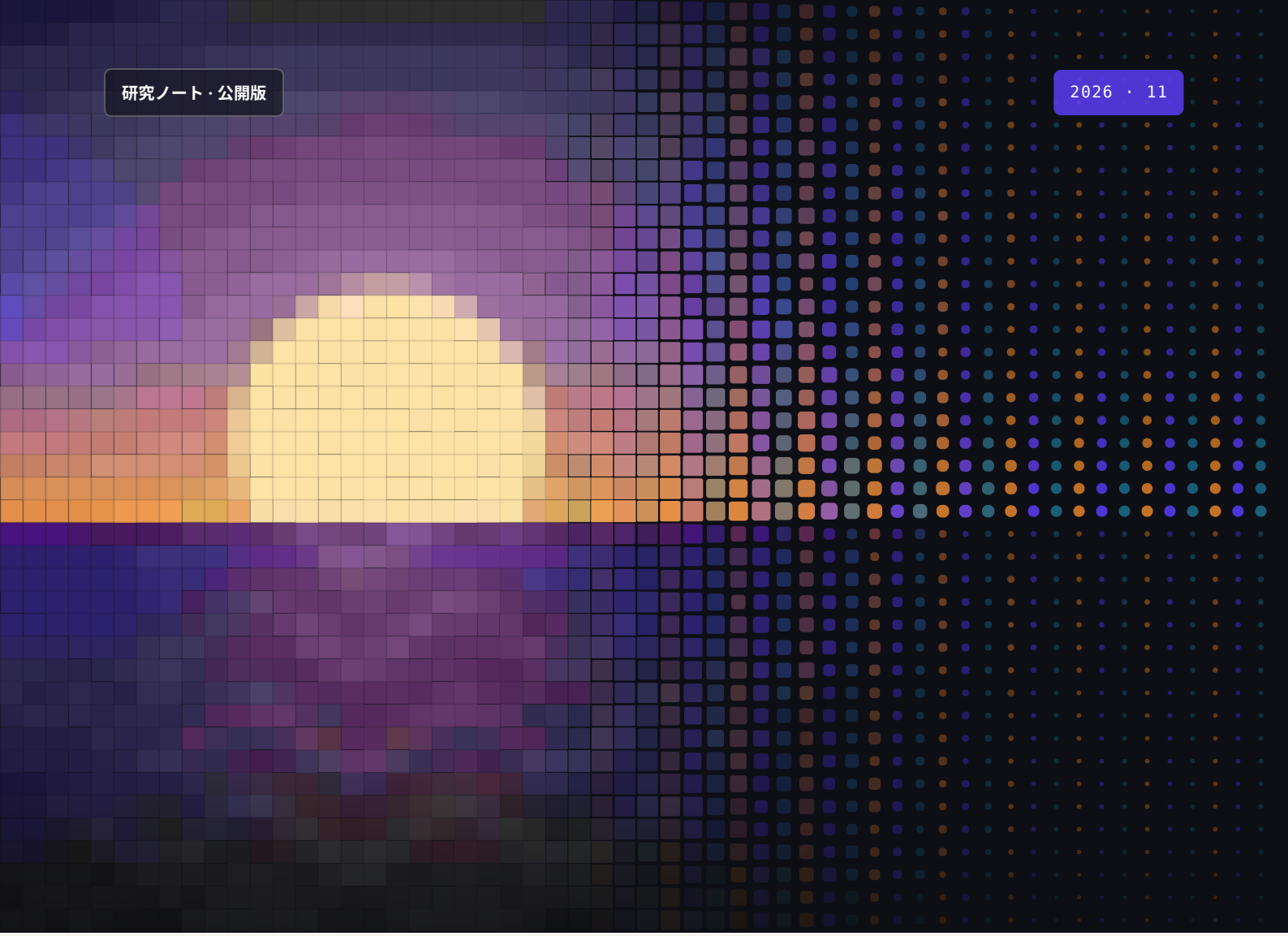


研究ノート · 公開版

2026 · 11

学習型圧縮 · 再生 × シリコン

2026年のニューラル符号化

コーデックの代わりではなく、その前後で実用化が進む

再生 · 写真 · 通話 · 生成AI · ゲーム · シリコン

結論

世界の映像を運ぶコーデックの両端では、すでにニューラルネットワークが働いている。エンドツーエンドのニューラルコーデックが実用化されているのは閉じた系だけで、一般の画面はH.267を待っている。

タン・シュアイ

テクノロジスト、創業者 · AI研究開発者

tanshuai.com · i@tanshuai.com

データは2026年11月12日時点
金額は米ドル

目次

目次

◆ **エグゼクティブサマリー** 04
すでに主流、ただしコーデックの前後で

01 **対象と方法** 07
本稿でいう「ニューラル符号化」と、私の調べ方

02 **再生とストリーミング** 09
再生とストリーミング：両端はニューラル、中間は標準

03 **写真と画像** 12
写真と画像：カメラの中はニューラル、ウェブは従来型

04 **通話と会議** 13
通話と会議：エンドツーエンドのニューラルコーデックが実際に実用化されている場

05 **生成AI** 15
生成AI：誰もコーデックと呼ばない最大の導入例

06 **ゲームとグラフィックス** 16
ゲームとグラフィックス：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

07 **特殊用途リンク** 18
特殊用途リンク：衛星、ドローン、ロボット

08 **シリコン** 18
シリコンの視点：エネルギー、演算量、メモリ、厳密性

09 **標準と特許** 22
標準と特許：ニューラルツールがライセンス対象になる場

10 **誰が収益を得るか** 24
誰が収益を得るか、企業はどこで勝負できるか

11 **主張と根拠** 26
よく聞く主張と、証拠が示すこと

12 **注目すべき節目** 27
2027年から2030年の注目点

13 **未解決の問い** 28

◆ **著者について** 29
タン・シュアイ

◆ **出典** 30

図一覧

図一覧



図01	2026年、ニューラル符号化はどこで実用化されているか	06
図02	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	08
図03	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	10
図04	通話では、ニューラルコーデックがハードウェアHEVC比でビットレートを65–76%削減	14
図05	一社が技術スタック全体を握ると、ニューラル再構成は急速に普及する	16
図06	復号画素あたりのエネルギー：固定機能デコーダーがひと桁優位	19
図07	ニューラルツールは画素あたり数万MAC、ニューラルデコーダー全体では数十万MAC	20
図08	実用化されている分担：ビットストリームは固定機能ハードウェア、画素はニューラルエンジン	21
図09	2026年、ニューラル符号化で収益を得るのは誰か	24

本文中のタグ

- 推定 — 引用データに基づく著者の計算または推論
- ベンダー発表 — 製品を販売する企業が公表したもので、第三者による検証を経ていない
- 分析 — 証拠に対する著者の解釈

本書は英語版の翻訳であり、両者に相違がある場合は英語版が優先する。

データは2026年11月12日時点 · 金額は米ドル

エグゼクティブサマリー

すでに主流、ただしコーデックの前後で

結論

2026年、ニューラル符号化はすでに主流だが、標準コーデックの代わりではなく、その前後で働いている

約75%
2025年末のモバイルデータトラフィックに占める動画の割合（Ericsson）

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する

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。実測HEVCデコーダーチップ（DRAM込み）との比較

約22 TOPS
DCVC-RT級のニューラルデコーダーが1080p30で必要とする持続演算性能 推定

75.5%
通話の360pで、MLVCがハードウェアHEVCより少なく済むビットレート。Microsoft自身のテストセット ベンダー発表

2029年末
H.267第1版。ニューラルツールをあらゆる画面に届ける可能性が高い最初の標準

実用化の場

。世界の映像のほぼすべてを運ぶコーデックの両端で、いまやニューラルネットワークが動いている。入口では映像を整える。スマートフォンのカメラに載るAI画像信号処理プロセッサーや、Netflixの符号化パイプラインにあるニューラルダウンスケーラーである。出口では映像を修復する。テレビ向けのYouTubeのSuper Resolution、Metaの超解像、数百のニューラルネットワークをうたうテレビ用プロセッサーである。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。ニューラル映像フォーマットをテレビやスマートフォンに配信しているストリーミングサービスはない。今日デバイス上でニューラル映像デコーダーが動いているとすれば、それは一社が自ら提供するアプリの内部である。

閉じた系

エンコーダーとデコーダーの両方を学習で得るエンドツーエンドのニューラルコーデックは、閉じた系でしか実用化されていない。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環境である。Microsoftは2021年から、TeamsとSkypeのすべての二者間通話でAI音声コーデックを使っている。さらに2026年7月にはニューラル映像コーデックMLVCをオープンソース化し、Teamsのピアツーピア通話で展開を進めている。360pでは、同じ評価品質を得るのに必要なビットレートがハードウェアHEVCより約75%少ない。音声や動画を生成するモデルはすべて、学習型コーデックの上で動いている。12.5 Hzの音声トークナイザーから、圧縮率1:192の動画VAEまでである。ニューラルテクスチャ圧縮は、AAAゲーム1作に搭載された。これらは実際の導入例であり、一つの条件を共有している。そのストリームを、ほかの誰も復号する必要がないという条件である。

時期

ニューラルビットストリームの一般向け再生は、標準とチップ開発サイクルを待つ。ニューラルツールを載せる可能性が高い最初の標準はVVCの後継で、私はこれをH.267と呼ぶ。2026年7月に提案募

集を発行し、2027年1月に応募提案を評価して、2029年末に第1版の完成を目指している。AV1は、仕様の完成からNetflix視聴の30%に達するまで8年近くを要した。同じパターンをたどるなら、最初のH.267チップは2031年ごろに登場し、主流化は2030年代半ばになる 推定 。その間に次に画面に届くコーデックはAV2である。2026年5月に完成したAV2にニューラルデコーダーは含まれておらず、IPベンダー3社が3週間以内にAV2デコーダーの顧客を発表した。

シリコン

理由はシリコンにある。唯一公開されているニューラル映像デコーダーのチップ設計は、画素あたり14.6 nJを要する。実測されたHEVCデコーダーチップのDRAM込みの値の約12倍、コア単体の約47倍である。Microsoftのリアルタイム研究用コーデックDCVC-RTは、論文自身のモジュール別の表を私が集計すると、復号画素あたり約17万5000回の積和演算を要する。1080p30では、およそ22 TOPSの持続演算性能に相当する 推定 。AI PCを定義する40 TOPSの約半分を、映像ストリーム1本に費やすことになる。ニューラルデコーダーは、どのデバイスでもビット単位で同一の結果を出さなければならない。そうでなければ、エントロピー復号が破綻する。MLVCが圧縮率を犠牲にしてNPU間で同一の復号を実現し、4Kではなく540pで動作しているのは、そのためである。

ビジネス

ビジネスの観点では、お金は単体のニューラルコーデックではなく、標準とプラットフォームに流れる。プラットフォームは、ニューラルな前処理と後処理による節約分を手元に残す。IPベンダーは、新しい標準ごとに固定機能デコーダーを、さらに画像処理向けのNPUをライセンスする。パテントプールはデバイスごとに徴収する。VVCの場合、米国などRegion 1の市場で、スマートフォン1台あたり0.50ドル、4Kテレビ1台あたり1.00–1.50ドルである。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。開かれた機会は、ニューラルな前処理と後処理、閉じた系の製品、固定機能コーデックとNPUを組み合わせたシリコン、そしてその間にある評価と標準化の仕事にある。

