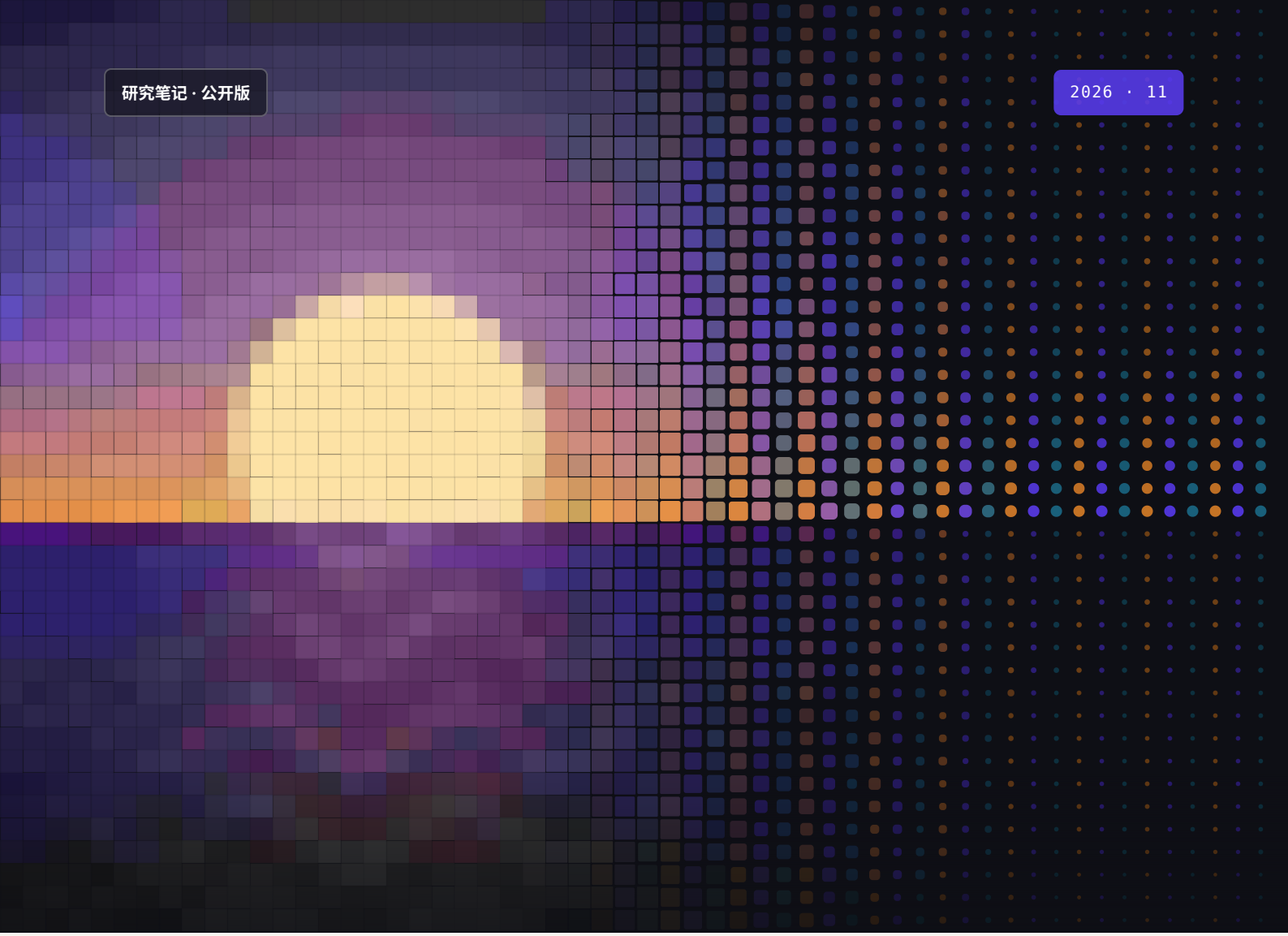


研究笔记 · 公开版

2026 · 11

学习型压缩 · 播放 × 芯片

2026 年的神经编码

围绕编解码器落地，而非取而代之

播放 · 拍照 · 通话 · 生成式 AI · 游戏 · 芯片

> **结论**
>
> **神经网络已经在承载全世界视频的编解码器两端工作。端到端神经编解码器只在闭环中上线，主流屏幕还要等 H.267。**

谭帅

技术专家和创业者 · AI 研发者

tanshuai.com · i@tanshuai.com

数据截至 2026 年 11 月 12 日

金额以美元为准，括注的人民币按 1 美元兑 6.72 元换算

目录

目录



◆ **摘要** 04
已是主流，但在编解码器外围

01 **范围与方法** 07
本文所说的“神经编码”，以及我是怎么研究的

02 **播放与流媒体** 08
播放与流媒体：两端是神经网络，中间是标准

03 **拍照与图像** 11
拍照与图像：相机里是神经网络，网页上是传统格式

04 **通话与会议** 12
通话与会议：端到端神经编解码器真正落地的地方

05 **生成式 AI** 14
生成式 AI：规模最大、却没人称之为编解码器的部署

06 **游戏与图形** 15
游戏与图形：神经重建已是主流，神经资源刚刚起步

07 **专业链路** 16
专业链路：卫星、无人机和机器人

08 **芯片** 17
芯片视角：能耗、算力、访存与精确性

09 **标准与专利** 20
标准与专利：神经工具在哪里变得可授权

10 **钱流向了谁** 21
钱流向了谁，公司可以在哪里切入

11 **说法与证据** 23
你会听到的说法，以及证据怎么说

12 **关键节点** 24
2027 至 2030 年值得关注什么

13 **开放问题** 25

◆ **关于作者** 26
谭帅

◆ **资料来源** 27

图表

图表



图 01	2026 年神经编码落地在哪里	06
图 02	神经网络进入编解码的三条路径	07
图 03	从标准定稿到主流屏幕，要花掉大半个十年	09
图 04	在通话中，神经编解码器比硬件 HEVC 少用 65–76% 码率	12
图 05	一家公司掌控整套技术栈时，神经重建普及得很快	15
图 06	每解码一个像素的能耗：固定功能解码器领先一个数量级	17
图 07	标准内的神经工具每像素数万次乘加，完整的神经解码器每像素数十万次	18
图 08	已经落地的分工：码流交给固定功能硬件，像素交给神经引擎	19
图 09	2026 年，神经编码的钱流向了谁	22

正文中的标注

估算	作者根据引用数据所做的计算或推断
厂商说法	由销售该产品的公司发布，未经独立验证
分析	作者对证据的解读

数据截至 2026 年 11 月 12 日 · 金额以美元为准，括注的人民币按 1 美元兑 6.72 元换算

摘要

已是主流，但在编解码器外围

结论

神经编码在 2026 年已经是主流，但它围绕标准编解码器工作，而不是取代它们。

≈75%
移动数据流量中视频所占的比例，2025 年底（爱立信）

12.6 亿
2025 年智能手机出货量，每一部都用固定功能硬件解码视频

≈12 倍
唯一公开的神经解码芯片设计的每像素能耗，对比实测 HEVC 解码芯片（含 DRAM）

≈22 TOPS
DCVC-RT 级神经解码器在 1080p30 下所需的持续算力 估算

75.5%
MLVC 在 360p 通话中比硬件 HEVC 节省的码率，基于微软自己的测试集 厂商说法

2029 年底
H.267 第一版，第一个可能把神经工具带到每一块屏幕上的标准

落地在哪里

如今，神经网络运行在承载全世界几乎所有视频的编解码器两端。编码之前，它们塑造画面：手机摄像头里的 AI 图像信号处理器，Netflix 编码流程里的神经下采样器。解码之后，它们修复画面：YouTube 电视端的超分、Meta 的超分，以及宣称内置数百个神经网络的电视画质芯片。中间的码流仍是 AV1、HEVC、AVC 或 VVC，由固定功能硬件解码。没有哪家流媒体服务向电视或手机发送神经视频格式；今天凡是在终端上运行的神经视频解码器，都运行在某一家公司自己的应用里。

闭环

端到端神经编解码器的编码器和解码器都由学习得到，它们只在闭环中上线：一家公司掌控两端，并且可以退回标准编解码器。微软自 2021 年起在所有 Teams 和 Skype 双人通话中使用一款 AI 音频编解码器；2026 年 7 月，它又开源了神经视频编解码器 MLVC，目前正在 Teams 点对点通话中推广：在 360p 下达到相同的主观评分，所需码率比硬件 HEVC 少约 75%。每个语音和视频生成模型都运行在学习得到的编解码器上，从 12.5 Hz 的语音 tokenizer，到压缩比 1:192 的视频 VAE。神经纹理压缩已在一款 3A 游戏中上线。这些都是真实的部署，并且有一个共同条件：没有别人需要解码这个码流。

时间表

主流播放的神经码流，要等一个标准和一轮芯片周期。 第一个可能带神经工具的标准是 VVC 的继任者，我称之为 H.267：它于 2026 年 7 月发布征集提案，2027 年 1 月评估提案，目标是在 2029 年底推出第一版。AV1 从标准定稿到占 Netflix 观看量的 30%，用了将近八年。照此规律，首批 H.267 芯片约在 2031 年出现，2030 年代中期才会普及 估算。与此同时，下一个走上屏幕的编解码器是 AV2：它于 2026 年 5 月定稿，其中没有神经解码器；三周之内，三家 IP 供应商就宣布了 AV2 解码器客户。

芯片

原因在芯片上。 唯一公开的神经视频解码芯片设计，每像素需要 14.6 纳焦，约为实测 HEVC 解码芯片含 DRAM 时的 12 倍，约为其芯片核心本身的 47 倍。微软的实时研究型编解码器 DCVC-RT，按我对它自己分模块数据表的加总，每解码一个像素约需 17.5 万次乘加，1080p30 约需 22 TOPS 的持续算力 估算：约为定义 AI PC 的 40 TOPS 的一半，只花在一路视频上。神经解码器还必须在每台设备上算出

逐比特相同的结果，否则熵解码就会崩溃。这就是为什么 MLVC 宁可牺牲压缩率，也要在不同 NPU 上解码出完全相同的结果，而且只跑 540p，不跑 4K。

商业

对企业来说，钱跟着标准和平台走，而不是跟着独立的神经编解码器走。 平台留住神经前处理和后处理节省下来的钱。IP 供应商为每个新标准授权固定功能解码器，外加用于成像的 NPU。专利池按台收费：在美国等一类地区，VVC 每部手机 0.50 美元（约 3.4 元），每台 4K 电视 1.00–1.50 美元（约 6.7–10.1 元）。最知名的三家神经编解码创业公司 WaveOne、iSIZE 和 Deep Render，分别被苹果、索尼和 InterDigital 收购。尚待把握的机会在于：神经前处理和后处理，闭环产品，把固定功能编解码器与 NPU 配在一起的芯片，以及介于其间的评测和标准工作。

