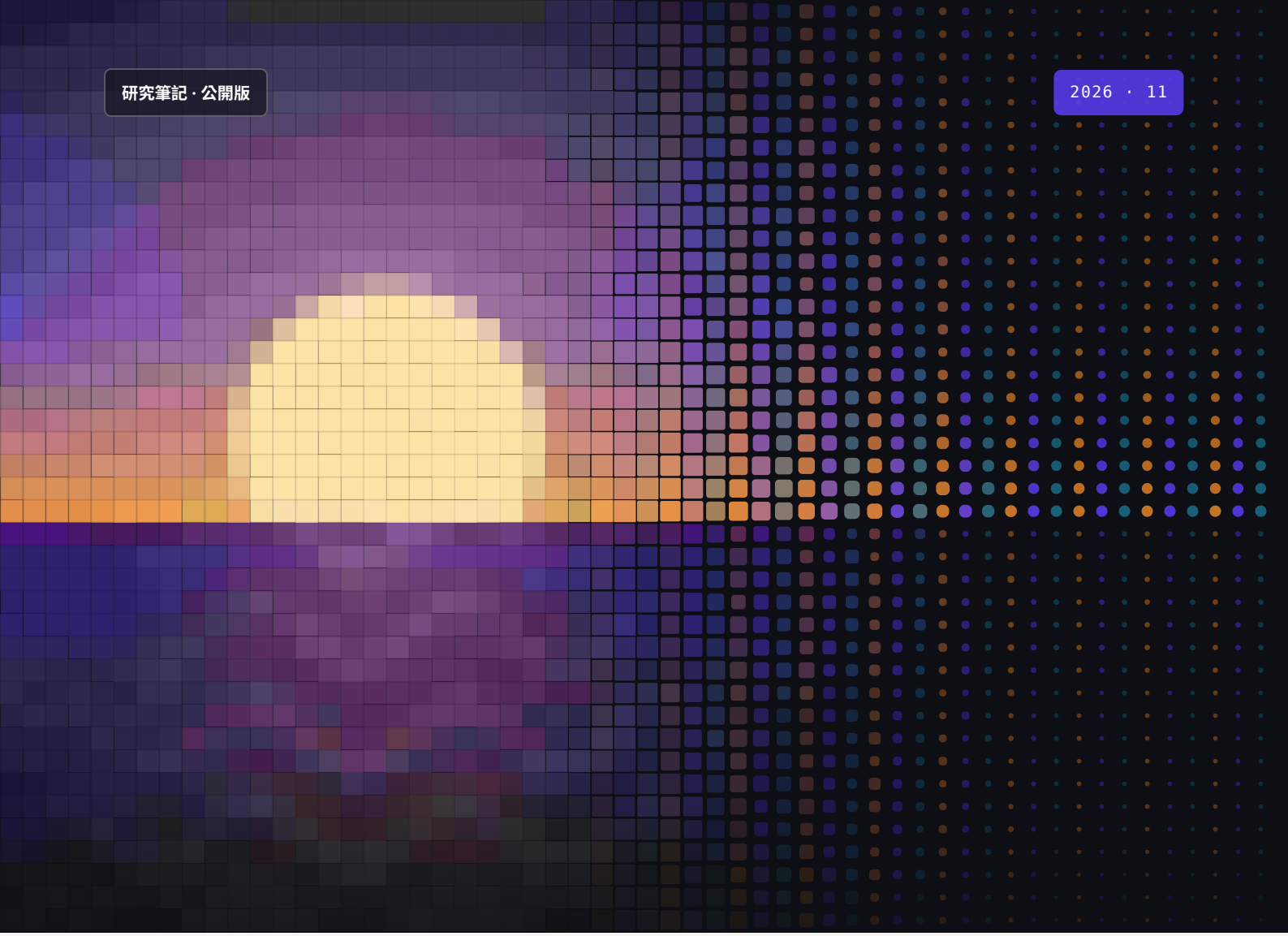


研究筆記 · 公開版

2026 · 11

學習式壓縮 · 播放 × 晶片

2026 年的神經編碼

在編解碼器周邊落地，而非取而代之

播放 · 攝影 · 通話 · 生成式 AI · 遊戲 · 晶片

結論

神經網路已在承載全球影片的編解碼器兩端運作。端對端神經編解碼器只在封閉迴路中落地，主流螢幕則要等 H.267。

譚帥

技術專家與創辦人 · AI 研發者

tanshuai.com · i@tanshuai.com

資料截至 2026 年 11 月 12 日

金額以美元計

目錄

目錄

◆ **執行摘要** 04
已是主流，但在編解碼器周邊

01 **範圍與方法** 07
本文所說的「神經編碼」，以及我的研究方法

02 **播放與串流** 08
播放與串流：兩端是神經網路，中間是標準

03 **攝影與影像** 11
攝影與影像：相機內是神經網路，網頁上是傳統格式

04 **通話與會議** 12
通話與會議：端對端神經編解碼器真正落地之處

05 **生成式 AI** 14
生成式 AI：規模最大、卻沒人稱之為編解碼器的部署

06 **遊戲與繪圖** 15
遊戲與繪圖：神經重建已是主流，神經素材剛起步

07 **專用鏈路** 16
專用鏈路：衛星、無人機與機器人

08 **晶片** 17
晶片視角：能耗、算力、記憶體與精確性

09 **標準與專利** 20
標準與專利：神經工具在哪裡成為可授權的技術

10 **誰在賺錢** 21
誰在賺錢，企業可以在哪裡切入

11 **說法與證據** 23
你會聽到的說法，以及證據怎麼說

12 **觀察指標** 24
2027 至 2030 年的觀察重點

13 **未解問題** 25

◆ **關於作者** 26
譚帥

◆ **資料來源** 27

圖表

圖表

圖 01	2026 年神經編碼在哪裡落地	06
圖 02	神經網路進入編解碼器的三種方式	07
圖 03	從標準定案到主流螢幕，要花上大半個十年	09
圖 04	在通話中，神經編解碼器比硬體 HEVC 少用 65–76% 位元率	12
圖 05	一家公司掌控整套技術堆疊時，神經重建普及得很快	15
圖 06	每解碼一個像素的能耗：固定功能解碼器領先一個數量級	17
圖 07	神經工具每像素需數萬次 MAC；完整的神經解碼器需數十萬次	18
圖 08	已落地的分工：位元串流交給固定功能硬體，像素交給神經網路引擎	19
圖 09	2026 年，誰從神經編碼賺到錢	22

文中標籤

- 估算 — 作者根據引用資料所做的計算或推論
- 廠商說法 — 由銷售該產品的公司發布，未經獨立驗證
- 分析 — 作者對證據的解讀

本文譯自英文版；內容如有出入，以英文版為準。

資料截至 2026 年 11 月 12 日 · 金額以美元計

執行摘要

已是主流，但在編解碼器周邊

結論

神經編碼在 2026 年已是主流，但它在標準編解碼器周邊運作，而不是取代它們

≈75%
2025 年底行動數據流量中影片的占比（愛立信）

12.6 億
2025 年智慧型手機出貨量，每一支都以固定功能硬體解碼影片

≈12×
唯一公開的神經解碼晶片設計的每像素能耗，相對於實測 HEVC 解碼晶片（含 DRAM）

≈22 TOPS
DCVC-RT 等級的神經解碼器在 1080p30 下所需的持續算力 估算

75.5%
通話中 MLVC 在 360p 下比硬體 HEVC 節省的位元率，以微軟自己的測試集計
廠商說法

2029 年底
H.267 第一版，第一個可能把神經工具帶到每一塊螢幕的標準

落地情況

。神經網路如今在承載全球幾乎所有影片的編解碼器兩端運作。在輸入端，它們塑造畫面：手機相機中的 AI 影像訊號處理器，以及 Netflix 編碼流程中的神經網路降取樣器。在輸出端，它們修復畫面：YouTube 電視版的超解析度、Meta 的超解析度，以及標榜搭載數百個神經網路的電視處理器。兩者之間的位元串流仍是 AV1、HEVC、AVC 或 VVC，由固定功能硬體解碼。沒有任何串流服務向電視或手機傳送神經視訊格式；如今裝置上若有神經視訊解碼器在執行，也只是在某家公司自己的 App 裡。

封閉迴路

端對端神經編解碼器（編碼器與解碼器都經過學習）只在封閉迴路中落地：由一家公司掌控兩端，並可退回標準編解碼器。微軟自 2021 年起在所有 Teams 與 Skype 雙人通話中使用 AI 音訊編解碼器；2026 年 7 月，它又開源了神經視訊編解碼器 MLVC，目前正在 Teams 點對點通話中推出：在 360p 下，達到相同主觀評分所需的位元率比硬體 HEVC 少約 75%。每個語音與影片生成模型都建立在學習式編解碼器之上，從 12.5 Hz 的語音 tokenizer，到壓縮比 1:192 的影片 VAE。神經紋理壓縮已在一款 AAA 遊戲中上市。這些都是真實的部署，而且有一個共同條件：沒有別人需要解碼這些串流。

時程

神經位元串流要大規模用於播放，得等一項標準與一輪晶片週期。第一個可能納入神經工具的標準，也就是 VVC 的後繼者（我稱之為 H.267），已於 2026 年 7 月發布徵求提案，2027 年 1 月評估回應，目標在 2029 年底推出第一版。AV1 從規格定案到占 Netflix 觀看量的 30%，花了將近八年。照這個模式，首批 H.267 晶片約在 2031 年問世，2030 年代中期才會成為主流 估算。與此同時，下一個走上螢幕的編解碼器是 AV2；它在 2026 年 5 月定案，裡面沒有神經解碼器，而三家 IP 廠商在三週內就宣布了 AV2 解碼器客戶。

晶片

原因在晶片上。唯一公開的神經視訊解碼晶片設計，每像素需要 14.6 nJ，約為實測 HEVC 解碼晶片含 DRAM 時的 12 倍、僅計核心時的 47 倍。微軟的即時研究型編解碼器 DCVC-RT，依我對其分模組數據表的計算，每解碼一個像素約需 17.5 萬次乘加運算，相當於 1080p30 下約 22 TOPS 的持續算力

估算：大約是定義 AI PC 的 40 TOPS 的一半，全花在一路視訊串流上。神經解碼器還必須在每台裝置上產生逐位元相同的結果，否則熵解碼就會崩潰。這就是 MLVC 為什麼要犧牲壓縮率，以便在各種 NPU 上解出相同的結果，而且只跑 540p，而不是 4K。

商業

從商業角度看，錢跟著標準與平台走，而不是跟著獨立的神經編解碼器走。平台把神經前處理與後處理省下的成本留給自己。IP 廠商為每個新標準授權固定功能解碼器，外加用於影像處理的 NPU。專利池按裝置收費：VVC 在美國及其他 Region 1 市場，每支手機 0.50 美元，每台 4K 電視 1.00–1.50 美元。最知名的三家神經編解碼新創公司 WaveOne、iSIZE 與 Deep Render，分別被蘋果、Sony 與 InterDigital 收購。仍有空間的機會在於神經前處理與後處理、封閉迴路產品、把固定功能編解碼器與 NPU 搭在一起的晶片，以及介於其間的評測與標準工作。