問い	判断	主要な数値（出典は各章）
① **ニューラル符号化は、今日どこで大規模に実用化されているか？**	コーデックの前後：カメラのISP、符号化側のダウンスケール、再生側の超解像。エンドツーエンドは閉じた系だけ：通話、ゲームのアセット、生成AIのトークナイザー	動画はモバイルデータの約75%。2025年に出荷されたスマートフォン12億6000万台とテレビ約2億台は、すべて固定機能ハードウェアで復号する。DLSS 4は250以上のゲームとアプリに対応
② **ニューラルコーデックは、再生でAV1、HEVC、VVCを置き換えるか？　それはいつか？**	標準とデバイスの更新サイクルを経るまでは置き換えない。現実的な最初の担い手はH.267で、従来型とニューラルのツールが混在する可能性が高い	H.267の提案募集は2026年7月、評価は2027年1月、第1版は2029年末。主流化は2030年代半ば 推定
③ **シリコンでのコストは？**	現在の固定機能による復号と比べて画素あたりのエネルギーがひと桁大きく、さらにメモリトラフィックとビット単位の一致という要件が加わる	画素あたり14.6 nJ（ニューラル）対1.19 nJ（HEVC、DRAM込み）。復号画素あたり約175 kMAC、1080p30で約22 TOPSに相当 推定
④ **標準と特許は何を許すか？**	ニューラルポストフィルターは、すでにVVCのストリームで指定できる。JPEG AIはISO標準である。どのニューラルツールがライセンス対象になるかは、H.267が決める	ECM（従来型）はVVC比約28%削減。ニューラルループ内フィルターは10–22+ kMAC／画素で6–14%。VVC後継の分野でByteDanceは645の特許ファミリーを保有
⑤ **誰が収益を得て、企業はどこで勝負できるか？**	プラットフォーム、IPベンダー、特許権者。純粋なニューラルコーデックのスタートアップは買収された。前処理と後処理、閉じた系、コーデックとNPUを組み合わせたシリコン、評価で勝負する	VVCのプール料率はスマートフォン1台あたり0.50ドル。AV2デコーダーIPは仕様完成から数週間でライセンス。買収は3件、条件は非公開

問い	判断	主要な数値（出典は各章）
⑥ どの主張を割り引いて聞くべきか？	古いコーデックとの比較、低解像度での値、復号コストを伴わない値、発売済みの製品ではなくデモの値として示された改善	MLVC：H.264比では88%減だが、ハードウェアHEVC比では65–76%減。テクスチャ：デモではVRAMを85%削減、発売済みのゲームでは約30%

図01 2026年、ニューラル符号化はどこで実用化されているか

主流のコーデックの両端では、すでにニューラルネットワークが動いている。エンドツーエンドのニューラルコーデックが実用化されているのは、エンコーダーとデコーダーの両方を一社が握る場合だけである。

	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と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	—

出典: 本レポートで引用した企業、標準化団体、開発者向けの文書。状況は2026年11月12日時点。行は各領域が届く人数の多い順に並べた。

第01章

01 対象と方法
本稿でいう「ニューラル符号化」と、私の調べ方

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

「ニューラル符号化」という言葉は三つの異なるものを指しており、これをめぐる議論の多くは、それらを混同することから生じている。

01 **コーデックの前後**。ネットワークは符号化の前か復号の後で画素を処理し、ビットストリームは標準のままである。例としては、カメラのAI画像信号処理プロセッサー（ISP）、エンコーダーの前段に置かれたNetflixのニューラルダウンスケーラー、テレビ上の超解像がある。ストリームはどのデバイスでも再生でき、ネットワークが変えるのは入力か出力だけである。

02 **標準の内部**。ネットワークが標準コーデックのツールになる。たとえば、すべてのデコーダーが実行しなければならないニューラルループ内フィルターやイントラ予測器であり、厳密に、通常は整数演算で仕様化する必要がある。より軽い形態として、ビットストリームがメタデータで記述し、デコーダーが適用するかどうかを選べるニューラルポストフィルターがある。

03 **コーデックの代わり**。エンコーダーとデコーダーの両方を、エンドツーエンドで学習する。ビットストリームは量子化された潜在値の集合で、対応するネットワークでしか復号できない。例としては、TeamsでのMicrosoftのMLVC、音声モデルの音声トークナイザー、ゲームのニューラルテクスチャがある。

図02 **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**

2026年に実用化されているものの大半は、標準コーデックの前後で働く。標準の内部に入るニューラルツールは仕様策定中である。エンドツーエンドのニューラルコーデックがコーデックに取って代わるのは、一社が両端を握る場合だけである。



出典: 本レポートの出典を著者がまとめたもの。状況は2026年11月12日時点。

一つ目は誰の許可も要らない。二つ目には標準が要る。三つ目には両端を握ることが要る。人々に届く順番もこのとおりであり、以下の内容の大半はこれで説明できる。

研究の関心ではなく、現実の規模で並べる

研究論文が主に扱うのは三つ目である。圧縮性能の記録が生まれるのがそこだからである。本レポートでは、お金とエンジニアリングの行き先を決める尺度として、各領域が届く人数の多い順に並べた。再生とストリーミングを最初に置き、写真、通話、生成AI、ゲームと続け、特殊用途リンク（衛星、ドローン、ロボット）を最後にした。Ericssonの集計によれば、2025年末の時点で、動画はモバイルデータトラフィック全体の約75%を占める。2025年に業界はスマートフォンを約12億6000万台（IDC）、テレビをおよそ1億9600万–2億900万台（TrendForceとOmdiaの推定）出荷した。そのすべてが、固定機能ハードウェアで動画を復号する。

調査の方法

これはデスクリサーチであり、チップや製品計画をレビューするのと同じやり方で進めた。標準文書と会合報告（JVET、MPEG、JPEG、AOMedia、IETF）、企業のエンジニアリングブログとリリースノート、チップの論文、パテントプールの料率表、輸出管理の条文、報道を読んだ。すべての数値には日付を付けた。論文の表から画素あたりの演算量を求める場合のように、私が自分で数値を導いたときはその旨を記し、 推定 と表示した。一次資料と照合できなかった主張には、 ベンダー発表 と表示した。金額は米ドルである。状況は2026年11月12日時点のものである。

可能な限り同じ条件どうしで比較し、それができない場合はその旨を記した。圧縮性能の改善は、四つの情報を伴わなければほとんど意味を持たない。比較の基準（何に対する改善か）、指標（PSNR、VMAF、あるいは視聴者パネル）、コンテンツと解像度、そして復号にかかるコストである。この四つを明示すれば、ニューラルコーデックをめぐる意見の相違の大半は消える。

第02章

02

再生とストリーミング

再生とストリーミング：両端はニューラル、中間は標準

視聴者が受け取るのは、今も標準のビットストリーム

最大手の動画サービスはコーデックの移行を一つ終え、次の移行を始めているが、どちらにもニューラルデコーダーは関わっていない。Netflixは2025年12月、AV1が視聴の約30%を担い、AVCに次いでよく使われるコーデックであると報告した。AV1がテレビに届いたのは、ハードウェアデコーダーが届いてからだった。Netflixは2019年にAV1をデバイス認証に加え、2021年後半からテレビ向けに配信した。同社によれば、2021年から2025年の間に認証に提出された大画面デバイスの88%がAV1に対応しており、そのほぼすべてが2023年以降のものである。YouTubeは、2026年3月19日時点で、視聴時間で重み付けしたカタログの90%以上がAV1で視聴できるとした。

Alliance for Open MediaによるAV1の後継AV2は、2026年5月28日にバージョン1.0.0に達した。AOMediaの客観評価では、圧縮効率はAV1より約30%高い。YouTubeは、ウェブでのAV2配信を「早ければ2027年初め」に始める計画である。シリコンもすぐに追随した。Chips&Mediaは5月にAV2ハードウェアデコーダーを完成させ、6月に北米の大手テクノロジー企業（社名非公表）にライセンスした。Allegro DVTは6月16日までにAV2デコーダーIPを3社の顧客に納入し、VeriSiliconは6月10日までにAV2対応の映像IPを複数の顧客にリリースしていた。AV2は従来型のコーデックである。改善は、より優れたブロック分割、予測、変換、ループ内フィルター（その一つは「データ駆動型」のガイド付きディテールフィルター）から生まれており、エンドツーエンドのネットワークによるものではない。

現行のITU/ISO標準であるVVCは、最初の放送の場をブラジルに得た。同国のDTV+（TV 3.0）は、ワールドカップに合わせて2026年6月に商用運用に移行した。2026年7月時点で、ブラジルで販売されているテレビにDTV+受信機を内蔵したものはなく、最初の製品は2026年後半から2027年初めにかけて登場する見込みである。

エンコーダー側：Netflixのニューラルダウンスケーラー

再生におけるニューラル符号化の最もわかりやすい例は2022年から稼働しており、視聴者のデバイスがすべきことは何も変えていない。Netflixは各タイトルを複数の解像度で符号化する。4Kのマスターから1080p、720p、540pを作るダウンスケールの工程には、従来はLanczosのような固定フィルターが使われていた。Netflixはこれを小さなニューラルネットワークに置き換えた。前処理ブロックとリサイズブロックからなり、デバイス側の標準的なアップスケーラーを通した後の画質が向上するように学習させたものである。Netflixはこれをドロップイン型の解決策と説明している。符号化側にもクライアント側にもほかの変更はなく、「Netflixのストリーミングに対応する何百万台ものデバイスが自動的に恩恵を受ける」。Netflixの選好テストでは、視聴者の約77%がニューラルダウンスケーラーの出力を選んだ ベンダー発表。Netflixの規模でコストを許容範囲に収めるには、「ニューラルネットワークの層は数層で十分だった」という。

覚えておくべきパターンはこれである。ネットワークはデータセンターで、タイトルごとに一度だけ動く。そこでは計算資源が安く、実行のタイミングも調整できる。そして出力は、通常のAV1やHEVCのストリームである。

画面側：アプリとテレビの中の超解像

もう一方の端では、コーデックと元の素材が失ったものをネットワークが修復する。YouTubeは2025年10月29日、テレビアプリ向けにSuper Resolutionを発表した。1080p未満（240pから720p）でアップロードされた動画を機械学習でHD

にアップスケールするもので、4Kを目標に掲げている。アップスケールされた版にはラベルが付き、視聴者は元の動画に戻すことができ、クリエイターは対象から外れることを選べる。Metaは2025年10月、アップロード時にサーバー上で（Intelと協力し、標準的なx86 CPUで）、また再生時にクライアント上で動画の超解像を実行する方法を説明し、両者が補完し合うことを確認した。テレビメーカーは同じことをシリコンで行う。Samsungのフラッグシップ8Kプロセッサー NQ8 AI Gen3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。

