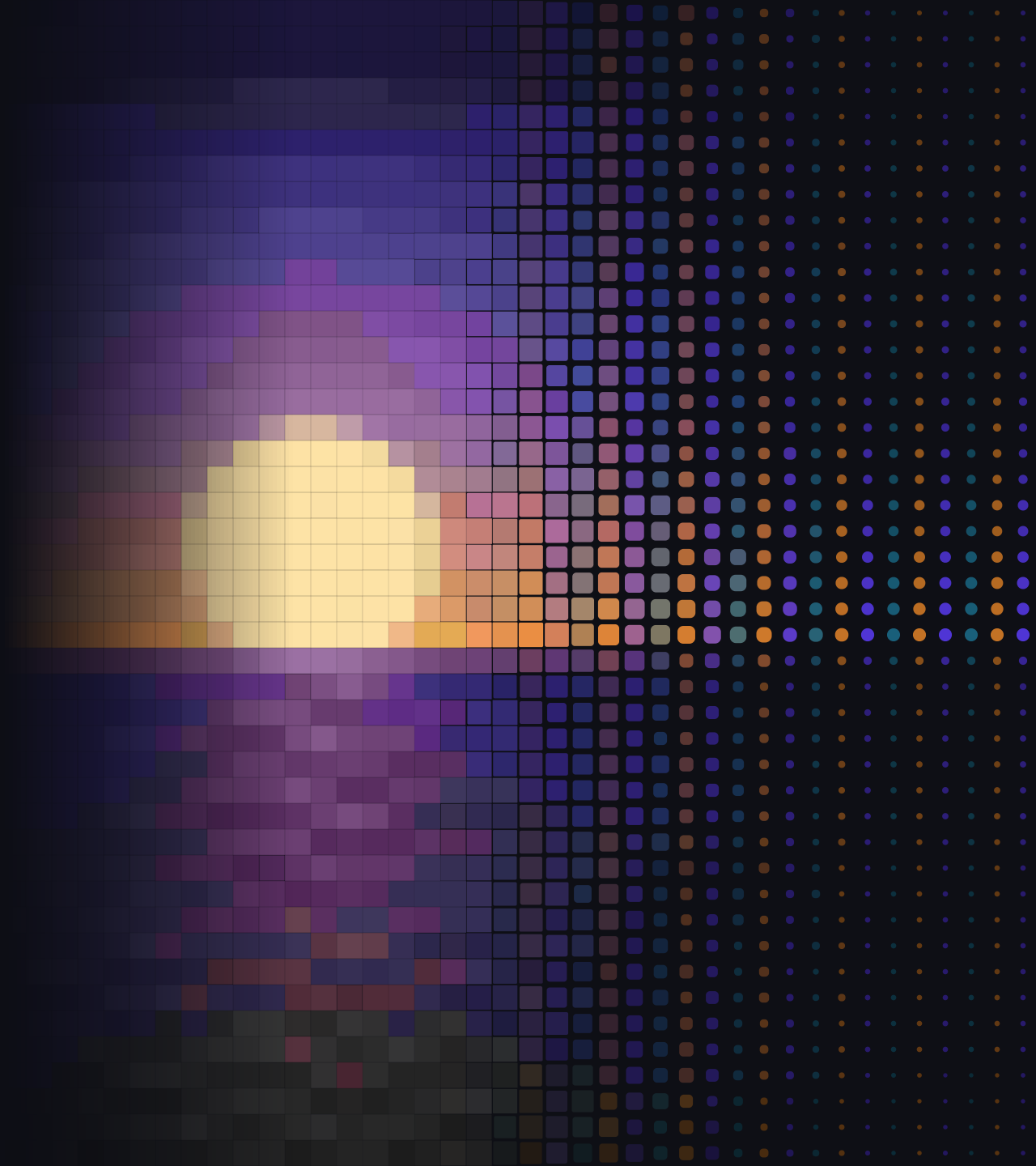
研究ノート · 2026年11月

2026年のニューラル符号化

コーデックの代わりではなく、その前後で



再生 · 写真 · 通話 · 生成AI · ゲーム · シリコン

タン・シュアイ

tanshuai.com · i@tanshuai.com

結論

すでに主流、ただしコーデックの前後で

01 ニューラルネットワークは標準コーデックの両端ですでに主流

カメラのISP、符号化側のダウンスケール、再生側の超解像。その間のビットストリームは今もAV1、HEVC、AVC、VVCで、固定機能ハードウェアが復号する。

02 エンドツーエンドのニューラルコーデックは閉じた系でのみ実用化

通話、生成AIのトークナイザー、そしてAAAゲーム1作。いずれも一社が両端を握り、標準コーデックにフォールバックできる。

03 一般の画面でのニューラルビットストリームはH.267待ち

第1版は2029年末の予定。初のチップは2031年ごろ、主流化は2030年代半ば 推定 。

約75%

モバイルデータトラフィックに占める動画の割合（2025年末）

12.6億台

2025年のスマートフォン出荷台数。すべてが固定機能ハードウェアで動画を復号

約12倍

唯一公開されたニューラルデコーダーチップ設計の画素あたりエネルギー（ハードウェアHEVCデコーダー比）

定義

ニューラルネットワークがコーデックに入る三つの経路



コーデックの前後なら、誰の許可も要らない。標準の内部には、標準が要る。コーデックの代わりには、両端を握ることが要る。

実用化の場

2026年、ニューラル符号化はどこで実用化されているか

	標準コーデックの前後 前処理、画質改善、アップスケール	エンドツーエンドのニューラルコーデック 学習型のエンコーダーとデコーダー	標準の内部 標準のビットストリームやAPIに入るニューラルツール
01 再生とストリーミング	● 大規模に実用化 NetflixのダウンスケーラーYouTubeとMetaの超解像、テレビのAIアップスケーラー	● 未導入 ニューラル映像フォーマットを配信するストリーミングサービスはない。テレビとスマートフォンは標準コーデックをハードウェアで復号する	● 標準化中 H.267第1版は2029年末の予定。ニューラルツールも候補に入っている
02 写真と画像	● 大規模に実用化 フラッグシップスマートフォンのAI ISP。CES 2026では完全なニューラルISPがIPとして披露された	● デモ・試験導入 JPEG AIはISO標準。スマートフォンでのデモはあるが、製品はない	● 未導入 普及が進むのは従来型のフォーマット。Chrome 155でJPEG XLがデフォルトで有効に
03 通話と会議	● 大規模に実用化 受信側での超解像とニューラルノイズ抑制	● 大規模に実用化 2021年からTeamsの二者間通話でニューラル音声。MLVCによる映像はTeamsで展開中	● 限定的に実用化 Opus 1.5がオプションで後方互換のMLツールを追加
04 生成AI	—	● 大規模に実用化 あらゆる動画・音声モデルが学習型トークナイザーを使う。例：圧縮率1:192の動画VAE	—
05 ゲームとグラフィックス	● 大規模に実用化 ニューラルアップスケールとフレーム生成。DLSS 4だけで250以上のゲームとアプリが対応	● 限定的に実用化 ニューラルテクスチャがAAAゲーム1作に搭載。NVIDIAのSDKはベータ版	● 標準化中 DirectXがシェーダー向けの行列演算を追加中（プレビュー）
06 特殊用途リンク：衛星、ロボット	● デモ・試験導入 低ビットレート映像のサーバー側での復元	● デモ・試験導入 損失に強いニューラルコーデックが研究・試験段階	—

行は各領域が届く人数の多い順に並べた。

両端はニューラル、中間は標準

- **エンコーダー側**

  Netflixのニューラルダウンスケーラーは2022年から稼働し、どのデバイスにも変更を求めない。Netflixのテストでは視聴者の約77%がこちらを好んだ ベンダー発表 。