問題	判斷	關鍵數字（來源見各章）
① **神經編碼如今在哪裡大規模落地？**	在編解碼器周邊：相機 ISP、編碼端降取樣、播放端超解析度。端對端只在封閉迴路中：通話、遊戲素材、生成式 AI 的 tokenizer	影片占行動數據 ≈75%；2025 年出貨 12.6 億支手機與 ≈2 億台電視，全都以固定功能硬體解碼；DLSS 4 支援 250 多款遊戲與應用程式
② **神經編解碼器會在播放上取代 AV1、HEVC 與 VVC 嗎？何時？**	要先有一項標準與一輪裝置週期。H.267 是第一個現實可行的載體，很可能混用傳統與神經工具	H.267 於 2026 年 7 月徵求提案，2027 年 1 月評估，2029 年底推出第一版；2030 年代中期成為主流 估算
③ **在晶片上要付出多少代價？**	每像素能耗比如今的固定功能解碼高一個數量級，還有記憶體存取流量與逐位元一致的要求	14.6 nJ/像素（神經）對 1.19（HEVC，含 DRAM）；每解碼像素 ≈175 kMAC ≈ 1080p30 下 22 TOPS 估算
④ **標準與專利允許什麼？**	VVC 串流已可標示神經後處理濾波器；JPEG AI 已是 ISO 標準；H.267 將決定哪些神經工具成為可授權的技術	ECM（傳統）比 VVC 省 ≈28%；神經網路迴路濾波器在 10–22+ kMAC/像素下省 6–14%；字節跳動在 VVC 後繼標準中擁有 645 個專利家族
⑤ **誰在賺錢？企業可以在哪裡切入？**	平台、IP 廠商、專利權人。純神經編解碼新創公司都被收購了。可切入前處理／後處理、封閉迴路、編解碼器加 NPU 的晶片、評測	VVC 專利池每支手機 0.50 美元；AV2 解碼器 IP 在規格定案後數週內就完成授權；三起收購，條件均未公開
⑥ **哪些說法要打折扣？**	相對舊編解碼器、在低解析度下、未計解碼成本，或來自展示而非已上市產品的增益數字	MLVC：比 H.264 少 88%，但比硬體 HEVC 少 65–76%；紋理：展示中省下 85% VRAM，已上市遊戲約 30%

圖 01

2026 年神經編碼在哪裡落地

神經網路已在主流編解碼器的兩端運作。端對端神經編解碼器只在一家公司同時掌控編碼器與解碼器的地方落地。

	標準編解碼器周邊 前處理、畫質增強、升頻	端對端神經編解碼器 學習式編碼器與解碼器	標準內部 標準位元串流或 API 中的神經工具
01 播放與串流	● 大規模部署 Netflix 降取樣器；YouTube 與 Meta 超解析度；電視 AI 升頻晶片	● 未部署 沒有串流服務傳送神經視訊格式；電視與手機以硬體解碼標準編解碼器	● 標準化中 H.267 第一版預定 2029 年底完成；神經工具名列候選
02 攝影與影像	● 大規模部署 旗艦手機的 AI ISP；CES 2026 展出以 IP 形式提供的全神經網路 ISP	● 展示與試行 JPEG AI 已是 ISO 標準：有手機展示，沒有產品	● 未部署 正在普及的是傳統格式：Chrome 155 預設啟用 JPEG XL
03 通話與會議	● 大規模部署 接收端超解析度與神經網路噪音抑制	● 大規模部署 Teams 雙人通話自 2021 年起使用神經音訊；MLVC 視訊正在 Teams 推出	● 有限部署 Opus 1.5 加入可選用、向下相容的機器學習工具
04 生成式 AI	—	● 大規模部署 每個影片與語音模型都執行學習式 tokenizer，例如壓縮比 1:192 的影片 VAE	—
05 遊戲與繪圖	● 大規模部署 神經網路升頻與畫格生成：光是 DLSS 4 就支援 250 多款遊戲與應用程式	● 有限部署 一款 AAA 遊戲已採用神經紋理；輝達的 SDK 仍為測試版	● 標準化中 DirectX 正為著色器加入矩陣運算（預覽版）
06 專用鏈路：衛星、機器人	● 展示與試行 在伺服器端修復低位元率畫面	● 展示與試行 抗遺失神經編解碼器處於研究與試行階段	—

資料來源: 報告引用的公司、標準與開發者文件；狀態截至 2026 年 11 月 12 日。各列依各領域觸及的人數排序。

第 01 章

01 範圍與方法 本文所說的「神經編碼」，以及我的研究方法

神經網路進入編解碼器的三種方式

「神經編碼」其實涵蓋三件不同的事，關於它的爭論大多源自把三者混為一談。

01 **在編解碼器周邊**。神經網路在編碼前或解碼後處理像素，位元串流維持標準格式。例如：相機中的 AI 影像訊號處理器（ISP）、Netflix 在編碼器之前的神經網路降取樣器、電視上的超解析度。任何裝置都能播放這樣的串流；神經網路只改變輸入或輸出的內容。

02 **在標準內部**。神經網路成為標準編解碼器中的一項工具，例如每個解碼器都必須執行的神經網路迴路濾波器或畫面內預測器，因此必須被精確規範，通常以整數運算定義。較輕量的形式是神經後處理濾波器：由位元串流以中繼資料描述，解碼器可自行選擇是否套用。

03 **取代編解碼器**。編碼器與解碼器都經過端對端學習。位元串流是一組量化後的潛在值，只有配對的神經網路才能解碼。例如：Teams 中的微軟 MLVC、語音模型中的語音 tokenizer、遊戲中的神經紋理。

圖 02 **神經網路進入編解碼器的三種方式**

2026 年已落地的，大多在標準編解碼器周邊運作。標準內部的神經工具正在制定中。端對端神經編解碼器只在一家公司掌控兩端時才取代編解碼器。



資料來源: 作者綜整本報告引用的資料；狀態截至 2026 年 11 月 12 日。

第一種不需要任何人許可。第二種需要一項標準。第三種需要掌控兩端。這正是它們觸及大眾的先後順序，也解釋了後文的大部分內容。

依現實世界排序，而不是依研究議程

研究論文大多研究第三種，因為壓縮紀錄都出在那裡。本報告依各領域觸及的人數排序，因為這才決定資金與工程資源流向哪裡：先是播放與串流，接著是攝影、通話、生成式 AI、遊戲，最後是專用鏈路（衛星、無人機、機器人）。依愛立信在 2025 年底的統計，影片約占所有行動數據流量的 75%。2025 年，業界出貨約 12.6 億支智慧型手機（IDC），以及約 1.96–2.09 億台電視（TrendForce 與 Omdia 的估計）。每一台都以固定功能硬體解碼影片。

我的研究方式

這是一份文獻研究，做法就像我審查一顆晶片或一份產品計畫。我讀了標準文件與會議報告（JVET、MPEG、JPEG、AOMedia、IETF）、企業工程部落格與發布說明、晶片論文、專利池費率表、出口管制條文與新聞報導。每個數字都附上日期。凡是我自行推導的數字，例如從論文表格算出的每像素運算量，我都會說明，並標上 估算。無法對照原始出處查證的說法，標為 廠商說法。金額以美元計。狀態截至 2026 年 11 月 12 日。

能做同類比較時，我盡量做同類比較；做不到時，我會說明。壓縮增益若沒有附帶四件事，就沒有什麼意義：比較基準（相對什麼的增益）、指標（PSNR、VMAF 或觀眾評分小組）、內容與解析度，以及解碼成本。只要把這四項講清楚，關於神經編解碼器的爭議大多就會消失。

第 02 章

02 播放與串流：兩端是神經網路，中間是標準

播放與串流

觀眾收到的仍是標準位元串流

最大的幾家影音服務剛完成一次編解碼器換代，又開始了下一次，兩次都與神經解碼器無關。2025 年 12 月，Netflix 報告 AV1 約占其觀看量的 30%，是使用量僅次於 AVC 的編解碼器。AV1 是在硬體解碼器進入電視之後，才跟著進入電視：Netflix 在 2019 年把 AV1 納入裝置認證，2021 年底起向電視串流，並表示 2021 年至 2025 年間送交認證的大螢幕裝置有 88% 支援它，而且幾乎都是 2023 年以後的機種。YouTube 表示，截至 2026 年 3 月 19 日，按觀看時數加權，其內容庫有超過 90% 已提供 AV1 版本。

AV2 是開放媒體聯盟推出的 AV1 後繼者，於 2026 年 5 月 28 日推出 1.0.0 版；依 AOMedia 的客觀量測，壓縮效率比 AV1 高約 30%。YouTube 計畫「最快在 2027 年初」於網頁上提供 AV2 串流。晶片隨即跟上：Chips&Media 在 5 月完成 AV2 硬體解碼器，6 月授權給一家未具名的北美大型科技公司；Allegro DVT 在 6 月 16 日前已向三家客戶交付 AV2 解碼器 IP；VeriSilicon 在 6 月 10 日前已向數家客戶推出支援 AV2 的視訊 IP。AV2 是傳統編解碼器：它的增益來自更好的區塊分割、預測、轉換與迴路濾波器（其中之一是「資料驅動」的引導式細節濾波器），而不是端對端神經網路。

VVC 是現行的 ITU/ISO 標準，它的第一個廣播應用落腳巴西：巴西的 DTV+（TV 3.0）系統於 2026 年 6 月正式商轉，時間點配合世界盃。截至 2026 年 7 月，巴西市售電視都還沒有內建 DTV+ 接收器；第一批預計在 2026 年底到 2027 年初之間上市。

在編碼端：Netflix 的神經網路降取樣器

播放領域中最乾淨俐落的神經編碼例子，從 2022 年運作至今，而且觀眾的裝置不必做任何改變。Netflix 會把每部作品編碼成多種解析度；從 4K 母帶產生 1080p、720p 與 540p 的降取樣步驟，傳統上使用 Lanczos 之類的固定濾波器。Netflix 改用一個小型神經網路取代它：一個前處理區塊，後接一個縮放區塊，訓練目標是提升影片在裝置上經標準升頻器放大後的

觀感品質。Netflix 形容這是一個可直接替換的方案：編碼端與用戶端都不需要其他改動，因此「數百萬台支援 Netflix 串流的裝置都自動受益」。在 Netflix 的偏好測試中，約 77% 的觀眾選擇了神經網路降取樣器的輸出 廠商說法。為了在 Netflix 的規模下把成本控制在可接受範圍，「只需要幾層神經網路就夠了」。

這就是值得記住的模式。神經網路每部作品只在資料中心執行一次，那裡的算力便宜、可以排程，而它的輸出是一般的 AV1 或 HEVC 串流。

在螢幕端：App 與電視中的超解析度

在另一端，神經網路修復編解碼器與片源所損失的部分。2025 年 10 月 29 日，YouTube 宣布在電視版 App 推出超解析度功能：解析度低於 1080p（240p 到 720p）的上傳影片，會以機器學習升頻到 HD，並以 4K 為既定目標。升頻後的版本會加上標示，觀眾可以切換回原始版本，創作者也可以選擇退出。Meta 在 2025 年 10 月說明了它如何執行影片超解析度：一方面在上傳時於伺服器端執行（與英特爾合作，使用標準 x86 CPU），另一方面在播放時於用戶端執行，並發現兩者相輔相成。電視廠商則在晶片中做同樣的事：三星的旗艦 8K 處理器 NQ8 AI Gen3，標榜以 768 個神經網路進行升頻。