问题	判断	关键数字（来源见各章）
① **神经编码今天在哪里大规模落地?**	在编解码器外围：摄像头 ISP、编码端下采样、播放端超分。端到端只在闭环里：通话、游戏资源、生成式 AI 的 tokenizer	视频约占移动数据流量的 75%；2025 年出货 12.6 亿部手机和约 2 亿台电视，全部用固定功能硬件解码；DLSS 4 覆盖 250 多款游戏和应用
② **神经编解码器会在播放中取代 AV1、HEVC 和 VVC 吗？何时?**	要先有一个标准和一轮设备周期。H.267 是第一个现实的载体，而且很可能混合传统工具和神经工具	H.267 于 2026 年 7 月征集提案，2027 年 1 月评估，2029 年底推出第一版；2030 年代中期普及 估算
③ **芯片上的代价多大?**	每像素能耗比今天的固定功能解码高一个数量级，还有访存流量和逐比特一致的要求	每像素 14.6 纳焦（神经）对 1.19 纳焦（HEVC，含 DRAM）；每解码一个像素约 17.5 万次乘加，相当于 1080p30 下约 22 TOPS 估算
④ **标准和专利允许什么?**	VVC 码流已能标示神经后处理滤波；JPEG AI 已是 ISO 标准；H.267 将决定哪些神经工具变得可授权	ECM（传统）比 VVC 省约 28%；神经环路滤波省 6–14%，每像素 1 万至 2.2 万次以上乘加；字节跳动在超越 VVC 的工作中拥有 645 个专利族
⑤ **钱流向了谁，公司可以在哪里切入?**	平台、IP 供应商、专利权人。纯神经编解码创业公司都被收购了。切入点：前处理和后处理、闭环、编解码器加 NPU 的芯片、评测	VVC 专利池每部手机 0.50 美元（约 3.4 元）；AV2 定稿几周内解码器 IP 就已授权；三起收购，条款均未披露
⑥ **哪些说法要打折扣?**	相对老编解码器、在低分辨率下、不计解码代价，或来自演示而非已上市产品的增益	MLVC：比 H.264 少 88%，但只比硬件 HEVC 少 65–76%；纹理：演示中省 85% 显存，已上市游戏中约 30%

图 01

2026 年神经编码落地在哪里

神经网络已经运行在主流编解码器的两端；端到端神经编解码器只在一家公司同时掌控编码端和解码端的地方真正上线。

	标准编解码器外围 预处理、增强、超分	端到端神经编解码器 编码器和解码器都由学习得到	标准内部 标准码流或接口中的神经工具
01 播放与流媒体	● 大规模部署 Netflix 神经下采样；YouTube、Meta 超分；电视 AI 画质芯片	● 未部署 没有流媒体服务向电视或手机发送神经视频格式；终端都用硬件解码标准编解码器	● 标准化中 H.267 第一版定于 2029 年底，神经工具在候选之列
02 拍照与图像	● 大规模部署 旗舰手机的 AI ISP；CES 2026 发布全神经网络 ISP 的 IP	● 演示与试点 JPEG AI 已是 ISO 标准：有手机演示，没有产品	● 未部署 正在普及的是传统格式：Chrome 155 默认支持 JPEG XL
03 通话与会议	● 大规模部署 接收端超分、神经降噪	● 大规模部署 Teams 两方通话自 2021 年用神经音频；MLVC 视频正在 Teams 中推出	● 有限部署 Opus 1.5 加入可选、向后兼容的机器学习工具
04 生成式 AI	—	● 大规模部署 每个视频和语音模型都运行一个学习得到的 tokenizer，例如压缩比 1:192 的视频 VAE	—
05 游戏与图形	● 大规模部署 神经超分与插帧：仅 DLSS 4 就覆盖 250 多款游戏和应用	● 有限部署 一款 3A 游戏已上线神经纹理；英伟达 SDK 仍是 beta	● 标准化中 DirectX 正在为着色器加入矩阵运算（预览）
06 专业链路：卫星、机器人	● 演示与试点 在服务器端修复低码率画面	● 演示与试点 抗丢包神经编解码器处于研究与试点	—

来源: 报告引用的公司、标准与开发者文件，状态截至 2026 年 11 月 12 日。各行按所覆盖的人群规模排序。

第 01 章

01 范围与方法 本文所说的“神经编码”，以及我是怎么研究的

神经网络进入编解码的三条路径

“神经编码”其实是三件不同的事，大多数争论都源于把它们混为一谈。

01 **编解码器外围。** 神经网络在编码前或解码后处理像素，码流保持标准。例子：相机里的 AI 图像信号处理器（ISP），Netflix 编码器之前的神经下采样器，电视上的超分。任何设备都能播放这个码流；神经网络只改变输入的内容或输出的内容。

02 **标准内部。** 神经网络成为标准编解码器中的一个工具，例如每个解码器都必须运行的神经环路滤波器或帧内预测器，因此必须被精确规定，通常采用整数运算。较轻的形式是神经后处理滤波器：码流用元数据描述它，解码器可以选择是否执行。

03 **取代编解码器。** 编码器和解码器都通过端到端学习得到。码流是一组量化后的潜变量，只有与之配对的网络才能解码。例子：微软在 Teams 中使用的 MLVC，语音模型里的语音 tokenizer，游戏里的神经纹理。

图 02 **神经网络进入编解码的三条路径**

2026 年已大规模落地的，大多在标准编解码器外围；标准内部的神经工具正在制定；端到端神经编解码器只在一家公司掌控两端时才取代传统编解码器。



来源: 作者根据本报告引用资料综合整理；状态截至 2026 年 11 月 12 日。

第一种不需要任何人许可，第二种需要一个标准，第三种需要掌控两端。这也正是它们触达用户的先后顺序，下文的大部分内容都可以由此得到解释。

按现实世界排序，而不是按研究议程

研究论文大多研究第三种，因为压缩纪录都出在那里。本报告按各领域覆盖的人数排序，因为这决定了资金和工程投入流向哪里：先是播放和流媒体，然后是拍照、通话、生成式 AI、游戏，最后是专业链路（卫星、无人机、机器人）。据爱立信统计，2025 年底视频约占全部移动数据流量的 75%。2025 年，全行业出货约 12.6 亿部智能手机（IDC），以及约 1.96 亿至 2.09 亿台电视（TrendForce 和 Omdia 的估计）。每一台都用固定功能硬件解码视频。

我的研究方法

这是一项案头研究，做法和我评审一颗芯片或一份产品规划一样。我读了标准文件和会议报告（JVET、MPEG、JPEG、AOMedia、IETF）、公司工程博客和发布说明、芯片论文、专利池费率表、出口管制文本和新闻报道。每个数字都注明日期。凡是我自己推算的数字，例如根据论文表格算出的每像素计算量，我都会说明，并标为 估算。无法对照一手来源核实的说法，标为 厂商说法。金额以美元为准，括注的人民币按 1 美元兑 6.72 元换算。状态截至 2026 年 11 月 12 日。

能同类相比的地方，我尽量同类相比；做不到的地方，我会说明。压缩增益如果不附上四样东西，就没有多大意义：比较基准（相对什么的增益）、指标（PSNR、VMAF 还是观众评审）、内容和分辨率，以及解码代价。只要把这四点讲清楚，围绕神经编解码器的大部分分歧就会消失。

第 02 章

02 播放与流媒体：两端是神经网络，中间是标准

播放与流媒体

观众收到的仍是标准码流

最大的几家视频服务刚完成一次编解码器换代，又开始了下一次，两次都与神经解码器无关。2025 年 12 月，Netflix 报告 AV1 约占其观看量的 30%，是仅次于 AVC 的第二大编解码器。AV1 是随着硬件解码器一起进入电视的：Netflix 于 2019 年把 AV1 纳入设备认证，2021 年底开始向电视推送 AV1 流，并称 2021 至 2025 年间提交认证的大屏设备中有 88% 支持 AV1，而且几乎都是 2023 年以后的设备。YouTube 表示，截至 2026 年 3 月 19 日，按观看时长加权，其内容库 90% 以上已有 AV1 版本。

AV2 是开放媒体联盟（AOMedia）推出的 AV1 继任者，于 2026 年 5 月 28 日发布 1.0.0 版；按 AOMedia 的客观测量，压缩效率比 AV1 高约 30%。YouTube 计划“最早在 2027 年初”在网页端推出 AV2 流。芯片随即跟进：Chips&Media 在 5 月完成了 AV2 硬件解码器，6 月授权给一家未具名的北美大型科技公司；到 6 月 16 日，Allegro DVT 已向三家客户交付 AV2 解码器 IP；到 6 月 10 日，芯原已向多家客户发布支持 AV2 的视频 IP。AV2 是传统编解码器：它的增益来自更好的块划分、预测、变换和环路滤波（其中一个是“数据驱动”的引导式细节滤波器），而不是来自端到端网络。

现行的 ITU/ISO 标准 VVC，在巴西找到了第一个广播落脚点：巴西的 DTV+（TV 3.0）系统于 2026 年 6 月投入商业运营，时间正好赶上世界杯。截至 2026 年 7 月，巴西在售的电视还没有一台内置 DTV+ 接收器；第一批预计在 2026 年底至 2027 年初之间上市。

编码端：Netflix 的神经下采样器

播放领域最清晰的神经编码例子，自 2022 年起一直在运行，而且观众的设备什么都不用改。Netflix 把每部作品编码成多种分辨率，其中从 4K 母版生成 1080p、720p 和 540p 的下采样步骤，传统上使用 Lanczos 之类的固定滤波器。Netflix 用一个小型神经网络取代了它：一个前处理模块后接一个缩放模块，训练目标是提升设备用标准放大算法放大之后的观看质

量。Netflix 称之为即插即用的方案：编码端和客户端都无需其他改动，因此“数以百万计支持 Netflix 流媒体的设备都会自动受益”。在 Netflix 的偏好测试中，约 77% 的观众选择了神经下采样器的输出 厂商说法。为了在 Netflix 的规模下把成本控制在可接受的范围，“只用几层神经网络就够了”。

这个模式值得记住。神经网络在数据中心里对每部作品只运行一次，那里的算力便宜，而且可以调度；它的输出是普通的 AV1 或 HEVC 码流。

屏幕端：应用里和电视里的超分

在另一端，神经网络修复编解码器和片源损失掉的东西。2025 年 10 月 29 日，YouTube 为其电视应用推出超分（Super Resolution）功能：低于 1080p 上传的视频，从 240p 到 720p，都用机器学习超分到高清，并把 4K 作为既定目标。超分后的版本带有标注，观众可以切换回原版，创作者也可以选择退出。Meta 在 2025 年 10 月介绍了它运行视频超分的方式：既在上传时于服务器端运行，与英特尔合作使用标准 x86 CPU，也在播放时于客户端运行，并发现两者可以互补。电视厂商在芯片里做同样的事：三星的旗舰 8K 处理器 NQ8 AI Gen3，宣称有 768 个用于超分的神经网络。