- **画面側**

  YouTubeのSuper Resolutionは2025年10月からテレビで提供。Metaはサーバーとスマートフォンの両方で超解像を動かす。

- **テレビの中**

  SamsungのNQ8 AI Gen3プロセッサーは、アップスケール用に768のニューラルネットワークをうたう。

ビットストリームは標準のまま

30%

Netflixの視聴に占めるAV1の割合（2025年12月）

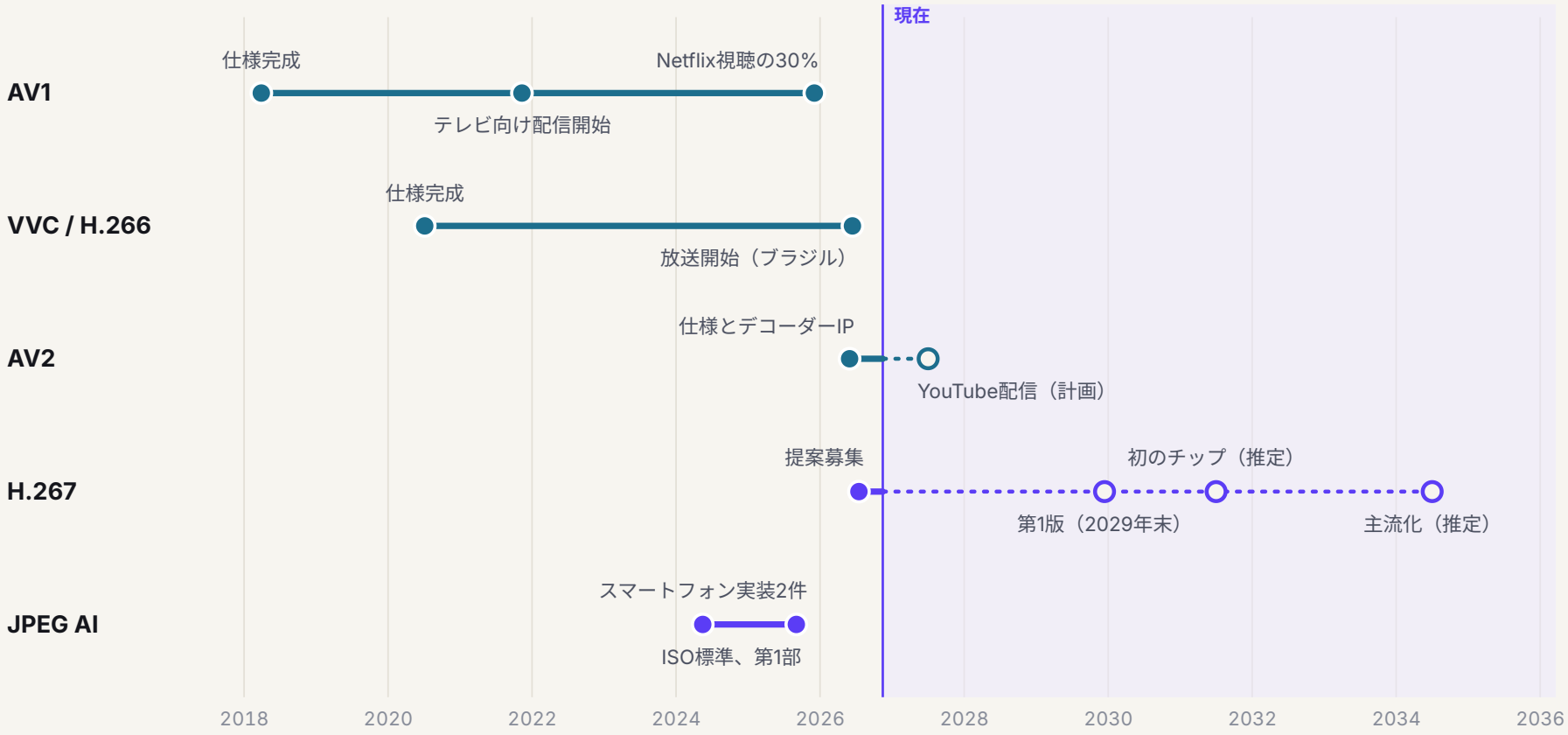
90%以上

AV1で視聴できるYouTubeカタログの割合（視聴時間ベース、2026年3月）

3週間

AV2バージョン1.0.0（2026年5月28日）から、3社がデコーダーIPの顧客を得るまで

標準の完成から一般の画面に届くまで、10年の大半を要する



AV1は仕様の完成（2018年）からNetflix視聴の30%（2025年12月）に達するまで、8年近くかかった。白抜きの印は目標または私の推定である。

ニューラルツールがあらゆる画面に届くかは、H.267が決める

2025年10月
エビデンス募集：16の企業・研究機関から5件の応募。ニューラルの効果は8K HDRで最も明確