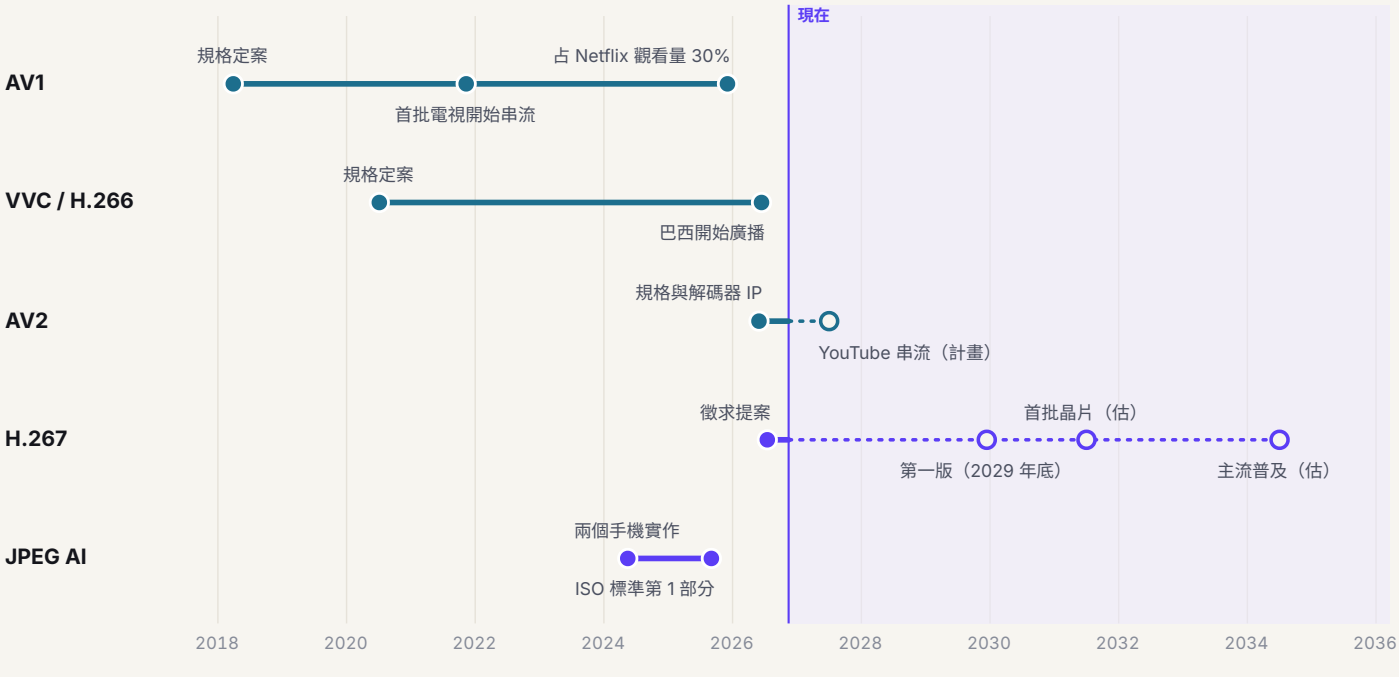
經濟效益依神經網路在哪裡執行而不同。伺服器端處理，每處理一部影片，平台都要付出算力成本。裝置端處理則不花平台任何成本；它用的是買家已經付過錢的 NPU。這就是為什麼 2026 年最大的效益同時來自兩端。

為什麼沒有電視解碼神經位元串流

播放格式必須能在日後所有要播放它的裝置上解碼，而這些裝置一用就是好幾年。因此，即使是傳統編解碼器，從標準定案到主流螢幕的路也很漫長。AV1 的規格於 2018 年 3 月定案；dav1d 軟體解碼器在幾個月內推出；電視從 2021 年底開始串流；到 2025 年 12 月，它約占 Netflix 觀看量的 30%，距規格定案將近八年。VVC 於 2020 年 7 月定案，2026 年 6 月才迎來第一個廣播服務。

圖 03 **從標準定案到主流螢幕，要花上大半個十年**

播放需要裝置內建硬體解碼器。AV1 從規格定案到占 Netflix 觀看量的 30%，花了將近八年；第一個可能納入神經工具的標準預定 2029 年底完成。



資料來源: Netflix 技術部落格（2025 年 12 月）；AOMedia；Allegro DVT；TV Technology；Fraunhofer IIS（巴西 DTV+）；MPEG-155；JPEG 委員會；ISO。空心標記為目標或作者估算。

神經位元串流還多了兩個傳統編解碼器早已解決的問題：每像素能耗，以及逐位元一致的解碼。兩者都在第 8 章討論。這兩個問題加在一起，意味著在標準明確告訴廠商該做什麼之前，沒有任何電視或手機廠商會在每台裝置裡放進神經視訊解碼器。

什麼能改變現狀：H.267

能把神經工具帶到每一塊螢幕的標準，是 VVC 的後繼者，由 ITU-T 與 ISO/IEC 在聯合視訊專家小組（JVET）中共同開發。它的名稱尚未定案；我跟業界許多人一樣，稱之為 H.267。

2025 年 10 月，JVET 評估了徵求證據的回應：來自 16 家公司與研究實驗室的五份回應，共測試 20 種編解碼器組態。委員會發現了壓縮效率優於 VVC 的證據，並認為改良後的傳統技術與神經網路技術都「值得關注」。神經網路方法展現出明顯的感知增益，在 8K HDR 內容上尤其如此，在遊戲內容上的增益則較小。回應者之一 InterDigital 提交了兩個設計：一個以傳統的增強壓縮模型（ECM）為基礎，另一個則在 VVC 上疊加 AI 工具。

正式的徵求提案在 2026 年 7 月的日內瓦會議上發布。主要提案的截止日是 2026 年 10 月底，補充資料包於 12 月中旬截止，所有回應將在 2027 年 1 月於布里斯本舉行的 JVET 會議上評估。JVET 預期在 2027 年下半年推出第一個測試模型，2029 年底前推出標準第一版，並形容這次工作「可能是有史以來為啟動一項新標準所投入的最大心力」。

傳統路線設下了很高的門檻。以傳統訊號處理打造的 ECM，在隨機存取組態下比 VVC 節省約 28%。JVET 的神經網路探索報告指出，迴路濾波器在每像素 1 萬次乘加運算下約省 6%，2.2 萬次下省 8%，在高複雜度操作點下省 14%。我的解讀是：H.267 最可能以傳統工具為核心，只在節省的位元值得付出解碼成本之處加入少數神經工具，而 2027 年 1 月的評估將顯示會有幾項 分析。

其實已經有一座橋梁。VVC 與 HEVC 串流可以攜帶補充增強資訊（SEI），描述供解碼端套用的神經後處理濾波器；2025 年 10 月完成的 VVC 新版及其 SEI 標準（VSEI）新版，擴充了這些訊息。串流可以指明解碼後要執行哪個神經網路，而不必強制每個解碼器都執行它。

這對播放領域的企業意味著什麼

現在就賣在編解碼器周邊可行的東西：為影音服務提供編碼端前處理，為電視與機上盒晶片提供後處理區塊與調校。在 2030 年代之前，不要把播放業務建立在必須於第三方裝置上解碼的神經位元串流上。如果你做晶片，近期的營收來自 AV2 與 VVC 解碼器；神經的部分，是在它們旁邊為畫質增強提供一條 NPU 路徑（第 8 章）。

第 03 章

03 攝影與影像：相機內是神經網路，網頁上是傳統格式

相機處理流程已經神經網路化

近年旗艦手機拍下的每一張照片，在壓縮之前都經過了神經網路。高通的 Snapdragon 8 Gen 2 用於 2023 年起的手機，導入了「Cognitive ISP」，在拍照與錄影時執行即時語意分割，分別調整臉部、頭髮、天空與背景；三星 Galaxy S23 Ultra 是第一款在前鏡頭上使用它的手機。2026 年 1 月，Chips&Media 與 Visionary.ai 在 CES 上展示了他們所稱的第一個完全以 AI 為基礎的 ISP：一條端對端的影像處理流程，以 16 位元浮點完全在 NPU（Chips&Media 的 WAVE-N）上執行，不需要固定功能 ISP，並可透過 OTA 更新。Visionary.ai 執行長說：「現有的處理流程是把神經網路區塊外掛在固定功能 ISP 上。我們則是完全取代傳統 ISP。」展後，這套方案已提供給製造商評估。

之後，結果被存成 JPEG 或 HEIF，也就是任何裝置都能開啟的傳統格式。模式與播放相同：單一公司的處理流程內部用神經網路，邊界上則用標準格式。

JPEG AI：已定案的神經標準，但還沒有產品

JPEG AI 是 ISO/IEC 與 ITU-T 推出的第一個學習式影像編碼標準（ISO/IEC 6048、ITU-T T.840）。第 1 部分，也就是核心編碼系統，於 2025 年發布。JPEG 委員會報告，相對 VVC 畫面內編碼的壓縮增益為 12.5% 到 27.9%，解碼複雜度範圍為每像素 7 千到 21.6 萬次乘加運算。它的設計包含兩個編碼器與三個解碼器，經過聯合訓練，因此接收端只需實作一個解碼器，仍能讀取每一個串流。到 2024 年 5 月，委員會已報告兩個智慧型手機上的實作。與此同時，IEEE 於 2024 年 12 月發布了 IEEE 1857.11-2024，這是一項為 NPU 設計的神經網路影像編碼標準。

截至 2026 年 11 月，我沒有找到任何預設寫入或讀取 JPEG AI 的手機相機、照片 App 或瀏覽器。

網頁轉而採用的：JPEG XL

2026 年真正擴大版圖的影像格式是傳統格式。2026 年 10 月 6 日，Google 在 Chrome 155 中預設啟用 JPEG XL 解碼，使用以 Rust 撰寫的新記憶體安全解碼器；距它移除實驗性支援已有四年。Safari 自 2023 年起支援解碼 JPEG XL，Mozilla 也朝同一方向前進。Google 支持它的理由是：壓縮效率優於 JPEG、可無損再壓縮既有的 JPEG 檔案、支援 HDR 與漸進式解碼。

這個教訓可以推而廣之。必須在任何地方都能開啟的格式，靠的是快速、確定性、能在任何 CPU 上執行且可做到記憶體安全的解碼器。神經影像解碼器需要 NPU 或 GPU，還需要跨裝置的精確運算。神經編碼勝出的地方，是裝置自身的處理流程內部，以及由同一家公司負責編碼與解碼的系統，例如從不把壓縮檔交給別人的雲端照片服務。

這對影像領域的企業意味著什麼

錢在拍攝處理流程裡：AI ISP 的 IP、為影像處理定規模的 NPU，以及相機廠商願意付費的調校工作。在儲存與傳遞方面，學習式影像編解碼器如今是封閉系統內的工具，而不是值得押注的格式。

第 04 章

04 通話與會議：端對端神經編解碼器真正落地之處

音訊先行

微軟在 2021 年 2 月推出以 AI 驅動的語音編解碼器 Satin，並表示它已用於 Teams 與 Skype 的所有雙人通話。Satin 在 6 kbps 下提供超寬頻語音，音訊頻寬最高 16 kHz；Skype 先前的 Silk 編解碼器在同樣的位元率下只有窄頻。省下的位元率被用在冗餘上，讓通話更能承受封包遺失。

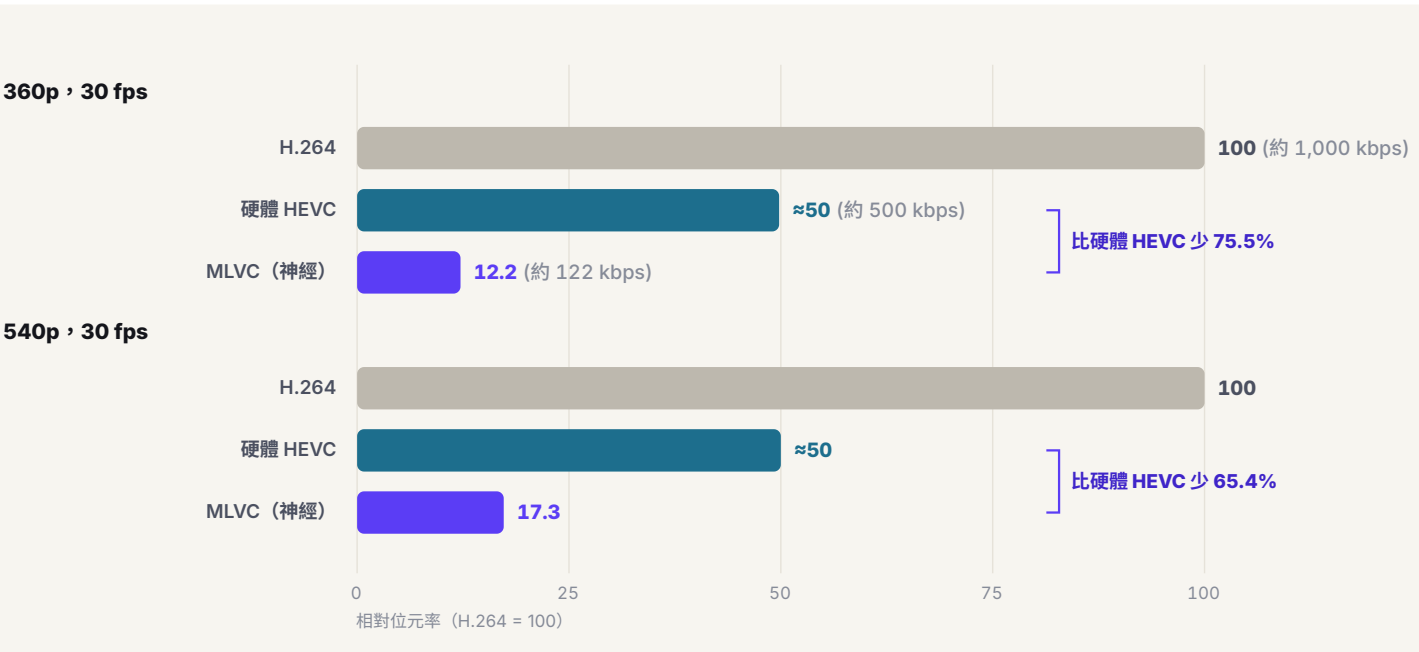
MLVC：用於正式通話的神經視訊編解碼器

2026 年 7 月 24 日，微軟以 MIT 授權開源了它的機器學習視訊編解碼器 MLVC，內容包括模型程式碼、訓練好的權重、訓練腳本，以及針對不同 NPU 的轉換工具。MLVC 是微軟研究院自 2021 年起陸續發表的 DCVC 系列的產品化版本。微軟表示，它正在 Teams 點對點通話中推出 MLVC，並搭配即時遙測、A/B 測試，以傳統編解碼器作為備援。