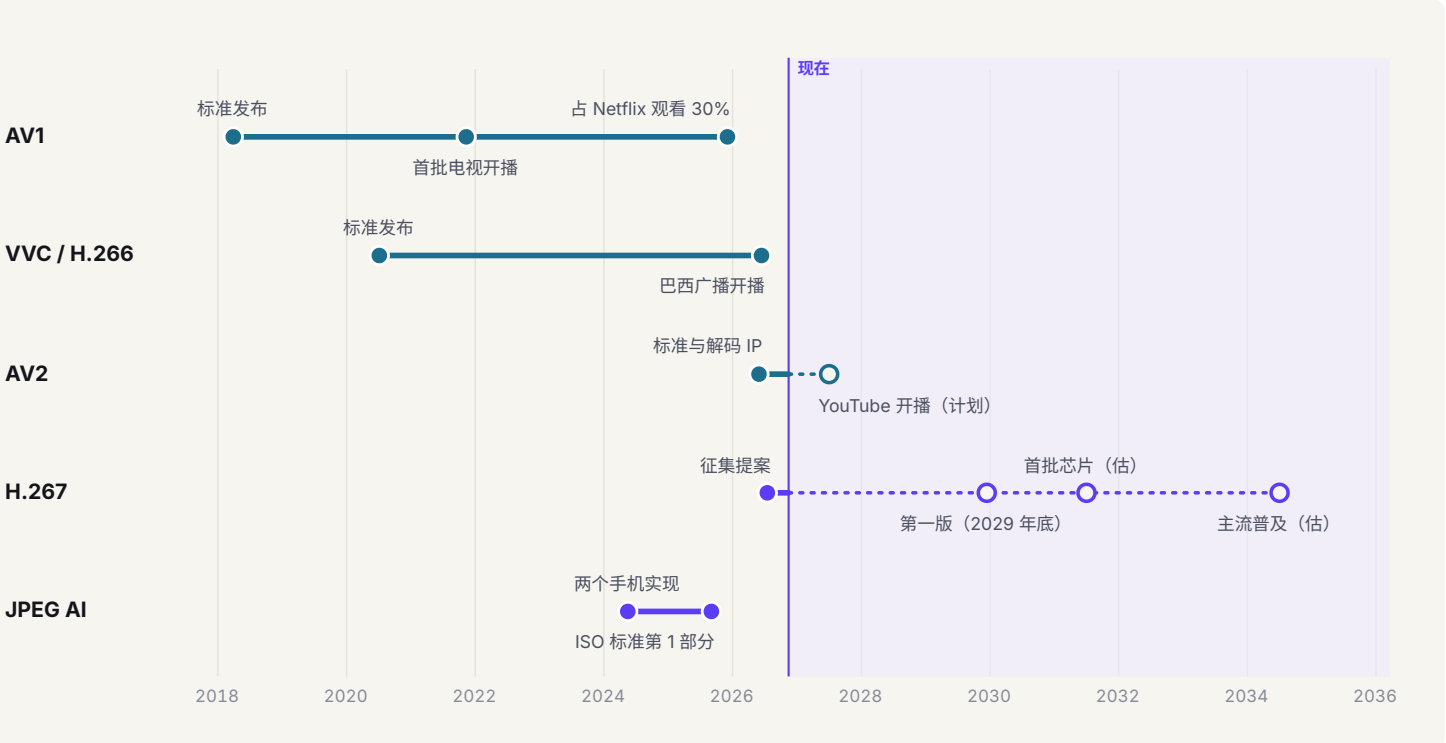
经济账取决于神经网络在哪里运行。在服务器端处理，平台每处理一个视频都要花算力。在终端侧处理，平台不花一分钱，用的是买家已经付过钱的 NPU。所以 2026 年最大的收益来自两者兼用。

为什么没有电视解码神经码流

播放格式必须能在今后所有要播放它的设备上解码，而这些设备一用就是好几年。因此，即便是传统编解码器，从标准定稿到主流屏幕的路也很慢。AV1 的标准于 2018 年 3 月定稿；dav1d 软件解码器几个月后推出；电视从 2021 年底开始播放 AV1 流；到 2025 年 12 月，AV1 约占 Netflix 观看量的 30%，距标准定稿将近八年。VVC 于 2020 年 7 月定稿，到 2026 年 6 月才迎来第一个广播服务。

图 03 **从标准定稿到主流屏幕，要花掉大半个十年**

播放需要设备里有硬件解码器。AV1 从标准定稿到占 Netflix 观看量的 30%，用了将近八年；第一个可能带神经工具的标准要到 2029 年底。



来源: Netflix 技术博客（2025 年 12 月）；AOMedia；Allegro DVT；TV Technology；Fraunhofer IIS（巴西 DTV+）；MPEG-155；JPEG 委员会；ISO。空心标记为目标或作者估算。

神经码流还会带来两个传统编解码器早已解决的问题：每像素能耗和逐比特一致的解码。第 8 章会讨论这两点。两者合在一起意味着，在标准明确规定要做什么之前，没有哪家电视或手机厂商会在每台设备里都放一个神经视频解码器。

什么能改变这一点：H.267

能把神经工具带到每一块屏幕上的标准，是 VVC 的继任者，由 ITU-T 和 ISO/IEC 在联合视频专家组（JVET）中共同制定。它的名称尚未确定；和业内许多人一样，我称之为 H.267。

2025 年 10 月，JVET 评估了征集证据的响应：来自 16 家公司和研究机构的五份响应，共测试了 20 种编解码器配置。委员会找到了压缩效率优于 VVC 的证据，并认为改进的传统技术和神经网络技术都“值得关注”。神经方法表现出明显的感知质量增益，在 8K HDR 内容上尤为明显，在游戏内容上增益较小。响应方之一 InterDigital 提交了两个方案：一个基于传统的增强压缩模型（ECM），另一个在 VVC 上叠加 AI 工具。

正式的征集提案于 2026 年 7 月在日内瓦会议上发布。主要提案的提交截止日期是 2026 年 10 月底，补充材料定于 12 月中旬提交，提案将在 2027 年 1 月布里斯班的 JVET 会议上评估。JVET 预计在 2027 年下半年推出第一个测试模型，到 2029 年底推出标准第一版，并称这项工作“可能是有史以来为启动一个新标准而付出的最大努力”。

传统技术路线设下了很高的门槛。用传统信号处理构建的 ECM，在随机接入配置下比 VVC 节省约 28%。据 JVET 神经网络探索的报告，环路滤波器在每像素 1 万次乘加时节省约 6%，2.2 万次时节省 8%，在高复杂度工作点上节省 14%。我的判断是：H.267 最可能是传统内核，加上少数节省的码率足以抵消解码代价的神经工具；2027 年 1 月的评估会显示这样的工具有多少 分析。

过渡的桥梁已经有了。VVC 和 HEVC 码流可以携带补充增强信息（SEI），描述一个交由解码端执行的神经后处理滤波器；2025 年 10 月完成的 VVC 及其 SEI 标准（VSEI）新版，扩展了这些消息。码流可以指明解码后运行哪个网络，而不必让每个解码器都运行它。

对播放领域企业的意义

现在就卖在编解码器外围管用的东西：给视频服务做编码端前处理，给电视和机顶盒芯片做后处理模块和调校。在 2030 年代之前，不要把播放业务建立在必须由第三方设备解码的神经码流上。如果你做芯片，近期的收入来自 AV2 和 VVC 解码器；神经部分是它们旁边一条用于增强处理的 NPU 通路（第 8 章）。

第 03 章

03

拍照与图像

拍照与图像：相机里是神经网络，网页上是传统格式

相机的处理流程已经用上神经网络

近几年旗舰手机拍的每张照片，在压缩之前都经过了神经网络。2023 年起上市的手机所搭载的高通 Snapdragon 8 Gen 2，引入了“Cognitive ISP”，在拍照和录像时实时做语义分割，对人脸、头发、天空和背景分别调整；三星 Galaxy S23 Ultra 是第一款在前置摄像头上使用它的手机。2026 年 1 月，Chips&Media 和 Visionary.ai 在 CES 上展示了它们所称的第一个完全基于 AI 的 ISP：一条端到端的成像流程，完全运行在 Chips&Media 的 NPU WAVE-N 上，采用 16 位浮点，不需要固定功能 ISP，并且可以通过 OTA 升级。“现有的流程是把神经模块外挂在固定功能 ISP 上。我们则完全取代传统 ISP。”Visionary.ai 的 CEO 说。展会之后，这套方案已提供给制造商评估。

然后，结果被存成 JPEG 或 HEIF，这些传统格式任何设备都能打开。模式和播放一样：一家公司自己的流程内部用神经网络，边界上用标准。

JPEG AI：标准已完成，产品还没有

JPEG AI 是 ISO/IEC 和 ITU-T 推出的第一个学习型图像编码标准（ISO/IEC 6048，ITU-T T.840）。第 1 部分，即核心编码系统，于 2025 年发布。JPEG 委员会报告，相对 VVC 帧内编码的压缩增益为 12.5% 至 27.9%，对应的解码复杂度从每像素 7 千到 21.6 万次乘加不等。它的设计包含两个编码器和三个解码器，联合训练，因此接收端只实现一个解码器，也能读取所有码流。到 2024 年 5 月，委员会报告已有两个智能手机实现。与此同时，IEEE 于 2024 年 12 月发布了 IEEE 1857.11-2024，这是一个为 NPU 设计的、基于神经网络的图像编码标准。

截至 2026 年 11 月，我没有找到任何默认写入或读取 JPEG AI 的手机相机、照片应用或浏览器。

网页采用的是另一种格式：JPEG XL

2026 年真正站稳脚跟的图像格式是传统的。2026 年 10 月 6 日，谷歌在 Chrome 155 中默认启用 JPEG XL 解码，使用一个用 Rust 编写、内存安全的新解码器，距它移除实验性支持已过去四年。Safari 自 2023 年起就能解码 JPEG XL，Mozilla 也在朝同一方向走。谷歌给出的理由是：压缩率比 JPEG 更高，能对现有 JPEG 文件做无损再压缩，并支持 HDR 和渐进式解码。

这个教训具有普遍性。必须到处都能打开的格式，靠的是一个快速、确定、能在任何 CPU 上运行、并且可以做到内存安全的解码器。神经图像解码器需要 NPU 或 GPU，还需要跨设备的精确运算。神经编码胜出的地方，是设备自己的处理流程，以及由同一家公司编码和解码的系统，比如从不把压缩文件交给别人的云相册服务。

对影像领域企业的意义

钱在拍摄流程里：AI ISP 的 IP、按成像需求确定规模的 NPU，以及相机厂商购买的调校服务。在存储和分发上，学习型图像编解码器今天是封闭系统里的工具，不是值得押注的格式。

第 04 章

04 通话与会议：通话与会议：端到端神经编解码器真正落地的地方

音频先行

微软于 2021 年 2 月推出 AI 语音编解码器 Satin，并称它已用于 Teams 和 Skype 的所有双人通话。Satin 在 6 kbps 下提供超宽带语音，音频带宽最高 16 kHz；而 Skype 之前的 Silk 编解码器在这个码率下只能做到窄带。省下来的码率用于冗余，因此通话更能扛住丢包。