2026年7月
ジュネーブで提案募集を発行

2027年1月
ブリスベンで応募提案を評価

2027年後半
最初のテストモデル

2029年末
標準の第1版

約28%
従来型のECM路線によるVVC比の削減

6–14%
ニューラルループ内フィルターによる削減（画素あたり1万から2万2000超のMAC）

SEI
VVCのストリームは、デコーダーが適用するニューラルポストフィルターをすでに指定できる

私の見立て：従来型のコアに、削減できるビットがデコーダーのコストに見合う少数のニューラルツールを加えたもの 分析 **。**

カメラの中はニューラル、ウェブは従来型

- **カメラ**

  QualcommのCognitive ISPは2023年からスマートフォンでリアルタイムのセグメンテーションを実行している。CES 2026では、NPU上で動く完全なAI ISPが披露された。

- **JPEG AI**

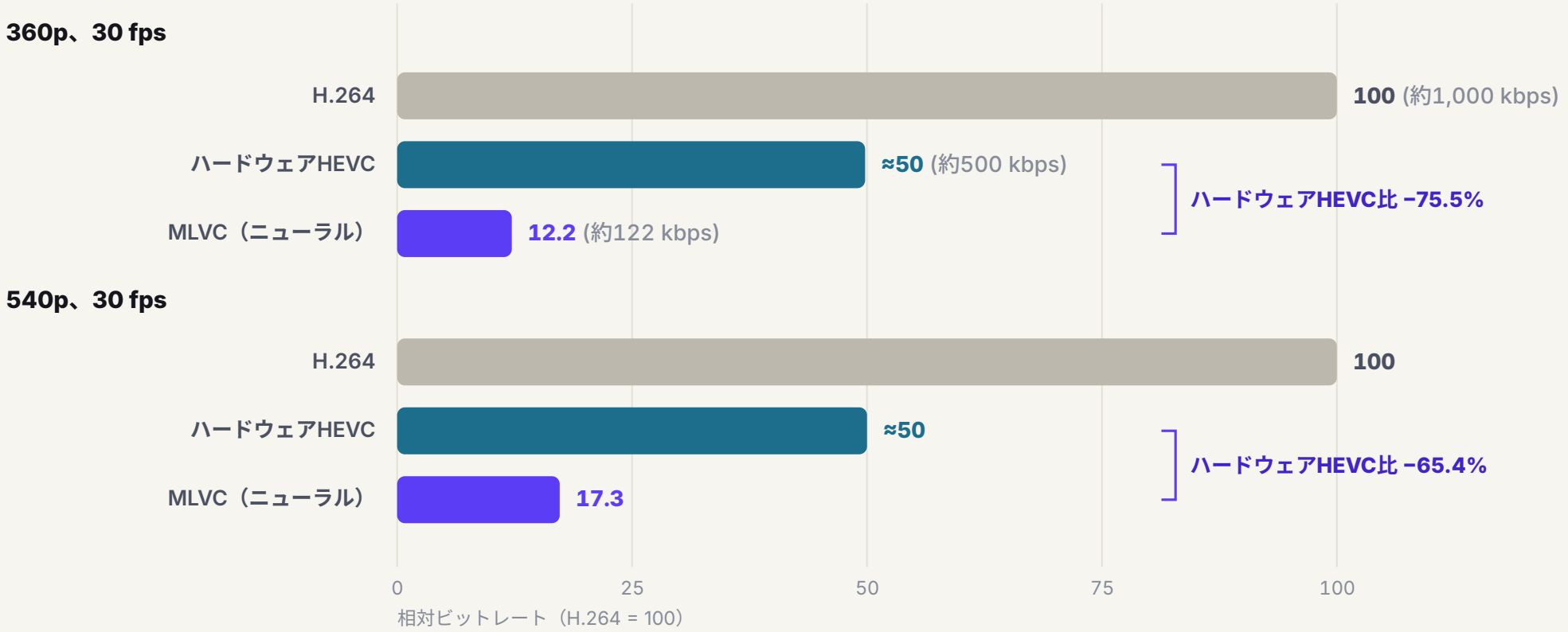
  2025年に発行。画素あたり7000から21万6000のMACで、VVCイントラ比12.5–27.9%の改善。私が調べた限り、デフォルトで使う製品はない。