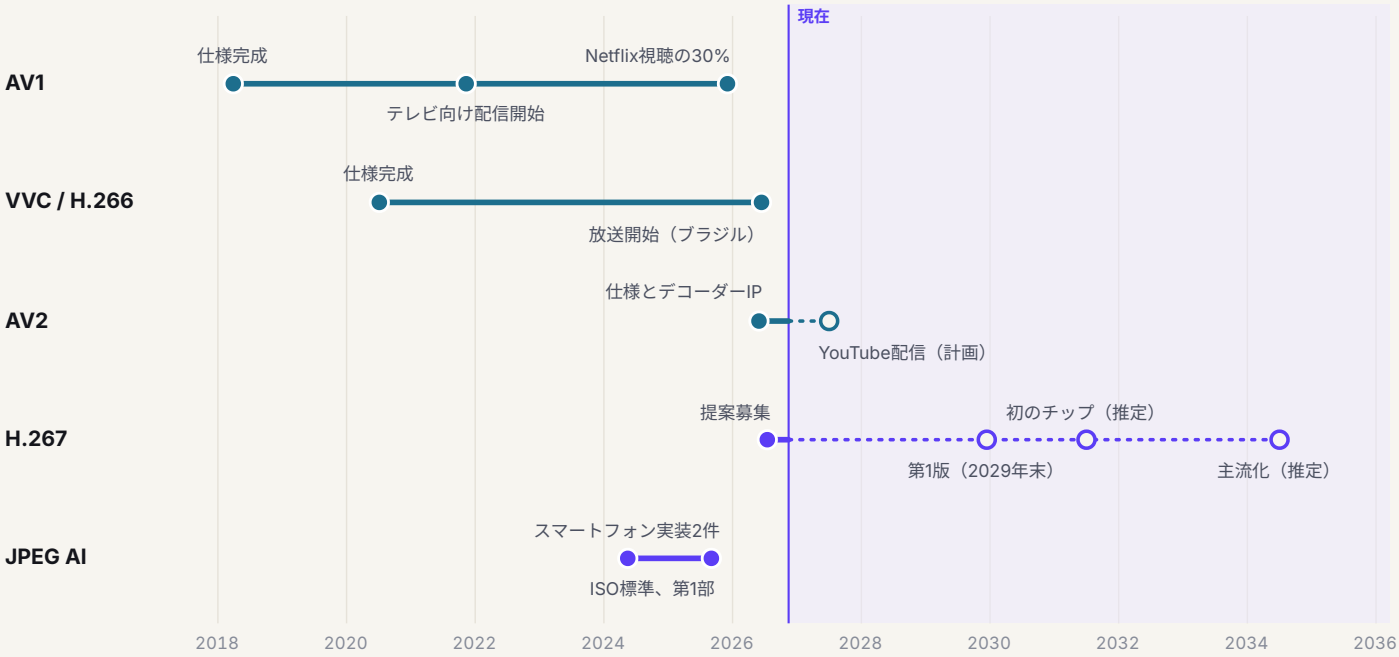
経済性は、ネットワークをどこで動かすかによって異なる。サーバー側の処理では、プラットフォームは扱う動画のすべてについて計算コストを負担する。デバイス側の処理ではプラットフォームの負担はなく、購入者がすでに代金を払ったNPUを使う。2026年に最大の効果が両方の組み合わせから生まれているのは、そのためである。

なぜニューラルビットストリームを復号するテレビがないのか

再生フォーマットは、それを再生しうるすべてのデバイスで復号できなければならず、そうしたデバイスは何年も使われる。そのため、標準の完成から一般の画面に届くまでの道のりは、従来型のコーデックでさえ長い。AV1の仕様は2018年3月に完成した。数カ月のうちにソフトウェアデコーダーdav1dが続き、2021年後半にはテレビ向けの配信が始まった。そして仕様の完成から8年近く経った2025年12月に、Netflix視聴の約30%に達した。VVCは2020年7月に完成し、2026年6月に最初の放送サービスに採用された。

図03 **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**

再生にはデバイス内のハードウェアデコーダーが必要である。AV1は仕様の完成からNetflix視聴の30%に達するまで8年近くかかった。ニューラルツールを載せる可能性が高い最初の標準は、2029年末ごろに完成予定である。



出典: Netflix Tech Blog（2025年12月）、AOMedia、Allegro DVT、TV Technology、Fraunhofer IIS（ブラジルのDTV+について）、MPEG-155、JPEG委員会、ISO。白抜きの印は目標または著者の推定。

ニューラルビットストリームには、従来型のコーデックがとうに解決した二つの問題が加わる。画素あたりのエネルギーと、ビット単位で一致する復号である。どちらも第8章で扱う。この二つがあるため、何をつくるべきかを標準が厳密に示すまでは、テレビメーカーもスマートフォンメーカーも、すべてのデバイスにニューラル映像デコーダーを載せることはない。

状況を変えうるもの：H.267

ニューラルツールをあらゆる画面に届けうる標準は、ITU-TとISO/IECがJoint Video Experts Team（JVET）で共同開発しているVVCの後継である。名称はまだ確定していない。業界の多くにならい、私はこれをH.267と呼ぶ。

JVETは2025年10月、エビデンス募集への応募を評価した。16の企業と研究機関から5件の応募があり、20のコーデック構成が試験された。委員会はVVCを上回る圧縮性能の証拠を認め、改良された従来技術とニューラルネットワーク技術の両方を「興味深い」と評価した。ニューラル方式は、とくに8K HDRのコンテンツで知覚上の明確な改善を示し、ゲームのコンテンツでは改善が小さかった。応募者の一つであるInterDigitalは、従来型のEnhanced Compression Model（ECM）に基づく設計と、VVCにAIツールを重ねた設計を一つずつ提出した。

正式な提案募集は、2026年7月のジュネーブ会合で発行された。主な提出物の締め切りは2026年10月末、補足パッケージの締め切りは12月中旬で、応募提案は2027年1月のブリスベンでのJVET会合で評価される。JVETは、2027年後半に最初のテストモデル、2029年末までに標準の第1版を見込んでおり、この取り組みを「新しい標準を立ち上げるための、おそらくこれまでで最大の取り組み」と表現している。

従来型の路線は、高いハードルを設定している。従来の信号処理で構築されたECMは、ランダムアクセス構成でVVC比およそ28%の削減を達成している。JVETのニューラルネットワーク探索的検討は、画素あたり1万回の積和演算で約6%、2万2000回で8%、高複雑度の動作点で14%を削減するループ内フィルターを報告している。私の見立てでは、H.267は、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になる可能性が最も高い。それがいくつになるかは、2027年1月の評価で明らかになるだろう 分析 。

橋渡しは、すでにある。VVCとHEVCのストリームは、デコーダー側で適用するニューラルポストフィルターを記述した補足拡張情報（SEI）を運ぶことができる。2025年10月に完成したVVCとそのSEI標準（VSEI）の新版は、これらのメッセージを拡張した。ストリームは、すべてのデコーダーに実行を強いることなく、復号後にどのネットワークを動かすべきかを示せる。

再生分野のビジネスにとっての意味

いま売るべきは、コーデックの前後で機能するものである。サービス事業者向けの符号化側の前処理と、テレビ用およびセットトップボックス用チップ向けの後処理ブロックとチューニングである。2030年代より前に、他社のデバイスで復号しなければならないニューラルビットストリームの上に再生ビジネスを築くべきではない。シリコンをつくるなら、当面の収益源はAV2とVVCのデコーダーであり、ニューラルの部分はその横に置く画質改善用のNPU経路である（第8章）。

第03章

03 写真と画像：カメラの中はニューラル、ウェブは従来型

写真と画像

カメラのパイプラインは、すでにニューラル

最近のフラッグシップスマートフォンで撮った写真は、圧縮される前にすべてニューラルネットワークを通っている。2023年のスマートフォンに搭載されたQualcommのSnapdragon 8 Gen 2は、「Cognitive ISP」を導入した。写真や動画の撮影中にリアルタイムでセマンティックセグメンテーションを実行し、顔、髪、空、背景を個別に調整する。SamsungのGalaxy S23 Ultraは、これを自撮りカメラに使った最初のスマートフォンである。2026年1月、Chips&MediaとVisionary.aiはCESで、初の完全なAIベースISPと称するものを披露した。NPU（Chips&MediaのWAVE-N）上ですべてを16ビット浮動小数点で処理するエンドツーエンドの画像処理パイプラインで、固定機能ISPを使わず、無線でアップデートできる。「既存のパイプラインは、固定機能ISPにニューラルブロックを後付けしている。私たちは従来のISPを丸ごと置き換える」とVisionary.aiのCEOは述べた。展示会の後、メーカー向けに評価用として提供された。

そのうえで、結果はJPEGやHEIFで保存される。どのデバイスでも開ける従来型のフォーマットである。パターンは再生と同じで、一社のパイプラインの内部はニューラル、境界では標準である。

JPEG AI：完成したニューラル標準、製品はまだない

JPEG AIは、ISO/IECとITU-Tによる初の学習型画像符号化標準である（ISO/IEC 6048、ITU-T T.840）。コア符号化システムを定める第1部は、2025年に発行された。JPEG委員会は、画素あたり7000から21万6000回の積和演算というデコーダー複雑度の範囲で、VVCのイントラ符号化に対して12.5%から27.9%の圧縮性能の改善を報告している。設計には共同で学習された二つのエンコーダーと三つのデコーダーがあり、受信側はデコーダーを一つだけ実装すれば、すべてのストリームを読むことができる。委員会は2024年5月までに、スマートフォンでの実装を2件報告した。並行してIEEEは2024年12月、NPU向けに設計されたニューラルネットワークベースの画像符号化の標準、IEEE 1857.11-2024を発行した。

2026年11月時点で、JPEG AIをデフォルトで書き出したり読み込んだりするスマートフォンのカメラ、写真アプリ、ブラウザーは、私が調べた限り見つからなかった。

ウェブが代わりに採用したもの：JPEG XL

2026年に実際に普及が進んだ画像フォーマットは、従来型である。Googleは2026年10月6日、Rustで書かれた新しいメモリ安全なデコーダーを使い、Chrome 155でJPEG XLの復号をデフォルトで有効にした。試験的なサポートを削除してから4年後のことである。Safariは2023年からJPEG XLを復号しており、Mozillaも同じ方向に動いている。Googleが挙げる利点は、JPEGより優れた圧縮、既存のJPEGファイルの可逆な再圧縮、HDR、プログレッシブ復号である。

この教訓は一般化できる。どこでも開ける必要があるフォーマットは、どのCPUでも動き、メモリ安全にできる、高速で決定論的なデコーダーによって勝つ。ニューラル画像デコーダーには、NPUかGPUと、デバイス間で厳密に一致する演算が必要である。ニューラル符号化が勝つのは、デバイス自身のパイプラインの内部と、同じ企業が符号化と復号を行うシステムである。圧縮したファイルを決して他者に渡さないクラウド写真サービスが、その例である。

画像分野のビジネスにとっての意味

お金は撮影パイプラインにある。AI ISPのIP、画像処理に見合う規模のNPU、そしてカメラメーカーが買うチューニング作業である。保存と配信について言えば、学習型画像コーデックは今日では閉じた系のための道具であり、賭けるべきフォーマットではない。

第04章

通話と会議

04 通話と会議：エンドツーエンドのニューラルコーデックが実際に実用化されている場

音声が先行した

Microsoftは2021年2月、AIを活用した音声コーデックSatinを発表し、TeamsとSkypeのすべての二者間通話ですでに使っていると述べた。Satinは6 kbpsで、最大16 kHzの音声帯域を持つ超広帯域の音声を届ける。Skypeの以前のコーデックSilkは、同じビットレートでは狭帯域だった。節約したビットレートは冗長化に充てられ、通話はパケットロスに強くなる。