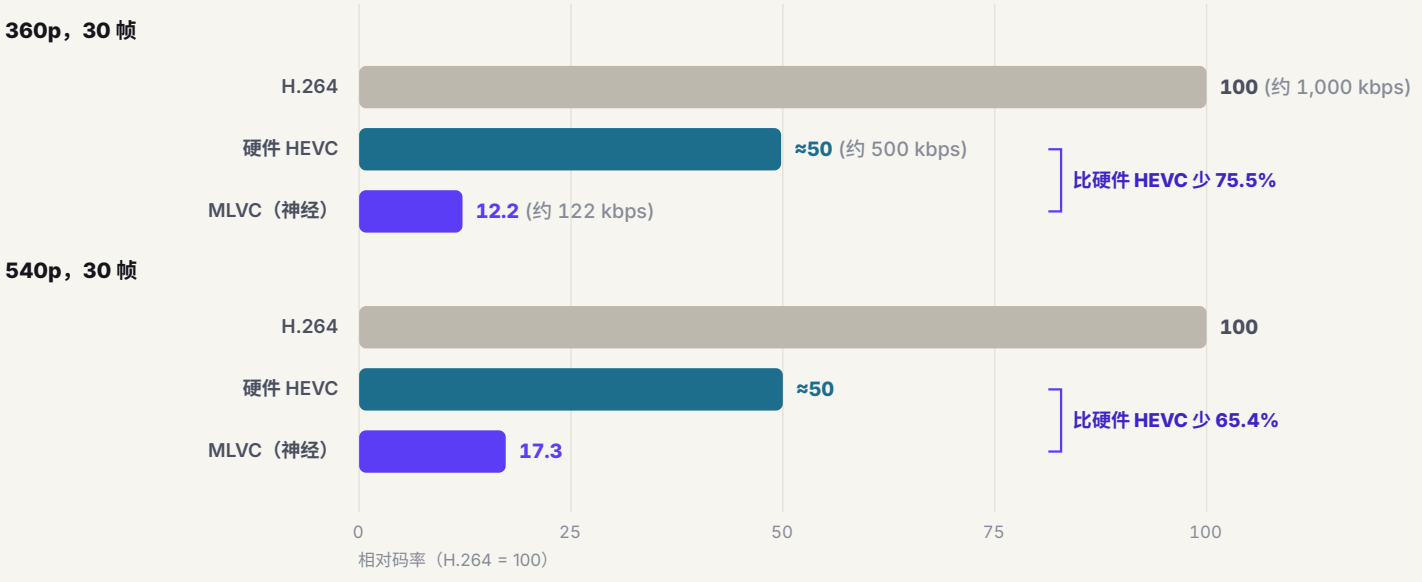
MLVC：已用于实际通话的神经视频编解码器

2026 年 7 月 24 日，微软以 MIT 许可证开源了它的机器学习视频编解码器 MLVC，包括模型代码、训练好的权重、训练脚本，以及面向不同 NPU 的转换工具。MLVC 是微软研究院自 2021 年起发表的 DCVC 系列的产品版。微软表示，正在 Teams 点对点通话中推广 MLVC，并配有实时遥测和 A/B 测试，以传统编解码器兜底。

MLVC 以每秒 30 帧运行在 540p 和 360p，并经过专门设计：在苹果的神经网络引擎以及高通和英特尔的 NPU 上，占用都不超过 NPU 的一半，工作由 CPU 和 NPU 分担。在微软自己的视频会议数据集上按 ITU-T P.910 做的主观测试中，360p 下它比 H.264 少用 87.8% 码率，比硬件 HEVC 少用 75.5%；540p 下分别少用 82.7% 和 65.4% 厂商说法。在 360p 下，这大约是 122 kbps，而 H.264 需要约 1 Mbps。

图 04　在通话中，神经编解码器比硬件 HEVC 少用 65–76% 码率

在微软视频会议测试集上达到相同主观评分所需的码率，以 H.264 = 100 计。MLVC 运行在笔记本和手机的 NPU 上，最高 540p30，占用不超过一半 NPU。



来源： 微软 MLVC 发布（2026 年 7 月 24 日）：在其视频会议数据集上按 ITU-T P.910 做的主观测试。HEVC 柱由公布的节省比例推算。均为厂商说法，非独立测量。

有两点要给这个标题数字降降温。88% 是相对 H.264；相对同一台笔记本里已有的硬件 HEVC，是 65–76%，仍然很大。另外，MLVC 为了能够部署而放弃了一部分压缩率：一份独立分析指出，按感知质量训练的 DCVC-RT 在同一指标上达到 81.9%，而 MLVC 是 75.5%。在不属于微软的硬件上解码出完全相同的结果，代价约为六个百分点。

为什么通话是天然的落脚点

通话是闭环。两端运行同一个应用，在通话开始时协商各自的能力，通话中途也能退回标准编解码器。分辨率不高，网络往往不好，而两端的笔记本电脑如今都配有 NPU，手机更是早已如此。微软把 40 TOPS 的 NPU 定为其 AI PC 的门槛，苹果 M4 的神经网络引擎标称 38 TOPS。在播放中阻碍神经编解码器的每一个条件，在这里都不存在。

反例：Opus 和 MLow

并非每个通话平台都做了同样的选择，其中的原因很有启发。

Opus 是 WebRTC 和大多数通话应用使用的开放音频编解码器，它在 1.5 版（2024 年 3 月）中加入了深度学习，这是它第一次用神经网络来处理或生成信号本身。但这些都以选项的形式加入：无需改动码流的解码端增强（LACE 和 NoLACE）、神经丢包隐藏，以及深度冗余方案（DRED）；全部默认关闭，以免给老设备增加负担。Opus 1.6 又加入了神经带宽扩展和一个缩小到三分之一的 DRED 模型，并且与 2012 年的标准保持完全兼容。这种做法是“在标准内部，可选，向后兼容”。

Meta 退得更远。2024 年 6 月，它推出了 MLow，一款基于传统信号处理而非神经网络的低码率语音编解码器，如今已用于所有 Instagram 和 Messenger 通话，并正在 WhatsApp 上推广。Meta 的理由在于用户的设备：超过 20% 的通话是在 ARMv7 手机上进行的，每天还有数千万次 WhatsApp 通话跑在用了十年以上的手机上。Meta 写道，神经编解码器，包括 Meta 自己的 EnCodec，只能在昂贵的手机上可靠运行。MLow 在 6 kbps 下的音质约为 Opus 的两倍，计算量却少 10%。

对用户从旗舰笔记本一直延伸到十年前手机的服务来说，起决定作用的是复杂度预算。神经编解码器最先落地在存量设备新、厂商掌控两端的地方。

对通信领域企业的意义

企业会议、远程呈现、远程协助、回传链路和专网，是端到端神经编解码器今天就能做成产品的地方，前提是两端都由你掌控。从一开始就要为三件事做好规划：一个兜底编解码器，在你支持的每一种 NPU 上做解码测试，以及用你自己的内容而不是厂商的数据集做质量评估。

第 05 章

05 生成式 AI：规模最大、却没人称之为编解码器的部署

每个语音和视频模型都运行在学习得到的编解码器上

端到端神经压缩规模最大的部署，不在任何编解码标准里，而在生成式 AI 内部。语音模型并不直接听波形，也不直接生成波形：它读写的是一个学习型音频编解码器的离散 token。Kyutai 于 2024 年发布的 Mimi，把语音转换成每秒 12.5 帧、1.1 kbps 的 token，码率低于任何电话编解码器，并在一条码流中同时包含语义信息和声学信息。Meta 于 2022 年发表的 EnCodec 是另一个例子；Meta 自己的工程师后来形容这类编解码器在低码率下表现出色，但对老手机来说太重（第 4 章）。

视频生成也是同样的方式。扩散模型或 Transformer 模型在一个学习得到的视频自编码器（VAE）的潜空间里生成内容，再由 VAE 的解码器把结果变成像素。LTX-Video 的 VAE 把视频压缩到 1:192：每个潜变量 token 代表 8 帧上一个 32×32 像素的块，有 128 个通道。它的解码器还承担最后一步去噪，因此它在一定程度上是在生成细节，而不只是重建细节。

为什么它们不是分发格式

语音或视频模型每收到一次请求，这些编解码器就运行一次，但没有一个是别人会去解码的格式。每个模型都有自己的潜空间；没有码率控制，没有一致性测试，也不保证换一个解码器能得到同样的画面。它们为生成质量而设计，而不是为给定码率下的保真度而设计，并且在数据中心 GPU 上解码。它们的成功表明，当一家公司同时拥有编码器、解码器和硬件时，学习型压缩可以大规模运行。

对企业的意义

有两点很实际。第一，对提供生成式视频或语音服务的公司来说，tokenizer 的效率直接就是一项成本：对于在所有 token 上做注意力计算的 Transformer，计算量的增长快于 token 数量的增长，因此压缩率更高的潜变量能让生成成本超比例下降 分析。懂潜变量设计、熵模型和解码器效率的编解码工程师，对生成式 AI 团队很有价值。第二，边界正在移动：标准组织已经开始研究面向机器而非面向人的编码。MPEG 的面向机器的视频编码（Video Coding for Machines，ISO/IEC 23888-2）于 2026 年初进入国际标准草案阶段，计划在 2026 年底定稿；JPEG AI 从一开始就被设计为既服务人眼观看，也服务计算机视觉任务。

第 06 章

06 游戏与图形：神经重建已是主流，神经资源刚刚起步

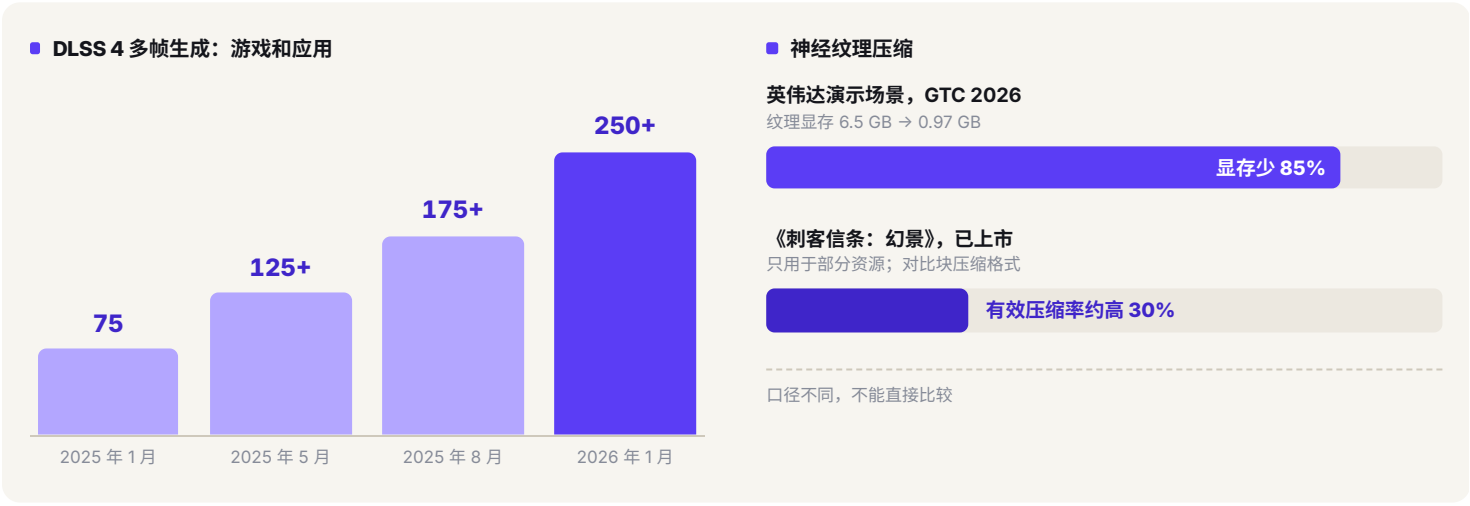
游戏与图形

少渲染一些像素，让神经网络补齐

在实时图形中，玩家看到的大部分内容已经由神经网络重建。英伟达的 DLSS 渲染更少的像素和更少的帧，其余由神经网络生成。支持多帧生成的 DLSS 4 于 2025 年 1 月推出，首发支持 75 款游戏和应用；按英伟达的统计，到 2025 年 5 月超过 125 款，8 月超过 175 款，到 2026 年 1 月超过 250 款，英伟达在那时发布了 DLSS 4.5，其超分采用第二代 Transformer 模型。这除了名字之外，就是一个解码器：“码流”是一帧低分辨率画面加上运动数据，神经网络据此重建完整图像。

