตัวอย่างเช่น หน่วยประมวลผลสัญญาณภาพ (ISP) แบบ AI ในกล้อง ตัวลดขนาดภาพแบบนิวรัลของ Netflix ที่อยู่ก่อนตัวเข้ารหัส และ super-resolution บนทีวี อุปกรณ์ใดก็เล่นสตรีมนั้นได้ โครงข่ายเพียงเปลี่ยนสิ่งที่ป้อนเข้าหรือสิ่งที่ออกมา

02 **ในมาตรฐาน** โครงข่ายกลายเป็นเครื่องมือหนึ่งในโคเดกมาตรฐาน เช่น ฟิลเตอร์ในลูปหรือตัวทำนายภายในเฟรมแบบนิวรัลที่ตัวถอดรหัสทุกตัวต้องรัน จึงต้องกำหนดสเปกไว้อย่างแม่นยำ โดยทั่วไปด้วยเลขคณิตจำนวนเต็ม รูปแบบที่เบากว่าคือฟิลเตอร์หลังถอดรหัสแบบนิวรัลที่บิตสตรีมอธิบายไว้ในเมทาดาทา และตัวถอดรหัสจะเลือกใช้หรือไม่ก็ได้

03 **แทนโคเดก** ทั้งตัวเข้ารหัสและตัวถอดรหัสเรียนรู้จากข้อมูลแบบ end-to-end บิตสตรีมคือชุดค่า latent ที่ผ่านการควอนไทซ์ ซึ่งถอดรหัสได้ด้วยโครงข่ายที่เข้าคู่กันเท่านั้น ตัวอย่างเช่น MLVC ของ Microsoft ใน Teams, tokenizer เสียงพูดในโมเดลเสียง และเท็กซ์เจอร์นิวรัลในเกม

ภาพที่ 02 **สามทางที่โครงข่ายประสาทเทียมเข้าสู่โคเดก**

สิ่งที่ใช้งานจริงในปี 2026 ส่วนใหญ่ทำงานรอบโคเดกมาตรฐาน เครื่องมือนิวรัลในมาตรฐานกำลังอยู่ระหว่างการกำหนดสเปก โคเดกนิวรัลแบบ end-to-end เข้ามาแทนโคเดกเฉพาะในที่ที่บริษัทเดียวควบคุมทั้งสองฝั่ง



ที่มา: ผู้เขียนสังเคราะห์จากแหล่งข้อมูลในรายงานนี้ สถานะ ณ วันที่ 12 พฤศจิกายน 2026

แบบแรกไม่ต้องขออนุญาตใคร แบบที่สองต้องมีมาตรฐาน แบบที่สามต้องควบคุมทั้งสองฝั่ง นี่คือลำดับที่ทั้งสามแบบเข้าถึงผู้คน และอธิบายเนื้อหาส่วนใหญ่ที่ตามมา

เรียงตามโลกจริง ไม่ใช่ตามวาระการวิจัย

บทความวิจัยส่วนใหญ่ศึกษาแบบที่สาม เพราะสถิติการบีบอัดอยู่ตรงนั้น รายงานนี้เรียงด้านต่าง ๆ ตามจำนวนคนที่เข้าถึง เพราะสิ่งนี้เป็นตัวกำหนดว่าเงินและงานวิศวกรรมจะไปทางไหน: การเล่นวิดีโอและสตรีมมิงก่อน ตามด้วยการถ่ายภาพ การโทร Generative AI เกม และลิงก์เฉพาะทาง (ดาวเทียม โดรน หุ่นยนต์) เป็นลำดับสุดท้าย วิดีโอคิดเป็นราว 75% ของทราฟฟิกข้อมูลมือถือทั้งหมด ตามการนับของ Ericsson ณ สิ้นปี 2025 ในปี 2025 อุตสาหกรรมจัดส่งสมาร์ตโฟนราว 1.26 พันล้านเครื่อง (IDC) และทีวีราว 196–209 ล้านเครื่อง (ตัวเลขประมาณการของ TrendForce และ Omdia) ทุกเครื่องถอดรหัสวิดีโอด้วยฮาร์ดแวร์ฟังก์ชันตายตัว

วิธีทำงานของผม

นี่คือการวิจัยจากเอกสาร (desk research) ที่ทำแบบเดียวกับที่ผมรีวิวชิปหรือแผนผลิตภัณฑ์ ผมอ่านเอกสารมาตรฐานและรายงานการประชุม (JVET, MPEG, JPEG, AOMedia, IETF) บล็อกวิศวกรรมและบันทึกการออกรุ่นของบริษัท บทความวิจัยด้านชิป ตารางอัตราค่าสิทธิของกลุ่มสิทธิบัตร ตัวบทกฎหมายควบคุมการส่งออก และข่าว ตัวเลขทุกตัวมีวันที่กำกับ ตัวเลขที่ผมคำนวณเอง เช่น การประมวลผลต่อพิกเซลจากตารางในบทความวิจัย ผมจะบอกไว้และกำกับว่า ประมาณการ คำกล่าวอ้างที่ผมตรวจสอบกับแหล่งข้อมูลปฐมภูมิไม่ได้ กำกับว่า ผู้ขายอ้าง จำนวนเงินเป็นดอลลาร์สหรัฐ สถานะเป็นข้อมูล ณ วันที่ 12 พฤศจิกายน 2026

ผมเปรียบเทียบสิ่งที่เทียบกันได้เมื่อทำได้ และบอกไว้เมื่อทำไม่ได้ ค่าการบีบอัดที่ดีขึ้นแทบไม่มีความหมายหากไม่มีสี่สิ่งนี้กำกับ: จุดอ้างอิง (anchor: ดีขึ้นเมื่อเทียบกับอะไร) ตัวชี้วัด (PSNR, VMAF หรือคณะผู้ชม) เนื้อหาและความละเอียด และต้นทุนการถอดรหัส ข้อขัดแย้งส่วนใหญ่เรื่องโคเดกนิวรัลจะหายไปเมื่อระบุสี่สิ่งนี้ครบ

บทที่ 02

02 การเล่นวิดีโอและสตรีมมิง: นิวรัลที่สองปลาย มาตรฐานตรงกลาง

การเล่นวิดีโอและสตรีมมิง

สิ่งที่ผู้ชมได้รับยังเป็นบิตสตรีมมาตรฐาน

บริการวิดีโอรายใหญ่ที่สุดเปลี่ยนผ่านโคเดกรอบหนึ่งเสร็จแล้วและกำลังเริ่มรอบถัดไป และไม่มีรอบใดเกี่ยวกับตัวถอดรหัสนิวรัล ในเดือนธันวาคม 2025 Netflix รายงานว่า AV1 รองรับการรับชมราว 30% ของทั้งหมด เป็นโคเดกที่ใช้มากเป็นอันดับสองรองจาก AVC AV1 ไปถึงทีวีได้ก็ต่อเมื่อตัวถอดรหัสฮาร์ดแวร์ไปถึงทีวีแล้ว: Netflix เพิ่ม AV1 เข้าในการรับรองอุปกรณ์ในปี 2019 เริ่มสตรีมไปยังทีวีตั้งแต่ปลายปี 2021 และระบุว่า 88% ของอุปกรณ์จอใหญ่ที่ส่งเข้ารับรองระหว่างปี 2021 ถึง 2025 รองรับโคเดกนี้ ซึ่งเกือบทั้งหมดเป็นอุปกรณ์ตั้งแต่ปี 2023 YouTube ระบุว่า ณ วันที่ 19 มีนาคม 2026 คลังวิดีโอของตนกว่า 90% เมื่อถ่วงน้ำหนักด้วยเวลารับชม มีให้บริการในรูปแบบ AV1

AV2 ผู้สืบทอด AV1 จาก Alliance for Open Media ออกเวอร์ชัน 1.0.0 เมื่อวันที่ 28 พฤษภาคม 2026 โดยบีบอัดได้ดีกว่า AV1 ราว 30% ตามการวัดเชิงวัตถุวิสัยของ AOMedia YouTube วางแผนสตรีม AV2 บนเว็บ “เร็วที่สุดในต้นปี 2027” ฝั่งซิลิคอนตามมาทันที: Chips&Media พัฒนาตัวถอดรหัสฮาร์ดแวร์ AV2 เสร็จในเดือนพฤษภาคม และขายไลเซนส์ให้บริษัทเทคโนโลยีรายใหญ่ในอเมริกาเหนือที่ไม่เปิดเผยชื่อในเดือนมิถุนายน Allegro DVT ส่งมอบ IP ตัวถอดรหัส AV2 ให้ลูกค้าสามรายภายในวันที่ 16 มิถุนายน VeriSilicon ออก IP วิดีโอที่รองรับ AV2 ให้ลูกค้าหลายรายภายในวันที่ 10 มิถุนายน AV2 เป็นโคเดกแบบดั้งเดิม

ประสิทธิภาพที่ดีขึ้นมาจากการแบ่งบล็อก การทำนาย การแปลง และฟิลเตอร์ในลูปที่ดีขึ้น (หนึ่งในนั้นเป็นฟิลเตอร์รายละเอียดแบบชี้นำที่ “ขับเคลื่อนด้วยข้อมูล”) ไม่ได้มาจากโครงข่ายแบบ end-to-end

VVC มาตรฐานปัจจุบันของ ITU/ISO ได้บ้านแรกในงานออกอากาศที่บราซิล ซึ่งระบบ DTV+ (TV 3.0) เริ่มให้บริการเชิงพาณิชย์ในเดือนมิถุนายน 2026 ให้ทันฟุตบอลโลก ณ เดือนกรกฎาคม 2026 ยังไม่มีทีวีที่วางขายในบราซิลรุ่นใดมีเครื่องรับ DTV+ ในตัว รุ่นแรกคาดว่าจะออกมาระหว่างปลายปี 2026 ถึงต้นปี 2027

ที่ตัวเข้ารหัส: ตัวลดขนาดภาพแบบนิวรัลของ Netflix

ตัวอย่างการเข้ารหัสแบบนิวรัลในการเล่นวิดีโอที่ชัดเจนที่สุดทำงานมาตั้งแต่ปี 2022 และไม่เปลี่ยนสิ่งใดที่อุปกรณ์ของผู้ชมต้องทำ Netflix เข้ารหัสแต่ละเรื่องไว้หลายความละเอียด และขั้นตอนลดขนาดภาพที่สร้าง 1080p, 720p และ 540p จากต้นฉบับ 4K เดิม ใช้ฟิลเตอร์ตายตัว เช่น Lanczos Netflix แทนที่ขั้นตอนนี้ด้วยโครงข่ายประสาทเทียมขนาดเล็ก ซึ่งประกอบด้วยบล็อกประมวลผลล่วงหน้าตามด้วยบล็อกปรับขนาด และฝึกให้ยกระดับคุณภาพที่เห็นหลังผ่านตัวอัปสเกลมาตรฐานบนอุปกรณ์ Netflix อธิบายว่าเป็นโซลูชันแบบ drop-in คือไม่ต้องเปลี่ยนอะไรอื่นทั้งฝั่งเข้ารหัสและฝั่งไคลเอนต์ ดังนั้น “อุปกรณ์หลายล้านเครื่องที่รองรับการสตรีมของ Netflix ได้ประโยชน์โดยอัตโนมัติ” ในการทดสอบความชอบของ Netflix ผู้ชมราว 77% เลือกภาพจากตัวลดขนาดภาพแบบนิวรัล ผู้ขายอ้าง และเพื่อให้ต้นทุนรับได้ในระดับของ Netflix “โครงข่ายประสาทเทียมเพียงไม่กี่ชั้นก็เพียงพอ”

นี่คือแบบแผนที่ควรจำ โครงข่ายรันเพียงครั้งเดียวต่อหนึ่งเรื่องในศูนย์ข้อมูล ซึ่งเป็นที่ที่การประมวลผลมีราคาถูกและจัดตารางได้ และผลลัพธ์ที่ได้คือสตรีม AV1 หรือ HEVC ธรรมดา

ที่หน้าจอ: super-resolution ในแอปและในทีวี

ที่อีกปลายหนึ่ง โครงข่ายช่วยซ่อมสิ่งที่โคเดกและต้นฉบับทำหายไป เมื่อวันที่ 29 ตุลาคม 2025 YouTube ประกาศฟีเจอร์ Super Resolution สำหรับแอปบนทีวี: วิดีโอที่อัปโหลดด้วยความละเอียดต่ำกว่า 1080p ตั้งแต่ 240p ถึง 720p จะถูกอัปสเกลเป็น HD ด้วยแมชชีนเลิร์นนิง โดยระบุเป้าหมายไว้ที่ 4K เวอร์ชันที่อัปสเกลแล้วจะมีป้ายกำกับ ผู้ชมสลับกลับไปดูต้นฉบับได้ และครีเอเตอร์เลือกไม่เข้าร่วมได้ Meta อธิบายไว้ในเดือนตุลาคม 2025 ว่าบริษัทรัน super-resolution สำหรับวิดีโอทั้งบนเซิร์ฟเวอร์ตอนอัปโหลด บน CPU x86 มาตรฐานโดยร่วมมือกับ Intel และบนไคลเอนต์ตอนเล่น และพบว่าทั้งสองทางเสริมกัน ผู้ผลิตทีวีทำแบบเดียวกันในซิลิคอน: NQ8 AI Gen3 ชิปประมวลผล 8K รุ่นเรือธงของ Samsung โฆษณาว่ามีโครงข่ายประสาทเทียม 768 ชุดสำหรับการอัปสเกล

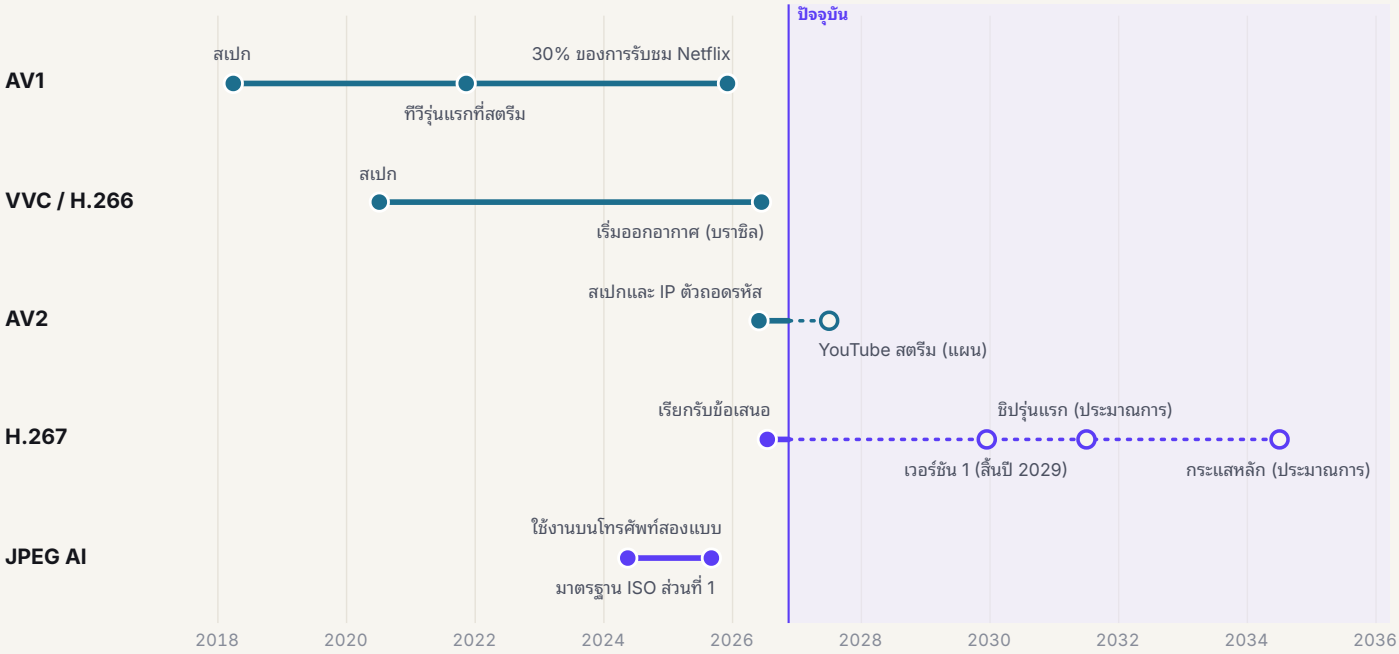
เศรษฐศาสตร์ต่างกันตามตำแหน่งที่โครงข่ายรัน การประมวลผลฝั่งเซิร์ฟเวอร์ทำให้แพลตฟอร์มเสียค่าประมวลผลกับทุกวิดีโอที่ผ่านมือ การประมวลผลฝั่งอุปกรณ์ไม่มีต้นทุนสำหรับแพลตฟอร์ม เพราะใช้ NPU ที่ผู้ซื้อจ่ายเงินไปแล้ว นี่คือเหตุที่ประโยชน์สูงสุดในปี 2026 มาจากการใช้ทั้งสองทาง