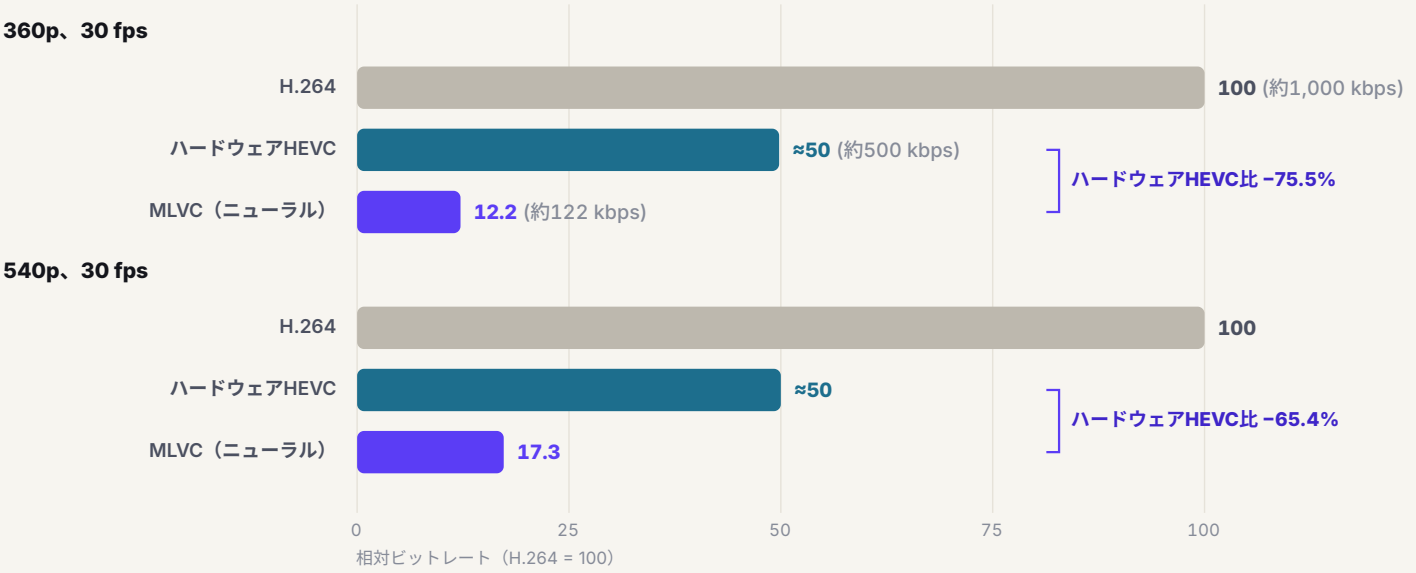
MLVC：本番の通話で動くニューラル映像コーデック

Microsoftは2026年7月24日、機械学習による映像コーデックMLVCを、MITライセンスのオープンソースとして公開した。モデルのコード、学習済みの重み、学習スクリプト、各種NPU向けの変換ツールが含まれる。MLVCは、Microsoft Researchが2021年から発表してきたDCVCファミリーの製品版である。Microsoftによれば、Teamsのピアツーピア通話でMLVCを展開しており、リアルタイムのテレメトリーとA/Bテストを行い、従来のコーデックをフォールバックとして用意している。

MLVCは540pと360pを毎秒30フレームで処理し、AppleのNeural EngineやQualcomm、IntelのNPUで、NPUの半分以下しか使わないように設計されている。処理はCPUとNPUで分担する。Microsoft自身のVideo Conferencing Datasetを用いたITU-T P.910準拠の主観評価では、360pでH.264より87.8%、ハードウェアHEVCより75.5%少ないビットレートで済み、540pではそれぞれ82.7%と65.4%少なかった ベンダー発表 。360pでは、H.264が約1 Mbpsを要するところを、およそ122 kbpsで済ませることになる。

図04 **通話では、ニューラルコーデックがハードウェアHEVC比でビットレートを65–76%削減**

Microsoftのビデオ会議テストセットで、同じ評価品質を得るのに必要なビットレート（H.264 = 100）。MLVCはノートPCやスマートフォンのNPU上で最大540p30で動作し、使うのはNPUの半分以下である。



出典: Microsoft、MLVCのリリース（2026年7月24日）：同社のVideo Conferencing DatasetでのITU-T P.910準拠の主観評価。HEVCの棒は公表された削減率から算出。ベンダー発表であり、独立した測定ではない。

見出しの数字は、二つの点で割り引いて読む必要がある。第一に、88%はH.264との比較である。同じノートPCにすでに載っているハードウェアHEVCと比べると65–76%で、それでも大きい。第二に、MLVCは実際に導入できるよう、圧縮率を犠牲にしている。ある独立した分析によれば、知覚品質で学習したDCVC-RTは、MLVCが75.5%を記録する同じ指標で81.9%に達する。Microsoftが所有しないハードウェアで同一の復号結果を得るための代償は、約6ポイントだった。

通話が自然な居場所である理由

通話は閉じた系である。両端のエンドポイントは同じアプリケーションを動かし、通話の開始時に互いの能力を取り決め、通話の途中でも標準コーデックにフォールバックできる。解像度は控えめで、ネットワークはしばしば劣悪であり、両端のノートPCには、スマートフォンが何年も前からそうであるように、いまやNPUが搭載されている。MicrosoftはAI PCの基準としてNPUの40 TOPSを定め、AppleのM4のNeural Engineは38 TOPSとされる。再生でニューラルコーデックを阻む条件は、ここには一つもない。

反例：OpusとMLow

すべての通話プラットフォームが同じ選択をしたわけではない。その理由は示唆に富む。

WebRTCとほとんどの通話アプリで使われるオープンな音声コーデックOpusは、バージョン1.5（2024年3月）で深層学習を取り入れた。信号そのものの処理や生成にネットワークを使うのは、これが初めてである。ただし、それらはオプションとして加えられた。ビットストリームの変更を必要としないデコーダー側の音質改善（LACEとNoLACE）、ニューラルなパケットロス補償、深層冗長化方式（DRED）である。古いデバイスに負担をかけないよう、いずれもデフォルトでは無効になっている。Opus 1.6はニューラルな帯域拡張と、3分の1の大きさのDREDモデルを加え、2012年の標準との完全な互換性を保っている。そのアプローチは「標準の内部で、オプションかつ後方互換」である。

Metaは、さらに従来型へと立ち返った。2024年6月、ネットワークではなく従来の信号処理に基づく低ビットレート音声コーデックMLowを投入し、現在はInstagramとMessengerのすべての通話で使われ、WhatsAppでも展開が進んでい

る。Metaの理由はデバイスにあった。同社の通話の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。Metaによれば、同社自身のEnCodecを含むニューラルコーデックが安定して動くのは、高価な端末だけである。MLowは6 kbpsでOpusの約2倍の品質を実現しつつ、計算量は10%少ない。

利用者がフラッグシップのノートPCから10年前のスマートフォンまでにまたがるサービスでは、許容できる計算量が決め手になる。ニューラルコーデックが最初に実用化されるのは、既存端末が新しく、ベンダーが両端を握る場所である。

コミュニケーション分野のビジネスにとっての意味

企業向けのビデオ会議、テレプレゼンス、遠隔支援、中継回線、プライベートネットワークは、両端が自社のものである限り、エンドツーエンドのニューラルコーデックが今日すでに製品になる場である。最初から三つのことを計画しておくべきである。フォールバック用のコーデック、サポートするすべてのNPUでの復号テスト、そしてベンダーのデータセットではなく自社のコンテンツでの品質評価である。

第05章

05

生成AI

生成AI：誰もコーデックと呼ばない最大の導入例

あらゆる音声・動画モデルは、学習型コーデックの上で動く

エンドツーエンドのニューラル圧縮の最大の導入例は、どのコーデック標準の中にもない。それは生成AIの内部にある。音声モデルは、波形を直接聞いたり生成したりはしない。学習型音声コーデックの離散トークンを読み書きする。2024年に公開されたKyutaiのMimiは、音声を毎秒12.5フレーム、1.1 kbpsのトークンに変換する。どの電話用コーデックよりも低いビットレートで、意味情報と音響情報を一つのストリームにまとめる。2022年に発表されたMetaのEnCodecもその一つである。Metaのエンジニアは後に、こうしたコーデックは低ビットレートでは優れているが、古いスマートフォンには重すぎると述べている（第4章）。

動画生成も同じ仕組みである。拡散モデルやTransformerモデルは、学習型の動画オートエンコーダー（VAE）の潜在空間で生成を行い、VAEのデコーダーが結果を画素に変換する。LTX-VideoのVAEは、動画を1:192に圧縮する。各潜在トークンは8フレームにわたる32×32画素のブロックを表し、128チャネルを持つ。そのデコーダーは最後のノイズ除去ステップも担うため、細部を再構成するだけでなく、部分的には生成している。

なぜこれらは配信フォーマットではないのか

これらのコーデックは音声・動画モデルへのリクエストのたびに動くが、どれも他者が復号するフォーマットではない。モデルごとに独自の潜在空間があり、レート制御も、適合性試験も、別のデコーダーが同じ映像を出力する保証もない。設計の目標は生成品質であって、所定のビットレートでの忠実度ではない。復号はデータセンターのGPUで行われる。その成功が示すのは、一社がエンコーダー、デコーダー、ハードウェアを所有していれば、学習型圧縮は大規模に機能するということである。

ビジネスにとっての意味

実務上の要点は二つある。第一に、生成動画や生成音声を提供する事業者にとって、トークナイザーの効率は直接のコスト項目である。すべてのトークンに対してアテンションを計算するTransformerでは、演算量はトークン数より速く増えるため、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 。潜在表現の設計、エントロピーモデル、デコーダーの効率を理解するコーデック技術者は、生成AIのチームにとって貴重である。第二に、境界は動いている。標準化団体は、人ではなく機械のための符号化に取り組み始めた。MPEGのVideo Coding for Machines（ISO/IEC 23888-2）は2026年初めに国際標準案に達し、2026年後半に最終化される予定である。JPEG AIも当初から、人間による視聴だけでなく、コンピュータービジョンのタスクにも使えるよう設計されている。

第06章

ゲームとグラフィックス

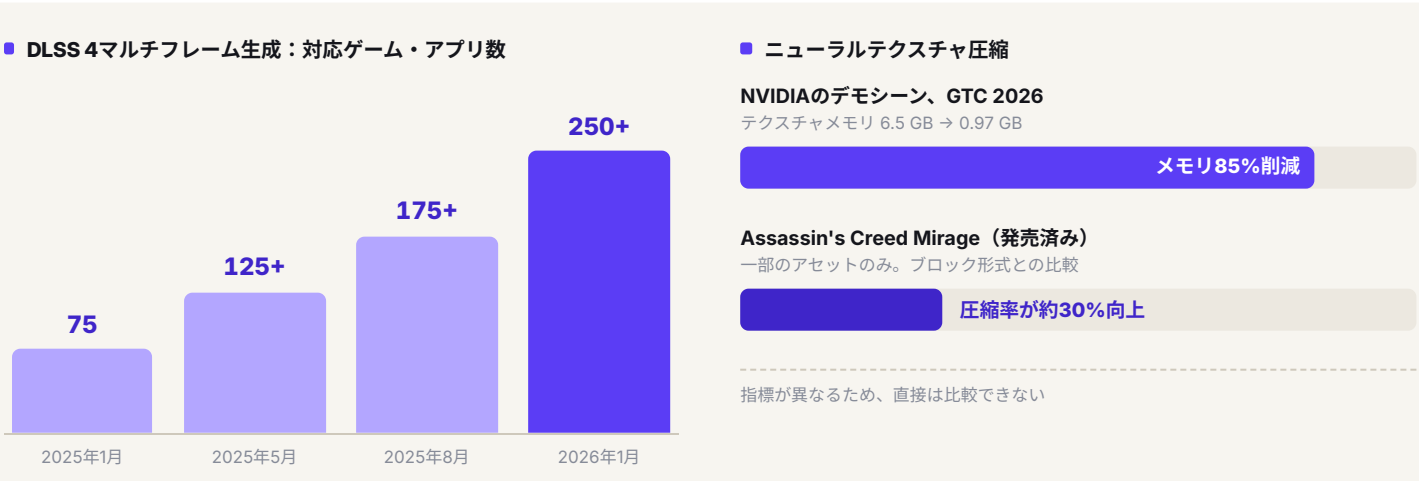
06 ゲームとグラフィックス：ニューラル再構成は主流に、ニューラルアセットは始まったばかり