MLVC 以每秒 30 格執行 540p 與 360p，經過工程設計，在蘋果的神經網路引擎以及高通與英特爾的 NPU 上，最多只占用一半的 NPU，並把工作分配給 CPU 與 NPU。在微軟自家視訊會議資料集上依 ITU-T P.910 進行的主觀測試中，它在 360p 下所需的位元率比 H.264 少 87.8%，比硬體 HEVC 少 75.5%；在 540p 下則分別少 82.7% 與 65.4% 廠商說法。在 360p 下，這大約是 122 kbps，而 H.264 約需 1 Mbps。

圖 04 **在通話中，神經編解碼器比硬體 HEVC 少用 65–76% 位元率**

在微軟視訊會議測試集上達到相同主觀評分所需的位元率，以 H.264 = 100 計。MLVC 在筆電與手機的 NPU 上執行，最高 540p30，最多占用一半的 NPU。



資料來源: 微軟 MLVC 發布（2026 年 7 月 24 日）：在其視訊會議資料集上依 ITU-T P.910 進行的主觀測試。HEVC 長條由公布的節省比例推得。均為廠商說法，非獨立量測。

有兩點會沖淡這個標題數字。88% 是相對 H.264；相對同一台筆電裡既有的硬體 HEVC，是 65–76%，依然可觀。此外，MLVC 為了能實際部署而放棄了一部分壓縮率：一份獨立分析指出，以感知品質訓練的 DCVC-RT 在同一項衡量上達到 81.9%，MLVC 則是 75.5%。要在不屬於微軟的硬體上解出完全相同的結果，代價約是六個百分點。

為什麼通話是天然的落腳處

通話是封閉迴路。兩個端點執行同一個應用程式，在通話開始時協商彼此的能力，並可在通話途中退回標準編解碼器。解析度不高，網路往往不佳，而兩端的筆電如今都內建 NPU，手機更是早已如此。微軟把 40 TOPS 的 NPU 訂為 AI PC 的門檻，蘋果 M4 的神經網路引擎則標示為 38 TOPS。在播放領域阻擋神經編解碼器的每一個條件，在這裡都不存在。

反例：Opus 與 MLow

並非每個通話平台都做了同樣的選擇，而原因值得借鏡。

Opus 是 WebRTC 與多數通話 App 所用的開放音訊編解碼器，在 1.5 版（2024 年 3 月）加入了深度學習，這是它第一次用神經網路來處理或生成訊號本身。但這些都以選項形式加入：不需更動位元串流的解碼端增強（LACE 與 NoLACE）、神經網路封包遺失隱藏，以及深度冗餘機制（DRED）；全都預設關閉，以免加重舊裝置的負擔。Opus 1.6 加入了神經網路頻寬擴展，以及縮小為原來三分之一的 DRED 模型，並與 2012 年的標準保持完全相容。這種做法是「在標準內部、可選用且向下相容」。

Meta 則退得更遠。2024 年 6 月，它推出了 MLow，一個以傳統訊號處理而非神經網路打造的低位元率語音編解碼器，目前已用於所有 Instagram 與 Messenger 通話，並正在 WhatsApp 推出。Meta 的理由在於使用者的裝置：它有超過 20% 的通話是在 ARMv7 手機上進行，每天還有數千萬通 WhatsApp 通話在用了十年以上的手機上進行。Meta 寫道，包括它自己的 EnCodec 在內，神經編解碼器只能在昂貴的手機上穩定執行。MLow 在 6 kbps 下的品質約為 Opus 的兩倍，運算量卻少 10%。

對使用者橫跨旗艦筆電與十年老手機的服務來說，複雜度預算決定一切。神經編解碼器會先在既有裝置較新、且廠商掌控兩端的地方落地。

這對通訊領域的企業意味著什麼

企業視訊會議、遠端臨場、遠端協助、節目回傳鏈路與專用網路，是端對端神經編解碼器今天就能成為產品的地方，前提是兩端都由你掌控。從一開始就要規劃三件事：備援編解碼器、在你支援的每一款 NPU 上做解碼測試，以及用你自己的內容而非廠商的資料集做品質評估。

第 05 章

05

生成式 AI

生成式 AI：規模最大、卻沒人稱之為編解碼器的部署

每個語音與影片模型都建立在學習式編解碼器之上

端對端神經壓縮最大的部署，不在任何編解碼器標準裡，而在生成式 AI 裡。語音模型並不直接聆聽或產生波形：它讀寫的是學習式音訊編解碼器的離散 token。Kyutai 於 2024 年發布的 Mimi，以每秒 12.5 個音框、1.1 kbps 把語音轉成 token，低於任何電話語音編解碼器，並在同一個串流中結合語意與聲學資訊。Meta 於 2022 年發表的 EnCodec 是另一個例子；Meta 自己的工程師後來形容這類編解碼器在低位元率下表現出色，但對舊手機來說太吃重（第 4 章）。

影片生成也是同樣的道理。擴散模型或 Transformer 模型在學習式影片自編碼器（VAE）的潛在空間中生成，再由 VAE 的解碼器把結果轉成像素。LTX-Video 的 VAE 把影片壓縮為 1:192：每個潛在 token 代表 8 個影格中 32×32 像素的區塊，共 128 個通道。它的解碼器還負責最後一步去噪，因此它有一部分是在生成細節，而不只是重建細節。

為什麼它們不是發行格式

每一次對語音或影片模型的請求都會執行這些編解碼器，但沒有一個是別人會去解碼的格式。每個模型都有自己的潛在空間；沒有位元率控制、沒有一致性測試，也無法保證換一個解碼器會產生相同的畫面。它們是為生成品質而設計，而不是為特定位元率下的保真度而設計，並在資料中心的 GPU 上解碼。它們的成功證明，只要一家公司同時擁有編碼器、解碼器與硬體，學習式壓縮就能大規模運作。

這對企業意味著什麼

有兩個實務重點。第一，對提供生成式影片或語音服務的業者而言，tokenizer 的效率直接就是一項成本：對於會對所有 token 做注意力運算的 Transformer，運算量的成長快於 token 數，因此壓縮率更高的潛在表示，能讓生成成本以超過等比例的幅度下降 分析。懂得潛在表示設計、熵模型與解碼器效率的編解碼器工程師，對生成式 AI 團隊很有價值。第二，邊界正在移動：標準組織已開始研究為機器而非為人編碼。MPEG 的 Video Coding for Machines（ISO/IEC 23888-2）於 2026 年初進入國際標準草案階段，預計 2026 年底定案；JPEG AI 則從一開始就設計成既服務人眼觀看，也服務電腦視覺任務。

第 06 章

06 遊戲與繪圖：神經重建已是主流，神經素材剛起步

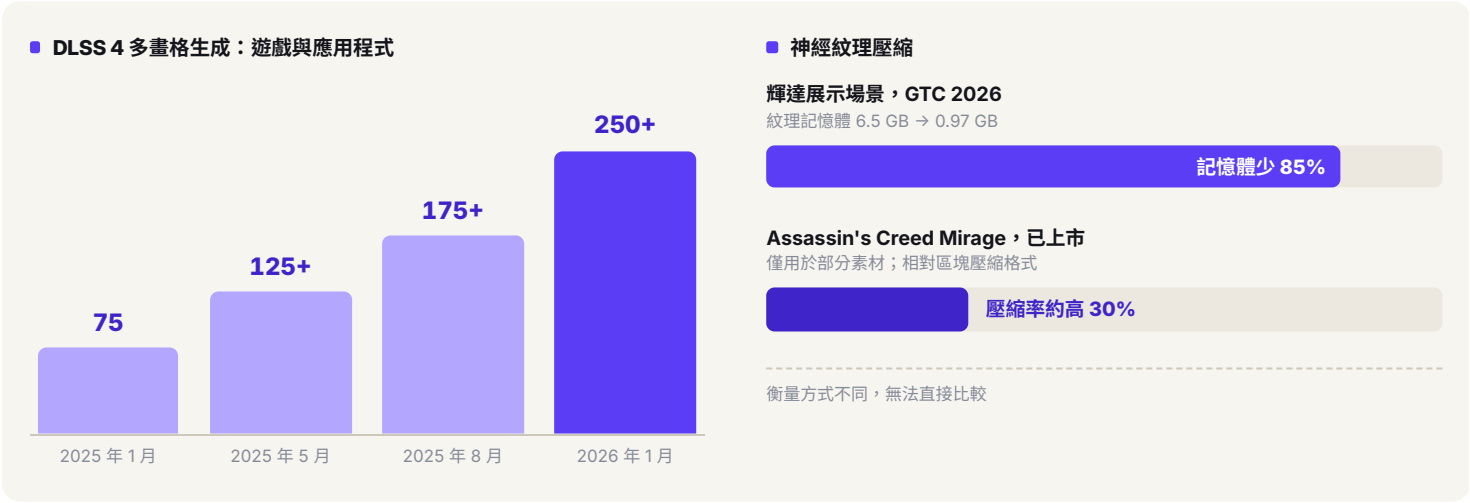
遊戲與繪圖

少算繪一些像素，其餘交給神經網路補上

在即時繪圖中，玩家看到的大部分畫面已經由神經網路重建。輝達的 DLSS 只算繪較少的像素與畫格，其餘由神經網路產生。搭載多畫格生成的 DLSS 4 於 2025 年 1 月推出，當時有 75 款遊戲與應用程式支援；依輝達統計，2025 年 5 月超過 125 款，8 月超過 175 款，2026 年 1 月超過 250 款，輝達也在當月發表了 DLSS 4.5，以第二代 Transformer 模型做超解析度。這除了名稱之外，根本就是一個解碼器：「位元串流」是一張低解析度畫格加上運動資料，由神經網路重建出完整影像。

圖 05 **一家公司掌控整套技術堆疊時，神經重建普及得很快**

左：支援輝達 DLSS 4 多畫格生成的遊戲與應用程式。右：神經紋理壓縮在展示場景與已上市遊戲中的成效；後者因執行成本，只用在部分素材上。



資料來源: 輝達（2025 年 1 月至 2026 年 1 月；GTC 2026 展示）；Ubisoft La Forge（2026 年 2 月）。均為廠商說法；兩項紋理數據的衡量對象不同。