图 05 **一家公司掌控整套技术栈时，神经重建普及得很快**

左：支持英伟达 DLSS 4 多帧生成的游戏和应用数量。右：神经纹理压缩在演示场景和已上市游戏中的效果；后者因运行开销只用在部分资源上。



来源: 英伟达（2025 年 1 月至 2026 年 1 月；GTC 2026 演示）；育碧 La Forge（2026 年 2 月）。均为厂商说法；两项纹理数据口径不同。

神经纹理已经上线，但有选择地使用

纹理用固定的块压缩格式（BCn）压缩，由 GPU 硬件解码。神经纹理压缩把一组材质纹理换成一个紧凑的潜变量表示，外加一个在需要时重建纹素的小型网络。2026 年 2 月，育碧 La Forge 实验室介绍了在 2023 年的游戏《刺客信条：幻景》中上线神经纹理压缩的经验：在相同视觉质量下，有效压缩率比传统格式高约 30% 厂商说法。它并没有全面使用：“使用受到运行时性能的限制”，因此育碧只把它用在实例数多、内存压力大的一部分资源上。

英伟达在 GitHub 上以 beta 版发布的 RTX Neural Texture Compression SDK，展示了上限。在 GTC 2026 的一次演示中，一个别墅场景用 BCn 时占用 6.5 GB 纹理显存，改用神经纹理后只需 970 MB，少了 85% 厂商说法。这个 SDK 提供两条路径：在加载时解码成 BCn，能节省磁盘空间和下载量，但不省显存；或者在采样时解码，能省显存，但每帧都要花着色器时间。

API 正在到位

图形 API 正在加入神经解码在着色器内所需的矩阵运算。微软于 2026 年 2 月正式发布 Shader Model 6.9；它在 2025 年预览的协作向量（cooperative vector）功能，被并入一个面向后续着色器模型的统一线性代数设计，DirectX Linear Algebra 则于 2026 年 4 月进入公开预览。等到每个驱动都支持它，神经纹理解码器就能用同一条代码路径在任何厂商的 GPU 上运行。

对企业的意义

游戏能上线神经解码，是因为闭环的条件成立：一个引擎，一条资源管线，GPU 里有张量硬件，还有兜底格式。尚待开拓的工作在于工具、能编码神经材质的资源管线、跨厂商运行时，尤其是移动端，那里起决定作用的是每解码一个纹素的能耗。

第 07 章

07

专业链路

专业链路：卫星、无人机和机器人

从纸面上看，这些是最适合端到端神经编码的场景。链路两端由同一个所有者掌控，带宽稀缺且容易丢包，内容范围窄到可以专门训练，而且画面只要在感知上乃至语义上忠实，往往就够了。面向这类链路的抗丢包神经编解码器是一个活跃的研究领域，在服务器端修复低码率画面的做法也已经在用。

与消费视频相比，它们也是一个规模小、监管日益严格的市场 分析。在中国，商务部会同另外两个部门于 2026 年 10 月 19 日发布 2026 年第 49 号公告，自 11 月 18 日起施行，修订了 13 类无人机相关物项的出口管制。据行业媒体报道，技术条目中有一项涉及无人地面平台，其界定部分依据 75% 的语义图像压缩、150 毫秒内完成图像重建等指标。语义压缩本身已经成为一项出口管制参数。在美国，联邦通信委员会于 2025 年 12 月 22 日把外国制造的无人机及其关键部件列入涵盖清单（Covered List），新的外国机型从此无法获得设备授权。

结论：对于拥有工业或政府客户、并有能力处理两边出口合规的专业公司，这是一个可靠的细分市场；但它不是主流市场，也不是检验神经编码能否大规模运行的地方。

第 08 章

08 芯片视角：能耗、算力、访存与精确性

这是我自己的背景最用得上的一章。我有多年无晶圆厂芯片研发和管理经验，包括全定制 ASIC 数据通路设计，以及视频编解码加速器等多媒体摄像头芯片项目；芯片团队对一个新编解码器提出的问题，并不是 BD-rate，而是：每像素多少焦耳，每像素多少字节 DRAM 流量，多少门，以及世界上每个解码器能否算出完全相同的比特？在这四点上，2026 年的神经解码器，相当于硬件 HEVC 第一批芯片出来之前的状态：软件上很有希望，芯片上很昂贵。

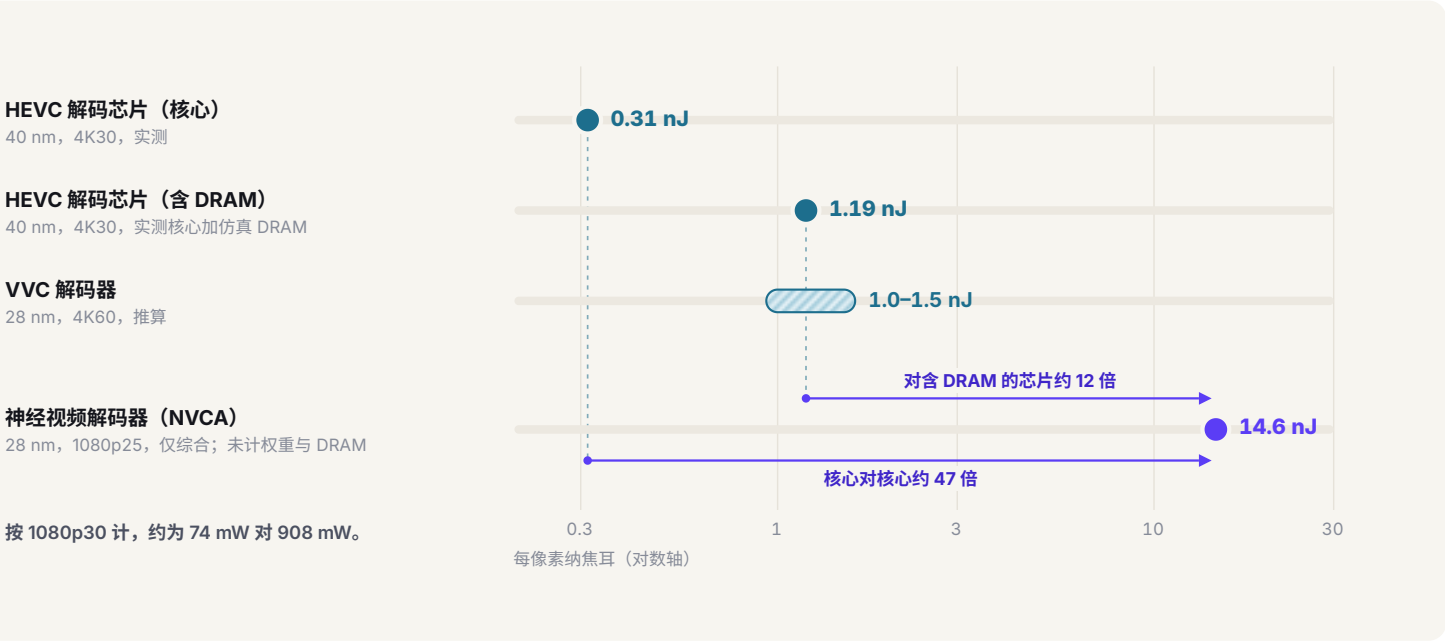
每像素能耗：相差一个数量级

做成芯片的神经解码器很少，所以比较只能依靠少数几个公开设计。传统一侧的基准是 MIT 的 HEVC 解码芯片（Tikekar 等，ISSCC 2013 和 IEEE JSSC 2014）：40 nm，1.77 mm²，在 200 MHz 下以每秒 30 帧解码 4K。它的核心功耗 78 毫瓦，即每像素 0.31 纳焦；加上仿真的 DRAM 功耗，系统每像素需要 1.19 纳焦。VVC 的量产数据不公开；Electronic Design 在 2026 年的一篇分析，把 MIT 芯片按比例缩放到 28 nm，并计入 VVC 更高的复杂度，推算一个约 100 万门的 4K60 解码器每像素需要 1–1.5 纳焦。

神经一侧，按 Electronic Design 的统计，唯一公开的端到端神经视频解码 ASIC 设计是 NVCA（Zhang 等，DATE 2024），采用台积电 28 nm 工艺综合，501 万门，不含放在片外的权重。它以每秒 25 帧解码 1080p，功耗 760 毫瓦，即每像素 14.6 纳焦。

图 06 **每解码一个像素的能耗：固定功能解码器领先一个数量级**

唯一公开的神经视频解码芯片设计，每像素能耗约为实测 HEVC 解码芯片（含 DRAM）的 12 倍、芯片核心本身的 47 倍；它自己的权重和 DRAM 访问还没算进去。



来源： Tikekar 等，IEEE JSSC 2014（MIT）；NVCA，Zhang 等，DATE 2024（arXiv 2312.10716）；VVC 区间为 Electronic Design 的推算（2026）。HEVC 芯片用的是更老的 40 nm 工艺（对比 28 nm），这对神经设计有利。

这约为含 DRAM 的 HEVC 芯片的 12 倍，约为其芯片核心本身的 47 倍。这个比较在三个方面对 NVCA 有利：HEVC 芯片用的是更老的工艺节点，NVCA 的数字没有计入权重和 DRAM 流量，而且 NVCA 是综合结果，不是实测芯片。按 1080p30 算，差别是约 74 毫瓦对大约 900 毫瓦：一个是没人会注意到的解码器，另一个会在手机的电池和散热预算里显现出来。