ทำไมยังไม่มีทีวีเครื่องใดถอดรหัสบิตสตรีมนิวรัล

ฟอร์แมตสำหรับการเล่นวิดีโอต้องถอดรหัสได้บนทุกอุปกรณ์ที่จะเล่นมัน และอุปกรณ์เหล่านั้นใช้งานกันหลายปี เส้นทางจากมาตรฐานที่เสร็จสมบูรณ์สู่จอกระแสหลักจึงช้า แม้แต่กับโคเดกแบบดั้งเดิม สเปกของ AV1 เสร็จในเดือนมีนาคม 2018 ตัวถอดรหัสซอฟต์แวร์ dav1d ตามมาภายในไม่กี่เดือน ทีวีเริ่มสตรีมตั้งแต่ปลายปี 2021 และขึ้นมาถึงราว 30% ของการรับชม Netflix ในเดือนธันวาคม 2025 เกือบแปดปีหลังสเปกเสร็จ VVC เสร็จในเดือนกรกฎาคม 2020 และได้ใช้ในบริการออกอากาศครั้งแรกในเดือนมิถุนายน 2026

ภาพที่ 03 **จากมาตรฐานที่เสร็จสมบูรณ์ถึงจอกระแสหลักใช้เวลาเกือบทศวรรษ**

การเล่นวิดีโอต้องมีตัวถอดรหัสฮาร์ดแวร์ในอุปกรณ์ AV1 ใช้เวลาเกือบแปดปีจากสเปกที่เสร็จสมบูรณ์จนถึง 30% ของการรับชม Netflix ส่วนมาตรฐานแรกที่น่าจะมีเครื่องมือนิวรัลมีกำหนดปลายปี 2029



ที่มา: Netflix Tech Blog (ธ.ค. 2025); AOMedia; Allegro DVT; TV Technology; Fraunhofer IIS เรื่อง DTV+ ของบราซิล; MPEG-155; คณะกรรมการ JPEG; ISO เครื่องหมายแบบกลวงคือเป้าหมายหรือค่าประมาณการของผู้เขียน

บิตสตรีมนิวรัลเพิ่มปัญหาสองข้อที่โคเดกแบบดั้งเดิมแก้ได้นานแล้ว คือพลังงานต่อพิกเซลและการถอดรหัสให้ตรงกันทุกบิต ทั้งสองเรื่องอยู่ในบทที่ 8 เมื่อรวมกันแล้ว หมายความว่าจะไม่มีผู้ผลิตทีวีหรือโทรศัพท์รายใดใส่ตัวถอดรหัสวิดีโอนิวรัลลงในอุปกรณ์ทุกเครื่อง ก่อนที่จะมีมาตรฐานบอกอย่างชัดเจนว่าต้องสร้างอะไร

สิ่งที่จะเปลี่ยนสถานการณ์: H.267

มาตรฐานที่อาจนำเครื่องมือนิวรัลไปสู่ทุกหน้าจอคือมาตรฐานรุ่นถัดจาก VVC ซึ่ง ITU-T และ ISO/IEC พัฒนาร่วมกันใน Joint Video Experts Team (JVET) ชื่อของมาตรฐานยังไม่เป็นทางการ ผมจะเรียกว่า H.267 ตามที่คนในอุตสาหกรรมจำนวนมากเรียก

ในเดือนตุลาคม 2025 JVET ประเมินผลงานที่ยื่นตอบประกาศเรียกรับหลักฐาน (call for evidence): ผลงานห้าชิ้นจาก 16 บริษัทและห้องแล็บวิจัย โดยทดสอบการตั้งค่าโคเดก 20 แบบ คณะกรรมการพบหลักฐานว่าบีบอัดได้ดีกว่า VVC และเห็นว่าทั้งเทคโนโลยีแบบดั้งเดิมที่ปรับปรุงแล้วและเทคโนโลยีโครงข่ายประสาทเทียม “น่าสนใจ” แนวทางนิวรัลให้ผลดีขึ้นเชิงการรับรู้อย่างชัดเจน โดยเฉพาะกับเนื้อหา 8K HDR และได้ผลน้อยกว่ากับเนื้อหาเกม InterDigital ผู้ยื่นรายหนึ่ง ส่งแบบหนึ่งที่สร้างบน Enhanced Compression Model (ECM) แบบดั้งเดิม และอีกแบบที่วางเครื่องมือ AI ซ้อนบน VVC

ประกาศเรียกรับข้อเสนอ (call for proposals) อย่างเป็นทางการออกในการประชุมเดือนกรกฎาคม 2026 ที่เจนีวา กำหนดส่งข้อเสนอหลักคือสิ้นเดือนตุลาคม 2026 ชุดข้อมูลเพิ่มเติมกำหนดส่งกลางเดือนธันวาคม และจะประเมินข้อเสนอในการประชุม JVET ที่บริสเบนในเดือนมกราคม 2027 JVET คาดว่าจะได้โมเดลทดสอบ (test model) แรกในครึ่งหลังของปี 2027 และมาตรฐานเวอร์ชันแรกภายในสิ้นปี 2029 และเรียกงานนี้ว่า “น่าจะเป็นความพยายามที่ใหญ่กว่าทุกครั้งที่เคยทำมาเพื่อเริ่มต้นมาตรฐานใหม่”

แนวทางแบบดั้งเดิมตั้งมาตรฐานไว้สูง ECM ซึ่งสร้างด้วยการประมวลผลสัญญาณแบบเดิม ประหยัดได้ราว 28% เมื่อเทียบกับ VVC ในการตั้งค่าแบบ random access งานสำรวจด้านโครงข่ายประสาทเทียมของ JVET รายงานฟิลเตอร์ในลูปที่ประหยัดได้ราว 6% ที่การคูณสะสม 10,000 ครั้งต่อพิกเซล 8% ที่ 22,000 ครั้ง และ 14% ที่จุดทำงานความซับซ้อนสูง ผมอ่านว่า H.267 น่าจะเป็นแกนแบบดั้งเดิมที่มีเครื่องมือนิวรัลไม่กี่ตัว เฉพาะในจุดที่บิตที่ประหยัดได้คุ้มกับต้นทุนของตัวถอดรหัส และการประเมินในเดือนมกราคม 2027 จะบอกว่ามีกี่ตัว บทวิเคราะห์

มีสะพานเชื่อมอยู่แล้ว สตรีม VVC และ HEVC แนบข้อมูลเสริม SEI (supplemental enhancement information) ที่อธิบายฟิลเตอร์หลังถอดรหัสแบบนิวรัลให้ฝั่งตัวถอดรหัสนำไปใช้ได้ และฉบับใหม่ของ VVC และของมาตรฐาน SEI ของ VVC (VSEI) ที่เสร็จในเดือนตุลาคม 2025 ได้ขยายข้อความเหล่านั้น สตรีมจึงบอกได้ว่าควรรันโครงข่ายใดหลังถอดรหัส โดยไม่บังคับให้ตัวถอดรหัสทุกตัวต้องรัน

ความหมายต่อธุรกิจด้านการเล่นวิดีโอ

ขายสิ่งที่ใช้ได้รอบโคเดกตอนนี้: การประมวลผลล่วงหน้าฝั่งเข้ารหัสสำหรับผู้ให้บริการ และบล็อกประมวลผลภายหลังพร้อมงานปรับจูนสำหรับชิปทีวีและกล่องรับสัญญาณ อย่าสร้างธุรกิจการเล่นวิดีโอบนบิตสตรีมนิวรัลที่ต้องถอดรหัสบนอุปกรณ์ของบุคคลที่สามก่อนทศวรรษ 2030 หากคุณทำชิลิคอน รายได้ระยะใกล้คือตัวถอดรหัส AV2 และ VVC ส่วนที่เป็นนิวรัลคือเส้นทาง NPU สำหรับการปรับปรุงภาพที่วางคู่กัน (บทที่ 8)

บทที่ 03

03 การถ่ายภาพและภาพนิ่ง: นิวรัลในกล้อง แบบดั้งเดิมบนเว็บ

การถ่ายภาพและภาพนิ่ง

ไปป์ไลน์ของกล้องเป็นนิวรัลแล้ว

ภาพถ่ายทุกภาพจากโทรศัพท์เรือธงรุ่นใหม่ ๆ ผ่านโครงข่ายประสาทเทียมมาก่อนจะถูกบีบอัด Snapdragon 8 Gen 2 ของ Qualcomm ในโทรศัพท์ตั้งแต่ปี 2023 เปิดตัว “Cognitive ISP” ที่ทำการแบ่งส่วนภาพเชิงความหมาย (semantic segmentation) แบบเรียลไทม์ระหว่างถ่ายภาพและวิดีโอ เพื่อปรับใบหน้า เส้นผม ท้องฟ้า และฉากหลังแยกกัน Galaxy S23 Ultra ของ Samsung เป็นโทรศัพท์รุ่นแรกที่ใช้ฟีเจอร์นี้กับกล้องเซลฟี ในเดือนมกราคม 2026 Chips&Media และ Visionary.ai จัดแสดงที่ CES สิ่งที่ทั้งสองบริษัทเรียกว่า ISP แบบ AI เต็มรูปแบบตัวแรก: ไปป์ไลน์ภาพแบบ end-to-end ที่รันทั้งหมดบน NPU คือ WAVE-N ของ Chips&Media ด้วยเลขทศนิยมลอยตัว 16 บิต โดยไม่มี ISP ฟังก์ชันตายตัว และอัปเดตผ่านเครือข่ายไร้สาย (over the air) ได้ “ไปป์ไลน์ที่มีอยู่เอาบล็อกนิวรัลไปติดเสริมบน ISP ฟังก์ชันตายตัว เราแทนที่ ISP แบบเดิมทั้งหมด” ซีอีโอของ Visionary.ai กล่าว หลังงานจบ ระบบนี้เปิดให้ผู้ผลิตนำไปประเมิน

จากนั้นผลลัพธ์ถูกบันทึกเป็น JPEG หรือ HEIF ซึ่งเป็นฟอร์แมตแบบดั้งเดิมที่อุปกรณ์ใดก็เปิดได้ แบบแผนเหมือนกับในการเล่นวิดีโอ: นิวรัลอยู่ภายในไปป์ไลน์ของบริษัทเดียว มาตรฐานอยู่ที่รอยต่อ

JPEG AI: มาตรฐานนิวรัลที่เสร็จแล้ว แต่ยังไม่มีผลิตภัณฑ์

JPEG AI เป็นมาตรฐานการเข้ารหัสภาพแบบเรียนรู้ฉบับแรกจาก ISO/IEC และ ITU-T (ISO/IEC 6048, ITU-T T.840) ส่วนที่ 1 ซึ่งเป็นระบบเข้ารหัสแกนหลัก เผยแพร่ในปี 2025 คณะกรรมการ JPEG รายงานว่าบีบอัดได้ดีกว่า VVC แบบ intra 12.5% ถึง 27.9% ในช่วงความซับซ้อนของตัวถอดรหัสตั้งแต่ 7,000 ถึง 216,000 การคูณสะสมต่อพิกเซล การออกแบบมีตัวเข้ารหัสสองตัว

และตัวถอดรหัสสามตัวที่ฝึกร่วมกัน ผู้รับจึงทำตัวถอดรหัสเพียงตัวเดียวก็ยังอ่านได้ทุกสตรีม ภายในเดือนพฤษภาคม 2024 คณะกรรมการรายงานการใช้งานบนสมาร์ตโฟนสองแบบ ในช่วงเดียวกัน IEEE ได้เผยแพร่ IEEE 1857.11-2024 มาตรฐานการเข้ารหัสภาพด้วยโครงข่ายประสาทเทียมที่ออกแบบสำหรับ NPU ในเดือนธันวาคม 2024

ณ เดือนพฤศจิกายน 2026 ผมไม่พบกล้องโทรศัพท์ แอปภาพถ่าย หรือเบราว์เซอร์ใดที่เขียนหรืออ่าน JPEG AI เป็นค่าเริ่มต้น

สิ่งที่เว็บเลือกใช้แทน: JPEG XL

ฟอร์แมตภาพที่ได้รับความนิยมเพิ่มขึ้นจริงในปี 2026 เป็นแบบดั้งเดิม เมื่อวันที่ 6 ตุลาคม 2026 Google เปิดการถอดรหัส JPEG XL เป็นค่าเริ่มต้นใน Chrome 155 โดยใช้ตัวถอดรหัสใหม่ที่ปลอดภัยด้านหน่วยความจำ (memory-safe) เขียนด้วย Rust สี่ปีหลังจากที่ถอดการรองรับแบบทดลองออกไป Safari ถอดรหัส JPEG XL ได้ตั้งแต่ปี 2023 และ Mozilla ก็กำลังไปทางเดียวกัน เหตุผลที่ Google สนับสนุนคือบีบอัดได้ดีกว่า JPEG บีบอัดไฟล์ JPEG เดิมซ้ำได้โดยไม่สูญเสียข้อมูล รองรับ HDR และการถอดรหัสแบบ progressive

บทเรียนนี้ใช้ได้กว้าง ฟอร์แมตที่ต้องเปิดได้ทุกที่จะชนะด้วยตัวถอดรหัสที่เร็ว ให้ผลแน่นอน รันได้บน CPU ใดก็ได้ และทำให้ปลอดภัยด้านหน่วยความจำได้ ตัวถอดรหัสภาพแบบนิวรัลต้องใช้ NPU หรือ GPU และต้องคำนวณได้ตรงกันแม่นยำข้ามอุปกรณ์ การเข้ารหัสแบบนิวรัลชนะภายในไปป์ไลน์ของอุปกรณ์เอง และในระบบที่บริษัทเดียวกันทั้งเข้ารหัสและถอดรหัส เช่น บริการเก็บภาพบนคลาวด์ที่ไม่เคยส่งไฟล์ที่บีบอัดแล้วให้ใครอื่น

ความหมายต่อธุรกิจด้านภาพ

เงินอยู่ในไปป์ไลน์การถ่ายภาพ: IP ของ AI ISP, NPU ที่มีขนาดเหมาะกับงานภาพ และงานปรับจูนที่ผู้ผลิตกล้องซื้อ สำหรับการจัดเก็บและการส่งมอบ วันนี้โคเดกภาพแบบเรียนรู้เป็นเครื่องมือของระบบปิด ไม่ใช่ฟอร์แมตที่ควรเดิมพัน

บทที่ 04



การโทรและการประชุม

04 การโทรและการประชุม: ที่ที่โคเดกนิวรัลแบบ end-to-end ใช้งานจริง

เสียงมาก่อน

Microsoft เปิดตัว Satin โคเดกเสียงพูดที่ขับเคลื่อนด้วย AI ในเดือนกุมภาพันธ์ 2021 และระบุว่าใช้อยู่แล้วในการโทรสองฝ่ายทั้งหมดบน Teams และ Skype Satin ให้เสียงพูดแบบ super-wideband ซึ่งมีแบนด์วิดท์เสียงสูงสุด 16 kHz ที่บิตเรต 6 kbps ขณะที่ Silk โคเดกรุ่นก่อนของ Skype ให้ได้เพียง narrowband ที่บิตเรตเท่ากัน บิตเรตที่ประหยัดได้นำไปใช้กับข้อมูลซ้ำซ้อน (redundancy) การโทรจึงทนต่อการสูญหายของแพ็กเก็ตได้ดีขึ้น