神經紋理已經上市，但有所取捨

紋理以固定的區塊壓縮格式（BCn）壓縮，由 GPU 以硬體解碼。神經紋理壓縮則以精簡的潛在表示，加上一個在需要時重建紋素的小型神經網路，取代一整組材質紋理。2026 年 2 月，Ubisoft 的 La Forge 實驗室說明了他們如何在 2023 年的遊戲 Assassin's Creed Mirage 中實際採用這項技術：在相同視覺品質下，有效壓縮率比傳統格式高約 30% 廠商說法。它並沒有全面使用：「使用範圍受到執行效能的限制」，因此 Ubisoft 只把它用在實例數量多、記憶體壓力大的部分素材上。

輝達的 RTX Neural Texture Compression SDK 目前在 GitHub 上以測試版提供，展現了這項技術的上限。在 GTC 2026 的展示中，一個別墅場景使用 BCn 時需要 6.5 GB 紋理記憶體，改用神經紋理後只需 970 MB，減少 85% 廠商說法。這套 SDK 提供兩種路徑：在載入時解碼成 BCn，可節省磁碟空間與下載容量，但不省顯示記憶體；或在取樣時解碼，可省下記憶體，但每個畫格都要耗費著色器時間。

API 即將到位

繪圖 API 正在加入神經解碼在著色器內所需的矩陣運算。微軟於 2026 年 2 月推出 Shader Model 6.9；它在 2025 年預覽的 cooperative vector 功能，已併入一套統一的線性代數設計，留待後續的 Shader Model 版本，而 DirectX Linear Algebra 則於 2026 年 4 月進入公開預覽。等到每個驅動程式都支援之後，神經紋理解碼器就能以同一套程式路徑，在任何廠商的 GPU 上執行。

這對企業意味著什麼

遊戲之所以能採用神經解碼，是因為封閉迴路的條件成立：一套引擎、一條素材管線、GPU 內的張量硬體，以及一種備援格式。尚待開發的工作在於工具、能編碼神經材質的素材管線、跨廠商的執行環境，以及最重要的行動裝置；在行動裝置上，每解碼一個紋素的能耗成本決定一切。

第 07 章

07 專用鏈路：衛星、無人機與機器人

專用鏈路

就紙面而言，這些是最適合端對端神經編碼的場景。一個擁有者掌控鏈路兩端，頻寬稀缺且容易遺失資料，內容範圍窄到足以針對性訓練，而且畫面只要在感知上、甚至只要在語意上忠實，往往就足夠了。針對這類鏈路的抗遺失神經編解碼器是活躍的研究領域，而在伺服器端修復低位元率畫面的做法也已有人使用。

與消費性影片相比，它們也是一個規模小、而且受管制日益增加的市場 分析。在中國，商務部會同另外兩個機關發布的 2026 年第 49 號公告，於 10 月 19 日發布、2026 年 11 月 18 日生效，修訂了 13 類無人機相關品項的出口管制。據產業媒體報導，在技術類條目中，有一項涵蓋無人地面平台，其定義部分取決於門檻指標，例如 75% 的語意影像壓縮，以及在 150 毫秒內完成影像重建。語意壓縮本身已經成為出口管制的參數。在美國，聯邦通訊委員會於 2025 年 12 月 22 日將外國製無人機及其關鍵元件列入其「涵蓋清單」（Covered List），因此新的外國機型再也無法取得設備授權。

結論：對擁有工業或政府客戶、並有能力處理雙方出口合規的專業公司來說，這是一個穩健的利基市場；但它不是主流市場，也不是驗證神經編碼能否大規模運作的地方。

第 08 章

08 晶片視角：能耗、算力、記憶體與精確性

這一章是我自己的背景最派得上用場的地方。我有多年無晶圓廠晶片研發與管理經驗，包括全客製化 ASIC 資料路徑設計，以及視訊編解碼加速器等多媒體相機晶片專案；而晶片團隊對新編解碼器提出的問題，與 BD-rate 無關。他們問的是：每像素多少焦耳、每像素多少位元組的 DRAM 存取流量、多少邏輯閘，以及世界上每個解碼器能否算出完全相同的位元？在這四點上，2026 年的神經解碼器，就像第一批晶片問世前的硬體 HEVC：在軟體上前景看好，在晶片上代價高昂。

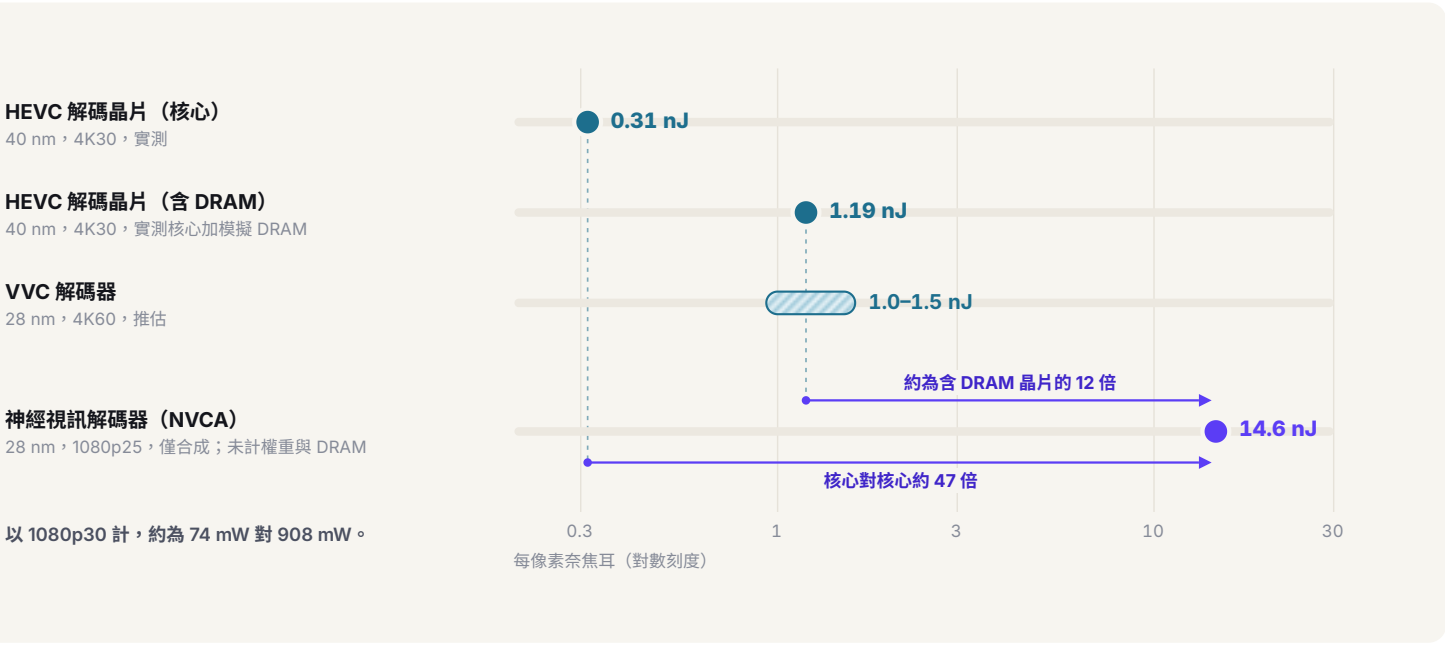
每像素能耗：相差一個數量級

做成晶片的神經解碼器很少，因此只能拿少數幾個已發表的設計來比較。傳統這一側的基準是 MIT 的 HEVC 解碼晶片（Tikekar 等人，ISSCC 2013 與 IEEE JSSC 2014）：40 nm、1.77 mm²，在 200 MHz 下以每秒 30 格解碼 4K。其核心功耗 78 mW，即每像素 0.31 nJ；若計入模擬的 DRAM 功耗，系統每像素需要 1.19 nJ。VVC 的量產數據屬於專有資訊；Electronic Design 在 2026 年的一篇分析，把 MIT 晶片微縮到 28 nm，再套用 VVC 較高的複雜度，推估一個約一百萬邏輯閘的 4K60 解碼器每像素需要 1–1.5 nJ。

神經這一側，依 Electronic Design 的統計，唯一公開的端對端神經視訊解碼器 ASIC 設計是 NVCA（Zhang 等人，DATE 2024），以台積電 28 nm 製程完成合成，共 501 萬個邏輯閘，還不包括放在晶片外的權重。它以每秒 25 格解碼 1080p，功耗 760 mW，即每像素 14.6 nJ。

圖 06 **每解碼一個像素的能耗：固定功能解碼器領先一個數量級**

唯一公開的神經視訊解碼晶片設計，每像素能耗約為實測 HEVC 解碼晶片含 DRAM 時的 12 倍、僅計核心時的 47 倍。神經設計本身的權重與 DRAM 存取流量都還沒計入。



資料來源: Tikekar 等人，IEEE JSSC 2014（MIT）；NVCA，Zhang 等人，DATE 2024（arXiv 2312.10716）；VVC 區間為 Electronic Design 的推估（2026）。HEVC 晶片採用較舊的製程（40 nm 對 28 nm），這對神經設計有利。

這約為 HEVC 晶片含 DRAM 時的 12 倍，僅計核心時的 47 倍。這個比較在三個方面對 NVCA 有利：HEVC 晶片採用較舊的製程；NVCA 的數字沒有計入權重與 DRAM 存取流量；而且 NVCA 是合成結果，不是實測晶片。以 1080p30 計，差距約是 74 mW 對 900 mW 左右：一個是沒人會注意到的解碼器，一個則會反映在手機的電池與散熱預算上。

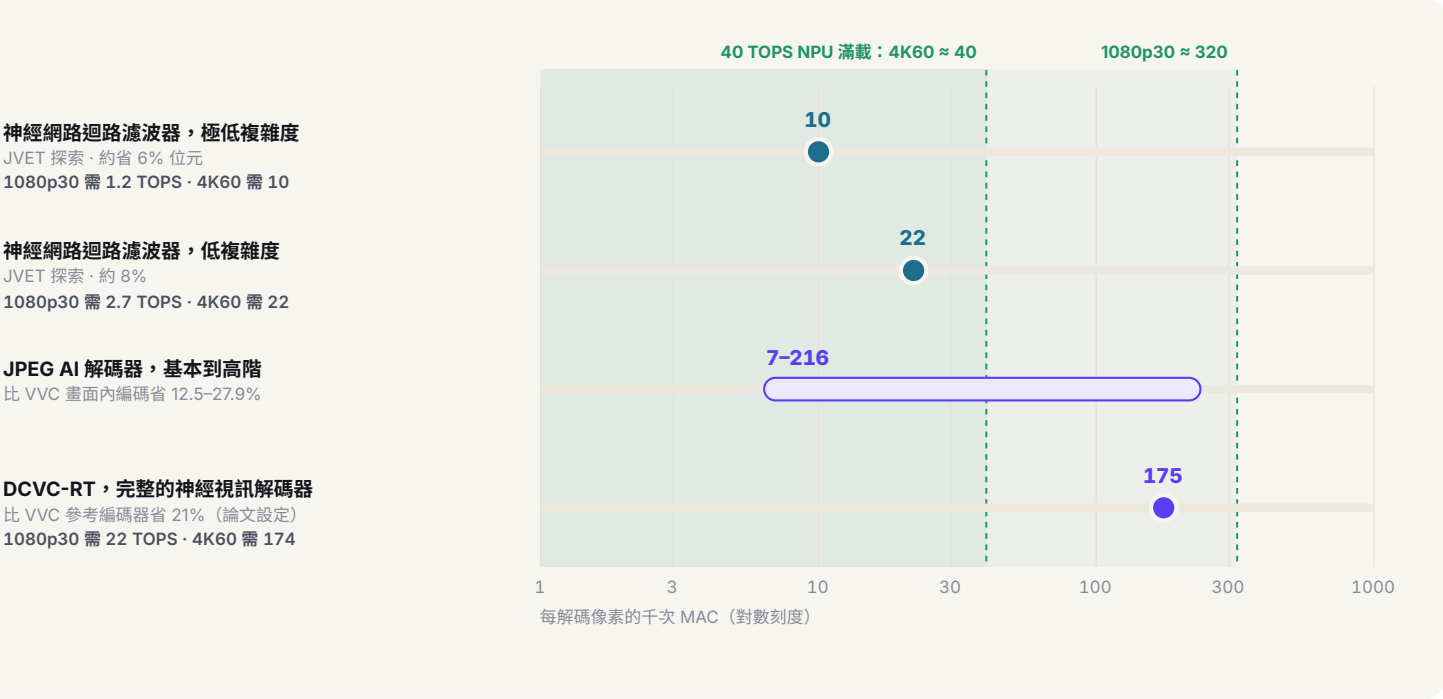
每像素運算量：神經解碼器對 NPU 的要求

我找到最快的已發表神經視訊編解碼器，是微軟研究院的 DCVC-RT（Jia 等人，CVPR 2025）。它在輝達 A100 GPU 上以每秒 125 格編碼、每秒 113 格解碼 1080p，同時在論文的測試設定下比 VVC 參考編碼器節省 21% 位元率，並採用整數運算，讓不同裝置能一致地解碼。它的補充資料列出了每個模組的複雜度。把解碼器必須執行的四個模組（解碼器、特徵擷取器、熵模型與重建）加總，我得到每解碼一個像素約 17.5 萬次乘加運算 估算。