- **ウェブ**

  Chrome 155は2026年10月6日、Rustで書かれたメモリ安全なデコーダーとともに、JPEG XLをデフォルトで有効にした。

> **どこでも開ける必要があるフォーマットは従来型のまま。ニューラル符号化が勝つのは、一社のパイプラインの内部である。**

エンドツーエンドのニューラルコーデックは通話で実用化



Satinは2021年から、TeamsとSkypeのすべての二者間通話で使われている。2026年7月24日にMITライセンスでオープンソース化されたMLVCは、Teamsのピアツーピア通話で最大540p30、NPUの半分以下の使用で展開中である ベンダー発表 。

どのプラットフォームもニューラルを選んだわけではない。理由は示唆に富む

標準の内部、オプション

Opus 1.5と1.6

ニューラルによる音質改善、パケットロス補償、深層冗長化、帯域拡張。いずれもデフォルトでは無効で、2012年の標準と完全な互換性がある。

あえて従来型

Meta MLow（2024年）

Metaの通話の20%以上はARMv7のスマートフォンで行われ、WhatsAppでは毎日数千万件の通話が10年以上前のスマートフォンで行われている。MLowは6 kbpsでOpusの約2倍の品質を、10%少ない計算量で実現する。

決め手は、既存端末が許容できる計算量である。ニューラルコーデックが最初に実用化されるのは、デバイスが新しく、一つのベンダーが両端を握る場所である。

誰もコーデックと呼ばない最大の導入例

- **どこにでもある学習型コーデック**
  あらゆる音声・動画モデルは、学習型コーデックのトークンや潜在表現を読み書きしている。
- **配信フォーマットではない**
  モデルごとに独自の潜在空間を持つ。レート制御も、適合性試験も、別のデコーダーもない。
- **コスト項目**
  生成動画を提供する事業者にとって、圧縮率の高い潜在表現は生成コストを比例以上に下げる 分析 。