MLVC: โคเดกวิดีโอนิวรัลในการโทรจริง

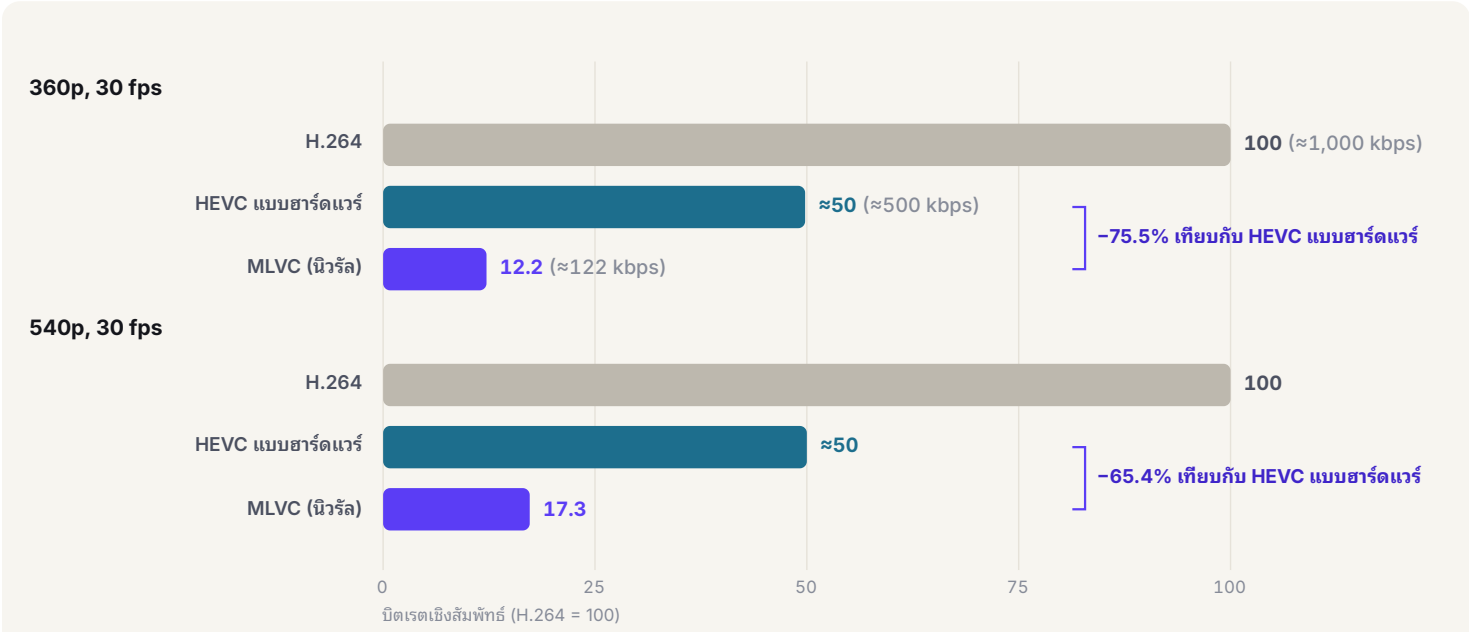
เมื่อวันที่ 24 กรกฎาคม 2026 Microsoft เปิดตัว MLVC โคเดกวิดีโอแบบแมชชีนเลิร์นนิงของบริษัท เป็นโอเพนซอร์สภายใต้สัญญาอนุญาต MIT พร้อมโค้ดโมเดล ค่าน้ำหนักที่ฝึกแล้ว สคริปต์การฝึก และเครื่องมือแปลงที่รองรับ NPU หลายแบบ MLVC คือเวอร์ชันผลิตภัณฑ์ของตระกูล DCVC ที่ Microsoft Research เผยแพร่มาตั้งแต่ปี 2021 Microsoft ระบุว่ากำลังทยอยใช้

MLVC ใน Teams สำหรับการโทรแบบ peer-to-peer พร้อมเทเลเมทรีแบบสด การทดสอบ A/B และใช้โคเดกแบบเดิมเป็นทางสำรอง

MLVC ทำงานที่ 540p และ 360p ที่ 30 เฟรมต่อวินาที และออกแบบให้ใช้ NPU ไม่เกินครึ่งหนึ่งบน Neural Engine ของ Apple และบน NPU ของ Qualcomm และ Intel โดยแบ่งงานระหว่าง CPU กับ NPU ในการทดสอบเชิงอัตวิสัยตาม ITU-T P.910 บน Video Conferencing Dataset ของ Microsoft เอง MLVC ใช้บิตเรตน้อยกว่า H.264 87.8% และน้อยกว่า HEVC แบบฮาร์ดแวร์ 75.5% ที่ 360p และน้อยกว่า 82.7% และ 65.4% ตามลำดับที่ 540p ผู้ขายอ้าง ที่ 360p คิดเป็นราว 122 kbps ในขณะที่ H.264 ต้องใช้ราว 1 Mbps

ภาพที่ 04 **ในการโทร โคเดกนิวรัลใช้บิตเรตน้อยกว่า HEVC แบบฮาร์ดแวร์ 65–76%**

บิตเรตที่ต้องใช้เพื่อให้ได้คุณภาพตามการให้คะแนนเท่ากัน บนชุดทดสอบการประชุมทางวิดีโอของ Microsoft โดย H.264 = 100 MLVC ทำงานบน NPU ของแล็ปท็อปและโทรศัพท์ที่สูงสุด 540p30 และใช้ NPU ไม่เกินครึ่งหนึ่ง



ที่มา: Microsoft, ประกาศเปิดตัว MLVC (24 กรกฎาคม 2026): การทดสอบเชิงอัตวิสัยตาม ITU-T P.910 บน Video Conferencing Dataset ของบริษัท แท่ง HEVC คำนวณจากค่าการประหยัดที่รายงาน เป็นตัวเลขที่ผู้ขายอ้าง ไม่ใช่การวัดโดยอิสระ

ตัวเลขพาดหัวมีข้อควรระวังสองข้อ ตัวเลข 88% เทียบกับ H.264 หากเทียบกับ HEVC แบบฮาร์ดแวร์ที่มีอยู่แล้วในแล็ปท็อปเครื่องเดียวกัน จะเหลือ 65–76% ซึ่งยังถือว่ามาก และ MLVC ยอมสละประสิทธิภาพการบีบอัดเพื่อให้นำไปใช้งานได้จริง: บทวิเคราะห์อิสระชิ้นหนึ่งระบุว่า DCVC-RT ที่ฝึกเชิงการรับรู้ได้ 81.9% ในตัวชี้วัดเดียวกันที่ MLVC ได้ 75.5% ราคาของการถอดรหัสให้ได้ผลเหมือนกันบนฮาร์ดแวร์ที่ Microsoft ไม่ได้เป็นเจ้าของคือราวหกจุด

ทำไมการโทรจึงเป็นบ้านโดยธรรมชาติ

การโทรคือวงปิด ปลายทางทั้งสองฝั่งรันแอปพลิเคชันเดียวกัน ตกลงความสามารถกันตอนเริ่มสาย และถอยกลับไปใช้โคเดกมาตรฐานได้กลางสาย ความละเอียดไม่สูงนัก เครือข่ายมักไม่ดี และแล็ปท็อปทั้งสองฝั่งวันนี้มาพร้อม NPU เช่นเดียวกับโทรศัพท์ที่มีมาหลายปีแล้ว Microsoft ตั้งเกณฑ์ NPU ไว้ที่ 40 TOPS สำหรับ AI PC ของตน และ Neural Engine ของ Apple M4 มีพิกัด 38 TOPS เงื่อนไขทุกข้อที่ขวางโคเดกนิวรัลในการเล่นวิดีโอไม่มีอยู่ที่นี่

ตัวอย่างแย้ง: Opus และ MLow

แพลตฟอร์มการโทรไม่ได้เลือกแบบเดียวกันทั้งหมด และเหตุผลก็ให้บทเรียนที่ดี

Opus โคเดกเสียงแบบเปิดใน WebRTC และแอปโทรส่วนใหญ่ เพิ่มดีปเลิร์นนิงในเวอร์ชัน 1.5 (มีนาคม 2024) ซึ่งเป็นครั้งแรกที่ใช้โครงข่ายประมวลผลหรือสร้างสัญญาณโดยตรง แต่เพิ่มเข้ามาเป็นตัวเลือก: การปรับปรุงเสียงฝั่งตัวถอดรหัส (LACE และ NoLACE) ที่ไม่ต้องเปลี่ยนบิตสตรีม การปกปิดการสูญหายของแพ็กเก็ตแบบนิวรัล และกลไกข้อมูลซ้ำซ้อนเชิงลึก (DRED) ทั้งหมดปิดไว้เป็นค่าเริ่มต้นเพื่อไม่ให้เป็นภาระกับอุปกรณ์รุ่นเก่า Opus 1.6 เพิ่มการขยายแบนด์วิดท์แบบนิวรัลและโมเดล DRED ที่เล็กลงสามเท่า และยังเข้ากันได้เต็มที่กับมาตรฐานปี 2012 แนวทางนี้คือ “ในมาตรฐาน เลือกใช้ได้ และเข้ากันได้ย้อนหลัง”

Meta ถอยไปไกลกว่านั้น ในเดือนมิถุนายน 2024 Meta เปิดตัว MLow โคเดกเสียงพูดบิตเรตต่ำที่สร้างบนการประมวลผลสัญญาณแบบดั้งเดิมแทนโครงข่ายประสาทเทียม ปัจจุบันใช้ในการโทรทั้งหมดบน Instagram และ Messenger และกำลังทยอยใช้บน WhatsApp เหตุผลของ Meta คืออุปกรณ์ของผู้ใช้: การโทรกว่า 20% ทำบนโทรศัพท์ ARMv7 และการโทรผ่าน WhatsApp หลายสิบล้านครั้งต่อวันทำบนโทรศัพท์ที่มีอายุเกินสิบปี Meta เขียนว่าโคเดกนิวรัล รวมถึง EnCodec ของ Meta เอง ทำงานได้อย่างเชื่อถือได้เฉพาะบนเครื่องราคาแพง MLow ให้คุณภาพราวสองเท่าของ Opus ที่ 6 kbps โดยใช้การประมวลผลน้อยกว่า 10%

สำหรับบริการที่ผู้ใช้มีตั้งแต่แล็ปท็อปเรือธงไปจนถึงโทรศัพท์อายุสิบปี งบความซับซ้อนคือตัวตัดสิน โคเดกนิวรัลจะใช้งานจริงก่อนในที่ที่ฐานอุปกรณ์ที่ใช้งานอยู่เป็นรุ่นใหม่ และผู้ขายควบคุมทั้งสองฝั่ง

ความหมายต่อธุรกิจด้านการสื่อสาร

การประชุมทางไกลในองค์กร เทเลเพรเซนซ์ การช่วยเหลือระยะไกล ลิงก์ส่งสัญญาณต้นทาง (contribution link) และเครือข่ายส่วนตัว คือที่ที่โคเดกนิวรัลแบบ end-to-end เป็นผลิตภัณฑ์ได้ในวันนี้ หากทั้งสองฝั่งเป็นของคุณ วางแผนสามเรื่องตั้งแต่ต้น: โคเดกสำรอง การทดสอบการถอดรหัสบน NPU ทุกตัวที่คุณรองรับ และการประเมินคุณภาพบนเนื้อหาของคุณเองแทนชุดข้อมูลของผู้ขาย

บทที่ 05

05 Generative AI

Generative AI: การใช้งานที่ใหญ่ที่สุดซึ่งไม่มีใครเรียกว่าโคเดก

โมเดลเสียงพูดและวิดีโอทุกตัวทำงานบนโคเดกแบบเรียนรู้

การใช้งานการบีบอัดนิวรัลแบบ end-to-end ที่ใหญ่ที่สุดไม่ได้อยู่ในมาตรฐานโคเดกใด แต่อยู่ภายใน Generative AI โมเดลเสียงพูดไม่ได้ฟังหรือสร้างรูปคลื่นโดยตรง แต่อ่านและเขียนโทเค็นแบบไม่ต่อเนื่อง (discrete token) ของโคเดกเสียงแบบเรียนรู้ Mimi ของ Kyutai ซึ่งเปิดตัวในปี 2024 แปลงเสียงพูดเป็นโทเค็นที่ 12.5 เฟรมต่อวินาทีและ 1.1 kbps ต่ำกว่าโคเดกโทรศัพท์ใด ๆ โดยรวมข้อมูลเชิงความหมายและเชิงเสียงไว้ในสตรีมเดียว EnCodec ของ Meta ซึ่งเผยแพร่ในปี 2022 เป็นอีกตัวหนึ่ง ภายหลังวิศวกรของ Meta เองอธิบายว่าโคเดกลักษณะนี้ยอดเยี่ยมที่บิตเรตต่ำ แต่หนักเกินไปสำหรับโทรศัพท์รุ่นเก่า (บทที่ 4)

การสร้างวิดีโอก็ทำงานแบบเดียวกัน โมเดล diffusion หรือ transformer สร้างผลลัพธ์ใน latent space ของ video autoencoder (VAE) แบบเรียนรู้ แล้วตัวถอดรหัสของ VAE จึงแปลงผลลัพธ์เป็นพิกเซล VAE ของ LTX-Video บีบอัดวิดีโอที่อัตรา 1:192: โทเค็น latent แต่ละตัวแทนบล็อก 32×32 พิกเซลตลอด 8 เฟรม โดยมี 128 แชนเนล ตัวถอดรหัสของ VAE นี้ยังทำขั้นตอนลดสัญญาณรบกวนขั้นสุดท้ายด้วย จึงสร้างรายละเอียดขึ้นมาเองส่วนหนึ่ง แทนที่จะเพียงสร้างคืน

ทำไมสิ่งเหล่านี้จึงไม่ใช่ฟอร์แมตสำหรับเผยแพร่

โคเดกเหล่านี้ทำงานในทุกคำขอที่ส่งไปยังโมเดลเสียงพูดหรือวิดีโอ แต่ไม่มีตัวใดเป็นฟอร์แมตที่คนอื่นถอดรหัส แต่ละโมเดลมี latent space ของตัวเอง ไม่มีการควบคุมบิตเรต ไม่มีการทดสอบความสอดคล้อง และไม่มีการรับประกันว่าตัวถอดรหัสตัวอื่นจะให้ภาพเดียวกัน โคเดกเหล่านี้ออกแบบมาเพื่อคุณภาพของการสร้าง ไม่ใช่เพื่อความเที่ยงตรงที่บิตเรตหนึ่ง ๆ และถอดรหัสบน GPU ในศูนย์ข้อมูล ความสำเร็จของโคเดกเหล่านี้แสดงว่าการบีบอัดแบบเรียนรู้ใช้ได้ในวงกว้างเมื่อบริษัทเดียวเป็นเจ้าของทั้งตัวเข้ารหัส ตัวถอดรหัส และฮาร์ดแวร์

ความหมายต่อธุรกิจ

มีสองประเด็นในทางปฏิบัติ ข้อแรก ประสิทธิภาพของ tokenizer เป็นรายการต้นทุนโดยตรงสำหรับใครก็ตามที่ให้บริการวิดีโอหรือเสียงจาก Generative AI: สำหรับ transformer ที่ทำ attention ครอบคลุมทุกโทเค็น การประมวลผลเพิ่มขึ้นเร็วกว่าจำนวนโทเค็น ค่า latent ที่บีบอัดได้มากขึ้นจึงลดต้นทุนการสร้างได้มากกว่าสัดส่วน บทวิเคราะห์ วิศวกรโคเดกที่เข้าใจการออกแบบ latent โมเดลเอนโทรปี และประสิทธิภาพของตัวถอดรหัส มีคุณค่าต่อทีม Generative AI ข้อสอง เส้นแบ่งกำลังเคลื่อน: องค์กรมาตรฐานเริ่มงานด้านการเข้ารหัสสำหรับเครื่องจักรแทนมนุษย์แล้ว Video Coding for Machines ของ MPEG (ISO/IEC 23888-2) ถึงขั้นร่างมาตรฐานสากลในต้นปี 2026 และมีแผนเป็นฉบับสมบูรณ์ในปลายปี 2026 ส่วน JPEG AI ออกแบบตั้งแต่ต้นให้รองรับงานคอมพิวเตอร์วิทัศน์ควบคู่กับการรับชมของมนุษย์

บทที่ 06

เกมและกราฟิก

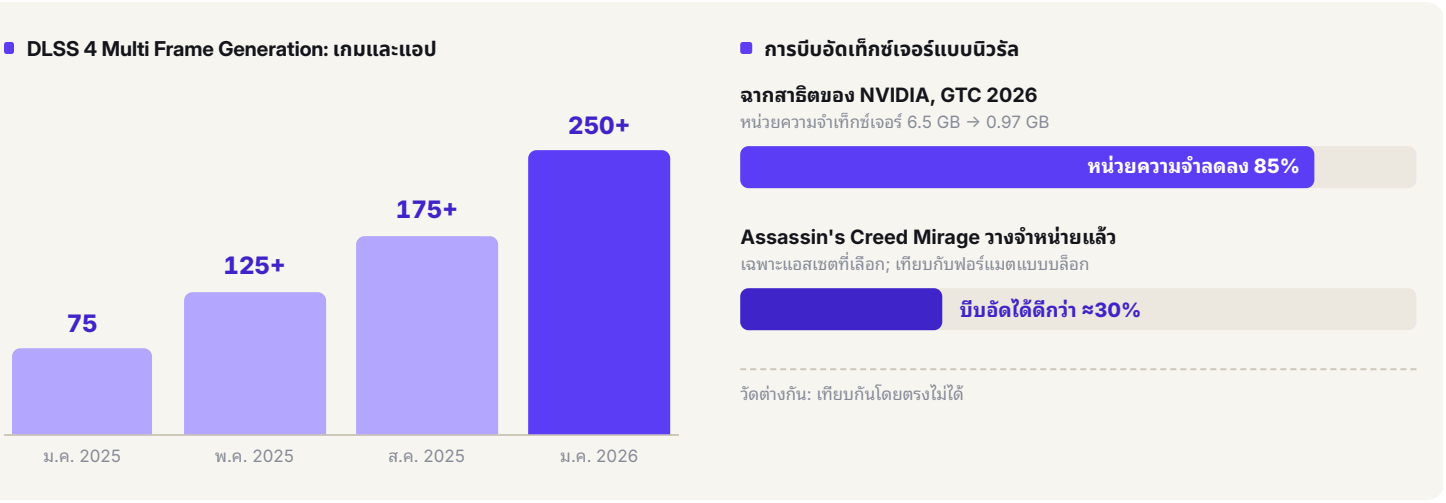
06 เกมและกราฟิก: การสร้างภาพขึ้นใหม่แบบนิวรัลเป็นกระแสหลักแล้ว แอสเซตนิวรัลกำลังเริ่มต้น