在 1080p30 下，這大約是每秒 11 兆次乘加運算；以每次 MAC 計兩次運算，約為 22 TOPS，而且這還沒算任何利用率損失；4K60 則需要八倍 估算。作為對照，微軟以至少 40 TOPS 的 NPU 來定義 AI PC，蘋果則把 M4 的神經網路引擎標示為 38 TOPS。這些都是峰值，而且要與裝置上其他所有工作共用。

圖 07 **神經工具每像素需數萬次 MAC；完整的神經解碼器需數十萬次**

解碼器每像素的乘加運算次數（對數刻度），對照一顆滿載運作的 40 TOPS NPU 在 4K60 與 1080p30 下能負擔的上限。各項增益的比較基準不同，無法直接比較。



資料來源: JVET（經 SC 29 主席，2026 年 4 月）；JPEG 委員會（2024）；DCVC-RT，CVPR 2025 補充資料，解碼模組由作者加總；微軟 Copilot+ PC 定義。作者的計算：每次 MAC 計為兩次運算，並假設完全利用。

圖 7 的階梯，是解讀這個領域的實用方式。每像素需要數萬次 MAC 的神經工具，例如 JVET 的迴路濾波器（1 萬次約省 6%，2.2 萬次省 8%），有機會放進未來的標準解碼器。JPEG AI 橫跨整個範圍，每像素 7 千到 21.6 萬次 MAC，比 VVC 畫面內編碼省 12.5–27.9%。每像素需要二十萬次上下 MAC 的完整神經視訊解碼器，則適合放在 NPU 上，以中等解析度在封閉迴路中執行。這正是 MLVC 的位置：540p30，最多占用一半的 NPU。

真正的限制往往是記憶體存取流量，而不是 MAC

DCVC-RT 的作者發現，限制速度的不是算術運算，而是「操作成本」：記憶體的輸入輸出，以及函式呼叫的次數。他們之所以移除顯式的運動模組與多尺度處理，主要就是這個原因。NVCA 的設計者打造了特殊的資料流，以減少運動特徵與殘差特徵造成的晶片外記憶體存取流量。晶片設計者會認得這個模式。在 MIT 的 HEVC 晶片中，即使已用快取把 DRAM 頻寬削

減 67%，DRAM 仍占系統能耗約四分之三（每像素 1.19 nJ 中的 0.88 nJ） 分析。在影格之間保留大型多通道特徵圖的神經解碼器，只會讓記憶體問題更糟，而不是更好。

精確性：每個解碼器都必須算出相同的位元

熵編碼只有在解碼器的機率模型與編碼器完全一致時才能運作。在傳統編解碼器中，這個模型是一張表或一小組整數規則。在神經編解碼器中，它是一個神經網路；兩台裝置之間只要有浮點差異，無論來自不同的運算核心、不同的累加順序或不同的 NPU，都可能讓解碼器讀錯符號，並從那一刻起失去整個串流。

凡是讓神經編解碼器上市或標準化的人，都得在工程上解決這個問題。DCVC-RT 把模型轉成整數，以實現一致的跨裝置編碼。MLVC 採用共享尺度的超先驗設計來承受跨平台的不一致，這很可能是它相對 DCVC-RT 犧牲壓縮率的部分原因 分析。JVET 的神經網路探索讓每個網路都透過 SADL 執行，這是一個以 16 位元整數推論的小型函式庫。JPEG AI 在潛在、殘差與像素三個域中定義了一致性檢查點。對晶片團隊來說，這與 HEVC 整數轉換是同一套紀律：標準必須精確定義算術運算，而這會限制網路設計，並排除各廠商自己的捷徑。

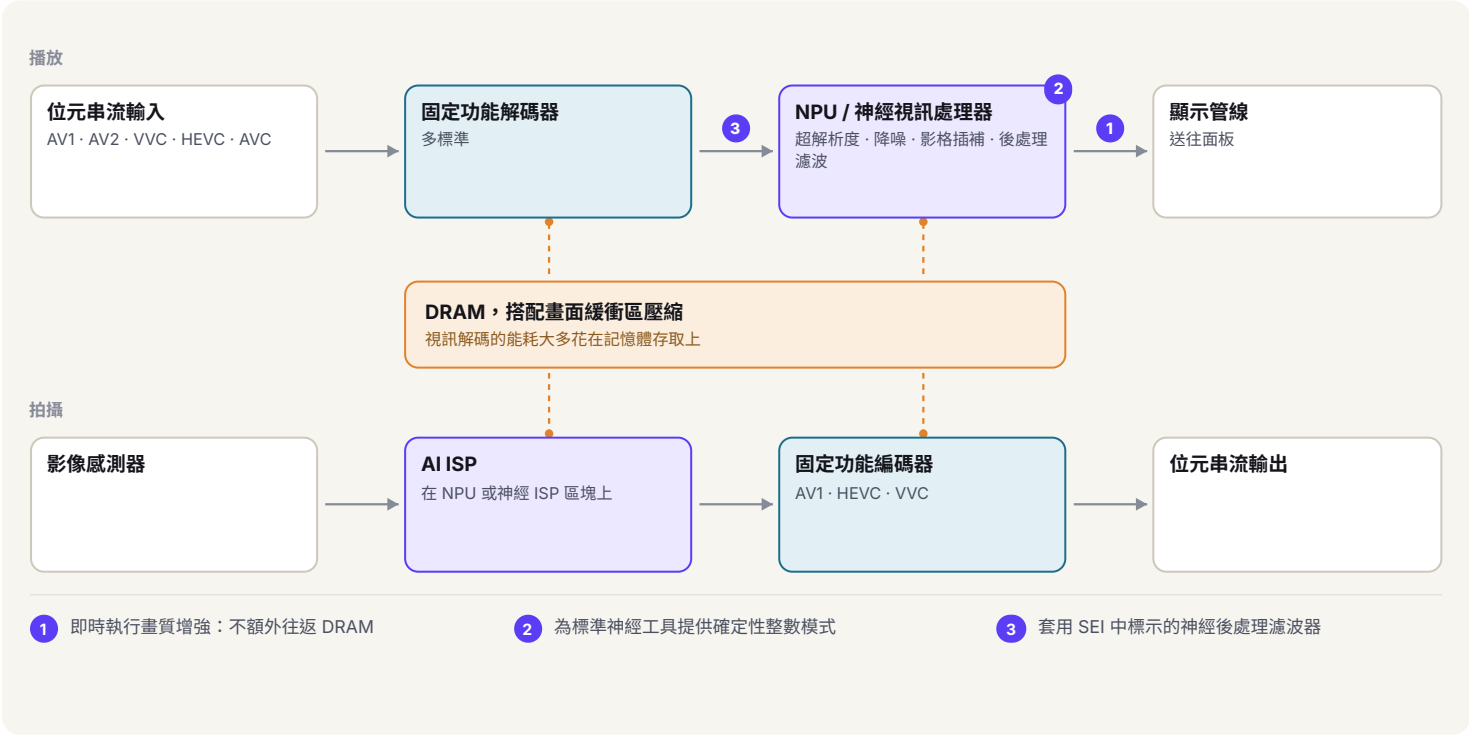
障礙不是面積，而是驗證與使用壽命

NVCA 的 500 萬個邏輯閘，對比推估 VVC 解碼器的約一百萬個，看起來很多；但在已經搭載數 TOPS 級 NPU 的現代手機或電視 SoC 上，幾百萬個邏輯閘並不是決定設計的關鍵。真正花錢的是：依一致性測試套件做驗證、在裝置的整個使用壽命中歷經軟體更新仍維持可程式化解碼器的正確性，以及 4K60 HDR 下的功耗。這些成本只有在標準確定了要做什麼之後，才值得投入。

現在勝出的架構：固定功能編解碼器加上 NPU

圖 08 **已落地的分工：位元串流交給固定功能硬體，像素交給神經網路引擎**

一顆 2026 年的媒體 SoC。編解碼區塊以每像素不到一奈焦耳的能耗解碼各種標準；NPU 負責神經網路擅長、標準又不加限制的工作。真正關鍵的成本是 DRAM 存取流量。



資料來源: 作者的架構觀點，參考騰訊（Canghai V2）、Chips&Media、Allegro DVT 與三星的產品揭露資訊（2026）。

2026 年賣得動的晶片，是把固定功能的多標準編解碼器與旁邊的神經網路引擎搭在一起。編解碼區塊以每像素不到一奈焦耳的能耗處理 AV1、AV2、VVC、HEVC 與 AVC。神經網路引擎負責神經網路擅長、標準又不加限制的工作：拍攝端的降噪與 ISP，顯示端的超解析度、影格插補與畫質增強。兩者共用畫面緩衝區壓縮與 DRAM 排程，因為真正關鍵的成本是記憶體存取流量。

市場印證了這種分工。騰訊自 2019 年起自研的 Canghai（滄海）視訊編解碼 ASIC，第一代晶片已在騰訊雲及其自家直播、隨選視訊與雲端遊戲服務中部署超過 10 萬顆；第二代於 2026 年 5 月宣布已完成點亮、進入量產，預計下半年投入服務，在單一架構中支援 H.265 與 H.266，壓縮效率分別比前一代提升超過 10% 與 30%，吞吐量約為兩倍，並內建超解析度等晶片內畫質增強。Chips&Media 同時銷售編解碼器 IP 與視訊導向的 NPU WAVE-N。Allegro DVT 銷售 AV2 解碼器，並推出了第一款神經視訊處理 IP。三星的電視處理器則把標準解碼器與數百個升頻神經網路搭在一起。

依角色來看，我會做什麼

SoC 與電視晶片團隊。為每個必須支援播放的標準保留固定功能解碼器，現在就要包括 AV2，廣播有需要的地方則包括 VVC。依顯示解析度與影格率下的後處理需求來決定 NPU 規模，並設計從解碼器到 NPU 再到顯示器的路徑，讓畫質增強不會多出一趟 DRAM 往返。支援神經後處理濾波器 SEI。在 NPU 中提供確定性整數模式。把 H.267 當作 2029 年或更晚才啟動的設計，並為它的神經工具保留一些可程式化能力。

IP 廠商。現在就銷售 AV2 與 VVC 解碼器，再加上神經處理在它們周邊所需的元件：後處理濾波加速器、畫面緩衝區壓縮與一致性工具。

影音平台。編碼端的神經前處理最先回本，因為每部作品只執行一次，而且不需要裝置做任何配合。在裝置條件允許時，採用用戶端超解析度。把端對端神經編解碼器留給你掌控的封閉迴路。