二つの例

12.5 Hz · 1.1 kbps

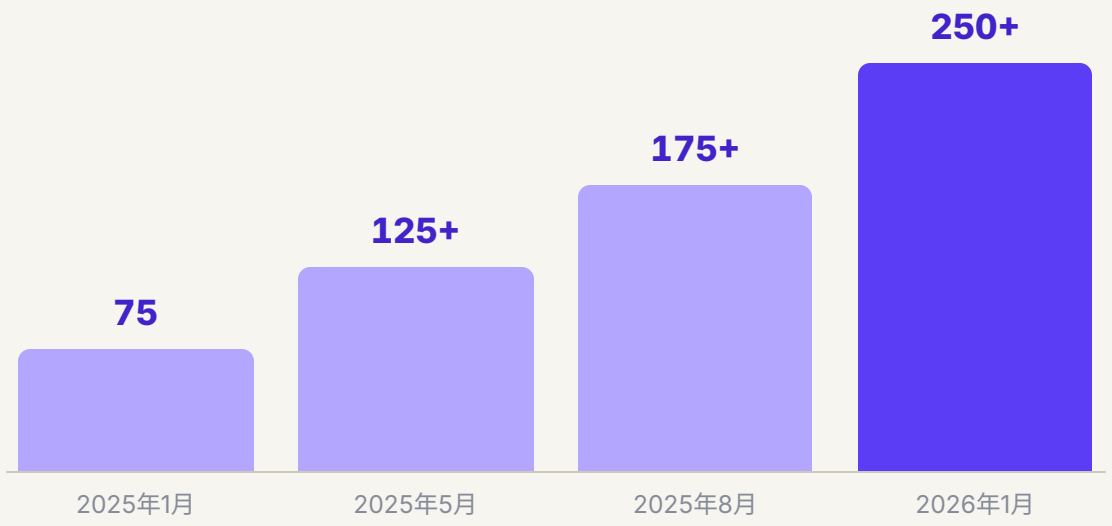
KyutaiのMimiの音声トークン。どの電話用コーデックよりも低い

1:192

LTX-Videoの動画VAE：8フレームにわたる32×32画素を一つの潜在トークンで表す

ニューラル再構成は主流に、ニューラルアセットは始まったばかり

DLSS 4マルチフレーム生成：対応ゲーム・アプリ数



ニューラルテクスチャ圧縮

NVIDIAのデモシーン、GTC 2026

テクスチャメモリ 6.5 GB → 0.97 GB

メモリ85%削減

Assassin's Creed Mirage（発売済み）

一部のアセットのみ。ブロック形式との比較

圧縮率が約30%向上

指標が異なるため、直接は比較できない

DirectX Linear Algebraは2026年4月にパブリックプレビューに入った。どのベンダーのGPUでも、シェーダー内のニューラル復号を一つのコードパスで実行できる。

衛星、ドローン、ロボット：最も適しているが、市場は小さい

- **適している理由**

  単一の所有者が両端を握り、帯域は乏しく、内容は学習で対応できるほど限定的である。

- **中国**

  11月18日施行の2026年第49号公告が、ドローンの輸出管理を改定する。その一項目は、セマンティック画像圧縮を基準値として用いている。

- **米国**

  FCCは2025年12月22日、外国製ドローンとその重要部品をCovered Listに加えた。

> **輸出管理に対応できる専門企業には堅実なニッチだが、主流の市場ではない。**

画素あたりのエネルギーはひと桁大きい

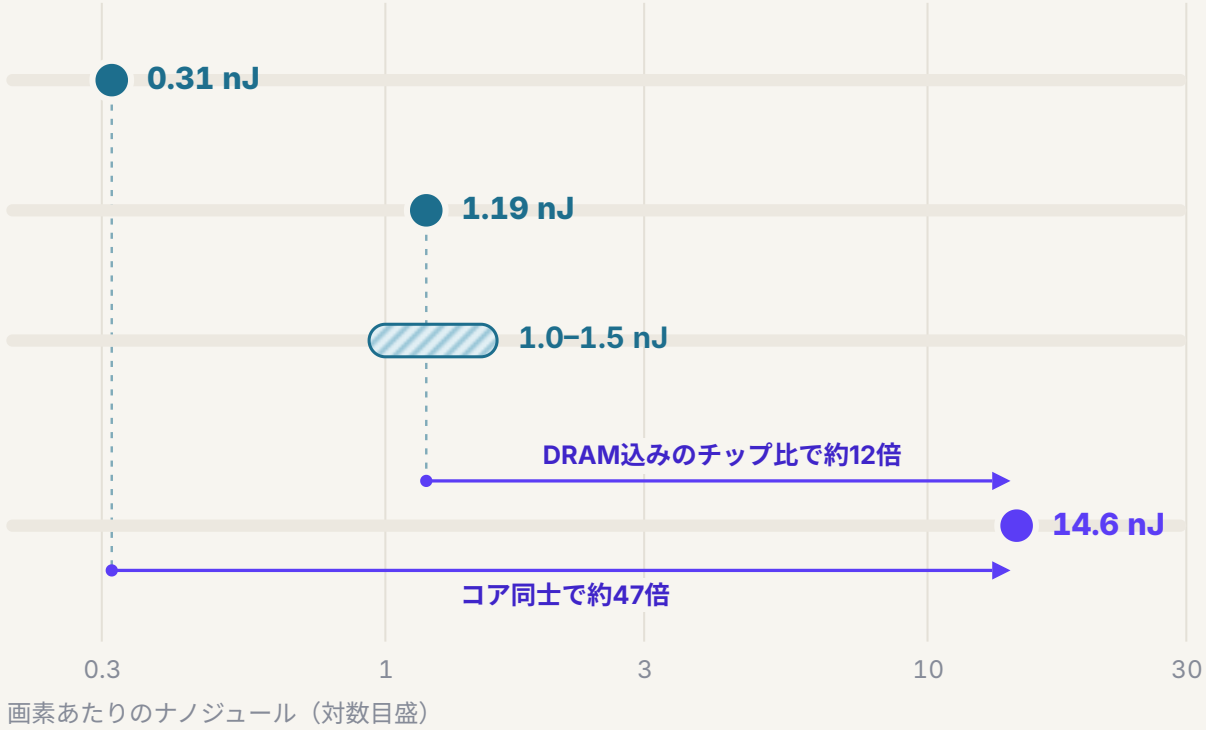
HEVCデコーダーチップ（コア）
40 nm、4K30、実測

HEVCデコーダーチップ（DRAM込み）
40 nm、4K30、コアは実測、DRAMはシミュレーション

VVCデコーダー
28 nm、4K60、予測

ニューラル映像デコーダー（NVCA）
28 nm、1080p25、論理合成のみ。重みとDRAMは含まない

1080p30に換算すると、約74 mW対908 mWである。



1080p30では、DRAM込みのHEVCチップの約74 mWに対し、NVCAはおよそ900 mW。HEVCチップは旧世代のプロセスノードで、NVCAは重みとDRAMを計上していない。この比較はニューラル設計に有利である。