เรนเดอร์พิกเซลน้อยลง แล้วให้โครงข่ายเติมส่วนที่เหลือ

ในกราฟิกแบบเรียลไทม์ โครงข่ายสร้างภาพส่วนใหญ่ที่ผู้เล่นเห็นขึ้นใหม่อยู่แล้ว DLSS ของ NVIDIA เรนเดอร์พิกเซลและเฟรมน้อยลง แล้วใช้โครงข่ายสร้างส่วนที่เหลือ DLSS 4 พร้อม Multi Frame Generation เปิดตัวในเดือนมกราคม 2025 โดยรองรับเกมและแอป 75 รายการ NVIDIA นับได้กว่า 125 รายการภายในเดือนพฤษภาคม 2025 กว่า 175 รายการภายในเดือนสิงหาคม และกว่า 250 รายการภายในเดือนมกราคม 2026 ซึ่งเป็นเดือนที่ประกาศ DLSS 4.5 พร้อมโมเดล transformer รุ่นที่สองสำหรับ super-resolution นี่คือตัวถอดรหัสในทุกแง่ยกเว้นชื่อ: “บิตสตรีม” คือเฟรมความละเอียดต่ำบวกข้อมูลการเคลื่อนไหว แล้วโครงข่ายสร้างภาพเต็มขึ้นใหม่

ภาพที่ 05 **เมื่อบริษัทเดียวควบคุมทั้งสแตก การสร้างภาพขึ้นใหม่แบบนิวรัลก็แพร่หลายเร็ว**

ซ้าย: เกมและแอปที่รองรับ DLSS 4 Multi Frame Generation ของ NVIDIA ขวา: การบีบอัดเท็กซ์เจอร์แบบนิวรัลในฉากสาธิต และในเกมที่วางจำหน่ายแล้ว ซึ่งใช้เฉพาะกับแอสเซตที่เลือกไว้เพราะต้นทุนขณะรัน



ที่มา: NVIDIA (มกราคม 2025 ถึงมกราคม 2026; การสาธิตที่ GTC 2026); Ubisoft La Forge (กุมภาพันธ์ 2026) เป็นตัวเลขที่ผู้ขายอ้าง และตัวเลขเท็กซ์เจอร์ทั้งสองวัดสิ่งที่ต่างกัน

เท็กซ์เจอร์นิวรัลใช้งานจริงแล้ว แบบเลือกใช้

เท็กซ์เจอร์ถูกบีบอัดในฟอร์แมตบล็อกแบบตายตัว (BCn) ที่ GPU ถอดรหัสด้วยฮาร์ดแวร์ การบีบอัดเท็กซ์เจอร์แบบนิวรัลแทนชุดเท็กซ์เจอร์ของวัสดุด้วยการแทนข้อมูลแบบ latent ที่กะทัดรัด และโครงข่ายขนาดเล็กที่สร้างเท็กเซลขึ้นใหม่เมื่อต้องใช้ ในเดือนกุมภาพันธ์ 2026 ห้องแล็บ La Forge ของ Ubisoft อธิบายการนำเทคนิคนี้มาใช้จริงใน Assassin's Creed Mirage เกมปี 2023 โดยบีบอัดได้มีประสิทธิผลสูงกว่าฟอร์แมตแบบเดิมราว 30% ที่คุณภาพภาพเท่ากัน ผู้ขายอ้าง แต่ไม่ได้ใช้ทุกที่: “การใช้งานถูกจำกัดด้วยประสิทธิภาพขณะรัน” Ubisoft จึงใช้กับแอสเซตบางส่วนที่มีจำนวนอินสแตนซ์สูงและสร้างแรงกดดันต่อหน่วยความจำมาก

RTX Neural Texture Compression SDK ของ NVIDIA ซึ่งเป็นเบตาบน GitHub แสดงเพดานที่ทำได้ ในการสาธิตที่ GTC 2026 ฉากวิลลาที่ใช้หน่วยความจำเท็กซ์เจอร์ 6.5 GB เมื่อใช้ BCn ต้องการเพียง 970 MB เมื่อใช้เท็กซ์เจอร์นิวรัล หรือน้อยลง 85% ผู้ขายอ้าง SDK มีสองทาง: ถอดรหัสเป็น BCn ตอนโหลด ซึ่งประหยัดพื้นที่ดิสก์และขนาดดาวน์โหลดแต่ไม่ประหยัดหน่วยความจำวิดีโอ หรือถอดรหัสตอนสุ่มตัวอย่าง (sampling) ซึ่งประหยัดหน่วยความจำแต่เสียเวลาเชเดอร์ทุกเฟรม

API กำลังมา

API กราฟิกกำลังเพิ่มการคำนวณเมทริกซ์ที่การถอดรหัสแบบนิวรัลต้องใช้ภายในเชเดอร์ Microsoft ออก Shader Model 6.9 ในเดือนกุมภาพันธ์ 2026 ส่วนฟีเจอร์ cooperative vector ที่บริษัทเปิดให้ทดลองในปี 2025 ถูกรวมเข้าเป็นการออกแบบพีชคณิตเชิงเส้นแบบเดียวสำหรับ shader model รุ่นหลัง โดย DirectX Linear Algebra เข้าสู่ public preview ในเดือนเมษายน 2026 เมื่อสิ่งนี้อยู่ในไดรเวอร์ทุกตัว ตัวถอดรหัสเท็กซ์เจอร์นิวรัลจะรันบน GPU ของผู้ผลิตรายใดก็ได้จากโค้ดชุดเดียว

ความหมายต่อธุรกิจ

เกมใช้การถอดรหัสแบบนิวรัลได้จริงเพราะเงื่อนไขของวงปิดครบถ้วน: เอนจินเดียว ไปป์ไลน์แอสเซตเดียว ฮาร์ดแวร์เทนเซอร์ใน GPU และฟอร์แมตสำรอง งานที่ยังเปิดอยู่คือเครื่องมือ ไปป์ไลน์แอสเซตที่เข้ารหัสวัสดุแบบนิวรัล รันไทม์ที่ใช้ได้ข้ามผู้ผลิต และเหนือสิ่งอื่นใดคืออุปกรณ์พกพา ซึ่งต้นทุนพลังงานต่อเท็กเซลที่ถอดรหัสเป็นตัวตัดสิน

บทที่ 07

07

ลิงก์เฉพาะทาง

ลิงก์เฉพาะทาง: ดาวเทียม โดรน และหุ่นยนต์

นี่คือกรณีที่การเข้ารหัสนิวรัลแบบ end-to-end เหมาะที่สุดบนกระดาษ เจ้าของรายเดียวควบคุมทั้งสองปลายของลิงก์ แบนด์วิดท์มีจำกัดและข้อมูลสูญหายง่าย เนื้อหาแคบพอที่จะฝึกโมเดลได้ และภาพที่เที่ยงตรงในเชิงการรับรู้ หรือแม้แต่เพียงในเชิงความหมายก็มักเพียงพอ โคเดกนิวรัลที่ทนต่อการสูญหายของข้อมูลสำหรับลิงก์แบบนี้เป็นสาขาวิจัยที่คึกคัก และมีการใช้การฟื้นฟูภาพจากฟีดบิตเรตต่ำฝั่งเซิร์ฟเวอร์แล้ว

ขณะเดียวกัน ลิงก์เหล่านี้ก็เป็นตลาดที่เล็กและถูกกำกับเข้มขึ้นเรื่อย ๆ เมื่อเทียบกับวิดีโอสำหรับผู้บริโภค บทวิเคราะห์ ในจีน ประกาศฉบับที่ 49 ปี 2026 ของกระทรวงพาณิชย์และหน่วยงานอื่นอีกสองแห่ง ซึ่งออกเมื่อวันที่ 19 ตุลาคมและมีผลวันที่ 18 พฤศจิกายน 2026 ปรับปรุงการควบคุมการส่งออกสินค้าเกี่ยวกับโดรน 13 หมวด ในรายการด้านเทคโนโลยี มีรายการหนึ่งครอบคลุมแพลตฟอร์มภาคพื้นดินไร้คนขับที่นิยามส่วนหนึ่งด้วยเกณฑ์อย่างการบีบอัดภาพเชิงความหมาย 75% และการสร้างภาพคืนภายใน 150 มิลลิวินาที ตามที่สื่อการค้ารายงาน การบีบอัดเชิงความหมายจึงกลายเป็นพารามิเตอร์หนึ่งของการควบคุมการส่งออกไปแล้ว ในสหรัฐฯ คณะกรรมาธิการการสื่อสารกลาง (FCC) เพิ่มโดรนที่ผลิตในต่างประเทศและชิ้นส่วนสำคัญเข้าใน Covered List เมื่อวันที่ 22 ธันวาคม 2025 โดรนรุ่นใหม่จากต่างประเทศจึงไม่สามารถได้รับการอนุญาตอุปกรณ์ (equipment authorization) อีกต่อไป

ข้อสรุป: เป็นตลาดเฉพาะที่ดีสำหรับบริษัทเฉพาะทางที่มีลูกค้าภาคอุตสาหกรรมหรือภาครัฐ และมีความสามารถจัดการการปฏิบัติตามกฎควบคุมการส่งออกของทั้งสองฝั่ง แต่ไม่ใช่ตลาดกระแสหลัก และไม่ใช่ที่สำหรับพิสูจน์ว่าการเข้ารหัสแบบนิวรัลใช้ได้ในวงกว้างหรือไม่

บทที่ 08

ซิลิคอน

08 มุมมองด้านซิลิคอน: พลังงาน การประมวลผล หน่วยความจำ และความตรงกันทุกบิต

นี่คือบทที่ภูมิหลังของผมมีความหมายมากที่สุด ผมมีประสบการณ์หลายปีในงานวิจัยและพัฒนาและการบริหารชิปแบบ fabless รวมถึงการออกแบบ datapath ของ ASIC แบบ full-custom และโครงการชิปกล้องมัลติมีเดีย เช่น ตัวเร่งโคเดกวิดีโอ และคำถามที่ทีมชิปถามเกี่ยวกับโคเดกใหม่ไม่ได้อยู่ที่ BD-rate แต่คือ: ใช้กี่จูลต่อพิกเซล ทราฟฟิก DRAM กี่ไบต์ต่อพิกเซล ใช้กี่เกต และตัวถอดรหัสทุกตัวในโลกคำนวณได้บิตเดียวกันทุกบิตหรือไม่? ในทั้งสี่ข้อ ตัวถอดรหัสนิวรัลในปี 2026 อยู่ในจุดเดียวกับที่ HEVC แบบฮาร์ดแวร์เคยอยู่ก่อนจะมีชิปรุ่นแรก: มีอนาคตบนซอฟต์แวร์ แต่แพงบนซิลิคอน

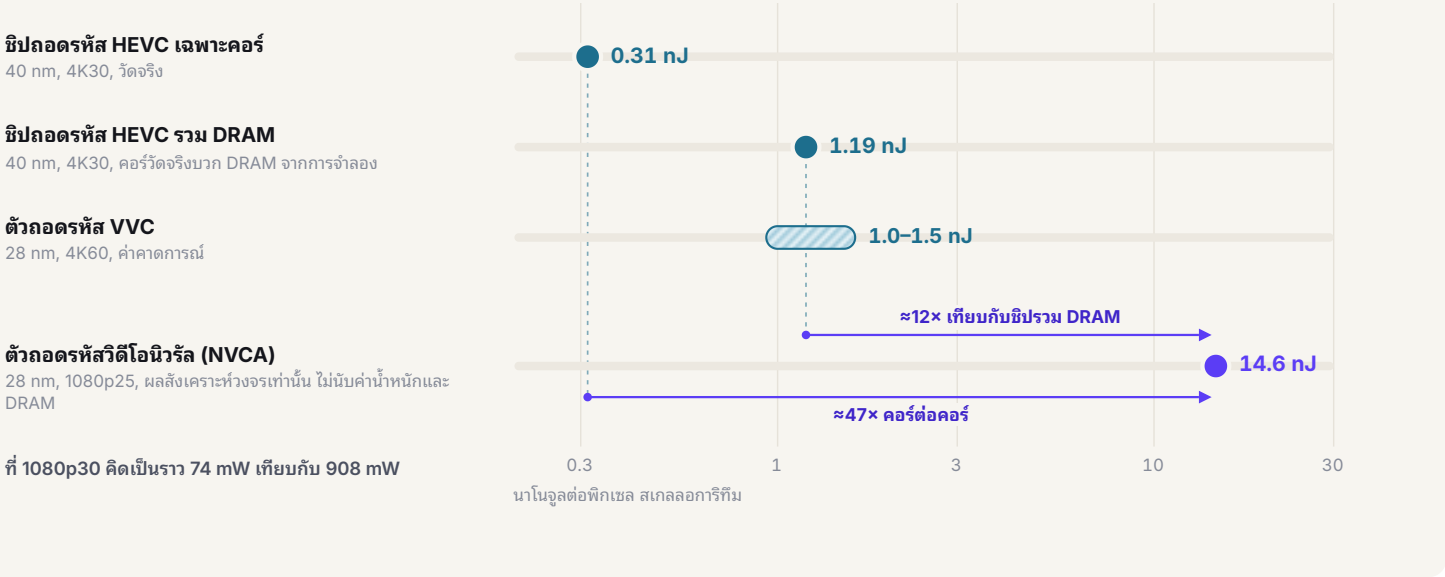
พลังงานต่อพิกเซล: ห่างกันระดับสิบเท่า

ตัวถอดรหัสนิวรัลที่ถูกออกแบบเป็นชิปมีอยู่น้อย การเปรียบเทียบจึงอาศัยการออกแบบที่เผยแพร่เพียงไม่กี่ชิ้น จุดอ้างอิงฝั่งแบบดั้งเดิมคือชิปถอดรหัส HEVC ของ MIT (Tikekar et al., ISSCC 2013 และ IEEE JSSC 2014): 40 nm, 1.77 mm² ถอดรหัส 4K ที่ 30 เฟรมต่อวินาทีที่ 200 MHz คอร์ใช้กำลังไฟฟ้า 78 mW หรือ 0.31 nJ ต่อพิกเซล เมื่อรวมกำลังไฟฟ้าของ DRAM จากการจำลอง ทั้งระบบใช้ 1.19 nJ ต่อพิกเซล สำหรับ VVC ตัวเลขของชิปที่ผลิตจริงเป็นความลับทางการค้า บทวิเคราะห์ใน Electronic Design ปี 2026 ซึ่งปรับขนาดชิปของ MIT ไปที่ 28 nm และคิดความซับซ้อนที่สูงขึ้นของ VVC คาดการณ์ไว้ที่ 1–1.5 nJ ต่อพิกเซล สำหรับตัวถอดรหัส 4K60 ขนาดราวหนึ่งล้านเกต

ฝั่งนิวรัล การออกแบบ ASIC ตัวถอดรหัสวิดีโอนิวรัลแบบ end-to-end เพียงแบบเดียวที่มีการเผยแพร่ ตามการนับของ Electronic Design คือ NVCA (Zhang et al., DATE 2024) ซึ่งสังเคราะห์วงจรบนกระบวนการ TSMC 28 nm ด้วย 5.01 ล้านเกต โดยไม่นับค่าน้ำหนักซึ่งเก็บไว้นอกชิป NVCA ถอดรหัส 1080p ที่ 25 เฟรมต่อวินาทีโดยใช้ 760 mW หรือ 14.6 nJ ต่อพิกเซล

ภาพที่ 06 **พลังงานต่อพิกเซลที่ถอดรหัส: ตัวถอดรหัสฟังก์ชันตายตัวนำอยู่ระดับสิบเท่า**