第 09 章

09

標準與專利

標準與專利：神經工具在哪裡成為可授權的技術

標準的進程

標準	組織	神經內容	2026 年 11 月狀態
H.267（暫定名稱；VVC 的後繼標準）	JVET（ITU-T 與 ISO/IEC）	未定。傳統工具（ECM）與神經工具（迴路濾波器、畫面內預測、超解析度）都是候選	2026 年 7 月徵求提案；2027 年 1 月評估回應；2027 年下半年推出第一個測試模型；2029 年底推出第一版
VVC 與 HEVC，搭配 VSEI	JVET	可在 SEI 中繼資料中描述神經後處理濾波器；解碼器可選擇是否套用	現行有效；2025 年 10 月的新版擴充了相關訊息
AV2	AOMedia	傳統工具，免權利金	2026 年 5 月 28 日推出 1.0.0 版；2026 年 6 月起授權解碼器 IP；YouTube 網頁串流預計 2027 年初推出
JPEG AI	JPEG（ISO/IEC 與 ITU-T）	端對端學習式影像編解碼器，兼顧人眼與機器視覺	第 1 部分於 2025 年發布；第二版討論中；未發現產品預設支援
IEEE 1857.11	IEEE	為 NPU 設計的神經網路影像編碼	2024 年 12 月發布

標準	組織	神經內容	2026 年 11 月狀態
Video Coding for Machines	MPEG	為機器分析而非人眼觀看而編碼	2026 年初進入國際標準草案階段；預計 2026 年底定案
Opus 機器學習擴充	IETF	可選用的深度冗餘、解碼端增強、頻寬擴展	已納入 libopus 1.5 與 1.6；深度冗餘正在標準化

有兩點值得注意。第一，神經內容先從邊緣進入標準，再進入核心：以可選用的後處理濾波器、可選用的增強功能與影像編碼的形式進入，這些都不會改變主要的解碼路徑。第二，唯一可能把神經工具放進每一塊螢幕解碼路徑的標準 H.267，距離第一版還有三年。

誰在專利上占好位置

標準中的工具會變成可授權的專利，因此 H.267 的制定過程也是一場專利競賽，神經工具也在其中。GreyB 在 2025 年針對 3,916 件 VVC 後繼標準專利所做的專利地圖分析，把字節跳動排在首位：645 個專利家族、251 件 JVET 提案；該分析也提到產學合作，例如小米與西安電子科技大學擁有超過 20 個專利家族。MLex 在 2026 年 3 月報導，騰訊、阿里巴巴與字節跳動正成為視訊編解碼專利池的積極參與者。2026 年 5 月，Sauvegarder Investment Management 向阿里巴巴買下 126 項視訊編解碼專利資產，涵蓋 AV1、AV2 與 Video Coding for Machines。InterDigital 在 2025 年 10 月收購 Deep Render，把 Deep Render 的 AI 視訊編碼專利納入其專利組合。

每台裝置的收費，從專利池的費率表就看得很清楚。Access Advance 的 VVC 費率（2026 年 7 月 1 日起，適用於合規且使用其商標的被授權方）為：在美國、歐盟與日本等 Region 1 市場，每支手機、平板或筆電 0.50 美元，Region 2 則為 0.25 美元；4K 電視在 Region 1 依售價為 1.00 至 1.50 美元；每家企業每年上限 7,500 萬美元。AOMedia 成員對 AV1 與 AV2 不收權利金。無論 H.267 包含什麼，它的工具都會以同樣的方式定價，這也是擁有神經編碼專利的公司現在就向 JVET 提交提案的原因 分析。

第 10 章

10

誰在賺錢

誰在賺錢，企業可以在哪裡切入

2026 年的錢

圖 09

2026 年，誰從神經編碼賺到錢

錢跟著標準與平台走。我沒有找到任何一家靠授權端對端神經視訊編解碼器維生的獨立公司。

已有收入

平台

留住省下的錢

Netflix 神經網路降取樣器；YouTube 與 Meta 超解析度；微軟以 MIT 授權開源 MLVC

已有收入

裝置與晶片廠商

銷售差異化

相機 ISP；電視升頻處理器；騰訊自研的 Canghai 編解碼晶片

已有收入

IP 廠商

每個新標準都授權一次

規格定案三週內，Chips&Media、Allegro DVT 與 VeriSilicon 都有了 AV2 解碼器客戶

已有收入

GPU 廠商

建立平台綁定

DLSS 4 支援 250 多款遊戲與應用程式；神經紋理 SDK 仍為測試版

按裝置收費

專利權人與專利池

按裝置收費

VVC：每支手機 0.50 美元，每台 4K 電視 1.00–1.50 美元；H.267 工具在 2029 年後

被併購出場

神經編解碼新創公司

被收購

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

資料來源: 報告引用的公司與專利池揭露資訊；狀態截至 2026 年 11 月 12 日。專利池費率：Access Advance，VVC，Region 1，合規並使用商標，2026 年 7 月 1 日起。

平台把省下的成本留給自己。Netflix、YouTube、Meta、微軟與騰訊都自行開發神經處理，並把效益留在自己手上：更低的頻寬、更少的儲存空間，以及每個位元更好的品質。若有助於推廣，它們會把程式碼免費送出：微軟就以 MIT 授權釋出了 MLVC。

裝置與晶片廠商銷售差異化。相機 ISP 與電視畫質處理器是神經網路接觸每一位使用者的地方，並以功能賣點的形式銷售：高通的 Cognitive ISP、三星的升頻處理器。

IP 廠商為每個新標準銷售解碼器，外加 NPU。AV2 規格定案後三週內，Chips&Media、Allegro DVT 與 VeriSilicon 都宣布了 AV2 解碼器客戶。這些廠商如今也在編解碼器旁一併銷售神經處理區塊。

GPU 廠商打造平台。DLSS 與神經紋理壓縮，讓輝達的張量硬體成為遊戲製作方式的一部分。

專利權人按裝置收費，透過 VVC 與 HEVC 的專利池收取，2029 年後還有 H.267。

神經編解碼新創公司都已被收購。WaveOne 在 2023 年歸了蘋果，iSIZE 在 2023 年歸了 Sony Interactive Entertainment，Deep Render 在 2025 年歸了 InterDigital，交易條件均未公開。每一次都是平台或授權方買下團隊與專利。我沒有找到任何一家在 2026 年靠授權端對端神經視訊編解碼器維生的獨立公司。

企業可以在哪裡切入

切入點	買方	為何是現在	主要風險
編碼端的神經前處理（降取樣、降噪、逐片調校）	串流與使用者影音平台、廣播業者	裝置不需改動；已在 Netflix 驗證	大平台會自己做

切入點	買方	為何是現在	主要風險
顯示端畫質增強 IP 與調校（超解析度、影格插補、畫質）	電視、機上盒與手機晶片廠商	裝置內已有 NPU；差異化肉眼可見	商品化；晶片廠商自行開發
封閉迴路神經編解碼器（視訊會議、遠端操作、專用鏈路）	企業通訊、工業與遠端臨場廠商	MLVC 證明可行，且程式碼開源	必須掌控兩端；NPU 碎片化
編解碼器加 NPU 的晶片或 IP（AV2、VVC 與後處理）	SoC 廠商	AV2 週期正要開始	既有大廠實力強
評測、一致性與確定性推論工具	所有出貨神經工具的業者	逐位元一致是硬性要求	除非搭售，否則市場小
向 H.267 提交標準提案	授權方、大平台	2027 年 1 月評估提案回應	週期長、成本高、結果不確定

對大多數企業來說，前兩列是 2027 年就能取得營收的地方。第三列對專注型的產品公司最有吸引力。最後一列則需要耐心與一套授權策略。

第 11 章

11 說法與證據 你會聽到的說法，以及證據怎麼說

說法	證據怎麼說
「神經編解碼器已經勝過 VVC。」	在實驗室裡，是的。在論文設定下，DCVC-RT 在 1080p 比 VVC 參考編碼器省 21%；JPEG AI 比 VVC 畫面內編碼省 12.5–27.9%。其解碼器每像素約需 17.5 萬次 MAC，也就是要付出一顆 NPU 的能耗，而固定功能硬體每像素不到一奈焦耳。
「AI 編解碼器幾年內就會取代 H.26x 與 AV1。」	播放領域不會。第一個能納入神經工具的標準目標在 2029 年底，晶片約在 2031 年跟進，2030 年代中期才成為主流 估算。現在到來的編解碼器 AV2 是傳統設計。
「MLVC 讓頻寬減少 88%。」	那是在 360p 下相對 H.264，而且用的是微軟自己的測試集。相對硬體 HEVC，360p 為 75.5%，540p 為 65.4%，而且是在微軟掌控兩端的通話中。
「神經紋理讓顯示記憶體減少 85–96%。」	那是在展示場景中。已上市的遊戲有效壓縮率約提升 30%，而且因為執行成本，只在部分素材上使用神經紋理。
「JPEG AI 是 ISO 標準，所以相機和瀏覽器都會採用。」	我沒有在產品中找到預設支援。2026 年網頁採用的格式，也就是 Chrome 155 中的 JPEG XL，是傳統格式。
「神經解碼對手機來說太昂貴。」	部分正確。MLVC 只用半顆 NPU 就能解碼 540p30；以 DCVC-RT 等級的模型解碼 4K60，則需要約 175 TOPS 的持續算力 估算。
「衛星與無人機鏈路是神經視訊的大市場。」	技術上最契合，但市場小，而且太平洋兩岸的出口管制都愈來愈嚴。
「大平台會向新創公司授權神經編解碼器。」	它們不是自己開發（Netflix、微軟、騰訊），就是直接買下新創公司（蘋果、Sony、InterDigital）。

有個簡單的檢驗可以看穿大多數的推銷說詞：問清楚比較基準、指標、解析度與解碼成本。交代不出這四項的說法，是研究成果，不是產品。

第 12 章

12

觀察指標

2027 至 2030 年的觀察重點

時間	觀察指標	為何重要
2027 年 1 月	JVET 在布里斯本評估 H.267 提案回應	有多少頂尖回應仰賴神經工具，每像素成本多少
2027 年初	YouTube 首批 AV2 網頁串流；已發表的 SoC 內建 AV2 解碼器	下一個傳統編解碼週期，將決定到 2030 年左右的裝置預算
2027 年下半年	H.267 的第一個測試模型	神經工具是進入核心設計，還是只停留在探索階段
2027 年	MLVC 走出 Teams 點對點通話；其他廠商採用其開源程式碼	開源神經編解碼器首度在原開發者之外接受考驗
2027–2028 年	第一個預設寫入或讀取 JPEG AI 的產品	第一個跨越公司邊界使用的神經格式
2027–2028 年	正式版驅動程式支援 DirectX Linear Algebra；更多遊戲與行動裝置採用神經紋理	繪圖中的跨廠商神經解碼
持續觀察	手機與 PC 的 NPU 預算	若持續算力達到約 100 TOPS，DCVC-RT 等級的 1080p30 解碼器只會占用約五分之一的 NPU，而不是一半 分析
2029 年底	H.267 第一版	神經工具進入主流播放的晶片週期就此展開

第 13 章

13 未解問題

本報告中有幾個問題還沒有公開的答案，也是我在技術上最感興趣的。

大規模下的精確性。在放棄大部分壓縮優勢之前，神經解碼器能在不同廠商的 NPU 上多接近逐位元一致、純整數的運算？MLVC 用六個百分點換來的取捨是一個資料點；標準需要的是通用的答案。

記憶體先於 MAC。神經後處理在硬體解碼器與顯示器之間即時執行時，每像素要耗用多少 DRAM 存取流量？它的工作集能縮到多小？

H.267 中神經工具的代價。在 4K60 解碼器中，神經網路迴路濾波器的每像素 MAC 數與奈焦耳數到多少，就不再值得它所節省的位元？

通話以外的封閉迴路。從雲端照片儲存、遊戲串流，到機器人與私有節目回傳鏈路，還有哪些封閉系統能在 H.267 到來之前採用端對端神經編解碼器？

會生成內容的解碼器。當解碼器像影片 VAE 已經在做的那樣，生成細節而非重建細節時，保真度該如何衡量與規範？

未來幾年，這些都是我們可以透過技術交流與技術合作一起探索的領域：i@tanshuai.com。

本報告的侷限

這是一份根據截至 2026 年 11 月 12 日公開文件所做的文獻研究。本版未進行一手訪談，也沒有獨立量測。數個關鍵數字是廠商或論文作者在自家測試集上提出的說法，例如 MLVC 的位元率節省與神經紋理壓縮率；這些都已如實標示。我對 DCVC-RT 的運算量估算，是加總其作者公布的各模組數據，並忽略利用率損失，因此實際的 NPU 負載會更高。能耗比較依據的是少數幾個採用不同製程節點的已發表設計；VVC 與 AV2 解碼器的量產數據屬於專有資訊。標準時程是各委員會自己的目標。H.267 在 2029 年之後的時間點，是我根據 AV1 與 VVC 的歷程所做的估計。