描画する画素を減らし、残りをネットワークに埋めさせる

リアルタイムグラフィックスでは、プレイヤーが目にするものの大半を、すでにネットワークが再構成している。NVIDIAのDLSSは、描画する画素とフレームを減らし、残りをネットワークで生成する。マルチフレーム生成を備えたDLSS 4は、2025年1月に75のゲームとアプリの対応とともに登場した。NVIDIAの集計では、対応数は2025年5月までに125以上、8月までに175以上、2026年1月までに250以上となった。NVIDIAはこの2026年1月に、超解像用の第2世代Transformerモデルを備えたDLSS 4.5を発表している。これは名前こそ違え、デコーダーそのものである。「ビットストリーム」は低解像度のフレームと動き情報で、ネットワークが完全な画像を再構成する。

図05 **一社が技術スタック全体を握ると、ニューラル再構成は急速に普及する**

左：NVIDIAのDLSS 4マルチフレーム生成に対応するゲームとアプリの数。右：デモシーンと発売済みゲームでのニューラルテクスチャ圧縮。後者は実行時コストのため、一部のアセットにしか使っていない。



出典: NVIDIA（2025年1月から2026年1月、GTC 2026のデモ）、Ubisoft La Forge（2026年2月）。ベンダー発表。二つのテクスチャの数値は異なるものを測っている。

ニューラルテクスチャは、選択的に実用化された

テクスチャは、GPUがハードウェアで復号する固定のブロック形式（BCn）で圧縮される。ニューラルテクスチャ圧縮は、一組のマテリアルテクスチャを、コンパクトな潜在表現と、必要なときにテクセルを再構成する小さなネットワークに置き換える。UbisoftのLa Forgeラボは2026年2月、2023年発売のゲームAssassin's Creed Mirageにこれを搭載したことを説明した。同じ見た目の品質で、従来の形式より約30%高い実効圧縮率だという ベンダー発表 。ただし、全面的には使われなかった。「使用は実行時の性能によって制約された」ため、Ubisoftはインスタンス数が多くメモリ負荷の大きい一部のアセットに適用した。

GitHubでベータ版が公開されているNVIDIAのRTX Neural Texture Compression SDKは、その上限を示している。GTC 2026でのデモでは、BCnで6.5 GBのテクスチャメモリを使っていた邸宅のシーンが、ニューラルテクスチャでは970 MBで済み、85%少なかった ベンダー発表 。SDKには二つの経路がある。ロード時にBCnへ復号する方法は、ディスク容量とダウンロードサイズを節約するが、ビデオメモリは節約しない。サンプリング時に復号する方法はメモリを節約するが、毎フレーム、シェーダーの時間を消費する。

APIが整いつつある

グラフィックスAPIは、シェーダー内でのニューラル復号に必要な行列演算を加えつつある。Microsoftは2026年2月にShader Model 6.9を出荷した。2025年にプレビューしたcooperative vector機能は、後のシェーダーモデルに向けた統一的な線形代数の設計に統合され、2026年4月にはDirectX Linear Algebraがパブリックプレビューに入った。これがすべてのドライバーに入れば、ニューラルテクスチャのデコーダーは一つのコードパスで、どのベンダーのGPUでも動くようになる。

ビジネスにとっての意味

ゲームがニューラル復号を実用化できるのは、閉じた系の条件がそろっているからである。一つのエンジン、一つのアセットパイプライン、GPU内のテンソル演算ハードウェア、そしてフォールバック用の形式である。残された仕事は、ツール、ニューラルマテリアルを符号化するアセットパイプライン、ベンダーをまたぐランタイム、そして何よりモバイルにある。モバイルでは、復号するテクセルあたりのエネルギーコストが決め手になる。

第07章

07 特殊用途リンク：衛星、ドローン、ロボット

机上では、エンドツーエンドのニューラル符号化が最もよく適合するのがこれらのケースである。単一の所有者がリンクの両端を握り、帯域は乏しく損失も多く、内容は学習で対応できるほど限定的で、知覚的に、あるいは意味的にさえ忠実な映像で足りることが多い。こうしたリンク向けの損失に強いニューラルコーデックは活発な研究分野であり、低ビットレート映像のサーバー側での復元は実際に使われている。

同時に、消費者向けの動画と比べれば、小さく、規制が強まりつつある市場でもある 分析 。中国では、商務部とほか2機関による2026年第49号公告が10月19日に公布され、2026年11月18日に施行される。ドローン関連品目13分類の輸出管理を改定するものである。業界紙の報道によれば、技術項目の一つは無人地上プラットフォームを対象とし、その定義の一部に、75%のセマンティック画像圧縮や150ミリ秒以内の画像再構成といった基準値が使われている。セマンティック圧縮そのものが、輸出管理のパラメーターになったのである。米国では、連邦通信委員会が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。このため、外国製の新機種はもはや機器認証を受けられない。

結論として、産業や政府の顧客を持ち、双方の輸出管理に対応できる専門企業には堅実なニッチだが、主流の市場ではなく、ニューラル符号化が大規模に機能するかを見極める場でもない。

第08章

08 シリコン：シリコンの視点：エネルギー、演算量、メモリ、厳密性

本章は、私自身の経歴が最も生きる章である。私はファブレス半導体の研究開発とマネジメントに長年携わり、フルカスタムASICのデータパス設計や、映像コーデックアクセラレーターをはじめとするマルチメディア・カメラチップのプロジェクトを手がけてきた。チップ設計チームが新しいコーデックについて問うのは、BD-rateではない。画素あたり何ジュールか、画素あたりのDRAMトラフィックは何バイトか、ゲート数はいくつか、そして世界中のすべてのデコーダーがまったく同じビットを計算できるか、である。四つのいずれにおいても、2026年のニューラルデコーダーは、最初のチップが出る前のハードウェアHEVCと同じ段階にある。ソフトウェアでは有望だが、シリコンでは高くつく。

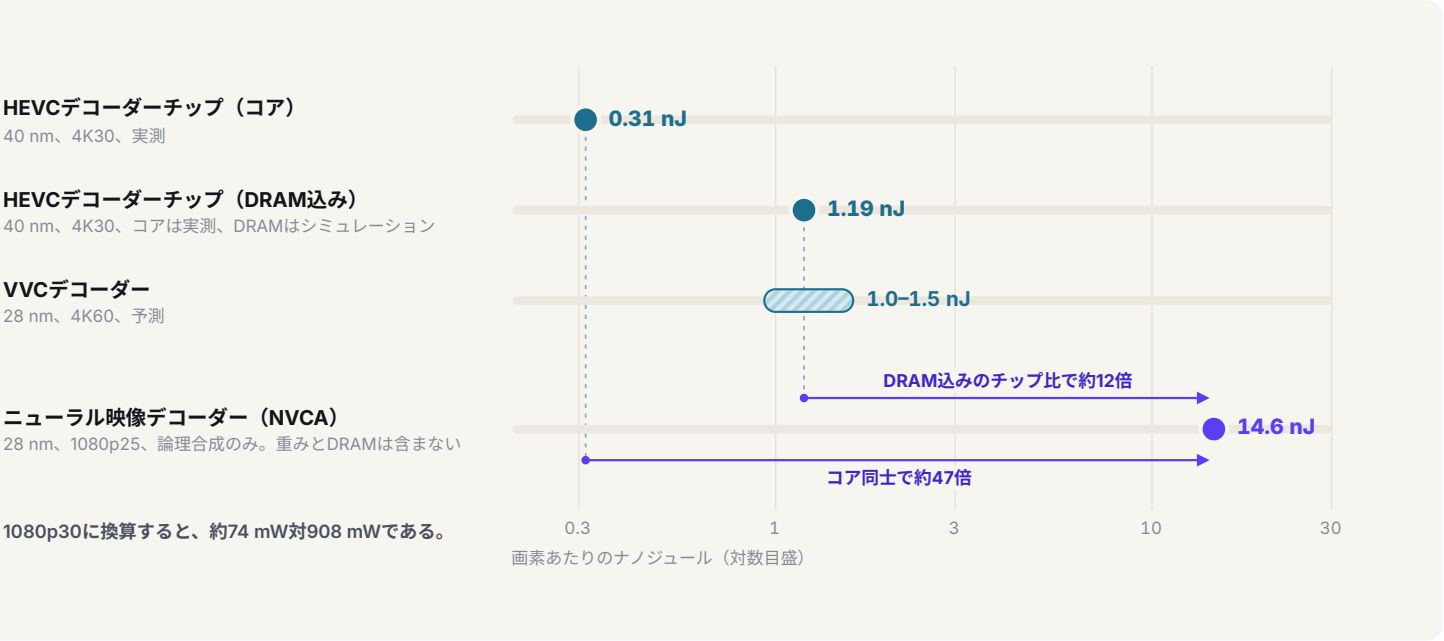
画素あたりのエネルギー：ひと桁の差

チップとして設計されたニューラルデコーダーは少ないため、比較は公開された一握りの設計に頼ることになる。従来型の側の基準は、MITのHEVCデコーダーチップ（Tikekarら、ISSCC 2013およびIEEE JSSC 2014）である。40 nm、1.77 mm²で、200 MHzで4Kを毎秒30フレームで復号する。コアの消費電力は78 mW、画素あたり0.31 nJで、シミュレーションによるDRAMの電力を含めると、システム全体で画素あたり1.19 nJになる。VVCについては、量産品の数値は非公開である。Electronic Designの2026年の分析は、MITのチップを28 nmにスケーリングし、VVCのより高い複雑度を反映して、約100万ゲートの4K60デコーダーで画素あたり1–1.5 nJと予測している。

ニューラルの側では、Electronic Designの数え方によれば、公開されているエンドツーエンドのニューラル映像デコーダーASICの設計はNVCA（Zhangら、DATE 2024）だけである。TSMCの28 nmで論理合成され、ゲート数は501万で、チップ外に置かれる重みは含まない。1080pを毎秒25フレーム、760 mWで復号し、画素あたり14.6 nJである。

図06 **復号画素あたりのエネルギー：固定機能デコーダーがひと桁優位**

唯一公開されているニューラル映像デコーダーのチップ設計は、画素あたりのエネルギーが、実測されたHEVCデコーダーチップ（DRAM込み）の約12倍、コア単体の約47倍に達する。ニューラル設計自身の重みとDRAMのトラフィックは含まれていない。



出典: Tikekarら、IEEE JSSC 2014（MIT）。NVCA：Zhangら、DATE 2024（arXiv 2312.10716）。VVCの範囲はElectronic Designによる予測（2026年）。HEVCチップは旧世代のプロセスノード（28 nmに対し40 nm）であり、比較はニューラル設計に有利に働く。

これは、DRAM込みのHEVCチップの約12倍、コア単体の約47倍にあたる。この比較は、三つの点でNVCAに有利である。HEVCチップは旧世代のプロセスノードであり、NVCAの数値は重みとDRAMのトラフィックを除外しており、NVCAは実測したシリコンではなく論理合成の結果である。1080p30では、差は約74 mW対およそ900 mWになる。誰も気づかないデコーダーと、スマートフォンのバッテリーと熱設計の予算に表れるデコーダーの差である。