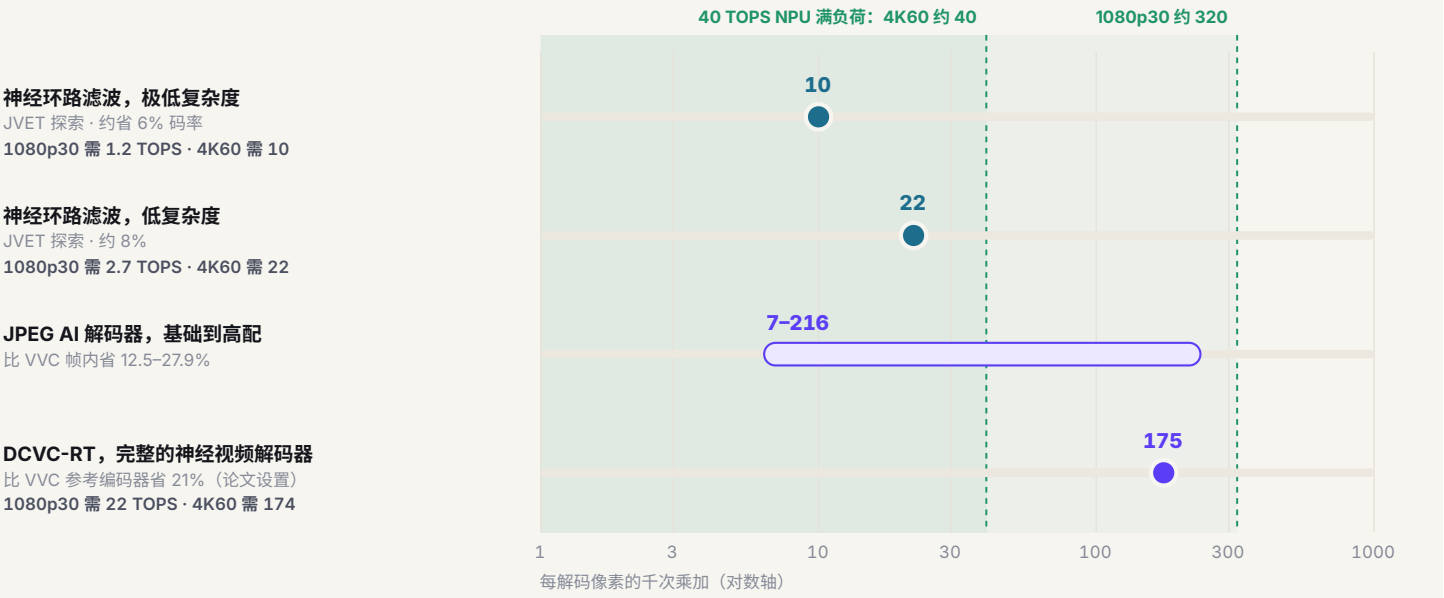
每像素算力：神经解码器对 NPU 的要求

我找到的已发表的最快神经视频编解码器，是微软研究院的 DCVC-RT（Jia 等，CVPR 2025）。在英伟达 A100 GPU 上，它对 1080p 的编码和解码速度分别为每秒 125 帧和 113 帧，在论文的测试设置下比 VVC 参考编码器节省 21% 码率，并使用整数运算，让不同设备的解码结果保持一致。它的补充材料列出了每个模块的复杂度。把解码时必须运行的四个模块（解码器、特征提取器、熵模型和重建）加起来，我得到每解码一个像素约 17.5 万次乘加 估算 。

按 1080p30 算，这是每秒约 11 万亿次乘加，按每次乘加两次运算计约为 22 TOPS，还没算任何利用率损失；4K60 需要八倍 估算 。作为参照，微软以至少 40 TOPS 的 NPU 来定义其 AI PC，苹果给 M4 神经网络引擎的标称值是 38 TOPS。这些都是峰值，还要与设备上的其他所有任务共享。

图 07　标准内的神经工具每像素数万次乘加，完整的神经解码器每像素数十万次

解码每个像素所需的乘加次数（对数轴），对照一颗 40 TOPS 的 NPU 满负荷时在 4K60 和 1080p30 下能负担的上限。各项增益的比较基准不同，不能直接比较。



来源： JVET（经 SC 29 主席，2026 年 4 月）；JPEG 委员会（2024）；DCVC-RT，CVPR 2025 补充材料，解码模块由作者加总；微软 Copilot+ PC 定义。作者计算：每次乘加按 2 次运算、利用率 100%。

图 7 中的阶梯，是理解这个领域的有用方式。每像素数万次乘加的神经工具，例如 JVET 的环路滤波器（1 万次时节省约 6%，2.2 万次时节省 8%），放进未来的标准解码器是可行的。JPEG AI 横跨整个区间，每像素 7 千到 21.6 万次乘加，相对 VVC 帧内节省 12.5–27.9%。每像素二十万次上下乘加的完整神经视频解码器，适合放在 NPU 上，以不高的分辨率运行在闭环中。这正是 MLVC 所在的位置：540p30，最多占用一半 NPU。

真正的限制往往是访存流量，而不是乘加

DCVC-RT 的作者发现，限制速度的不是运算，而是“运行开销”：内存的输入输出和函数调用次数。他们去掉显式运动模块和多尺度处理，很大程度上就是出于这个原因。NVCA 的设计者专门设计了一种数据流，以减少运动特征和残差特征带来

的片外访存流量。芯片设计者会认出这个规律。在 MIT 的 HEVC 芯片上，DRAM 约占系统能耗的四分之三（每像素 1.19 纳焦中占 0.88），而且这还是在一个把 DRAM 带宽降低了 67% 的缓存之后 分析。在帧与帧之间保留庞大多通道特征图的神经解码器，只会让访存问题更糟，而不是更好。

精确性：每个解码器都必须算出相同的比特

熵编码只有在解码器的概率模型与编码器完全一致时才能工作。在传统编解码器里，这个模型是一张表或一小组整数规则。在神经编解码器里，它是一个网络；两台设备之间的一点浮点差异，无论来自不同的内核、不同的累加顺序还是不同的 NPU，都可能让解码器读错符号，并从那一刻起丢失整个码流。

每个上线过或标准化过神经编解码器的团队，都不得不在工程上消除这个问题。DCVC-RT 把模型转换成整数，以实现跨设备一致的编解码。MLVC 采用共享尺度的超先验设计来扛住跨平台失配，这很可能是它相对 DCVC-RT 损失压缩率的部分原因 分析。JVET 的神经网络探索让每个网络都通过 SADL 运行，这是一个采用 16 位整数推理的小型库。JPEG AI 在潜变量域、残差域和像素域定义了一致性测试点。对芯片团队来说，这和 HEVC 整数变换是同一套规矩：标准必须精确定义运算，这会约束网络设计，也排除了厂商专有的捷径。

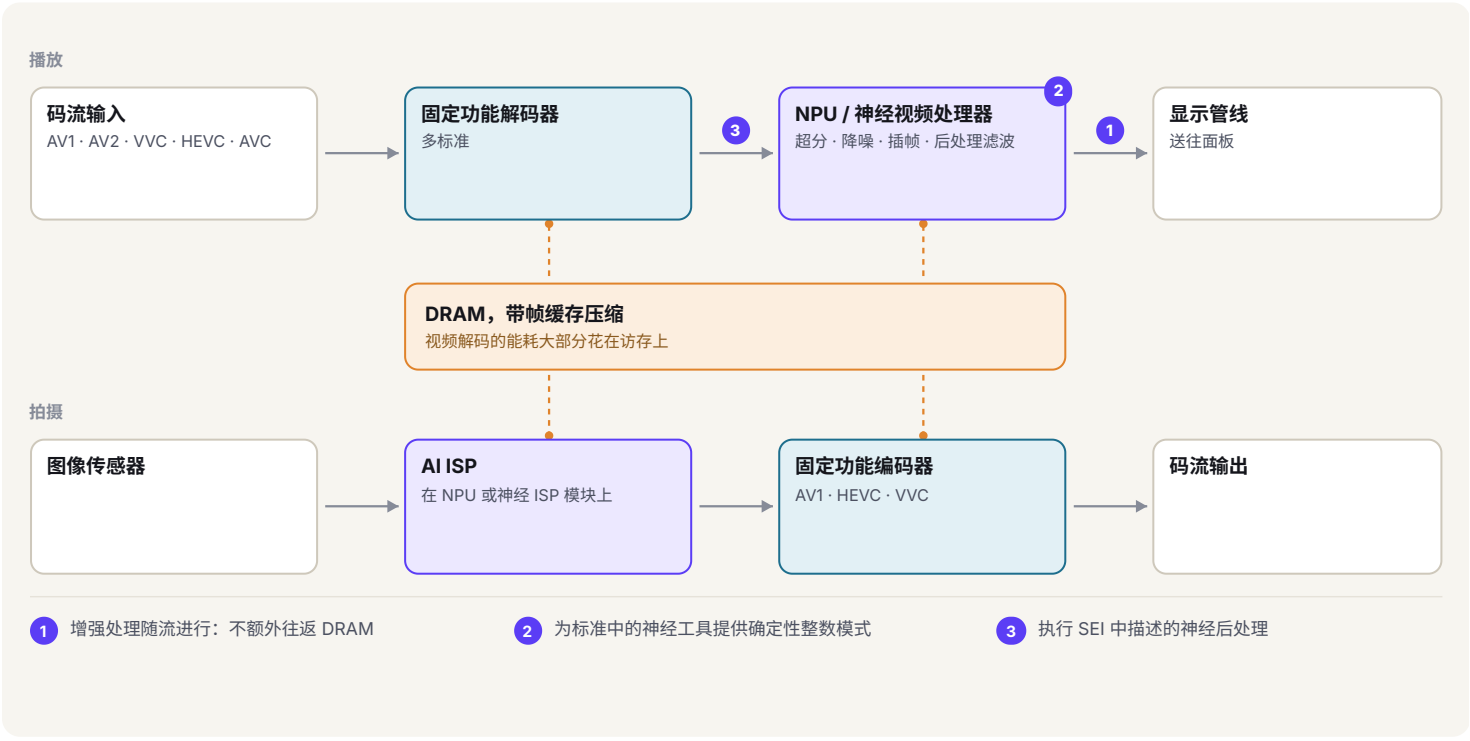
障碍不是面积，而是验证和生命周期

NVCA 的 500 万门，对比推算中约 100 万门的 VVC 解码器，看起来很多；但在已经配有数 TOPS 级 NPU 的现代手机或电视 SoC 上，几百万门并不是决定设计的因素。真正花钱的是：对照一致性测试套件做验证，在设备整个生命周期的软件更新中始终保证可编程解码器正确，以及 4K60 HDR 下的功耗。只有当标准确定了要做什么，这些成本才值得付出。