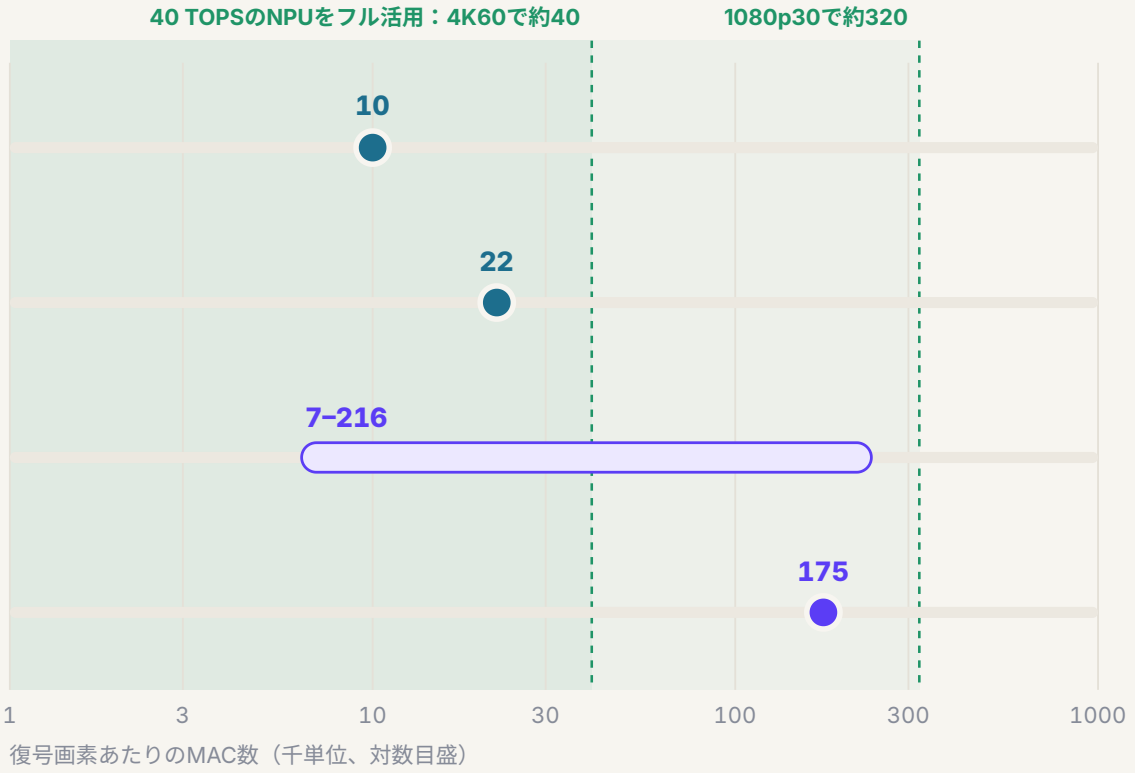
ニューラルデコーダーがNPUに求めるもの

ニューラルループ内フィルター（超低複雑度）
JVETの探索的検討 · 約6%のビット削減
1080p30で1.2 TOPS · 4K60で10

ニューラルループ内フィルター（低複雑度）
JVETの探索的検討 · 約8%
1080p30で2.7 TOPS · 4K60で22

JPEG AIデコーダー（ベースからハイまで）
VVCイントラ比12.5–27.9%

DCVC-RT（ニューラル映像デコーダー全体）
VVCのリファレンス比21%（論文の設定）
1080p30で22 TOPS · 4K60で174



DCVC-RTは画素あたり約17万5000 MACを要する。フル稼働で、1080p30なら約22 TOPS、4K60なら174 TOPSになる 推定 。40 TOPSのNPUをフルに使っても、4K60で賄えるのは約4万MACである。

メモリトラフィックとビット単位の一致こそが厳しい制約

DRAMが支配的

MACよりメモリ

MITのHEVCチップでは、DRAMがシステムのエネルギーの約4分の3を占める（画素あたり1.19 nJのうち0.88 nJ）。ニューラルデコーダーはフレーム間で大きな特徴マップを持ち回り、DCVC-RTの著者らは、速度を制約したのは演算ではなくメモリI/Oだったと結論づけた。