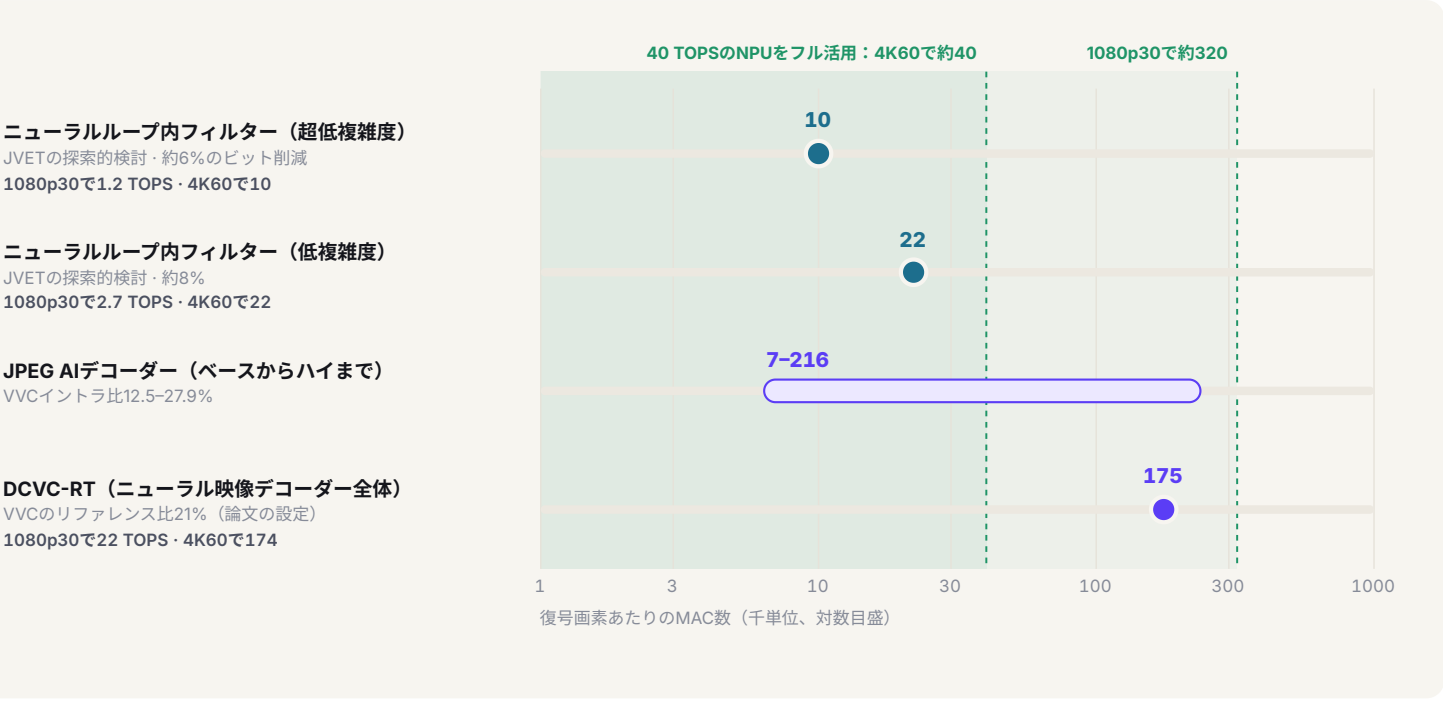
画素あたりの演算量：ニューラルデコーダーがNPUに求めるもの

私が見つけた中で最も高速な公開済みのニューラル映像コーデックは、Microsoft ResearchのDCVC-RT（Jiaら、CVPR 2025）である。NVIDIA A100 GPU上で、1080pを毎秒125フレームで符号化し、113フレームで復号する。論文の試験設定では、VVCのリファレンスエンコーダーに対して21%のビットレートを削減し、異なるデバイスでも一貫して復号できるよう整数演算を使う。補足資料には、各モジュールの複雑度が載っている。デコーダーが実行しなければならない四つのモジュール（デコーダー、特徴抽出器、エントロピーモデル、再構成）を合計すると、復号画素あたり約17万5000回の積和演算になる 推定 。

1080p30では毎秒約11兆回の積和演算となり、1 MACを2演算と数えれば、利用率の低下を考慮する前の段階で約22 TOPSになる。4K60ではその8倍が必要である 推定 。比較の目安として、MicrosoftはAI PCを40 TOPS以上のNPUで定義し、AppleはM4のNeural Engineを38 TOPSとしている。これらはピーク値であり、デバイスが行うほかのすべての処理と共有される。

図07 **ニューラルツールは画素あたり数万MAC、ニューラルデコーダー全体では数十万MAC**

デコーダーの画素あたりの積和演算数（対数目盛）と、40 TOPSのNPUをフルに使った場合に4K60と1080p30で賄える量。各手法の改善幅は比較の基準が異なり、直接は比較できない。



出典: JVET（SC 29議長経由、2026年4月）、JPEG委員会（2024年）、DCVC-RT：CVPR 2025補足資料のデコーダーモジュールを著者が合計、MicrosoftのCopilot+ PCの定義。著者の計算：MACあたり2演算、フル稼働。

図7に示した段階は、この分野を読み解く有用な手がかりになる。JVETのループ内フィルター（1万MACで約6%、2万2000 MACで8%の削減）のように画素あたり数万MACのニューラルツールなら、将来の標準デコーダーに入ることにも現実味がある。JPEG AIはその範囲にまたがり、画素あたり7000から21万6000 MACで、VVCイントラ比12.5–27.9%の改善を示す。画素あたり20万MAC前後の完全なニューラル映像デコーダーは、閉じた系の中で、控えめな解像度で、NPU上に置くのがふさわしい。MLVCが位置するのは、まさにそこである。540p30で、NPUの半分以下を使う。

MACよりも、メモリトラフィックが真の制約になることが多い

DCVC-RTの著者らは、速度を制約していたのは演算ではなく「オペレーションコスト」、つまりメモリの入出力と関数呼び出しの回数だったことを見いだした。明示的な動きモジュールとマルチスケール処理を取り除いたのは、主にそのためである。NVCAの設計者は、動きと残差の特徴量によるチップ外メモリへのトラフィックを削減するため、専用のデータフローを構築した。チップ設計者なら、このパターンに見覚えがあるだろう。MITのHEVCチップでは、DRAM帯域を67%削減するキャッシュを入れた後でも、DRAMがシステムのエネルギーの約4分の3（画素あたり1.19 nJのうち0.88 nJ）を占める 分析。フレーム間で多チャネルの大きな特徴マップを保持するニューラルデコーダーは、メモリの問題を改善するどころか悪化させる。

厳密性：すべてのデコーダーが同じビットを計算しなければならない

エントロピー符号化が機能するのは、デコーダーの確率モデルがエンコーダーのものと厳密に一致する場合だけである。従来型のコーデックでは、そのモデルは表か、少数の整数の規則である。ニューラルコーデックではそれがネットワークであり、異なるカーネル、異なる累算順序、異なるNPUから生じる二つのデバイス間の浮動小数点の差が、デコーダーに誤ったシンボルを読ませ、その時点以降のストリームを失わせることがある。

ニューラルコーデックを実用化した者も標準化した者も、皆この問題を工学的に解消しなければならなかった。DCVC-RTは、デバイス間で一貫した符号化のためにモデルを整数に変換する。MLVCは、プラットフォーム間の不一致に耐える

ためにスケール共有型のハイパープライアを採用しており、それがDCVC-RTに対して圧縮率で劣る理由の一部である可能性が高い 分析。JVETのニューラルネットワーク探索的検討では、すべてのネットワークを、16ビット整数推論を行う小さなライブラリSADLで動かす。JPEG AIは、潜在領域、残差領域、画素領域に適合性の検証点を定めている。チップ設計チームにとっては、HEVCの整数変換と同じ規律である。標準が演算を厳密に定義しなければならず、それがネットワークの設計を制約し、ベンダー固有の近道を排除する。

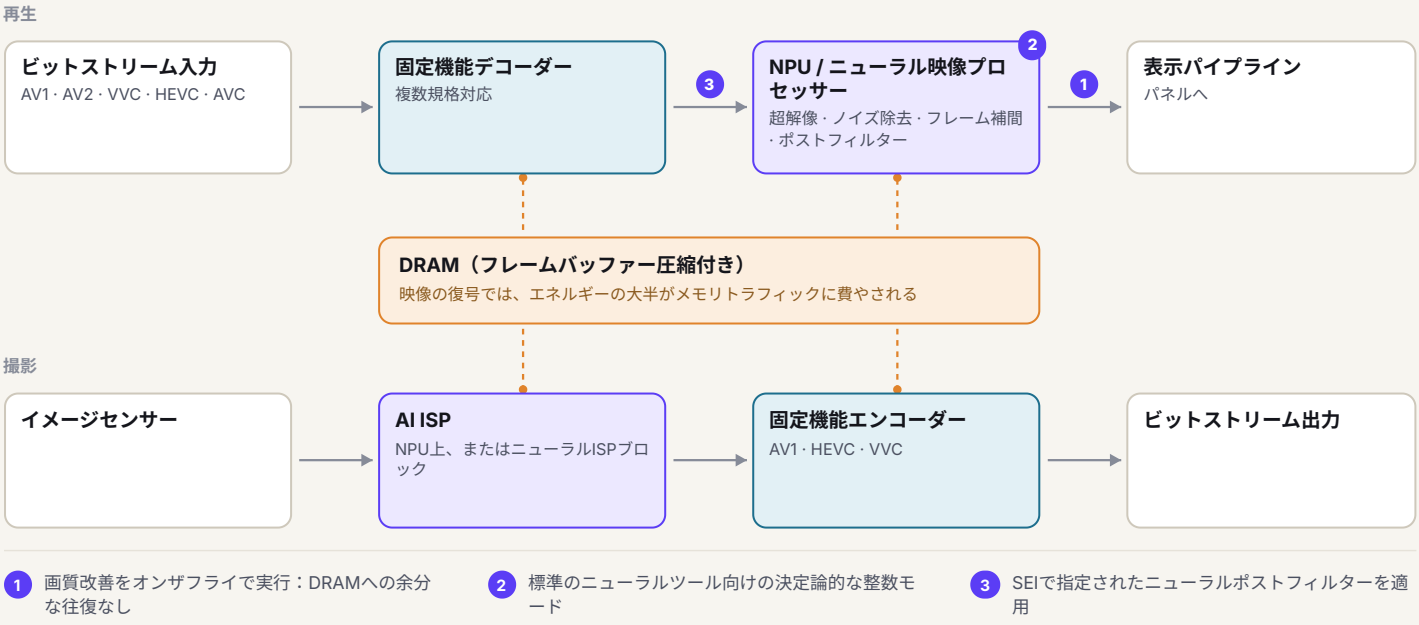
障害は面積ではなく、検証と製品寿命である

NVCAの500万ゲートは、予測されるVVCデコーダーの約100万ゲートに比べると大きく見える。しかし、すでに数TOPSのNPUを搭載している現代のスマートフォンやテレビのSoCでは、数百万ゲートは設計を左右する要素ではない。本当にお金がかかるのは、適合性試験一式に対する検証、デバイスの寿命を通じてソフトウェア更新のたびにプログラマブルなデコーダーの正しさを保つこと、そして4K60 HDRでの消費電力である。こうしたコストは、何をつくるべきかを標準が確定させて初めて意味を持つ。