การออกแบบชิปถอดรหัสวิดีโอนิวรัลเพียงแบบเดียวที่มีการเผยแพร่ใช้พลังงานต่อพิกเซลราว 12 เท่าของชิปถอดรหัส HEVC ที่วัดจริงเมื่อรวม DRAM และราว 47 เท่าของคอร์ของชิปนั้นอย่างเดียว ทั้งนี้ยังไม่นับทราฟฟิกค่าน้ำหนักและ DRAM ของการออกแบบนิวรัลเอง



ที่มา: Tikekar et al., IEEE JSSC 2014 (MIT); NVCA, Zhang et al., DATE 2024 (arXiv 2312.10716); ช่วงของ VVC เป็นการคาดการณ์ของ Electronic Design (2026) ชิป HEVC อยู่บนโหนดการผลิตที่เก่ากว่า (40 nm เทียบกับ 28 nm) ซึ่งเข้าข้างการออกแบบนิวรัล

นั่นคือราว 12 เท่าของชิป HEVC เมื่อรวม DRAM และราว 47 เท่าของคอร์อย่างเดียว การเปรียบเทียบนี้เข้าข้าง NVCA สามทาง: ชิป HEVC อยู่บนโหนดการผลิตที่เก่ากว่า ตัวเลขของ NVCA ไม่รวมทราฟฟิกของค่าน้ำหนักและ DRAM และ NVCA เป็นผลจากการสังเคราะห์วงจร ไม่ใช่ซิลิคอนที่วัดจริง ที่ 1080p30 ความต่างคือราว 74 mW เทียบกับราว 900 mW: ตัวถอดรหัสที่ไม่มีใครสังเกต กับตัวที่ส่งผลให้เห็นในงบแบตเตอรี่และความร้อนของโทรศัพท์

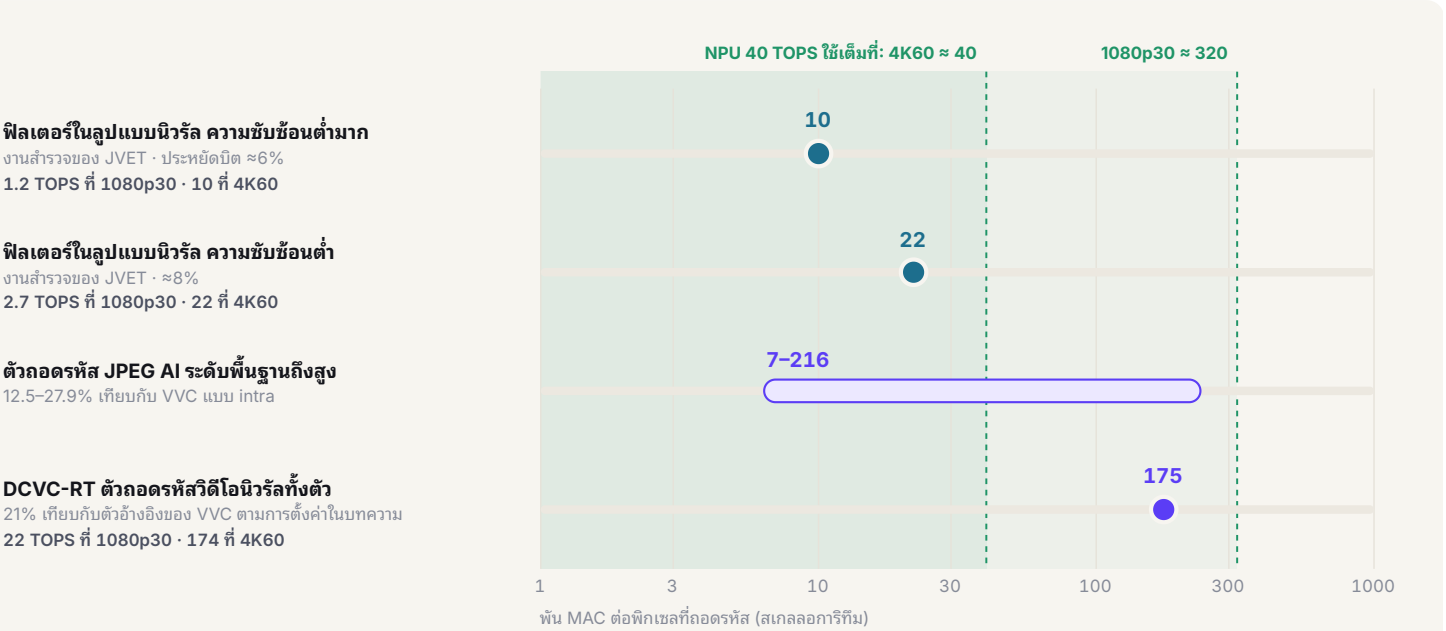
การประมวลผลต่อพิกเซล: สิ่งที่ตัวถอดรหัสนิวรัลเรียกร้องจาก NPU

โคเดกวิดีโอนิวรัลที่เร็วที่สุดเท่าที่ผมพบในงานที่เผยแพร่คือ DCVC-RT ของ Microsoft Research (Jia et al., CVPR 2025) ซึ่งเข้ารหัสและถอดรหัส 1080p ที่ 125 และ 113 เฟรมต่อวินาทีบน GPU NVIDIA A100 พร้อมประหยัดบิตเรต 21% เมื่อเทียบกับตัวเข้ารหัสอ้างอิงของ VVC ในการตั้งค่าทดสอบของบทความ และใช้เลขคณิตจำนวนเต็มเพื่อให้อุปกรณ์ต่างกันถอดรหัสได้ตรงกัน เอกสารเสริมของบทความระบุความซับซ้อนของแต่ละโมดูล เมื่อรวมสี่โมดูลที่ตัวถอดรหัสต้องรัน (ตัวถอดรหัส ตัวสกัดฟีเจอร์ โมเดลเอนโทรปี และส่วนสร้างภาพคืน) ผมได้ราว 175,000 การคูณสะสมต่อพิกเซลที่ถอดรหัส ประมาณการ

ที่ 1080p30 คิดเป็นราว 11 ล้านล้านการคูณสะสมต่อวินาที หรือราว 22 TOPS เมื่อนับสองการดำเนินการต่อ MAC ก่อนหักการสูญเสียอัตราการใช้งานใด ๆ ส่วน 4K60 ต้องใช้เป็นแปดเท่า ประมาณการ เพื่อให้เห็นภาพ Microsoft นิยาม AI PC ด้วย NPU อย่างน้อย 40 TOPS และ Apple ระบุพิกัด Neural Engine ของ M4 ไว้ที่ 38 TOPS ตัวเลขเหล่านี้เป็นค่าสูงสุด และต้องแบ่งใช้กับงานอื่นทั้งหมดของอุปกรณ์

ภาพที่ 07 **เครื่องมือนิวรัลหนึ่งตัวใช้หลายหมื่น MAC ต่อพิกเซล ตัวถอดรหัสนิวรัลทั้งตัวใช้หลายแสน**

จำนวนการคูณสะสม (MAC) ของตัวถอดรหัสต่อพิกเซล สเกลลอการิทึม เทียบกับที่ NPU 40 TOPS ซึ่งใช้งานเต็มที่จะรองรับได้ที่ 4K60 และที่ 1080p30 ค่าที่ประหยัดได้วัดเทียบกับจุดอ้างอิงต่างกัน จึงเทียบกันโดยตรงไม่ได้



ที่มา: JVET ผ่านประธาน SC 29 (เมษายน 2026); คณะกรรมการ JPEG (2024); DCVC-RT เอกสารเสริมของ CVPR 2025 ผลรวมโมดูลตัวถอดรหัสโดยผู้เขียน; นิยาม Copilot+ PC ของ Microsoft การคำนวณของผู้เขียน: สองการดำเนินการต่อ MAC ใช้งานเต็มที่

ขั้นบันไดในภาพที่ 7 เป็นวิธีอ่านภาพรวมของสาขานี้ที่มีประโยชน์ เครื่องมือนิวรัลที่ใช้หลายหมื่น MAC ต่อพิกเซล เช่น ฟิลเตอร์ในลูปของ JVET (ประหยัดราว 6% ที่ 10,000 และ 8% ที่ 22,000) มีความเป็นไปได้ที่จะอยู่ในตัวถอดรหัสมาตรฐานในอนาคต JPEG AI ครอบคลุมทั้งช่วง ตั้งแต่ 7,000 ถึง 216,000 MAC ต่อพิกเซล สำหรับการประหยัด 12.5–27.9% เมื่อเทียบกับ VVC แบบ intra ตัวถอดรหัสวิดีโอนิวรัลทั้งตัวที่ราวสองแสน MAC ต่อพิกเซลควรอยู่บน NPU ที่ความละเอียดไม่สูงนัก ในวงปิด ซึ่งตรงกับตำแหน่งของ MLVC พอดี: 540p30 โดยใช้ NPU ไม่เกินครึ่งหนึ่ง

ทราฟฟิกหน่วยความจำ ไม่ใช่ MAC ที่มักเป็นข้อจำกัดจริง

ผู้เขียน DCVC-RT พบว่าสิ่งที่จำกัดความเร็วไม่ใช่การคำนวณ แต่เป็น “ต้นทุนเชิงปฏิบัติการ” คือการรับส่งข้อมูลเข้าออกหน่วยความจำและจำนวนการเรียกฟังก์ชัน ผู้เขียนตัดโมดูลการเคลื่อนไหวแบบชัดแจ้ง (explicit motion) และการประมวลผลหลายสเกลออกไปก็ด้วยเหตุนี้เป็นหลัก ผู้ออกแบบ NVCA สร้าง dataflow พิเศษเพื่อลดทราฟฟิกหน่วยความจำนอกชิปจากฟีเจอร์ด้านการเคลื่อนไหวและส่วนเหลือ (residual) นักออกแบบชิปจะจำแบบแผนนี้ได้ ในชิป HEVC ของ MIT นั้น DRAM คิดเป็นราวสามในสี่ของพลังงานทั้งระบบ (0.88 จาก 1.19 nJ ต่อพิกเซล) แม้จะมีแคชที่ลดแบนด์วิดท์ DRAM ลง 67% แล้วก็ตาม บทวิเคราะห์ ตัวถอดรหัสนิวรัลที่เก็บ feature map ขนาดใหญ่หลายแชนเนลไว้ระหว่างเฟรม ทำให้ปัญหาหน่วยความจำแย่ลง ไม่ใช่ดีขึ้น

ความตรงกันทุกบิต: ตัวถอดรหัสทุกตัวต้องคำนวณได้บิตเดียวกัน

การเข้ารหัสเอนโทรปีทำงานได้ก็ต่อเมื่อโมเดลความน่าจะเป็นของตัวถอดรหัสตรงกับของตัวเข้ารหัสทุกประการ ในโคเดกแบบดั้งเดิม โมเดลนั้นคือตารางหรือกฎจำนวนเต็มชุดเล็ก ๆ ในโคเดกนิวรัล โมเดลนั้นคือโครงข่าย และความต่างของเลขทศนิยมลอยตัวระหว่างอุปกรณ์สองเครื่อง ไม่ว่าจะมาจากเคอร์เนลที่ต่างกัน ลำดับการสะสมผลที่ต่างกัน หรือ NPU ที่ต่างกัน ก็ทำให้ตัวถอดรหัสอ่านสัญลักษณ์ผิดและเสียสตรีมตั้งแต่จุดนั้นเป็นต้นไปได้

ทุกคนที่เคยนำโคเดกนิวรัลไปใช้จริงหรือทำเป็นมาตรฐานต้องใช้วิศวกรรมแก้ปัญหานี้ DCVC-RT แปลงโมเดลเป็นจำนวนเต็มเพื่อให้เข้ารหัสได้สม่ำเสมอข้ามอุปกรณ์ MLVC ใช้การออกแบบ hyperprior แบบ shared-scale เพื่อให้รอดจากความไม่ตรงกันข้ามแพลตฟอร์ม ซึ่งน่าจะเป็นส่วนหนึ่งของเหตุผลที่ MLVC บีบอัดได้น้อยกว่า DCVC-RT **บทวิเคราะห์** งานสำรวจด้านโครงข่ายประสาทเทียมของ JVET รันทุกโครงข่ายผ่าน SADL ไลบรารีขนาดเล็กที่อนุมานด้วยจำนวนเต็ม 16 บิต JPEG AI กำหนดจุดตรวจความสอดคล้องในโดเมน latent, residual และพิกเซล สำหรับทีมชิป นี่คือวินัยเดียวกับการแปลงแบบจำนวนเต็มของ HEVC: มาตรฐานต้องกำหนดการคำนวณไว้อย่างแม่นยำ ซึ่งจำกัดการออกแบบโครงข่ายและปิดทางลัดเฉพาะของผู้ขายแต่ละราย

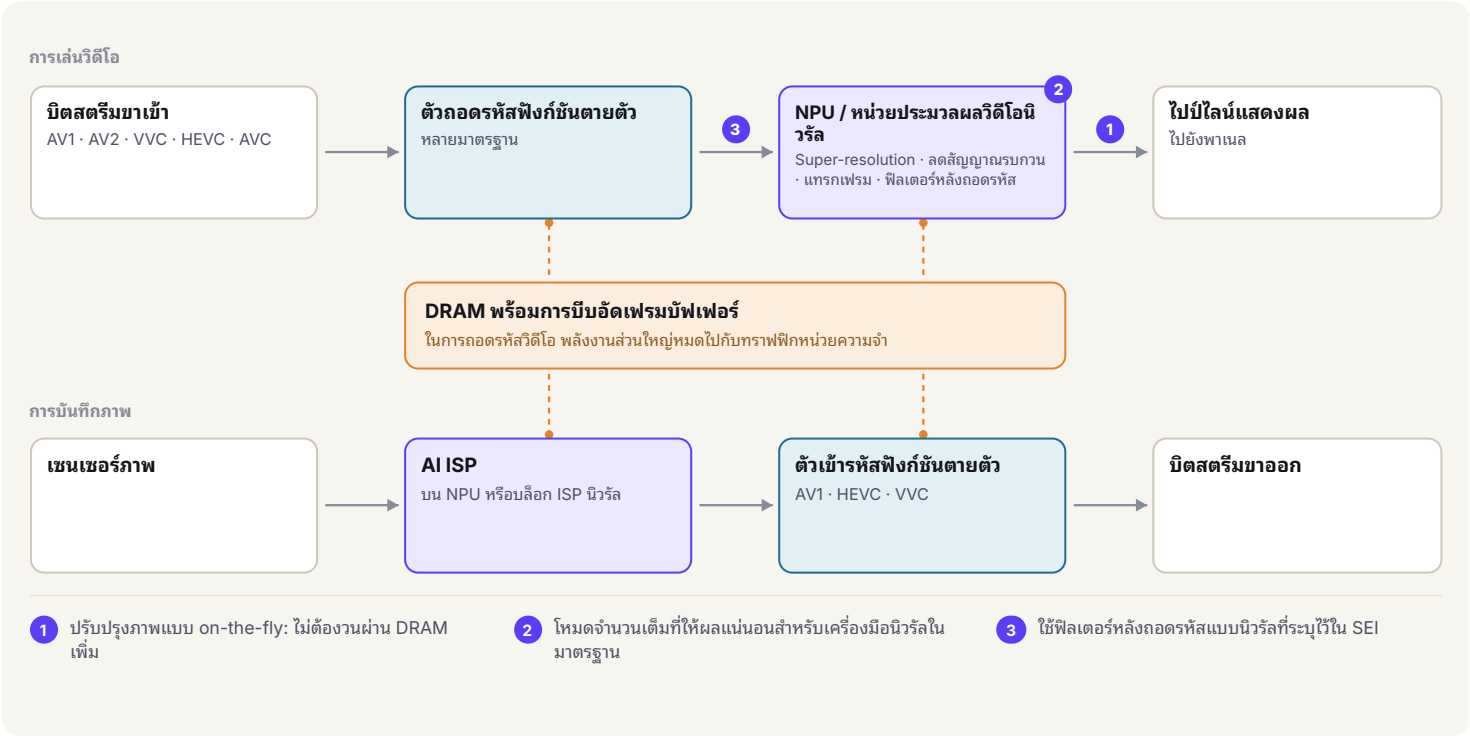
พื้นที่ชิปไม่ใช่อุปสรรค การตรวจสอบความถูกต้องและอายุการใช้งานต่างหาก

