RaySee RC1 — Smarter, Faster (VPU Chip) Product Overview

Overview

The RaySee RC1 is a VPU chip supporting AV1 hardware encoding. Adopting a dedicated ASIC pipeline architecture, it improves encoding density by 20 times compared to CPU-based solutions while maintaining broadcast-grade image quality. The chip integrates an independent AI Tensor Core (15–18 TOPS) to enable real-time ROI quality enhancement and content-adaptive encoding. Additionally, it incorporates a 2D video processing engine that can perform scaling, cropping, overlaying, and color space conversion without involving the host CPU. Whether in H.264, HEVC, or AV1 formats, the RC1 achieves sub-frame-level ultra-low latency encoding/decoding, providing the ultimate performance efficiency ratio for applications like live streaming, cloud gaming, and surveillance cameras.

Key Features

Basic Specifications

- Manufacturing Process: FinFET
- Interface: PCIe 4.0 x4 (PCIe 3.0/5.0 compatible), SR-IOV support (up to 8 virtual functions)
- Form Factor: U.2, AIC (HHHL), M.2 2260/2280, multi-card parallel expansion support
- Typical Power Consumption: 8W – 20W (depending on load and form factor), power states D0–D3 support
- Operating Temperature: 0°C to 45°C
- Reliability: 24/7 continuous operation, MTBF > 200,000 hours

Encode/Decode Throughput

- Encode: 32× 1080p30 / 8× 4Kp30 / 2× 8Kp30 (simultaneous)
- Decode: 48× 1080p30 / 12× 4Kp30 / 4× 8Kp30
- Transcode: 20× 1080p30 (HEVC ↔ AV1) or 10× 4Kp30 (HEVC ↔ AV1)
- Bitrate Range: 64 kbit/s to 700 Mbit/s (independently adjustable per stream)
- Resolution Range: 32×32 to 8192×5120 (supports arbitrary aspect ratios, even non-standard resolutions)
- Encoding Quality: Compared to x265 medium, VMAF is 2–5 points higher at equivalent bitrate; more than 20x faster than SVT-AV1

AI and Encoding Capabilities

- AI Compute Power: 15–18 TOPS (INT8), supports major frameworks (TensorFlow, PyTorch, ONNX)
- AI Applications: ROI region detection, scene classification, content-adaptive encoding, super-resolution pre-processing, noise detection
- Encoding Formats: H.264 (BP/MP/HP), HEVC (Main/Main10), AV1 (Main), JPEG, HEIF, AVIF
- Decoding Formats: H.264, HEVC, VP9, JPEG
- Color Depth/Sampling: 8bit/10bit, YUV 4:2:0 (encode); 4:4:4 optional (decode)
- HDR: HDR10, HDR10+, HLG, SEI metadata passthrough support
- Low Latency: Sub-frame-level latency, encode latency < 2ms (1080p), end-to-end < 10ms
- Bitrate Control: CBR, VBR, CRF, capped CRF, ABR ladder output
- GOP Structure: Flexible I/P/B frame configuration, forced IDR insertion, scene change detection

Advanced Video Processing Engine

- 2D Processing: Scaling (bilinear/bicubic/32-tap), cropping, padding, overlay, YUV and RGB conversion, color gamut conversion
- ROI Intelligent Encoding: AI-assisted detection of faces/license plates/motion regions, prioritized bitrate allocation, background compression saving 30%–50% bandwidth
- OSD Hardware: 8 independent OSD layers, text/image overlay
- Subtitles: CEA-608/708 subtitle passthrough and rendering
- Health Monitoring: SMART logs (temperature, power consumption, PCIe error count), temperature sensors x4

Software Integration and Ecosystem

- Framework Support: FFmpeg, GStreamer, LibXcoder API, NETINT Bitstreams™ firmware
- Operating System: Linux (Ubuntu, CentOS, Debian), Windows (10/11, Server 2019+)
- Virtualization: SR-IOV support, direct VPU instance assignment to each VM/container
- API: C/C++, Python bindings, integrated interface for encode/decode/scaling/AI inference
- Toolchain: Performance monitoring tool (nimble), firmware upgrade tool, quality comparison tool (PSNR/SSIM/VMAF)
- Plug and Play: Pre-loaded Bitstreams firmware, no additional driver configuration required

Target Applications

Live Streaming Platforms (20x CPU transcode density, broadcast-grade quality)

Cloud Gaming (Ultra-low latency + ASIC encoding + GPU rendering)

- Surveillance/Security (AI object detection + near lossless compression)
- Broadcasting/VR (8K HDR multi-bitrate ladder output)