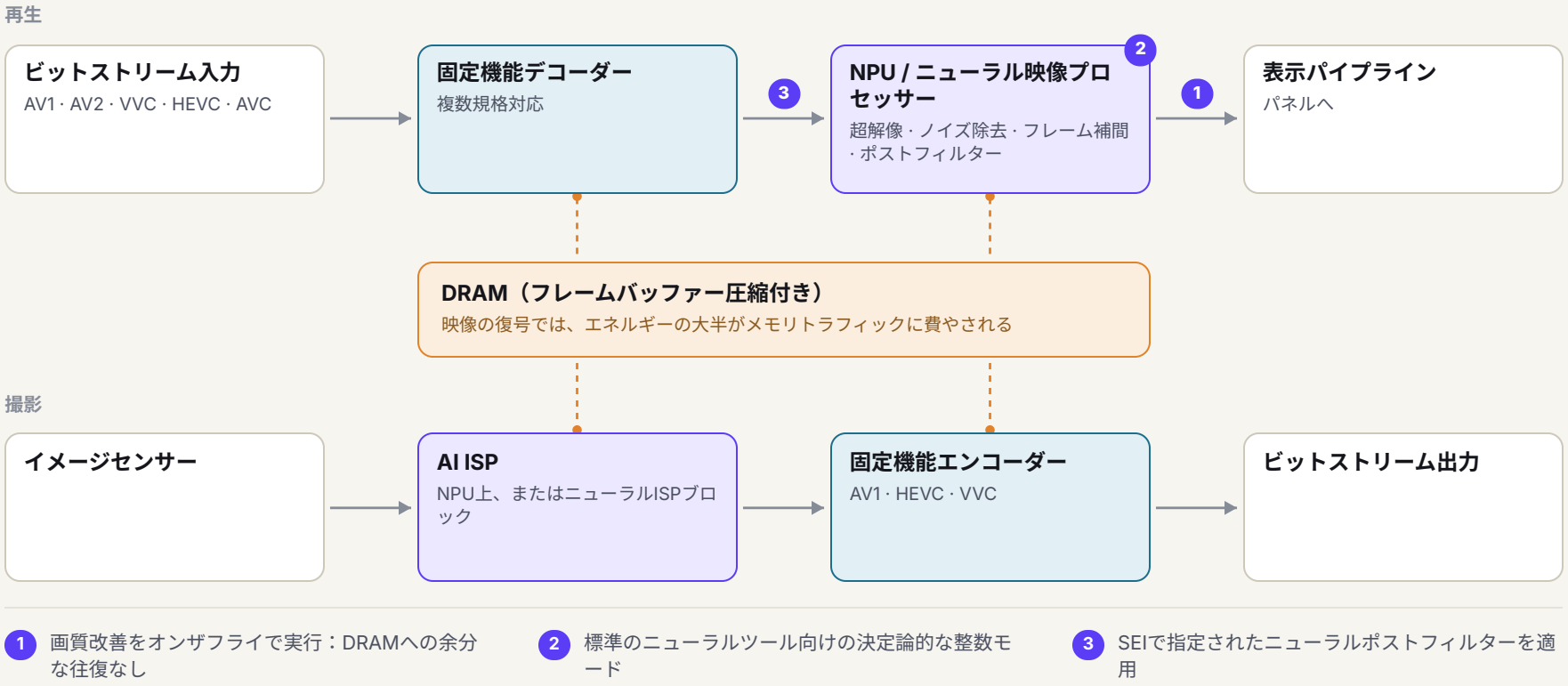
厳密性

全デコーダーが同じビットを計算しなければならない

二つのデバイスが確率を異なる値で計算すると、エントロピー復号は破綻する。対策は、整数化したモデル（DCVC-RT）、スケール共有型のハイパープライア（MLVC）、16ビット整数推論（JVETのSADL）、適合性の検証点（JPEG AI）。

チップ設計チームにとっては、HEVCの整数変換で求められた規律の再来である。標準が演算を厳密に定義しなければならない。

実用化されている分担：固定機能コーデックとNPU



1. 画質改善をオンザフライで実行：DRAMへの余分な往復なし
2. 標準のニューラルツール向けの決定論的な整数モード
3. SEIで指定されたニューラルポストフィルターを適用

> TencentのCanghai：第1世代チップは10万個以上が稼働し、第2世代はH.266とオンチップ超解像を加える。Chips&MediaとAllegro DVTは、コーデックIPとニューラル映像IPを並べて販売している。

私ならこうつくる：役割別

SoC・テレビ用チップのチーム

- 再生が必須のすべての標準に固定機能デコーダーを。AV2はいますぐ
- 表示解像度での画質改善に見合う規模のNPU。DRAMへの余分な往復なしで
- ニューラルポストフィルターのSEIと、決定論的な整数モード
- H.267の設計開始は2029年以降

IPベンダー

- AV2とVVCのデコーダーをいますぐ
- ポストフィルター用アクセラレーターとフレームバッファー圧縮
- ニューラル処理のための適合性検証ツール

動画プラットフォーム

- 最初に元が取れるのはエンコーダー側のニューラル前処理
- デバイスが許す範囲で、クライアント側の超解像
- エンドツーエンドのニューラルコーデックは、自社が握る閉じた系に限る

ニューラルツールがライセンス対象になる場

標準	ニューラルの要素	状況（2026年11月）
H.267（仮称）	未定：従来型（ECM）とニューラルツールがともに候補	応募提案の評価は2027年1月、第1版は2029年末
VVCとHEVC（VSEIを併用）	ニューラルポストフィルターをSEIで記述。適用はデコーダーの任意	有効。2025年10月に拡張
AV2	従来型、ロイヤルティフリー	バージョン1.0.0は2026年5月28日。YouTubeのウェブ配信は2027年初めに計画
JPEG AI	エンドツーエンドの学習型画像コーデック	第1部は2025年に発行。製品でのデフォルト対応は見つからず
IEEE 1857.11	NPU向けに設計されたニューラル画像符号化	2024年12月に発行
Video Coding for Machines	視聴ではなく、機械による解析のための符号化	2026年初めに国際標準案
OpusのML拡張	オプションかつ後方互換	libopus 1.5と1.6に搭載