5 ล้านเกตของ NVCA เทียบกับราวหนึ่งล้านเกตของตัวถอดรหัส VVC ที่คาดการณ์ไว้ดูเหมือนมาก แต่บน SoC ของโทรศัพท์หรือทีวีสมัยใหม่ที่มี NPU ระดับหลาย TOPS อยู่แล้ว เกตไม่กี่ล้านเกตไม่ใช่สิ่งที่ตัดสินการออกแบบ สิ่งที่มีต้นทุนจริงคือการตรวจสอบความถูกต้องกับชุดทดสอบความสอดคล้อง การดูแลให้ตัวถอดรหัสที่โปรแกรมได้ทำงานถูกต้องตลอดการอัปเดตซอฟต์แวร์ตลอดอายุการใช้งานของอุปกรณ์ และกำลังไฟฟ้าที่ 4K60 HDR ต้นทุนเหล่านี้จะสมเหตุสมผลก็ต่อเมื่อมาตรฐานกำหนดแน่ชัดแล้วว่าต้องสร้างอะไร

สถาปัตยกรรมที่ชนะในตอนนี้: โคเดกฟังก์ชันตายตัวบวก NPU

ภาพที่ 08 **การแบ่งงานที่ใช้จริง: ฮาร์ดแวร์ฟังก์ชันตายตัวสำหรับบิตสตรีม เอนจินนิวรัลสำหรับพิกเซล**

SoC มีเดียปี 2026 บล็อกโคเดกถอดรหัสได้ทุกมาตรฐานด้วยพลังงานเพียงเศษเสี้ยวของนาโนจูลต่อพิกเซล NPU รับงานที่โครงข่ายประสาทเทียมถนัดและมาตรฐานไม่ได้กำหนด ต้นทุนที่สำคัญคือทราฟฟิก DRAM



ที่มา: มุมมองสถาปัตยกรรมของผู้เขียน อ้างอิงข้อมูลผลิตภัณฑ์ที่ Tencent (Canghai V2), Chips&Media, Allegro DVT และ Samsung เปิดเผย (2026)

ซิลิคอนที่ขายได้ในปี 2026 จับคู่โคเดกฟังก์ชันตายตัวแบบหลายมาตรฐานกับเอนจินนิวรัลที่วางอยู่ข้างกัน บล็อกโคเดกรองรับ AV1, AV2, VVC, HEVC และ AVC ด้วยพลังงานเพียงเศษเสี้ยวของนาโนจูลต่อพิกเซล เอนจินนิวรัลรับงานที่โครงข่ายถนัดและมาตรฐานไม่ได้กำหนด: การลดสัญญาณรบกวนและ ISP ฝั่งการถ่ายภาพ super-resolution การแทรกเฟรม และการปรับปรุงภาพฝั่งแสดงผล ทั้งสองส่วนใช้การบีบอัดเฟรมบัฟเฟอร์และตารางการเข้าถึง DRAM ร่วมกัน เพราะทราฟฟิกหน่วยความจำคือต้นทุนที่สำคัญ

ตลาดสะท้อนการแบ่งงานนี้ ASIC โคเดกวิดีโอ Canghai (沧海) ของ Tencent ซึ่งพัฒนาเองตั้งแต่ปี 2019 มีชิปรุ่นแรกใช้งานแล้วกว่า 100,000 ชิ้นใน Tencent Cloud และในบริการไลฟ์ วิดีโอตามสั่ง และคลาวด์เกมมิงของบริษัทเอง ส่วนรุ่นที่สองซึ่งประกาศในเดือนพฤษภาคม 2026 ว่าเปิดเครื่องทำงานได้แล้วและกำลังเข้าสู่การผลิตจำนวนมากเพื่อให้บริการในครึ่งหลังของปี รองรับ H.265 และ H.266 ในสถาปัตยกรรมเดียว บีบอัดได้ดีกว่ารุ่นก่อนกว่า 10% และ 30% มีปริมาณงาน (throughput) ราวสองเท่า และมีการปรับปรุงภาพบนชิป เช่น super-resolution Chips&Media ขาย IP โคเดกควบคู่กับ NPU ที่เน้นงานวิดีโอคือ WAVE-N Allegro DVT ขายตัวถอดรหัส AV2 และได้เพิ่ม IP ประมวลผลวิดีโอแบบนิวรัลตัวแรกของบริษัท ชิปประมวลผลภาพในทีวีของ Samsung จับคู่ตัวถอดรหัสมาตรฐานกับโครงข่ายอัปสเกลหลายร้อยชุด

สิ่งที่ผมจะสร้าง แยกตามบทบาท

ทีม SoC และชิปทีวี คงตัวถอดรหัสฟังก์ชันตายตัวไว้สำหรับทุกมาตรฐานที่ต้องเล่นได้ รวมถึง AV2 ตั้งแต่ตอนนี้ และ VVC ในที่ที่งานออกอากาศกำหนด กำหนดขนาด NPU ให้รองรับการประมวลผลภายหลังที่ความละเอียดและอัตราเฟรมของจอแสดงผล และออกแบบเส้นทางจากตัวถอดรหัสสู่ NPU สู่จอแสดงผลให้การปรับปรุงภาพไม่ต้องวนผ่าน DRAM เพิ่มอีกรอบ รองรับ SEI สำหรับฟิลเตอร์หลังถอดรหัสแบบนิวรัล จัดให้มีโหมดจำนวนเต็มที่ให้ผลแน่นอนใน NPU มองว่า H.267 เป็นงานออกแบบที่จะเริ่มในปี 2029 หรือหลังจากนั้น และเผื่อความสามารถในการโปรแกรมไว้บางส่วนสำหรับเครื่องมือนิวรัลของมาตรฐานนี้

ผู้ขาย IP ขายตัวถอดรหัส AV2 และ VVC ตั้งแต่ตอนนี้ พร้อมชิ้นส่วนที่การประมวลผลแบบนิวรัลต้องใช้รอบ ๆ ตัวถอดรหัสเหล่านั้น: ตัวเร่งฟิลเตอร์หลังถอดรหัส การบีบอัดเฟรมบัฟเฟอร์ และเครื่องมือทดสอบความสอดคล้อง

แพลตฟอร์มวิดีโอ การประมวลผลล่วงหน้าแบบนิวรัลที่ตัวเข้ารหัสคืนทุนได้ก่อน เพราะรันครั้งเดียวต่อหนึ่งเรื่องและไม่ต้องการอะไรจากอุปกรณ์ ใช้ super-resolution ฝั่งไคลเอนต์ในที่ที่อุปกรณ์รองรับ และเก็บโคเดกนิวรัลแบบ end-to-end ไว้ใช้ในวงปิดที่คุณควบคุมได้

บทที่ 09

มาตรฐานและสิทธิบัตร

09 มาตรฐานและสิทธิบัตร: จุดที่เครื่องมือนิวรัลกลายเป็นสิ่งที่เก็บค่าสิทธิได้

ไปป์ไลน์ของมาตรฐาน

มาตรฐาน	องค์กร	เนื้อหาด้านนิวรัล	สถานะ ณ พฤศจิกายน 2026
H.267 (ชื่อชั่วคราว; beyond VVC)	JVET (ITU-T และ ISO/IEC)	ยังเปิดกว้าง ทั้งเครื่องมือแบบดั้งเดิม (ECM) และเครื่องมือนิวรัล (ฟิลเตอร์ในลูป การทำนายภายในเฟรม super-resolution) ต่างเป็นตัวเลือก	ประกาศเรียกรับข้อเสนอกรกฎาคม 2026; ประเมินข้อเสนอมกราคม 2027; โมเดลทดสอบแรกครึ่งหลังปี 2027; เวอร์ชัน 1 สิ้นปี 2029
VVC และ HEVC พร้อม VSEI	JVET	ระบุฟิลเตอร์หลังถอดรหัสแบบนิวรัลในเมทาดาทา SEI ได้ ส่วนการนำไปใช้เป็นทางเลือกของตัวถอดรหัส	มีผลบังคับใช้; ขยายข้อความในฉบับเดือนตุลาคม 2025
AV2	AOMedia	เครื่องมือแบบดั้งเดิม ปลอดค่าสิทธิ	เวอร์ชัน 1.0.0 เมื่อ 28 พฤษภาคม 2026; IP ตัวถอดรหัสขายไลเซนส์ตั้งแต่มิถุนายน 2026; YouTube วางแผนสตรีมบนเว็บต้นปี 2027

มาตรฐาน	องค์กร	เนื้อหาต้านนิวรัล	สถานะ ณ พฤศจิกายน 2026
JPEG AI	JPEG (ISO/IEC และ ITU-T)	โคเดกภาพแบบเรียนรู้ end-to-end สำหรับการมองเห็นของมนุษย์และเครื่องจักร	ส่วนที่ 1 เผยแพร่ปี 2025; เวอร์ชันที่สองอยู่ระหว่างหารือ; ไม่พบการรองรับเป็นค่าเริ่มต้นในผลิตภัณฑ์
IEEE 1857.11	IEEE	การเข้ารหัสภาพด้วยโครงข่ายประสาทเทียมที่ออกแบบสำหรับ NPU	เผยแพร่ธันวาคม 2024
Video Coding for Machines	MPEG	การเข้ารหัสเพื่อการวิเคราะห์โดยเครื่องจักรแทนการรับชม	ร่างมาตรฐานสากลต้นปี 2026; ฉบับสมบูรณ์ตามแผนปลายปี 2026
ส่วนขยาย ML ของ Opus	IETF	ข้อมูลซ้ำซ้อนเชิงลึก การปรับปรุงเสียงฝั่งตัวถอดรหัส และการขยายแบนด์วิดท์ ทั้งหมดเลือกใช้ได้	อยู่ใน libopus 1.5 และ 1.6; ข้อมูลซ้ำซ้อนเชิงลึกกำลังจัดทำเป็นมาตรฐาน

มีสองเรื่องที่โดดเด่น ข้อแรก เนื้อหาด้านนิวรัลเข้าสู่มาตรฐานทางขอบก่อนถึงแกนกลาง: ในรูปฟิลเตอร์หลังถอดรหัสที่เลือกใช้ได้ การปรับปรุงที่เลือกใช้ได้ และการเข้ารหัสภาพ ซึ่งทั้งหมดไม่เปลี่ยนเส้นทางการถอดรหัสหลัก ข้อสอง มาตรฐานเดียวที่อาจนำเครื่องมือนิวรัลเข้าไปในเส้นทางการถอดรหัสของทุกหน้าจอ คือ H.267 ยังอยู่ห่างจากเวอร์ชันแรกอีกสามปี

ใครวางตำแหน่งไว้แล้วในด้านสิทธิบัตร

เครื่องมือในมาตรฐานจะกลายเป็นสิทธิบัตรที่เก็บค่าสิทธิได้ กระบวนการของ H.267 จึงเป็นการแข่งขันด้านสิทธิบัตรด้วย และเครื่องมือนิวรัลก็เป็นส่วนหนึ่งของการแข่งขันนี้ ภาพรวมสิทธิบัตร beyond-VVC จำนวน 3,916 ฉบับที่ GreyB จัดทำในปี 2025 จัดให้ ByteDance อยู่แถวหน้า ด้วย 645 ตระกูลสิทธิบัตรและเอกสารข้อเสนอ (contribution) ต่อ JVET 251 ฉบับ และระบุความร่วมมือระหว่างมหาวิทยาลัยกับอุตสาหกรรม เช่น Xiaomi กับ Xidian University ที่มีกว่า 20 ตระกูล MLex รายงานในเดือนมีนาคม 2026 ว่า Tencent, Alibaba และ ByteDance กำลังกลายเป็นผู้เล่นที่แข็งขันในกลุ่มสิทธิบัตรโคเดกวิดีโอ ในเดือนพฤษภาคม 2026 Sauvegarder Investment Management ซื้อสินทรัพย์สิทธิบัตรโคเดกวิดีโอ 126 รายการจาก Alibaba ซึ่งครอบคลุม AV1, AV2 และ Video Coding for Machines การที่ InterDigital ซื้อ Deep Render ในเดือนตุลาคม 2025 ทำให้สิทธิบัตรการเข้ารหัสวิดีโอด้วย AI ของ Deep Render เข้ามาอยู่ในพอร์ตของ InterDigital

เงินต่อเครื่องดูได้ง่ายจากตารางอัตราค่าสิทธิของกลุ่มสิทธิบัตร อัตราของ Access Advance สำหรับ VVC (ณ วันที่ 1 กรกฎาคม 2026 สำหรับผู้รับอนุญาตที่ปฏิบัติตามเงื่อนไขและใช้เครื่องหมายการค้าของกลุ่ม) คือ 0.50 ดอลลาร์สหรัฐต่อโทรศัพท์มือถือ แท็บเล็ต หรือแล็ปท็อปในตลาด Region 1 เช่น สหรัฐฯ สหภาพยุโรป และญี่ปุ่น และ 0.25 ดอลลาร์สหรัฐใน Region 2 สำหรับทีวี 4K อยู่ที่ 1.00 ถึง 1.50 ดอลลาร์สหรัฐตามระดับราคาใน Region 1 โดยมีเพดานรายปีต่อองค์กร 75 ล้านดอลลาร์สหรัฐ AV1 และ AV2 ปลอดค่าสิทธิจากสมาชิกของ AOMedia ไม่ว่า H.267 จะมีอะไรอยู่ข้างใน เครื่องมือของมาตรฐานนี้จะถูกตั้งราคาแบบเดียวกัน นี่คือเหตุที่บริษัทที่มีสิทธิบัตรด้านการเข้ารหัสแบบนิวรัลกำลังส่งข้อเสนอให้ JVET ในตอนนี้ บทวิเคราะห์

บทที่ 10

10 ใครได้เงิน และบริษัทจะเข้าไปเล่นตรงไหนได้

ใครได้เงิน

เงินในปี 2026

ภาพที่ 09 **ใครได้เงินจากการเข้ารหัสแบบนิวรัลในปี 2026**

เงินไหลตามมาตรฐานและแพลตฟอร์ม ผมไม่พบบริษัทอิสระรายใดที่เลี้ยงตัวได้ด้วยการขายไลเซนส์โคเดกวิดีโอนิวรัลแบบ end-to-end

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)

ที่มา: ข้อมูลที่บริษัทและกลุ่มสิทธิบัตรเปิดเผยซึ่งอ้างถึงในรายงาน สถานะ ณ วันที่ 12 พฤศจิกายน 2026 อัตราของกลุ่มสิทธิบัตร: Access Advance, VVC, Region 1 กรณีปฏิบัติตามเงื่อนไขและใช้เครื่องหมายการค้า ณ วันที่ 1 กรกฎาคม 2026

แพลตฟอร์มเก็บส่วนที่ประหยัดได้ไว้เอง Netflix, YouTube, Meta, Microsoft และ Tencent สร้างการประมวลผลแบบนิวรัลเอง และเก็บประโยชน์ไว้ในรูปแบนด์วิดท์ที่ลดลง พื้นที่จัดเก็บที่ลดลง และคุณภาพต่อบิตที่ดีขึ้น เมื่อการแจกโค้ดช่วยให้มีผู้นำไปใช้มากขึ้น พวกเขาก็แจกโค้ดฟรี: Microsoft เปิด MLVC ภายใต้สัญญาอนุญาต MIT

ผู้ผลิตอุปกรณ์และชิปขายความแตกต่าง ISP ของกล้องและชิปประมวลผลภาพในทีวีคือจุดที่โครงข่ายประสาทเทียมสัมผัสผู้ใช้ทุกคน และขายในฐานะฟีเจอร์: Cognitive ISP ของ Qualcomm และชิปอัปสเกลของ Samsung

ผู้ขาย IP ขายตัวถอดรหัสสำหรับทุกมาตรฐานใหม่ พร้อม NPU ภายในสามสัปดาห์หลังสเปกของ AV2 ออก Chips&Media, Allegro DVT และ VeriSilicon ต่างประกาศว่ามีลูกค้าตัวถอดรหัส AV2 ผู้ขายกลุ่มเดียวกันนี้ตอนนี้ขายบล็อกประมวลผลแบบนิวรัลควบคู่กับโคเดกด้วย

ผู้ผลิต GPU สร้างแพลตฟอร์ม DLSS และการบีบอัดเท็กซ์เจอร์แบบนิวรัลทำให้ฮาร์ดแวร์เทนเซอร์ของ NVIDIA กลายเป็นส่วนหนึ่งของวิธีสร้างเกม

ผู้ถือสิทธิบัตรเก็บค่าสิทธิต่อเครื่อง ผ่านกลุ่มสิทธิบัตรสำหรับ VVC และ HEVC และหลังปี 2029 สำหรับ H.267

สตาร์ตอัปโคเดกนิวรัลถูกซื้อกิจการ WaveOne ไปอยู่กับ Apple ในปี 2023 iSIZE ไปอยู่กับ Sony Interactive Entertainment ในปี 2023 และ Deep Render ไปอยู่กับ InterDigital ในปี 2025 ทั้งหมดไม่เปิดเผยเงื่อนไข ในทุกกรณี ผู้ซื้อ