眼下胜出的架构：固定功能编解码器加 NPU

图 08　已经落地的分工：码流交给固定功能硬件，像素交给神经引擎

2026 年的媒体 SoC。编解码模块以每像素不到 1 纳焦的能耗解码各种标准；NPU 做神经网络擅长、标准又不限制的事。真正要算的成本是 DRAM 访问。



来源: 作者的架构观点，参考腾讯（沧海 V2）、Chips&Media、Allegro DVT 和三星的产品披露（2026）。

2026 年卖得出去的芯片，把固定功能的多标准编解码器和旁边的神经引擎配在一起。编解码模块以每像素不到 1 纳焦的能耗处理 AV1、AV2、VVC、HEVC 和 AVC。神经引擎负责神经网络擅长、标准又不限制的事：拍摄端的降噪和 ISP，显示端的超分、插帧和画质增强。两者共享帧缓存压缩和 DRAM 调度，因为访存流量才是关键成本。

市场印证了这种分工。腾讯自 2019 年起自研的沧海视频编解码 ASIC，第一代已有超过 10 万片部署在腾讯云及其自有的直播、点播和云游戏业务中；第二代于 2026 年 5 月宣布已点亮并进入量产，将于下半年投入服务，在同一架构中支持 H.265 和 H.266，压缩率比上一代分别提升 10% 和 30% 以上，吞吐量约为两倍，并有超分等片上画质增强。Chips&Media 同时销售编解码 IP 和面向视频的 NPU WAVE-N。Allegro DVT 销售 AV2 解码器，并推出了它的第一款神经视频处理 IP。三星的电视处理器把标准解码器和数百个超分网络配在一起。

按角色看，我会做什么

SoC 和电视芯片团队。 为每个必须支持的标准保留固定功能解码器，现在就要包括 AV2，广播有要求的地方还要包括 VVC。NPU 按显示分辨率和帧率下的后处理来定规模，并设计好从解码器到 NPU 再到显示的通路，让增强处理不额外往返 DRAM。支持神经后处理滤波 SEI。在 NPU 中提供确定性整数模式。把 H.267 当作 2029 年或更晚才开始的设计，并为它的神经工具保留一定的可编程性。

IP 供应商。 现在就卖 AV2 和 VVC 解码器，外加神经处理在它们周边需要的组件：后处理滤波加速器、帧缓存压缩和一致性工具。

视频平台。 编码端的神经前处理最先回本，因为它对每部作品只运行一次，对设备没有任何要求。在设备条件允许的地方使用客户端超分。端到端神经编解码器只留给你自己掌控的闭环。

第 09 章

09

标准与专利

标准与专利：神经工具在哪里变得可授权

标准进展

标准	组织	神经内容	状态（2026 年 11 月）
H.267（工作名；超越 VVC）	JVET（ITU-T 和 ISO/IEC）	未定。传统工具（ECM）和神经工具（环路滤波、帧内预测、超分）都是候选	2026 年 7 月征集提案；2027 年 1 月评估提案；2027 年下半年第一个测试模型；2029 年底第一版
VVC 和 HEVC，配合 VSEI	JVET	可在 SEI 元数据中描述神经后处理滤波器；解码端可选择是否执行	已生效；2025 年 10 月的新版扩展了相关消息
AV2	AOMedia	传统工具，免版税	2026 年 5 月 28 日发布 1.0.0 版；2026 年 6 月起已授权解码器 IP；YouTube 计划 2027 年初推出网页端流
JPEG AI	JPEG（ISO/IEC 和 ITU-T）	端到端学习型图像编解码器，服务人眼和机器视觉	第 1 部分于 2025 年发布；第二版正在讨论；未发现产品默认支持
IEEE 1857.11	IEEE	为 NPU 设计的神经网络图像编码	2024 年 12 月发布

标准	组织	神经内容	状态（2026 年 11 月）
面向机器的视频编码	MPEG	为机器分析而非人眼观看而编码	2026 年初进入国际标准草案阶段；计划 2026 年底定稿
Opus 机器学习扩展	IETF	可选的深度冗余、解码端增强、带宽扩展	已进入 libopus 1.5 和 1.6；深度冗余正在标准化

有两点很突出。第一，神经内容先从边缘进入标准，再进入核心：以可选的后处理滤波、可选的增强和图像编码的形式，这些都不改变主解码路径。第二，唯一可能把神经工具放进每块屏幕解码路径的标准 H.267，距离第一版还有三年。

专利布局中谁占了先机

标准工具会变成可授权的专利，所以 H.267 的制定过程也是一场专利竞赛，神经工具也在其中。GreyB 2025 年对 3,916 件超越 VVC 专利所做的全景分析显示，字节跳动排在最前，拥有 645 个专利族和 251 份 JVET 提案；分析还提到了产学合作，例如小米与西安电子科技大学，拥有超过 20 个专利族。MLex 在 2026 年 3 月报道，腾讯、阿里巴巴和字节跳动正在成为视频编解码专利池的活跃参与者。2026 年 5 月，Sauvegarder Investment Management 从阿里巴巴购买了 126 项视频编解码专利资产，覆盖 AV1、AV2 和面向机器的视频编码。InterDigital 于 2025 年 10 月收购 Deep Render，把 Deep Render 的 AI 视频编码专利纳入了自己的专利组合。

每台设备的钱，从专利池的费率表上一目了然。Access Advance 的 VVC 费率（2026 年 7 月 1 日起，适用于合规并使用其商标的被许可方）为：在美国、欧盟和日本等一类地区，每部手机、平板或笔记本电脑 0.50 美元（约 3.4 元），二类地区 0.25 美元（约 1.7 元）；一类地区的 4K 电视按价格为 1.00 至 1.50 美元（约 6.7 至 10.1 元）；每家企业每年封顶 7,500 万美元（约 5 亿元）。AOMedia 成员对 AV1 和 AV2 免收专利费。无论 H.267 包含什么，它的工具都会以同样的方式定价，这就是为什么拥有神经编码专利的公司现在就在向 JVET 提交提案 分析。

第 10 章

10

钱流向了谁

钱流向了谁，公司可以在哪里切入

2026 年钱在哪里

图 09 **2026 年，神经编码的钱流向了谁**

钱跟着标准和平台走。作者没有找到任何一家单靠授权端到端神经视频编解码器养活自己的独立公司。

现在就赚钱

平台

留住节省下来的钱

Netflix 神经下采样；YouTube、Meta 超分；微软以 MIT 许可开源 MLVC

现在就赚钱

终端和芯片厂商

卖差异化

摄像头 ISP；电视超分画质芯片；腾讯自研沧海编解码芯片

现在就赚钱

IP 供应商

每个新标准都授权一次

AV2 定稿三周内，Chips&Media、Allegro DVT、芯原都宣布了 AV2 解码客户

现在就赚钱

GPU 厂商

构建平台锁定

DLSS 4 覆盖 250 多款游戏和应用；神经纹理 SDK 处于 beta

按台收费

专利权人和专利池

按台收费

VVC：每部手机 0.50 美元，每台 4K 电视 1.00–1.50 美元；H.267 工具 2029 年后

以被收购退出

神经编解码创业公司

被收购

WaveOne → 苹果（2023）；iSIZE → 索尼（2023）；Deep Render → InterDigital（2025）

来源: 报告引用的公司和专利池披露；状态截至 2026 年 11 月 12 日。专利池费率：Access Advance，VVC，一类地区，合规并使用商标，2026 年 7 月 1 日起。

平台留住节省下来的钱。 Netflix、YouTube、Meta、微软和腾讯都自研神经处理，并把收益留在自己手里：更低的带宽、更少的存储和更高的每比特质量。在有助于推广的时候，它们会把代码送出去：微软以 MIT 许可证发布了 MLVC。

终端和芯片厂商卖差异化。 摄像头 ISP 和电视画质处理器，是神经网络触及每个用户的地方，它们被当作卖点来销售：高通的 Cognitive ISP，三星的超分处理器。

IP 供应商为每个新标准卖解码器，外加 NPU。 AV2 定稿三周之内，Chips&Media、Allegro DVT 和芯原都宣布了 AV2 解码器客户。这些供应商如今也在编解码器旁边销售神经处理模块。

GPU 厂商构建平台。 DLSS 和神经纹理压缩，让英伟达的张量硬件成为游戏制作方式的一部分。

专利权人按台收费，通过 VVC 和 HEVC 的专利池收取，2029 年以后还有 H.267。

神经编解码创业公司都被收购了。 WaveOne 于 2023 年被苹果收购，iSIZE 于 2023 年被索尼互动娱乐收购，Deep Render 于 2025 年被 InterDigital 收购，交易条款均未披露。每一起都是平台或专利许可方买下了团队和专利。我没有找到任何一家在 2026 年单靠授权端到端神经视频编解码器养活自己的独立公司。

公司可以在哪里切入

切入点	谁来买	为什么是现在	主要风险
编码端神经前处理（下采样、降噪、逐片调优）	流媒体和用户视频平台、广播机构	设备无需改动；已在 Netflix 得到验证	大平台自己做

切入点	谁来买	为什么是现在	主要风险
显示端增强 IP 和调校（超分、插帧、画质）	电视、机顶盒和手机芯片厂商	设备里已经有 NPU；差异化看得见	同质化；芯片厂商自研
闭环神经编解码器（会议、远程操作、专用链路）	企业通信、工业和远程呈现厂商	MLVC 证明了可行性，而且代码已开源	必须掌控两端；NPU 碎片化
编解码器加 NPU 的芯片或 IP（AV2、VVC 和后处理）	SoC 厂商	AV2 周期正在开始	在位者实力强
评测、一致性和确定性推理工具	所有交付神经工具的公司	逐比特一致是硬性要求	不捆绑销售则市场很小
向 H.267 贡献标准提案	专利许可方、大平台	提案将于 2027 年 1 月评估	周期长、成本高、结果不确定

对大多数公司来说，2027 年能拿到收入的是前两行。第三行对专注的产品公司最有意思。最后一行需要耐心和一套授权策略。

第 11 章

11 说法与证据

你会听到的说法，以及证据怎么说

说法	证据怎么说
“神经编解码器已经胜过 VVC。”	在实验室里，是的。在论文的设置下，DCVC-RT 在 1080p 上比 VVC 参考编码器节省 21%；JPEG AI 比 VVC 帧内节省 12.5–27.9%。代价是解码器每像素约 17.5 万次乘加，也就是要动用一颗 NPU 的能耗，而固定功能硬件每像素不到 1 纳焦。
“AI 编解码器几年内就会取代 H.26x 和 AV1。”	在播放领域不会。第一个能承载神经工具的标准目标是 2029 年底，芯片约在 2031 年跟进，2030 年代中期普及 估算 。眼下到来的编解码器 AV2 是传统的。
“MLVC 把带宽削减 88%。”	这是在 360p 下相对 H.264，基于微软自己的测试集。相对硬件 HEVC，360p 下是 75.5%，540p 下是 65.4%，而且是在微软掌控两端的通话中。
“神经纹理把显存减少 85–96%。”	那是演示场景。已上市的游戏有效压缩率提升约 30%，并且因为运行开销，只在部分资源上使用神经纹理。
“JPEG AI 是 ISO 标准，所以相机和浏览器都会采用。”	我没有找到任何产品默认支持。2026 年网页采用的格式，即 Chrome 155 中的 JPEG XL，是传统格式。
“神经解码对手机来说太贵了。”	部分正确。MLVC 用不超过一半的 NPU 就能解码 540p30；用 DCVC-RT 级模型解码 4K60，则需要约 175 TOPS 的持续算力 估算 。
“卫星和无人机链路是神经视频的大市场。”	技术上最契合，但市场小，而且在太平洋两岸都受到越来越多的出口管制。
“大平台会从创业公司那里授权神经编解码器。”	它们要么自己做（Netflix、微软、腾讯），要么把创业公司买下来（苹果、索尼、InterDigital）。