645

ByteDanceが保有するVVC後継関連の特許ファミリー（GreyB、2025年）

0.50ドル · 1.00–1.50ドル

VVCのプール料率（スマートフォンと4Kテレビの台あたり、Region 1）

0ドル

AOMediaのメンバーが求めるAV1・AV2のロイヤルティ

お金は標準とプラットフォームに流れる

収益化済み

プラットフォーム

節約分を手元に残す

NetflixのニューラルダウンスケーラーYouTubeとMetaの超解像、MicrosoftはMLVCをMITライセンスでオープンソース化

収益化済み

デバイス・チップメーカー

差別化を売る

カメラのISP、テレビのアップスケール用プロセッサー、Tencentが自社開発したCanghaiコーデックチップ

収益化済み

IPベンダー

新しい標準ごとにライセンス

仕様完成から3週間以内に、Chips&Media、Allegro DVT、VeriSiliconがAV2デコーダーの顧客を獲得

収益化済み

GPUベンダー

プラットフォームで囲い込む

DLSS 4が250以上のゲームとアプリに対応。ニューラルテクスチャSDKはベータ版

デバイス単位

特許権者とパテントプール

デバイスごとに徴収

VVC：スマートフォン0.50ドル／台、4Kテレビ1.00–1.50ドル／台。H.267のツールは2029年以降

買収でイグジット

ニューラルコーデックのスタートアップ

買収された

WaveOne → Apple（2023年）、iSIZE → ソニー（2023年）、Deep Render → InterDigital（2025年）

> 最もよく知られたニューラルコーデックのスタートアップ3社、WaveOne、iSIZE、Deep Renderは、それぞれApple、ソニー、InterDigitalに買収された。

六つの事業機会と、それぞれのリスク

事業機会	買い手	なぜ今か	主なリスク
エンコーダー側のニューラル前処理	ストリーミング・動画投稿プラットフォーム、放送事業者	デバイスの変更が不要。Netflixで実証済み	大手プラットフォームは自社で構築する
表示側の画質改善IPとチューニング	テレビ、STB、スマートフォンのチップメーカー	NPUはすでにデバイスに載っている	コモディティ化。社内チーム
閉じた系のニューラルコーデック	企業向けコミュニケーション、産業用、テレプレゼンスのベンダー	MLVCが機能することを証明し、コードも公開されている	両端を握る必要がある。NPUの断片化
コーデックとNPUのシリコンまたはIP	SoCメーカー	AV2のサイクルが始まりつつある	強力な既存勢力
評価、適合性検証、決定論的推論	ニューラルツールを出荷する全企業	ビット単位の一致は絶対条件	単体では市場が小さい
H.267への標準化寄与	ライセンサー、大手プラットフォーム	応募提案の評価は2027年1月	長く、費用がかさみ、不確実

四つの点を示すまで、主張は割り引いて聞く

主張	証拠が示すこと
「ニューラルコーデックはすでにVVCに勝る」	研究室では。DCVC-RTはVVCのリファレンスエンコーダー比で21%削減するが、デコーダーは画素あたり約17万5000 MACを要する。
「AIコーデックは数年以内にH.26xとAV1を置き換える」	再生ではそうならない。H.267第1版は2029年末、チップは2031年ごろ 推定 。いま登場しつつあるコーデックであるAV2は従来型である。
「MLVCは帯域を88%削減する」	360pでのH.264比である。ハードウェアHEVC比では、Microsoft自身のテストセットで360pが75.5%、540pが65.4%。
「ニューラルテクスチャはビデオメモリを85–96%削減する」	デモシーンでの話である。発売済みのゲームでは、一部のアセットに限って約30%の改善だった。
「JPEG AIはISO標準なので、カメラやブラウザーが採用する」	製品でのデフォルト対応は見つからない。ウェブが採用したのは、従来型フォーマットのJPEG XLである。

比較の基準、指標、解像度、復号コストを確かめること。この四つをすべて示せない主張は、製品ではなく研究成果である。

注目点

注目点と未解決の問い

注目すべき節目

2027年1月	H.267の応募提案を評価：ニューラルツールに頼る提案はいくつか、画素あたりのコストはどれだけか
2027年初め	YouTube初のAV2ウェブ配信。発表済みのSoCへのAV2デコーダー搭載
2027年後半	H.267の最初のテストモデル
2027年	Teamsのピアツーピア通話以外へのMLVCの広がり
2027–2028年	JPEG AIをデフォルトで使う最初の製品
2029年末	H.267第1版で、チップ開発サイクルが始まる

未解決の問い

01 大規模での厳密性：ベンダーの異なるNPU間でビット単位で一致する整数復号を、どれだけの圧縮率の損失で実現できるか

02 MACよりメモリ：ニューラル後処理の画素あたりDRAMトラフィック

03 H.267でニューラルツールが元を取れなくなる、画素あたりのコスト

04 通話以外のどの閉じた系が、最初にエンドツーエンドのニューラルコーデックを採用するか

05 デコーダーが細部を生成する場合、忠実度をどう測るか

これらは、技術交流と技術協力を通じて、ともに探究していける領域である。

i@tanshuai.com · tanshuai.com