คือแพลตฟอร์มหรือผู้ให้ไลเซนส์ที่ซื้อทั้งทีมและสิทธิบัตร ผมไม่พบบริษัทอิสระรายใดในปี 2026 ที่เลี้ยงตัวได้ด้วยการขายไลเซนส์โคเดกวิดีโอนิวรัลแบบ end-to-end

บริษัทจะเข้าไปเล่นตรงไหนได้

ช่องทาง	ใครซื้อ	ทำไมต้องตอนนี้	ความเสี่ยงหลัก
การประมวลผลล่วงหน้าแบบนิวรัลที่ตัวเข้ารหัส (ลดขนาดภาพ ลดสัญญาณรบกวน ปรับจูนรายเรื่อง)	แพลตฟอร์มสตรีมมิงและวิดีโอจากผู้ใช้ ผู้แพร่ภาพ	ไม่ต้องเปลี่ยนอุปกรณ์; พิสูจน์แล้วที่ Netflix	แพลตฟอร์มใหญ่สร้างเอง
IP และงานปรับจูนการปรับปรุงภาพฝั่งแสดงผล (super-resolution การแทรกเฟรม คุณภาพภาพ)	ผู้ผลิตชิปทีวี กล่องรับสัญญาณ และโทรศัพท์	NPU อยู่ในอุปกรณ์แล้ว; สร้างความแตกต่างที่มองเห็นได้	กลายเป็นสินค้าโภคภัณฑ์; ผู้ผลิตชิปทำเอง
โคเดกนิวรัลแบบวงปิด (การประชุมทางไกล การปฏิบัติงานระยะไกล ลิงก์ส่วนตัว)	ผู้ขายระบบสื่อสารองค์กร ระบบอุตสาหกรรม และเทเลเพรเซนซ์	MLVC พิสูจน์แล้วว่าทำได้ และโค้ดเปิดเผย	คุณต้องควบคุมทั้งสองฝั่ง; NPU กระจัดกระจายหลายแบบ
ซิลิคอนหรือ IP โคเดกบวก NPU (AV2, VVC และการประมวลผลภายหลัง)	ผู้ผลิต SoC	รอบของ AV2 กำลังเริ่ม	ผู้เล่นรายเดิมแข็งแกร่ง
เครื่องมือประเมินผล ทดสอบความสอดคล้อง และการอนุมานที่ให้ผลแน่นอน	ทุกรายที่นำเครื่องมือนิวรัลไปใช้จริง	ความตรงกันทุกบิตเป็นข้อกำหนดที่เลี่ยงไม่ได้	ตลาดเล็ก เว้นแต่ขายรวมกับอย่างอื่น
การส่งข้อเสนอเข้ามาตรฐาน H.267	ผู้ให้ไลเซนส์ แพลตฟอร์มใหญ่	ข้อเสนอจะถูกประเมินในเดือนมกราคม 2027	ใช้เวลานาน แพง ไม่แน่นอน

สำหรับบริษัทส่วนใหญ่ สองแถวแรกคือที่ที่มีรายได้ให้เก็บในปี 2027 แถวที่สามน่าสนใจที่สุดสำหรับบริษัทผลิตภัณฑ์ที่โฟกัสชัด ส่วนแถวสุดท้ายต้องอาศัยความอดทนและกลยุทธ์ด้านไลเซนส์

บทที่ 11

11 คำกล่าวอ้างและหลักฐาน
คำกล่าวอ้างที่คุณจะได้ยิน และหลักฐานว่าอย่างไร

คำกล่าวอ้าง	หลักฐานว่าอย่างไร
“โคเดกนิวรัลชนะ VVC แล้ว”	ในห้องแล็บถือว่าใช่ DCVC-RT ประหยัด 21% เมื่อเทียบกับตัวเข้ารหัสอ้างอิงของ VVC ที่ 1080p ในการตั้งค่าของบทความ และ JPEG AI ประหยัด 12.5–27.9% เมื่อเทียบกับ VVC แบบ intra ส่วนตัวถอดรหัสใช้ราว 175,000 MAC ต่อพิกเซล หรือพลังงานระดับ NPU เทียบกับเศษเสี้ยวของนาโนจูลในฮาร์ดแวร์ฟังก์ชันตายตัว
“โคเดก AI จะมาแทน H.26x และ AV1 ภายในไม่กี่ปี”	ไม่ใช่ในการเล่นวิดีโอ มาตรฐานแรกที่รองรับเครื่องมือนิวรัลได้ตั้งเป้าไว้ที่สิ้นปี 2029 ชิปจะตามมาราวปี 2031 และการใช้งานในกระแสหลักราวกลางทศวรรษ 2030 ประมาณการ โคเดกที่กำลังมาถึงตอนนี้คือ AV2 ซึ่งเป็นแบบดั้งเดิม
“MLVC ลดแบนด์วิดท์ได้ 88%”	เมื่อเทียบกับ H.264 ที่ 360p บนชุดทดสอบของ Microsoft เอง หากเทียบกับ HEVC แบบฮาร์ดแวร์จะเป็น 75.5% ที่ 360p และ 65.4% ที่ 540p ในการโทรที่ Microsoft ควบคุมทั้งสองฝั่ง
“เท็กซ์เจอร์นิวรัลลดหน่วยความจำวิดีโอได้ 85–96%”	ในฉากสาธิต เกมที่วางจำหน่ายแล้วได้การบีบอัดที่มีประสิทธิผลดีขึ้นราว 30% และใช้เท็กซ์เจอร์นิวรัลเฉพาะกับแอสเซตที่เลือกไว้เพราะต้นทุนขณะรัน
“JPEG AI เป็นมาตรฐาน ISO กล้องและเบราว์เซอร์จึงจะนำไปใช้”	ผมไม่พบการรองรับเป็นค่าเริ่มต้นในผลิตภัณฑ์ใด ฟอร์แมตที่เว็บเลือกใช้ในปี 2026 คือ JPEG XL ใน Chrome 155 ซึ่งเป็นแบบดั้งเดิม
“การถอดรหัสแบบนิวรัลแพงเกินไปสำหรับโทรศัพท์”	จริงบางส่วน MLVC ถอดรหัส 540p30 ได้ภายในครึ่งหนึ่งของ NPU ส่วน 4K60 ด้วยโมเดลระดับ DCVC-RT จะต้องใช้กำลังประมวลผลต่อเนื่องราว 175 TOPS ประมาณการ
“ลิงก์ดาวเทียมและโดรนคือตลาดใหญ่ของวิดีโอนิวรัล”	เหมาะที่สุดในทางเทคนิค แต่ตลาดเล็ก และถูกควบคุมการส่งออกเข้มขึ้นเรื่อย ๆ ทั้งสองฝั่งของมหาสมุทรแปซิฟิก
“แพลตฟอร์มใหญ่จะซื้อไลเซนส์โคเดกนิวรัลจากสตาร์ตอัป”	พวกเขาสร้างเอง (Netflix, Microsoft, Tencent) หรือซื้อสตาร์ตอัปไปเลย (Apple, Sony, InterDigital)

การทดสอบง่าย ๆ ข้อหนึ่งช่วยกรองคำเสนอขายส่วนใหญ่ได้: ขอให้ระบุจุดอ้างอิง ตัวชี้วัด ความละเอียด และต้นทุนการถอดรหัส คำกล่าวอ้างที่ระบุทั้งสี่ข้อไม่ได้คือผลการวิจัย ไม่ใช่ผลิตภัณฑ์

บทที่ 12

12

สัญญาณที่ต้องจับตา

สิ่งที่ต้องจับตา ปี 2027 ถึง 2030

เมื่อไร	สัญญาณ	ทำไมจึงสำคัญ
มกราคม 2027	JVET ประเมินข้อเสนอสำหรับ H.267 ที่บริสเบน	ข้อเสนอชิ้นนำกี่ฉบับที่พึ่งเครื่องมือนิวรัล และด้วยต้นทุนต่อพิกเซลเท่าไร
ต้นปี 2027	สตรีม AV2 บนเว็บครั้งแรกของ YouTube; ตัวถอดรหัส AV2 ใน SoC ที่ประกาศออกมา	รอบแบบดั้งเดิมรอบถัดไปจะกำหนดงบของอุปกรณ์ไปจนถึงราวปี 2030
ครึ่งหลังของปี 2027	โมเดลทดสอบแรกของ H.267	เครื่องมือนิวรัลอยู่ในการออกแบบแกนหลัก หรืออยู่แค่ในงานสำรวจ
2027	MLVC ขยายเกินการโทรแบบ peer-to-peer บน Teams; ผู้ขายรายอื่นนำโค้ดเปิดไปใช้	การทดสอบครั้งแรกของโคเดกนิวรัลแบบเปิดนอกมือเจ้าของ
2027–2028	ผลิตภัณฑ์แรกที่เข้ียนหรืออ่าน JPEG AI เป็นค่าเริ่มต้น	ฟอร์แมตนิวรัลแรกที่ใช้ข้ามขอบเขตของบริษัท
2027–2028	DirectX Linear Algebra ในไดรเวอร์ที่ออกใช้จริง; เท็กซ์เจอร์นิวรัลในเกมมากขึ้นและบนอุปกรณ์พกพา	การถอดรหัสแบบนิวรัลข้ามผู้ผลิตภายในงานกราฟิก
ต่อเนื่อง	งบ NPU ในโทรศัพท์และ PC	ที่ราว 100 TOPS แบบต่อเนื่อง ตัวถอดรหัส 1080p30 ระดับ DCVC-RT จะใช้ NPU ราวหนึ่งในห้าแทนที่จะเป็นครึ่งหนึ่ง บทวิเคราะห์
สิ้นปี 2029	H.267 เวอร์ชัน 1	จุดเริ่มต้นรอบการพัฒนาชิปสำหรับเครื่องมือนิวรัลในการเล่นวิดีโอกระแสหลัก

บทที่ 13

13 คำถามที่ยังเปิดอยู่

หลายคำถามในรายงานนี้ยังไม่มีคำตอบสาธารณะ และเป็นคำถามที่ผมเห็นว่าน่าสนใจที่สุดในทางเทคนิค

ความตรงกันทุกบิตในวงกว้าง ตัวถอดรหัสนิวรัลจะเข้าใกล้การคำนวณที่ตรงกันทุกบิตด้วยจำนวนเต็มล้วนบน NPU จากผู้ผลิตต่างรายได้แค่ไหน ก่อนจะเสียความได้เปรียบด้านการบีบอัดไปเป็นส่วนใหญ่? การแลกหกจุดของ MLVC เป็นข้อมูลอ้างอิงหนึ่งจุด แต่มาตรฐานต้องการคำตอบที่ใช้ได้ทั่วไป

หน่วยความจำมาก่อน MAC การประมวลผลภายหลังแบบนิวรัลมีต้นทุนทราฟฟิก DRAM ต่อพิกเซลเท่าไร เมื่อรันแบบ on-the-fly ระหว่างตัวถอดรหัสฮาร์ดแวร์กับจอแสดงผล และ working set ของมันจะทำให้เล็กลงได้แค่ไหน?

ราคาของเครื่องมือนิวรัลใน H.267 ที่กี่ MAC และกี่นาโนจูลต่อพิกเซล ฟิลเตอร์ในลูปแบบนิวรัลจึงจะไม่คุ้มกับบิตที่ประหยัดได้ในตัวถอดรหัส 4K60?

วงปิดนอกเหนือจากการโทร ระบบปิดอื่นใดบ้าง ตั้งแต่การเก็บภาพบนคลาวด์และการสตรีมเกม ไปจนถึงหุ่นยนต์และลิงก์ส่งสัญญาณต้นทางแบบส่วนตัว ที่จะนำโคเดกนิวรัลแบบ end-to-end ไปใช้ได้ก่อน H.267 มาถึง?

ตัวถอดรหัสที่สร้างเนื้อหาเอง เมื่อตัวถอดรหัสสร้างรายละเอียดขึ้นเองแทนที่จะสร้างคืน อย่างที่ VAE วิดีโอทำอยู่แล้ว ควรวัดและกำหนดความเที่ยงตรงอย่างไร?

เหล่านี้คือเรื่องที่เราสำรวจร่วมกันได้ในหลายปีข้างหน้า ผ่านการแลกเปลี่ยนและความร่วมมือทางเทคนิค: i@tanshuai.com

ข้อจำกัดของรายงานนี้

นี่คือการวิจัยจากเอกสารสาธารณะที่มีจนถึงวันที่ 12 พฤศจิกายน 2026 สำหรับฉบับนี้ ผมไม่ได้สัมภาษณ์เพื่อเก็บข้อมูลปฐมภูมิ และไม่ได้วัดผลด้วยตนเองอย่างอิสระ ตัวเลขสำคัญหลายตัวเป็นคำกล่าวอ้างของผู้ขายหรือของผู้เขียนบทความบนชุดทดสอบของตนเอง เช่น การประหยัดบิตเรตของ MLVC และอัตราการบีบอัดเท็กซ์เจอร์แบบนิวรัล ซึ่งได้กำกับไว้ตามนั้นแล้ว ค่าประมาณการด้านการประมวลผลของ DCVC-RT ที่ผมทำเป็นการรวมตัวเลขรายโมดูลที่ผู้เขียนเผยแพร่ และไม่คิดการสูญเสียอัตราการใช้งานภาระจริงบน NPU จึงจะสูงกว่านี้ การเปรียบเทียบพลังงานอาศัยการออกแบบที่เผยแพร่เพียงไม่กี่ชิ้นบนโหนดการผลิตที่ต่างกัน ส่วนตัวเลขของตัวถอดรหัส VVC และ AV2 ที่ผลิตจริงเป็นความลับทางการค้า กำหนดเวลาของมาตรฐานเป็นเป้าหมายของคณะกรรมการเอง วันที่ของ H.267 หลังปี 2029 เป็นการประมาณของผมจากประวัติของ AV1 และ VVC

เกี่ยวกับผู้เขียน



เกี่ยวกับผู้เขียน

ทัน ชว่าย เป็นนักเทคโนโลยีและผู้ก่อตั้ง มีประสบการณ์กว่ายี่สิบปีในงานด้านชิป ฮาร์ดแวร์ และซอฟต์แวร์ระบบ และมีสิทธิบัตรที่ได้รับอนุมัติแล้ว 20 ฉบับ เขามีประสบการณ์หลายปีในงานวิจัยและพัฒนาและการบริหารชิปแบบ fabless รวมถึงการออกแบบ datapath ของ ASIC แบบ full-custom และโครงการชิปกล้องมัลติมีเดีย เช่น ตัวเร่งโคเดกวิดีโอ ซึ่งเป็นมุมมองที่รายงานนี้ใช้: พลังงานต่อพิกเซล ทราฟฟิกหน่วยความจำ และสิ่งที่มาตรฐานยอมให้ตัวถอดรหัสตรึงไว้ในซิลิคอน งานของเขาพาไปทั้งซิลิคอนแวลลีย์ นิวยอร์ก ชิคาโก ไต้หวัน เขตเกรตเตอร์เบย์แอเรีย (เซินเจิ้นและฮ่องกง) เวียดนาม และมาเลเซีย ปัจจุบันเขาทำงานเป็นนักวิจัยและนักพัฒนา AI

พูดคุยเชิงเทคนิค: i@tanshuai.com · tanshuai.com

หมายเหตุและข้อสงวนสิทธิ์

รายงานนี้เป็นการวิจัยจากเอกสารอย่างอิสระ โดยอาศัยเอกสารสาธารณะที่มีอยู่จนถึงวันที่ 12 พฤศจิกายน 2026 จัดทำเพื่อให้ข้อมูลเท่านั้น และไม่ใช่คำแนะนำด้านการลงทุน กฎหมาย หรือวิศวกรรม ผลิตภัณฑ์ กำหนดเวลาของมาตรฐาน อัตราของกลุ่มสิทธิบัตร และกฎการส่งออกเปลี่ยนแปลงได้ โปรดตรวจสอบสิ่งที่คุณจะยึดถือกับแหล่งข้อมูลปฐมภูมิ ตัวเลขที่กำกับว่า ผู้ขายอ้าง, ประมาณการ หรือ บทวิเคราะห์ ควรอ่านตามนั้น ชื่อบริษัทและผลิตภัณฑ์เป็นของเจ้าของ และใช้เพื่อระบุตัวตนเท่านั้น ไม่มีบริษัทใดตรวจทานหรือสนับสนุนรายงานนี้