關於作者

譚帥是技術專家與創辦人，在晶片、硬體與系統軟體領域工作超過二十年，擁有 20 項獲准專利。他有多年無晶圓廠晶片研發與管理經驗，包括全客製化 ASIC 資料路徑設計，以及視訊編解碼加速器等多媒體相機晶片專案，這也是本報告的觀察視角：每像素能耗、記憶體存取流量，以及標準允許解碼器在晶片上固化什麼。他的工作足跡遍及矽谷、紐約、芝加哥、台灣、大灣區（深圳與香港）、越南與馬來西亞。如今他是一名 AI 研發者。

技術交流：i@tanshuai.com · tanshuai.com

說明與免責聲明

本報告是根據截至 2026 年 11 月 12 日可取得之公開文件所做的獨立文獻研究，僅供參考，不構成投資、法律或工程建議。產品、標準時程、專利池費率與出口規定都會改變；凡是你要依賴的資訊，請對照原始出處查證。標為 廠商說法 、 估算 或 分析 的數字，應依其標示解讀。公司與產品名稱屬於其所有者，僅用於識別；沒有任何公司審閱或贊助本報告。

© 2026 譚帥。歡迎在註明出處的前提下，原封不動地分享本報告。

資料來源

均於 2026 年 11 月 3 日至 11 月 12 日間查閱，依章節分組。

流量、裝置與範圍

- Ericsson，*Mobility Report*，2026 年 6 月（2025 年底影片約占行動數據流量的 75%）：https://www.ericsson.com/en/reports-and-papers/mobility-report/reports/june-2026 ；摘要：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，全球智慧型手機出貨量，2025 年第 4 季與全年（12.6 億支），2026 年 1 月 13 日：https://www.idc.com/resource-center/press-releases/q425mobilephonetop5/
- TrendForce，2025 年電視出貨量預測（1 億 9,571 萬台）：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，2025 年電視出貨預測（2 億 870 萬台）：https://iconnect007.com/article/145954/omdia-forecasts-flat-tv-shipments-in-2025-but-mini-led-poised-for-explosive-growth-at-829/145951/ein

播放與串流

- Netflix Technology Blog，「AV1 — Now Powering 30% of Netflix Streaming」，2025 年 12 月：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」，2022 年 11 月：https://netflixtechblog.com/for-your-eyes-only-improving-netflix-video-quality-with-neural-networks-5b8d032da09c
- AOMedia，AV2 Bitstream and Decoding Process Specification v1.0.0，2026 年 5 月 28 日：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」（AV1 在 YouTube 內容庫中的占比）：https://www.tvtechnology.com/platform/streaming/youtube-to-commence-av2-streaming-in-2027
- Google Research，「The in-loop filtering pipeline in AV2」，2026 年：https://research.google/pubs/the-in-loop-filtering-pipeline-in-av2/
- Fraunhofer IIS，「The Future of Broadcast in Brazil」（DTV+ 於 2026 年 6 月正式商轉）：https://mpegh.com/brazil/ ；CartaCapital 談電視內建接收器：https://www.cartacapital.com.br/?p=628337 ；Olhar Digital 談 TV 3.0 中的 VVC：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」，2025 年 10 月 29 日：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，2025 年 10 月 16 日：https://atscaleconference.com/how-meta-deployed-super-resolution-at-scale-to-transform-video-quality/
- Samsung，8K 電視特色介紹（NQ8 AI Gen3，768 個神經網路）：https://www.samsung.com/us/tvs/8k-tv/highlights/

攝影與影像

- 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 與 Visionary.ai，完全以 AI 為基礎的 ISP，2026 年 1 月 5 日：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/ ；PetaPixel 專訪：https://petapixel.com/2026/01/18/a-look-at-the-worlds-first-full-ai-based-image-signal-processor
- JPEG committee，第 103 次會議新聞稿（JPEG AI 的增益、複雜度與智慧型手機實作）：https://jpeg.org/items/20240507_press.html
- G. J. Sullivan，ISO/IEC JTC 1/SC 29 簡介，2026 年 4 月（JPEG AI 現況；ECM 與 NNVC 數據）：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/
- IEEE 1857.11 工作小組（基於神經網路的影像編碼）：https://sagroups.ieee.org/fvc
- Chrome 155 與 JPEG XL，2026 年 10 月 6 日：https://xenospectrum.com/en/chrome-155-jpeg-xl-rust/

通話與會議

- Microsoft，「Satin: Microsoft's latest AI-powered audio codec for real-time communications」，2021 年 2 月 17 日：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)」，2026 年 7 月 24 日：https://techcommunity.microsoft.com/blog/linuxandopensourceblog/announcing-the-open-source-release-of-ml-video-codec-mlvc/4539875 ；程式碼：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 ；iReplay 分析：https://ireplay.tv/blog/mlvc-microsoft-neural-video-codec-open-source-streaming/
- Opus 1.5 發行說明：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」，2024 年 6 月 13 日：https://engineering.fb.com/2024/06/13/web/mlow-metas-low-bitrate-audio-codec/
- Microsoft，Copilot+ PC（40+ TOPS NPU）：https://www.microsoft.com/en-us/windows/copilot-plus-pcs ；Apple，M4 發表（38 TOPS 神經網路引擎）：https://www.apple.com/newsroom/2024/05/apple-introduces-m4-chip/

生成式 AI

- Kyutai，Mimi 模型卡（12.5 Hz，1.1 kbps）：https://huggingface.co/kyutai/mimi
- Lightricks，「LTX-Video: Realtime Video Latent Diffusion」（1:192 影片 VAE）：https://arxiv.org/abs/2501.00103
- MPEG 第 153 次會議新聞稿（Video Coding for Machines 進入 DIS 階段）：https://www.mpeg.org/wp-content/uploads/2026/03/press_release_153.pdf

遊戲與繪圖

- NVIDIA，DLSS 4 採用情況：推出時與 Computex 2025（75 與 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」，2026 年 2 月：https://www.ubisoft.com/en-us/studio/laforge/news/415slmB3ZGzv8d2LhO9QT/shipping-neural-texture-compression-in-assassins-creed-mirage
- TweakTown 談輝達在 GTC 2026 的神經紋理展示（6.5 GB 降至 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 正式版，2026 年 2 月 26 日：https://devblogs.microsoft.com/directx/shader-model-6-9-retail-and-more/ ；關於 cooperative vector：https://devblogs.microsoft.com/directx/shader-model-6-9-and-the-future-of-cooperative-vector/ ；GDC 2026 規劃：https://windowsreport.com/microsoft-showcases-neural-rendering-and-dxr-2-0-plans-for-directx-at-gdc-2026/

專用鏈路

- Global Trade Alert 談中國 2026 年第 49 號公告：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（無人地面平台的門檻指標）：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 將外國製無人機及元件列入「涵蓋清單」，2025 年 12 月 22 日：https://dronelife.com/2025/12/22/fcc-adds-foreign-made-drones-and-components-to-covered-list-citing-national-security-risks/

晶片

- M. Tikekar 等人，「A 249-Mpixel/s HEVC Video-Decoder Chip for 4K Ultra-HD Applications」，IEEE JSSC 2014：https://dspace.mit.edu/handle/1721.1/93876 ；MIT 晶片頁面（4K30 下 78 mW）：https://eecsg.mit.edu/chips/hevc-video-decoder-for-4k-ultra-hd-applications/
- S. Zhang 等人，NVCA，DATE 2024：https://arxiv.org/abs/2312.10716
- Electronic Design，「Why Traditional VVC Will Outlive the Neural-Codec Hype on Silicon」，2026 年：https://www.electronicdesign.com/home/article/55414955/why-traditional-vvc-will-outlive-the-neural-codec-hype-on-silicon
- Z. Jia 等人，「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 ；補充資料（各模組複雜度）：https://openaccess.thecvf.com/content/CVPR2025/supplemental/Jia_Towards_Practical_Real-Time_CVPR_2025_supplemental.pdf
- NNVC 與 SADL 整數推論函式庫：https://arxiv.org/abs/2309.05846
- 騰訊 Canghai V2，2026 年 5 月 27 日（Shanghai Securities News，經 10jqka 轉載）：https://stock.10jqka.com.cn/20260527/c677016020.shtml ；功能摘要：https://www.besthub.dev/articles/tencent-s-canghai-v2-chip-enters-mass-production-what-the-new-capabilities-mean-for-video-encoding-bcfc18dbfa29
- Chips&Media AV2 解碼器 IP，2026 年 5 月 12 日：https://www.design-reuse.com/news/202530501-chips-media-completes-development-of-next-gen-av2-hw-decoder-ip/ ；AV2 授權，2026 年 6 月 17 日：https://www.edge-ai-vision.com/?p=58270
- Allegro DVT，AV2 解碼器 IP，2026 年 6 月 16 日：https://www.edge-ai-vision.com/2026/06/allegro-dvts-pulsar-decoder-ip-adds-support-for-av2-video-codec/
- VeriSilicon，支援 AV2 的 VC9800D，2026 年 6 月 10 日：https://www.businesswirechina.com/en/news/63176.html

標準與專利

- MPEG 第 155 次會議新聞稿（VVC 後繼標準的聯合徵求提案），2026 年 8 月 5 日：https://www.mpeg.org/wp-content/uploads/2026/08/ag3w254SN26474Press-Release-MPEG155.pdf
- MPEG 第 152 次會議新聞稿（徵求證據的結果；VVC 與 VSEI 新版），2025 年 10 月 30 日：https://www.mpeg.org/wp-content/uploads/2025/11/Press-Release-MPEG152.pdf ；ACM SIGMM MPEG 專欄：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，VVC 後繼標準（H.267）專利地圖：https://greyb.com/blog/exclusive-video-codec-landscape/ ；談 AV2 授權與阿里巴巴專利組合出售：https://greyb.com/blog/av2-licensing/
- MLex，「Fragmented codec licensing exposes need for clearer rules for Chinese tech majors」，2026 年 3 月 27 日：https://www.mlex.com/mlex/articles/2458402/fragmented-codec-licensing-exposes-need-for-clearer-rules-for-chinese-tech-majors
- Access Advance，MCBA 與 VVC Advance 權利金費率（2026 年 7 月 1 日起）：https://accessadvance.com/wp-content/uploads/2025/07/MCBA-Program-Overview-September-2026.pdf

資金與公司

- InterDigital 收購 Deep Render，2025 年 10 月 30 日：https://www.advanced-television.com/2025/10/30/interdigital-acquires-deep-render
- TechCrunch，「Apple acquired a startup using AI to compress videos」，2023 年 3 月 27 日：https://techcrunch.com/2023/03/27/apple-acquired-a-startup-using-ai-to-compress-videos
- Neowin，「Sony acquires AI video startup iSIZE」，2023 年 11 月：https://www.neowin.net/news/sony-acquires-ai-video-startup-isize-to-enhance-playstation-streaming/

學習式壓縮 · 播放 × 晶片

2026 年的神經編碼

在編解碼器周邊落地，而非取而代之

本報告的每個關鍵數字都附有來源與日期。作者自行推導的數字標為「估算」；只有賣方公布的數字標為「廠商說法」。

譚帥

技術專家與創辦人 · AI 研發者

tanshuai.com · i@tanshuai.com

公開版 · 資料截至 2026 年 11 月 12 日