いま勝つアーキテクチャ：固定機能コーデックとNPU

図08 **実用化されている分担：ビットストリームは固定機能ハードウェア、画素はニューラルエンジン**

2026年のメディアSoC。コーデックブロックはあらゆる標準を画素あたり1ナノジュール未満で復号し、NPUはネットワークが得意で標準が制約しない処理を担う。効いてくるコストはDRAMトラフィックである。



出典: Tencent（Canghai V2）、Chips&Media、Allegro DVT、Samsungの製品発表（2026年）をもとにした著者のアーキテクチャ観。

2026年に売れているシリコンは、固定機能の複数規格対応コーデックの横にニューラルエンジンを置いている。コーデックブロックは、AV1、AV2、VVC、HEVC、AVCを画素あたり1ナノジュール未満で処理する。ニューラルエンジンは、ネットワークが得意で標準が制約しない処理を担う。撮影側のノイズ除去とISP、表示側の超解像、フレーム補間、画質改善である。両者はフレームバッファー圧縮とDRAMのスケジュールを共有する。効いてくるコストは、メモリトラフィックだからである。

市場も、この分担を示している。Tencentが2019年から自社開発している映像コーデックASIC、Canghai（沧海）は、第1世代チップが10万個以上、Tencent Cloudと同社のライブ配信、オンデマンド配信、クラウドゲームのサービスで稼働し

ている。第2世代は2026年5月に、電源投入を終えて量産に入り、同年後半にサービスに投入すると発表された。一つのアーキテクチャでH.265とH.266に対応し、前世代に比べて圧縮効率はそれぞれ10%以上、30%以上向上し、スループットは約2倍で、超解像などのオンチップ画質改善を備える。Chips&Mediaは、コーデックIPと映像向けNPUのWAVE-Nを並べて販売している。Allegro DVTはAV2デコーダーを販売し、初のニューラル映像処理IPを加えた。Samsungのテレビ用プロセッサーは、標準デコーダーと数百のアップスケール用ネットワークを組み合わせている。

私ならこうつくる：役割別

SoCとテレビ用チップのチーム。再生が必須のすべての標準について、固定機能デコーダーを維持する。AV2はいまから、VVCは放送が必要とする場合に含める。NPUは表示解像度とフレームレートでの後処理に見合う規模にし、デコーダーからNPU、ディスプレイへの経路は、画質改善がDRAMとの往復を増やさないように設計する。ニューラルポストフィルターのSEIに対応する。NPUに決定論的な整数モードを用意する。H.267は2029年以降に設計を始めるものと位置づけ、そのニューラルツールのためにある程度のプログラマビリティを残しておく。

IPベンダー。AV2とVVCのデコーダーをいま売り、その周りでニューラル処理が必要とする部品も売る。ポストフィルター用アクセラレーター、フレームバッファー圧縮、適合性検証ツールである。

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である。タイトルごとに一度動かすだけで、デバイスには何も求めないからである。デバイスが許す範囲で、クライアント側の超解像を使う。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とどめる。

第09章

09

標準と特許

標準と特許：ニューラルツールがライセンス対象になる場

標準化のパイプライン

標準	団体	ニューラルの要素	状況（2026年11月）
H.267（仮称、VVCの後継）	JVET（ITU-TとISO/IEC）	未定。従来型ツール（ECM）とニューラルツール（ループ内フィルター、イントラ予測、超解像）がともに候補	提案募集は2026年7月、応募提案の評価は2027年1月、最初のテストモデルは2027年後半、第1版は2029年末
VVCとHEVC（VSEIを併用）	JVET	ニューラルポストフィルターをSEIメタデータで記述できる。適用するかどうかはデコーダーの任意	有効。2025年10月の版でメッセージを拡張
AV2	AOMedia	従来型ツール、ロイヤルティフリー	バージョン1.0.0は2026年5月28日。デコーダーIPは2026年6月からライセンス。YouTubeのウェブ配信は2027年初めに計画
JPEG AI	JPEG（ISO/IECとITU-T）	エンドツーエンドの学習型画像コーデック。人の視覚と機械の視覚の両方に対応	第1部は2025年に発行。第2版を検討中。製品でのデフォルト対応は見つからず
IEEE 1857.11	IEEE	NPU向けに設計されたニューラルネットワーク画像符号化	2024年12月に発行

標準	団体	ニューラルの要素	状況（2026年11月）
Video Coding for Machines	MPEG	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。2026年後半に最終化の予定
OpusのML拡張	IETF	オプションの深層冗長化、デコーダー側の音質改善、帯域拡張	libopus 1.5と1.6に搭載。深層冗長化は標準化中

二つの点が目を引く。第一に、ニューラルの要素はコアより先に、周縁から標準に入っている。オプションのポストフィルター、オプションの品質改善、画像符号化であり、いずれも主たる復号経路を変えない。第二に、あらゆる画面の復号経路にニューラルツールを入れうる唯一の標準H.267は、第1版まであと3年ある。

特許で有利な位置にいるのは誰か

標準のツールは、ライセンス可能な特許になる。そのためH.267のプロセスは特許競争でもあり、ニューラルツールもその一部である。GreyBが2025年にまとめた、VVC後継関連の特許3,916件のランドスケープ分析では、ByteDanceが645の特許ファミリーと251件のJVET寄書で首位に立つ。同分析は、Xiaomiと西安電子科技大学のように20以上のファミリーを持つ産学連携にも言及している。MLexは2026年3月、Tencent、Alibaba、ByteDanceが映像コーデックのパテントプールで積極的な参加者として台頭していると報じた。2026年5月には、Sauvegarder Investment ManagementがAlibabaから、AV1、AV2、Video Coding for Machinesにまたがる映像コーデック関連の特許資産126件を買い取った。InterDigitalは2025年10月にDeep Renderを買収し、そのAI映像符号化特許をポートフォリオに加えた。

デバイスあたりのお金は、プールの料率表を見ればよくわかる。Access AdvanceのVVCの料率（2026年7月1日時点、契約を順守し同社の商標を使用するライセンシーの場合）は、米国、EU、日本などRegion 1の市場で、携帯電話、タブレット、ノートPC 1台あたり0.50ドル、Region 2では0.25ドルである。4Kテレビは、Region 1で価格に応じて1.00ドルから1.50ドル、企業あたりの年間上限は7500万ドルである。AV1とAV2は、AOMediaのメンバーからはロイヤルティフリーで提供される。H.267に何が入るにせよ、そのツールは同じように値付けされる。ニューラル符号化の特許を持つ企業がいまJVETに寄書しているのは、そのためである 分析 。

第10章

10 誰が収益を得るか
誰が収益を得るか、企業はどこで勝負できるか

2026年のお金の流れ

図09 **2026年、ニューラル符号化で収益を得るのは誰か**

お金は標準とプラットフォームに流れる。エンドツーエンドのニューラル映像コーデックのライセンスで成り立っている独立企業は、私が調べた限り見当たらなかった。

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）

出典: 本レポートで引用した企業とパテントプールの開示情報。状況は2026年11月12日時点。プール料率：Access Advance、VVC、Region 1、契約を順守し商標を使用する場合、2026年7月1日時点。

プラットフォームは節約分を手元に残す。Netflix、YouTube、Meta、Microsoft、Tencentはニューラル処理を自社で構築し、帯域とストレージの削減、ビットあたりの品質向上という形で、利益を手元に残す。普及に役立つときは、コードを無償で公開する。MicrosoftはMLVCをMITライセンスで公開した。

デバイスメーカーとチップメーカーは差別化を売る。カメラのISPとテレビの画質プロセッサーは、ニューラルネットワークがすべてのユーザーに触れる場所であり、機能として販売される。QualcommのCognitive ISP、Samsungのアップスケール用プロセッサーがその例である。

IPベンダーは、新しい標準ごとのデコーダーとNPUを売る。AV2の仕様完成から3週間以内に、Chips&Media、Allegro DVT、VeriSiliconはいずれもAV2デコーダーの顧客を発表した。同じベンダーが、いまではコーデックの横に置くニューラル処理ブロックも販売している。

GPUベンダーはプラットフォームを築く。DLSSとニューラルテクスチャ圧縮によって、NVIDIAのテンソル演算ハードウェアは、ゲームのつくり方の一部になっている。

特許権者はデバイスごとに徴収する。その経路はVVCとHEVCのパテントプールであり、2029年以降はH.267のプールが加わる。

ニューラルコーデックのスタートアップは買収された。WaveOneは2023年にAppleへ、iSIZEは2023年にソニー・インタラクティブエンタテインメントへ、Deep Renderは2025年にInterDigitalへ渡り、いずれも条件は公表されていない。どの場合も、プラットフォームかライセンサーが、チームと特許を買い取った。2026年の時点で、エンドツーエンドのニューラル映像コーデックのライセンスで成り立っている独立企業は、私が調べた限り見当たらなかった。

企業はどこで勝負できるか

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（ダウンスケール、ノイズ除去、タイトルごとの最適化）	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（超解像、フレーム補間、画質）	テレビ、セットトップボックス、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。目に見える差別化	コモディティ化。チップメーカーの内製
閉じた系のニューラルコーデック（会議、遠隔操作、専用回線）	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が実現可能性を証明し、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を組み合わせたシリコンまたはIP（AV2、VVC、後処理）	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、決定論的推論のツール	ニューラルツールを出荷するすべての企業	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

ほとんどの企業にとって、2027年に収益を得られるのは最初の2行である。3行目は、特定の分野に集中する製品企業にとって最も興味深い。最後の行には、忍耐とライセンス戦略が要る。

第11章

11 主張と根拠
よく聞く主張と、証拠が示すこと

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、そのとおりである。DCVC-RTは、論文の設定の1080pで、VVCのリファレンスエンコーダーに対して21%削減する。JPEG AIは、VVCイントラに対して12.5–27.9%削減する。デコーダーのコストは画素あたり約17万5000 MACで、NPU並みのエネルギーを要する。固定機能ハードウェアなら1ナノジュール未満である。
「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。ニューラルツールを載せうる最初の標準は2029年末を目標とし、チップは2031年ごろ、主流化は2030年代半ばになる 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」	Microsoft自身のテストセットにおける、360pでのH.264比である。ハードウェアHEVC比では360pで75.5%、540pで65.4%であり、しかもMicrosoftが両端を握る通話での値である。
「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得られた実効圧縮率の改善は約30%で、実行時のコストのため、ニューラルテクスチャは一部のアセットにしか使われなかった。
「JPEG AIはISO標準なので、カメラやブラウザーが採用する」	私が調べた限り、製品でのデフォルト対応は見つからなかった。2026年にウェブが採用したフォーマット、すなわちChrome 155のJPEG XLは従来型である。
「ニューラル復号はスマートフォンには高すぎる」	部分的には正しい。MLVCはNPUの半分以内で540p30を復号する。DCVC-RT級のモデルで4K60を処理するには、約175 TOPSの持続演算性能が必要になる 推定 。
「衛星とドローンのリンクが、ニューラル映像の大きな市場になる」	技術的には最も適しているが、市場は小さく、太平洋の両側で輸出管理が強まっている。
「大手プラットフォームは、スタートアップからニューラルコーデックのライセンスを受ける」	大手は自社で開発したか（Netflix、Microsoft、Tencent）、スタートアップを買収した（Apple、ソニー、InterDigital）。

単純な試金石で、たいていの売り込みは見極められる。比較の基準、指標、解像度、復号コストを尋ねることである。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。