© 2026 ทัน ชว่าย คุณแชร์รายงานนี้ต่อได้โดยไม่ดัดแปลงและระบุที่มา

แหล่งข้อมูล

แหล่งข้อมูล

เข้าถึงทั้งหมดระหว่างวันที่ 3 ถึง 12 พฤศจิกายน 2026 จัดกลุ่มตามบท

ทราฟฟิก อุปกรณ์ และขอบเขต

- Ericsson, *Mobility Report*, มิถุนายน 2026 (วิดีโอราว 75% ของทราฟฟิกข้อมูลมือถือ ณ สิ้นปี 2025): https://www.ericsson.com/en/reports-and-papers/mobility-report/reports/june-2026 ; สรุป: https://telecomlead.com/5g/ericsson-mobility-report-2026-global-mobile-data-traffic-to-reach-515-eb-by-2031-as-india-leads-smartphone-data-consumption-126290
- IDC, ยอดจัดส่งสมาร์ตโฟนทั่วโลก ไตรมาส 4 และทั้งปี 2025 (1.26 พันล้านเครื่อง), 13 มกราคม 2026: https://www.idc.com/resource-center/press-releases/q425mobilephonetop5/
- TrendForce, การคาดการณ์ยอดจัดส่งทีวีปี 2025 (195.71 ล้านเครื่อง): https://iconnect007.com/article/146406/early-shipments-by-tv-brands-push-global-1h25-volume-up-2-but-raise-concerns-for-secondhalf-peak-season/146403/ein ; Omdia, การคาดการณ์ทีวีปี 2025 (208.7 ล้านเครื่อง): https://iconnect007.com/article/145954/omdia-forecasts-flat-tv-shipments-in-2025-but-mini-led-poised-for-explosive-growth-at-829/145951/ein

การเล่นวิดีโอและสตรีมมิง

- Netflix Technology Blog, “AV1 — Now Powering 30% of Netflix Streaming”, ธันวาคม 2025: https://netflixtechblog.com/av1-now-powering-30-of-netflix-streaming-02f592242d80
- Netflix Technology Blog, “For your eyes only: improving Netflix video quality with neural networks”, พฤศจิกายน 2022: https://netflixtechblog.com/for-your-eyes-only-improving-netflix-video-quality-with-neural-networks-5b8d032da09c
- AOMedia, AV2 Bitstream and Decoding Process Specification v1.0.0, 28 พฤษภาคม 2026: https://av2.aomedia.org/v1.0.0/index.html ; “AV2: What the Final Specification Means for Implementers”: https://aomedia.org/blog%20posts/AV2-What-the-Final-Specification-Means-for-Implementers/
- TV Technology, “YouTube to Commence AV2 Streaming in 2027” (สัดส่วน AV1 ในคลังวิดีโอของ YouTube): https://www.tvtechnology.com/platform/streaming/youtube-to-commence-av2-streaming-in-2027
- Google Research, “The in-loop filtering pipeline in AV2”, 2026: https://research.google/pubs/the-in-loop-filtering-pipeline-in-av2/
- Fraunhofer IIS, “The Future of Broadcast in Brazil” (DTV+ เริ่มให้บริการเชิงพาณิชย์ มิถุนายน 2026): https://mpegh.com/brazil/ ; CartaCapital เรื่องเครื่องรับในทีวี: https://www.cartacapital.com.br/?p=628337 ; Olhar Digital เรื่อง VVC ใน TV 3.0: https://olhardigital.com.br/2026/05/01/pro/em-breve-a-tv-aberta-vai-mudar-no-brasil/
- 9to5Google, “YouTube adds AI-upscaled 'Super Resolution' for low-quality videos”, 29 ตุลาคม 2025: https://9to5google.com/2025/10/29/youtube-adds-ai-upscaled-super-resolution-for-low-quality-videos/
- Meta, “How Meta Deployed Super Resolution at Scale to Transform Video Quality”, @Scale, 16 ตุลาคม 2025: https://atscaleconference.com/how-meta-deployed-super-resolution-at-scale-to-transform-video-quality/
- Samsung, ไฮไลต์ทีวี 8K (NQ8 AI Gen3, โครงข่ายประสาทเทียม 768 ชุด): https://www.samsung.com/us/tvs/8k-tv/highlights/

การถ่ายภาพและภาพนิ่ง

- Qualcomm, “Meet the Snapdragon 8 Gen 2 Super Cameras” (Cognitive ISP): https://www.qualcomm.com/news/onq/2023/07/meet-the-snapdragon-8-gen-2-super-cameras
- Chips&Media และ Visionary.ai, ISP แบบ AI เต็มรูปแบบ, 5 มกราคม 2026: https://www.edge-ai-vision.com/2026/01/chipsmedia-and-visionary-ai-unveil-the-worlds-first-ai-based-full-image-signal-processor-redefining-the-future-of-image-quality/ ; บทสัมภาษณ์ใน PetaPixel: https://petapixel.com/2026/01/18/a-look-at-the-worlds-first-full-ai-based-image-signal-processor
- JPEG committee, ข่าวประชาสัมพันธ์การประชุมครั้งที่ 103 (ค่าการประหยัดของ JPEG AI ความซับซ้อน และการใช้งานบนสมาร์ตโฟน): https://jpeg.org/items/20240507_press.html
- G. J. Sullivan, บทนำ ISO/IEC JTC 1/SC 29, เมษายน 2026 (สถานะของ JPEG AI; ตัวเลขของ ECM และ NNVC): https://www.mpeg.org/introduction-to-iso-iec-jtc-1-sc-29-coding-of-audio-picture-multimedia-and-hypermedia-information-the-world-of-jpeg-and-mpeg/
- คณะทำงาน IEEE 1857.11 (การเข้ารหัสภาพด้วยโครงข่ายประสาทเทียม): https://sagroups.ieee.org/fvc
- Chrome 155 กับ JPEG XL, 6 ตุลาคม 2026: https://xenospectrum.com/en/chrome-155-jpeg-xl-rust/

การโทรและการประชุม

- Microsoft, “Satin: Microsoft's latest AI-powered audio codec for real-time communications”, 17 กุมภาพันธ์ 2021: https://techcommunity.microsoft.com/t5/microsoft-teams-blog/satin-microsoft-s-latest-ai-powered-audio-codec-for-real-time/ba-p/2141382
- Microsoft, “Announcing the Open-Source Release of ML Video Codec (MLVC)”, 24 กรกฎาคม 2026: https://techcommunity.microsoft.com/blog/linuxandopensourceblog/announcing-the-open-source-release-of-ml-video-codec-mlvc/4539875 ; โค้ด: https://github.com/microsoft/mlvc
- Streaming Learning Center, “Microsoft MLVC: Interoperability, Power Realities, and the Neural Compression Patent Landscape”: https://streaminglearningcenter.com/articles/microsoft_mlvc.html ; บทวิเคราะห์ของ iReplay: https://ireplay.tv/blog/mlvc-microsoft-neural-video-codec-open-source-streaming/

- บันทึกการออกรุ่น Opus 1.5: https://opus-codec.org/demo/opus-1.5 ; Opus 1.6: https://opus-codec.org/demo/opus-1.6
- Meta Engineering, “MLow: Meta's low bitrate audio codec”, 13 มิถุนายน 2024: https://engineering.fb.com/2024/06/13/web/mlow-metas-low-bitrate-audio-codec/
- Microsoft, Copilot+ PCs (NPU 40+ TOPS): https://www.microsoft.com/en-us/windows/copilot-plus-pcs ; Apple, ประกาศเปิดตัว M4 (Neural Engine 38 TOPS): https://www.apple.com/newsroom/2024/05/apple-introduces-m4-chip/

Generative AI

- Kyutai, model card ของ Mimi (12.5 Hz, 1.1 kbps): https://huggingface.co/kyutai/mimi
- Lightricks, “LTX-Video: Realtime Video Latent Diffusion” (VAE วิดีโอ 1:192): https://arxiv.org/abs/2501.00103
- ข่าวประชาสัมพันธ์การประชุม MPEG ครั้งที่ 153 (Video Coding for Machines ถึงขั้น DIS): https://www.mpeg.org/wp-content/uploads/2026/03/press_release_153.pdf

เกมและกราฟิก

- NVIDIA, การนำ DLSS 4 ไปใช้: ช่วงเปิดตัวและ Computex 2025 (75 และ 125+): https://www.nvidia.com/en-us/geforce/news/125-dlss-4-multi-frame-gen-games-more-announced-computex-2025/ ; Gamescom 2025 (175+): https://www.nvidia.com/en-us/geforce/news/gamescom-2025-dlss-4-full-ray-tracing-game-announcements/ ; CES 2026 (250+, DLSS 4.5): https://www.nvidia.com/en-us/geforce/news/dlss-4-rtx-path-tracing-game-announcements-ces-2026/
- Ubisoft La Forge, “Shipping Neural Texture Compression in Assassin's Creed Mirage”, กุมภาพันธ์ 2026: https://www.ubisoft.com/en-us/studio/laforge/news/415slmB3ZGzv8d2LhO9QT/shipping-neural-texture-compression-in-assassins-creed-mirage
- TweakTown เรื่องการสาธิตเท็กซ์เจอร์นิวรัลของ NVIDIA ที่ GTC 2026 (6.5 GB เหลือ 970 MB): https://www.tweaktown.com/news/110886/nvidias-neural-texture-compression-cuts-vram-usage-from-6-5gb-down-to-970mb/index.html
- Microsoft DirectX blog, Shader Model 6.9 รุ่นจำหน่ายจริง (retail), 26 กุมภาพันธ์ 2026: https://devblogs.microsoft.com/directx/shader-model-6-9-retail-and-more/ ; เรื่อง cooperative vector: https://devblogs.microsoft.com/directx/shader-model-6-9-and-the-future-of-cooperative-vector/ ; แผนงานที่เปิดเผยในงาน GDC 2026: https://windowsreport.com/microsoft-showcases-neural-rendering-and-dxr-2-0-plans-for-directx-at-gdc-2026/

ลิงก์เฉพาะทาง

- Global Trade Alert เรื่องประกาศฉบับที่ 49 ปี 2026 ของจีน: https://globaltradealert.org/intervention/161396-china-government-revises-export-controls-on-drone-related-items ; DroneXL: https://dronexl.co/2026/10/19/china-drone-export-controls-motors-props-november-18/ ; Army Recognition (เกณฑ์ของแพลตฟอร์มภาคพื้นดินไร้คนขับ): https://www.armyrecognition.com/news/army-news/2026/china-tightens-drone-export-controls-raising-supply-chain-risks-for-the-us-europe-and-ukraine
- DroneLife, FCC เพิ่มโดรนและชิ้นส่วนที่ผลิตในต่างประเทศเข้าใน Covered List, 22 ธันวาคม 2025: https://dronelife.com/2025/12/22/fcc-adds-foreign-made-drones-and-components-to-covered-list-citing-national-security-risks/

ซิลิคอน

- M. Tikekar et al., “A 249-Mpixel/s HEVC Video-Decoder Chip for 4K Ultra-HD Applications”, IEEE JSSC 2014: https://dspace.mit.edu/handle/1721.1/93876 ; หน้าชิปของ MIT (78 mW ที่ 4K30): https://eecsg.mit.edu/chips/hevc-video-decoder-for-4k-ultra-hd-applications/
- S. Zhang et al., NVCA, DATE 2024: https://arxiv.org/abs/2312.10716
- Electronic Design, “Why Traditional VVC Will Outlive the Neural-Codec Hype on Silicon”, 2026: https://www.electronicdesign.com/home/article/55414955/why-traditional-vvc-will-outlive-the-neural-codec-hype-on-silicon
- Z. Jia et al., “Towards Practical Real-Time Neural Video Compression” (DCVC-RT), CVPR 2025: https://openaccess.thecvf.com/content/CVPR2025/html/Jia_Towards_Practical_Real-Time_Neural_Video_Compression_CVPR_2025_paper.html ; เอกสารเสริม (ความซับซ้อนรายโมดูล): https://openaccess.thecvf.com/content/CVPR2025/supplemental/Jia_Towards_Practical_Real-Time_CVPR_2025_supplemental.pdf
- NNVC และไลบรารีอนุมานจำนวนเต็ม SADL: https://arxiv.org/abs/2309.05846
- Tencent Canghai V2, 27 พฤษภาคม 2026 (Shanghai Securities News ผ่าน 10jqka): https://stock.10jqka.com.cn/20260527/c677016020.shtml ; สรุปฟีเจอร์: https://www.besthub.dev/articles/tencent-s-canghai-v2-chip-enters-mass-production-what-the-new-capabilities-mean-for-video-encoding-bcfc18dbfa29
- Chips&Media, IP ตัวถอดรหัส AV2, 12 พฤษภาคม 2026: https://www.design-reuse.com/news/202530501-chips-media-completes-development-of-next-gen-av2-hw-decoder-ip/ ; ไลเซนส์ AV2, 17 มิถุนายน 2026: https://www.edge-ai-vision.com/?p=58270
- Allegro DVT, IP ตัวถอดรหัส AV2, 16 มิถุนายน 2026: https://www.edge-ai-vision.com/2026/06/allegro-dvts-pulsar-decoder-ip-adds-support-for-av2-video-codec/
- VeriSilicon, VC9800D ที่รองรับ AV2, 10 มิถุนายน 2026: https://www.businesswirechina.com/en/news/63176.html

มาตรฐานและสิทธิบัตร

- ข่าวประชาสัมพันธ์การประชุม MPEG ครั้งที่ 155 (ประกาศเรียกรับข้อเสนอร่วมสำหรับ beyond VVC), 5 สิงหาคม 2026: https://www.mpeg.org/wp-content/uploads/2026/08/ag3w254SN26474Press-Release-MPEG155.pdf
- ข่าวประชาสัมพันธ์การประชุม MPEG ครั้งที่ 152 (ผลของประกาศเรียกรับหลักฐาน; ฉบับใหม่ของ VVC และ VSEI), 30 ตุลาคม 2025: https://www.mpeg.org/wp-content/uploads/2025/11/Press-Release-MPEG152.pdf ; คอลัมน์ MPEG ของ ACM SIGMM: https://records.sigmm.org/2025/12/09/mpeg-column-152nd-mpeg-meeting/

- Streaming Media, “AI Helps InterDigital Reach Beyond VVC in Race to Develop Next-Gen Codec”: https://www.streamingmedia.com/Articles/News/Online-Video-News/AI-Helps-InterDigital-Reach-Beyond-VVC-in-Race-to-Develop-Next-Gen-Codec-171971.aspx
- GreyB, ภาพรวมสิทธิบัตร beyond-VVC (H.267): https://greyb.com/blog/exclusive-video-codec-landscape/ ; เรื่องไลเซนส์ AV2 และการขายพอร์ตสิทธิบัตรของ Alibaba: https://greyb.com/blog/av2-licensing/
- MLex, “Fragmented codec licensing exposes need for clearer rules for Chinese tech majors”, 27 มีนาคม 2026: https://www.mlex.com/mlex/articles/2458402/fragmented-codec-licensing-exposes-need-for-clearer-rules-for-chinese-tech-majors
- Access Advance, อัตราค่าสิทธิของ MCBA และ VVC Advance ณ วันที่ 1 กรกฎาคม 2026: https://accessadvance.com/wp-content/uploads/2025/07/MCBA-Program-Overview-September-2026.pdf

เงินและบริษัท

- InterDigital ซื้อกิจการ Deep Render, 30 ตุลาคม 2025: https://www.advanced-television.com/2025/10/30/interdigital-acquires-deep-render
- TechCrunch, “Apple acquired a startup using AI to compress videos”, 27 มีนาคม 2023: https://techcrunch.com/2023/03/27/apple-acquired-a-startup-using-ai-to-compress-videos
- Neowin, “Sony acquires AI video startup iSIZE”, พฤศจิกายน 2023: https://www.neowin.net/news/sony-acquires-ai-video-startup-isize-to-enhance-playstation-streaming/

การบีบอัดแบบเรียนรู้ · การเล่นวิดีโอ × ซิลิคอน

การเข้ารหัสแบบนิวรัลในปี 2026

ใช้งานจริงรอบโคเดก ไม่ใช่แทนโคเดก

ตัวเลขสำคัญทุกตัวในรายงานนี้ระบุแหล่งที่มาและวันที่ ตัวเลขที่ผู้เขียนคำนวณเองกำกับว่าเป็นประมาณการ ส่วนตัวเลขที่มีเพียงผู้ขายเป็นผู้รายงานกำกับว่าเป็นคำกล่าวอ้างของผู้ขาย

ทัน ชว่าย

นักเทคโนโลยีและผู้ก่อตั้ง · นักวิจัยและนักพัฒนา AI

tanshuai.com · i@tanshuai.com

ฉบับเผยแพร่สาธารณะ · ข้อมูล ณ วันที่ 12 พฤศจิกายน 2026