一个简单的检验就能看穿大多数推销：问清比较基准、指标、分辨率和解码代价。讲不全这四项的说法，是研究成果，而不是产品。

第 12 章

12

关键节点

2027 至 2030 年值得关注什么

时间	节点	为什么重要
2027 年 1 月	JVET 在布里斯班评估 H.267 提案	排名靠前的提案有多少依赖神经工具，每像素代价多大
2027 年初	YouTube 首批 AV2 网页流；已发布的 SoC 中出现 AV2 解码器	下一轮传统编解码周期将决定到 2030 年前后的设备预算
2027 年下半年	H.267 的第一个测试模型	神经工具是进入核心设计，还是只停留在探索阶段
2027 年	MLVC 走出 Teams 点对点通话；其他厂商采用开源代码	开源神经编解码器第一次在其所有者之外接受检验
2027–2028 年	第一个默认写入或读取 JPEG AI 的产品	第一个跨越公司边界使用的神经格式
2027–2028 年	DirectX Linear Algebra 进入正式发布的驱动；神经纹理用于更多游戏和移动端	图形领域的跨厂商神经解码
持续	手机和 PC 的 NPU 算力	持续算力达到约 100 TOPS 时，DCVC-RT 级的 1080p30 解码器将只占约五分之一 NPU，而不是一半 分析
2029 年底	H.267 第一版	神经工具进入主流播放的芯片周期由此开始

第 13 章

13 开放问题

本报告中的几个问题目前还没有公开答案，也是我在技术上最感兴趣的。

大规模下的精确性。 在不同厂商的 NPU 上，神经解码器能在多大程度上接近逐比特一致、纯整数的运算，而又不丢掉大部分压缩优势？MLVC 用六个百分点做的交换是一个数据点；标准需要一个普遍适用的答案。

先看访存，再看乘加。 神经后处理在硬件解码器和显示之间随流运行时，每像素要占多少 DRAM 流量？它的工作集能做到多小？

H.267 中神经工具的价格。 在 4K60 解码器里，神经环路滤波器的每像素乘加次数和纳焦数到多少时，就不再抵得上它节省的码率？

通话之外的闭环。 从云相册存储、游戏串流到机器人和专用回传链路，还有哪些封闭系统能在 H.267 到来之前采用端到端神经编解码器？

会生成的解码器。 当解码器生成细节而不是重建细节时（视频 VAE 已经在这样做），保真度该如何衡量和规定？

未来在这些领域，我们可以一起探讨，包括技术交流和技术合作：i@tanshuai.com。

本报告的局限

这是一项基于截至 2026 年 11 月 12 日公开文件的案头研究。本版没有做一手访谈，也没有做独立测量。若干关键数字是厂商或作者在自己测试集上的说法，例如 MLVC 的码率节省和神经纹理的压缩率；文中都已如实标注。我对 DCVC-RT 的算力估算，是把其作者公布的分模块数字加总，没有考虑利用率损失，因此实际的 NPU 负载会更高。能耗比较依据的是少数几个不同工艺节点上的公开设计；量产 VVC 和 AV2 解码器的数据不公开。标准时间表是各委员会自己的目标。2029 年以后的 H.267 时间点，是我根据 AV1 和 VVC 的历史做出的估计。

关于作者

谭帅，技术专家和创业者，在芯片、硬件和系统软件领域工作二十多年，拥有 20 项授权专利。他有多年无晶圆厂芯片研发和管理经验，包括全定制 ASIC 数据通路设计，以及视频编解码加速器等多媒体摄像头芯片项目，这也是本报告的视角：每像素能耗、访存流量，以及标准允许解码器在硅片上固化什么。他的工作经历遍及硅谷、纽约、芝加哥、台湾地区、粤港澳大湾区（深圳和香港）、越南和马来西亚。如今他是一名 AI 研发者。

技术交流：i@tanshuai.com · tanshuai.com

说明与免责声明

本报告是基于截至 2026 年 11 月 12 日可获得的公开文件所做的独立案头研究，仅供参考，不构成投资、法律或工程建议。产品、标准时间表、专利池费率和出口规则都会变化；凡你要依赖的内容，请对照一手来源核实。标为 厂商说法 、 估算 或 分析 的数字，应按其标注来理解。公司和产品名称归其所有者所有，仅用于识别；没有任何公司审阅或赞助本报告。

© 2026 谭帅。在注明出处的前提下，你可以原样分享本报告。

资料来源

全部于 2026 年 11 月 3 日至 11 月 12 日之间访问，按章节分组。

流量、设备与范围

- Ericsson，*Mobility Report*，2026 年 6 月（2025 年底视频约占移动数据流量的 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，全球智能手机出货量，2025 年第四季度及全年（12.6 亿部），2026 年 1 月 13 日：https://www.idc.com/resource-center/press-releases/q425mobilephonetop5/
- TrendForce，2025 年电视出货量预测（1.9571 亿台）：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 年电视出货量预测（2.087 亿台）：https://iconnect007.com/article/145954/omdia-forecasts-flat-tv-shipments-in-2025-but-mini-led-poised-for-explosive-growth-at-829/145951/ein

播放与流媒体

- 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”（YouTube 内容库中 AV1 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+ 商业运营，2026 年 6 月）：https://mpegh.com/brazil/；CartaCapital 关于电视内置接收器的报道：https://www.cartacapital.com.br/?p=628337；Olhar Digital 关于 TV 3.0 采用 VVC 的报道：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 电视亮点介绍（NQ8 AI Gen3，768 个神经网络）：https://www.samsung.com/us/tvs/8k-tv/highlights/

拍照与图像

- 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 和 Visionary.ai，完全基于 AI 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 的采访：https://petapixel.com/2026/01/18/a-look-at-the-worlds-first-full-ai-based-image-signal-processor
- JPEG 委员会，第 103 次会议新闻稿（JPEG AI 的增益、复杂度和智能手机实现）：https://jpeg.org/items/20240507_press.html
- G. J. Sullivan，ISO/IEC JTC 1/SC 29 介绍，2026 年 4 月（JPEG AI 状态；ECM 和 NNVC 数据）：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 工作组（基于神经网络的图像编码）：https://sagroups.ieee.org/fvc
- Chrome 155 与 JPEG XL，2026 年 10 月 6 日：https://xenospectrum.com/en/chrome-155-jpeg-xl-rust/

通话与会议

- 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；代码：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 的分析：https://ireplay.tv/blog/mlvc-microsoft-neural-video-codec-open-source-streaming/
- Opus 1.5 发布说明：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 以上的 NPU）：https://www.microsoft.com/en-us/windows/copilot-plus-pcs ；Apple，M4 发布（38 TOPS 神经网络引擎）：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 视频 VAE）：https://arxiv.org/abs/2501.00103
- MPEG 第 153 次会议新闻稿（面向机器的视频编码进入 DIS 阶段）：https://www.mpeg.org/wp-content/uploads/2026/03/press_release_153.pdf

游戏与图形

- NVIDIA，DLSS 4 的采用情况：发布时和 Computex 2025（75 款和 125 款以上）：https://www.nvidia.com/en-us/geforce/news/125-dlss-4-multi-frame-gen-games-more-announced-computex-2025/ ；Gamescom 2025（175 款以上）：https://www.nvidia.com/en-us/geforce/news/gamescom-2025-dlss-4-full-ray-tracing-game-announcements/ ；CES 2026（250 款以上，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 关于英伟达 GTC 2026 神经纹理演示的报道（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/ ；关于协作向量：https://devblogs.microsoft.com/directx/shader-model-6-9-and-the-future-of-cooperative-vector/ ；GDC 2026 计划：https://windowsreport.com/microsoft-showcases-neural-rendering-and-dxr-2-0-plans-for-directx-at-gdc-2026/

专业链路

- Global Trade Alert 关于中国 2026 年第 49 号公告的条目：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（无人地面平台的指标）：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 将外国制造的无人机及部件列入涵盖清单，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/

芯片

- M. Tikekar 等，“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 芯片页面（4K30 下 78 毫瓦）：https://eecsg.mit.edu/chips/hevc-video-decoder-for-4k-ultra-hd-applications/
- S. Zhang 等，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，“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 ；补充材料（分模块复杂度）：https://openaccess.thecvf.com/content/CVPR2025/supplemental/Jia_Towards_Practical_Real-Time_CVPR_2025_supplemental.pdf
- NNVC 与 SADL 整数推理库：https://arxiv.org/abs/2309.05846
- 腾讯沧海 V2，2026 年 5 月 27 日（上海证券报，经同花顺转载）：https://stock.10jqka.com.cn/20260527/c677016020.shtml ；功能概要：https://www.besthub.dev/articles/tencent-s-canghai-v2-chip-enters-mass-production-what-the-new-capabilities-mean-for-video-encoding-bcfc18dbfa29
- Chips&Media 的 AV2 解码器 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 授权，2026 年 6 月 17 日：https://www.edge-ai-vision.com/?p=58270
- Allegro DVT，AV2 解码器 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，支持 AV2 的 VC9800D，2026 年 6 月 10 日：https://www.businesswirechina.com/en/news/63176.html

标准与专利

- MPEG 第 155 次会议新闻稿（超越 VVC 的联合征集提案），2026 年 8 月 5 日：https://www.mpeg.org/wp-content/uploads/2026/08/ag3w254SN26474Press-Release-MPEG155.pdf
- MPEG 第 152 次会议新闻稿（征集证据的结果；VVC 和 VSEI 新版），2025 年 10 月 30 日：https://www.mpeg.org/wp-content/uploads/2025/11/Press-Release-MPEG152.pdf ；ACM SIGMM 的 MPEG 专栏：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，超越 VVC（H.267）的专利全景：https://greyb.com/blog/exclusive-video-codec-landscape/ ；关于 AV2 授权和阿里巴巴专利组合出售：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 和 VVC Advance 的专利费率，2026 年 7 月 1 日起：https://accessadvance.com/wp-content/uploads/2025/07/MCBA-Program-Overview-September-2026.pdf

资金与公司

- InterDigital 收购 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/

学习型压缩 · 播放 × 芯片

2026 年的神经编码

围绕编解码器落地，而非取而代之

本报告的每个关键数字都注明了来源和日期。作者自行推算的数字标为估算；只有卖方报告的数字标为厂商说法。

谭帅

技术专家和创业者 · AI 研发者

tanshuai.com · i@tanshuai.com

公开版 · 数据截至 2026 年 11 月 12 日