第12章

12

注目すべき節目

2027年から2030年の注目点

時期	節目	意味
2027年1月	JVETがブリスベンでH.267の応募提案を評価	上位の提案のうちいくつがニューラルツールに頼るか、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	次の従来型サイクルが、2030年ごろまでのデバイスの設計予算を決める
2027年後半	H.267の最初のテストモデル	ニューラルツールがコア設計に入るのか、探索的検討にとどまるのか
2027年	Teamsのピアツーピア通話以外へのMLVCの広がり。公開コードを採用する他のベンダー	オープンなニューラルコーデックが、所有者の外で試される最初の機会
2027–2028年	JPEG AIをデフォルトで読み書きする最初の製品	企業の境界を越えて使われる最初のニューラルフォーマット
2027–2028年	DirectX Linear Algebraが製品版のドライバーに搭載。ニューラルテクスチャが他のゲームとモバイルに拡大	グラフィックスにおける、ベンダーをまたぐニューラル復号
継続的	スマートフォンとPCのNPUの演算能力	持続性能が約100 TOPSになれば、DCVC-RT級の1080p30デコーダーが占めるのは、NPUの半分ではなく約5分の1になる 分析
2029年末	H.267第1版	一般向けの再生におけるニューラルツールのチップ開発サイクルの始まり

第13章

13 未解決の問い

未解決の問い

本レポートで扱った問いのいくつかには、まだ公開された答えがない。そして、それこそが私が技術的に最も興味を引かれる問いである。

大規模での厳密性。ニューラルデコーダーは、圧縮性能の優位の大半を手放す前に、異なるベンダーのNPU間で、ビット単位で一致する整数のみの演算にどこまで近づけるか。MLVCの6ポイントのトレードオフは、一つのデータ点にすぎない。標準には一般的な答えが必要である。

MACよりメモリ。ハードウェアデコーダーとディスプレイの間でオンザフライで動くとき、ニューラル後処理は画素あたりどれだけのDRAMトラフィックを要するか。そのワーキングセットは、どこまで小さくできるか。

H.267におけるニューラルツールの値段。4K60デコーダーにおいて、ニューラルループ内フィルターは画素あたり何MAC、何ナノジュールで、削減するビットに見合わなくなるか。

通話以外の閉じた系。クラウド写真ストレージやゲームストリーミングから、ロボットや専用の中継回線まで、H.267の登場より前にエンドツーエンドのニューラルコーデックを採用できるのは、ほかにどの閉じた系か。

生成するデコーダー。動画VAEがすでにそうしているように、デコーダーが細部を再構成するのではなく生成する場合、忠実度はどう測定し、どう規定すべきか。

これらは、技術交流と技術協力を通じて、今後数年にわたりともに探究していける領域である（i@tanshuai.com）。

本レポートの限界

本レポートは、2026年11月12日までの公開文書に基づくデスクリサーチである。この版のために、一次インタビューも独自の測定も行っていない。MLVCのビットレート削減率やニューラルテクスチャの圧縮率など、主要な数値のいくつかは、ベンダーや論文著者が自らのテストセットで示したものであり、その旨を表示している。DCVC-RTに関する私の演算量の推定は、著者らが公表したモジュール別の数値を合計したもので、利用率の低下を無視しているため、実際のNPU負荷はこれより高くなる。エネルギーの比較は、異なるプロセスノードで公開された少数の設計に依拠しており、量産品のVVCおよびAV2デコーダーの数値は非公開である。標準化のスケジュールは、各委員会自身の目標である。2029年以降のH.267の日程は、AV1とVVCの経緯から私が推定したものである。

著者について

タン・シュアイは、チップ、ハードウェア、システムソフトウェアにわたって20年以上の経験を持つテクノロジストであり創業者で、取得した特許は20件にのぼる。ファブレス半導体の研究開発とマネジメントに長年携わり、フルカスタムASICのデータパス設計や、映像コーデックアクセラレーターをはじめとするマルチメディア・カメラチップのプロジェクトを手がけてきた。その経験が本レポートのレンズになっている。すなわち、画素あたりのエネルギー、メモリトラフィック、そして標準がデコーダーに何をシリコンへ固定することを許すか、である。活動の場は、シリコンバレー、ニューヨーク、シカゴ、台湾、グレーターベイエリア（深圳と香港）、ベトナム、マレーシアに及ぶ。現在はAI研究開発者として活動している。

技術的な議論：i@tanshuai.com · tanshuai.com

注記と免責事項

本レポートは、2026年11月12日までに入手できた公開文書に基づく、独立したデスクリサーチである。情報提供のみを目的としており、投資、法律、エンジニアリングに関する助言ではない。製品、標準化のスケジュール、パテントプールの料率、輸出規則は変わりうる。依拠する情報は、必ず一次資料で確認されたい。 ベンダー発表 、 推定 、 分析 と表示した数値は、その前提で読まれたい。企業名と製品名はそれぞれの所有者に帰属し、識別のためにのみ使用している。本レポートをレビューした企業も、スポンサーとなった企業もない。

© 2026 タン・シュアイ。出典を明記すれば、本レポートを改変せずに共有できる。

出典

すべて2026年11月3日から11月12日の間に閲覧した。章ごとに分類している。

トラフィック、デバイス、対象範囲

- Ericsson, *Mobility Report*, 2026年6月（2025年末時点で動画がモバイルデータトラフィックの約75%）: https://www.ericsson.com/en/reports-and-papers/mobility-report/reports/june-2026 ; 要約: 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、2025年第4四半期および通年の世界スマートフォン出荷台数（12億6000万台）、2026年1月13日: https://www.idc.com/resource-center/press-releases/q425mobilephonetop5/
- TrendForce、2025年のテレビ出荷予測（1億9571万台）: 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年のテレビ予測（2億870万台）: https://iconnect007.com/article/145954/omdia-forecasts-flat-tv-shipments-in-2025-but-mini-led-poised-for-explosive-growth-at-829/145951/ein

再生とストリーミング

- 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のカタログに占めるAV1の割合）: 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+の商用運用、2026年6月）: https://mpegh.com/brazil/ ; CartaCapital（テレビに内蔵される受信機について）: https://www.cartacapital.com.br/?p=628337 ; Olhar Digital（TV 3.0におけるVVCについて）: 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 TV highlights（NQ8 AI Gen3、768のニューラルネットワーク）: https://www.samsung.com/us/tvs/8k-tv/highlights/

写真と画像

- 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とVisionary.ai、完全なAIベース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のインタビュー: https://petapixel.com/2026/01/18/a-look-at-the-worlds-first-full-ai-based-image-signal-processor
- JPEG委員会、第103回会合のプレスリリース（JPEG AIの改善幅、複雑度、スマートフォンでの実装）: https://jpeg.org/items/20240507_press.html
- G. J. Sullivan、ISO/IEC JTC 1/SC 29の紹介、2026年4月（JPEG AIの状況、ECMとNNVCの数値）: 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ワーキンググループ（ニューラルネットワークベースの画像符号化）: https://sagroups.ieee.org/fvc
- Chrome 155とJPEG XL、2026年10月6日: https://xenospectrum.com/en/chrome-155-jpeg-xl-rust/

通話と会議

- 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 ; コード: 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の分析: https://ireplay.tv/blog/mlvc-microsoft-neural-video-codec-open-source-streaming/
- Opus 1.5のリリースノート: 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+ PCs（40 TOPS以上のNPU）: https://www.microsoft.com/en-us/windows/copilot-plus-pcs ; Apple、M4の発表（38 TOPSのNeural Engine）: https://www.apple.com/newsroom/2024/05/apple-introduces-m4-chip/

生成AI

- Kyutai、Mimiのモデルカード（12.5 Hz、1.1 kbps）: https://huggingface.co/kyutai/mimi
- Lightricks, "LTX-Video: Realtime Video Latent Diffusion"（1:192の動画VAE）: https://arxiv.org/abs/2501.00103
- MPEG第153回会合のプレスリリース（Video Coding for MachinesがDISに）: https://www.mpeg.org/wp-content/uploads/2026/03/press_release_153.pdf

ゲームとグラフィックス

- NVIDIA、DLSS 4の普及：発表時とComputex 2025（75と125以上）: https://www.nvidia.com/en-us/geforce/news/125-dlss-4-multi-frame-gen-games-more-announced-computex-2025/ ; Gamescom 2025（175以上）: https://www.nvidia.com/en-us/geforce/news/gamescom-2025-dlss-4-full-ray-tracing-game-announcements/ ; CES 2026（250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（NVIDIAのGTC 2026ニューラルテクスチャデモについて、6.5 GBから970 MBへ）: 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の製品版、2026年2月26日: https://devblogs.microsoft.com/directx/shader-model-6-9-retail-and-more/ ; cooperative vectorについて: https://devblogs.microsoft.com/directx/shader-model-6-9-and-the-future-of-cooperative-vector/ ; GDC 2026での計画: https://windowsreport.com/microsoft-showcases-neural-rendering-and-dxr-2-0-plans-for-directx-at-gdc-2026/

特殊用途リンク

- Global Trade Alert（中国の2026年第49号公告について）: 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（無人地上プラットフォームの基準値）: 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が外国製ドローンと部品をCovered Listに追加、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/

シリコン

- 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 ; MITのチップのページ（4K30で78 mW）: 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年: 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 ; 補足資料（モジュール別の複雑度）: https://openaccess.thecvf.com/content/CVPR2025/supplemental/Jia_Towards_Practical_Real-Time_CVPR_2025_supplemental.pdf
- NNVCとSADL整数推論ライブラリ: https://arxiv.org/abs/2309.05846
- Tencent Canghai V2、2026年5月27日（Shanghai Securities News、10jqka経由）: https://stock.10jqka.com.cn/20260527/c677016020.shtml ; 機能の概要: https://www.besthub.dev/articles/tencent-s-canghai-v2-chip-enters-mass-production-what-the-new-capabilities-mean-for-video-encoding-bcfc18dbfa29
- Chips&MediaのAV2デコーダー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のライセンス、2026年6月17日: https://www.edge-ai-vision.com/?p=58270
- Allegro DVT、AV2デコーダー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、AV2対応のVC9800D、2026年6月10日: https://www.businesswirechina.com/en/news/63176.html

標準と特許

- MPEG第155回会合のプレスリリース（VVC後継の共同提案募集）、2026年8月5日: https://www.mpeg.org/wp-content/uploads/2026/08/ag3w254SN26474Press-Release-MPEG155.pdf
- MPEG第152回会合のプレスリリース（エビデンス募集の結果、VVCとVSEIの新版）、2025年10月30日: https://www.mpeg.org/wp-content/uploads/2025/11/Press-Release-MPEG152.pdf ; ACM SIGMMのMPEGコラム: 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後継（H.267）の特許ランドスケープ: https://greyb.com/blog/exclusive-video-codec-landscape/ ; AV2のライセンスとAlibabaのポートフォリオ売却について: 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とVVC Advanceのロイヤルティ料率（2026年7月1日時点）: https://accessadvance.com/wp-content/uploads/2025/07/MCBA-Program-Overview-September-2026.pdf

お金と企業

- InterDigitalによるDeep Renderの買収、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/

学習型圧縮 · 再生 × シリコン

2026年のニューラル符号化

コーデックの代わりではなく、その前後で実用化が進む

本レポートの主要な数値には、すべて出典と日付を付した。著者が導いた数値は「推定」、販売元だけが公表している数値は「ベンダー発表」と表示している。

タン・シュアイ

テクノロジスト、創業者 · AI研究開発者

tanshuai.com · i@tanshuai.com

公開版 · データは2026年11月12日時点