



AIVATAR: AI and video codec enhancements for realistic avatar telepresence

Challenge

XR telepresence demands **high-quality video and low latency on devices** with constrained processing power and limited network bandwidth. Existing solutions either compromise video quality (AVC/H.264) or require computationally intensive codecs (e.g., HEVC/H.265, AV1), unsuitable for real-time use on consumer devices.

Technical approach

→ Low-complexity video codec enhancement

- MPEG-5 Part 2 LCEVC applied as an enhancement layer on AVC (H.264).
- Achieves significant bitrate savings and compression efficiency without increasing encoding complexity.

→ Edge AI-Superresolution and noise removal model

- Custom superresolution model running in the client-side AIVATAR Player.
- Improves sharpness and detail in avatar video streams.
- Reduces compression artifacts and enhances color fidelity.

Conceptual architecture

The system is integrated into **SPIRIT's Real-Time Animation and Streaming of Realistic Avatars** use case. The LCEVC-enhanced encoder is deployed within the split rendering pipeline, while the decoder and AI enhancement run in a new AIVATAR player on consumer laptops, supporting **Fluendo's goal of making XR accessible in daily work environments**.

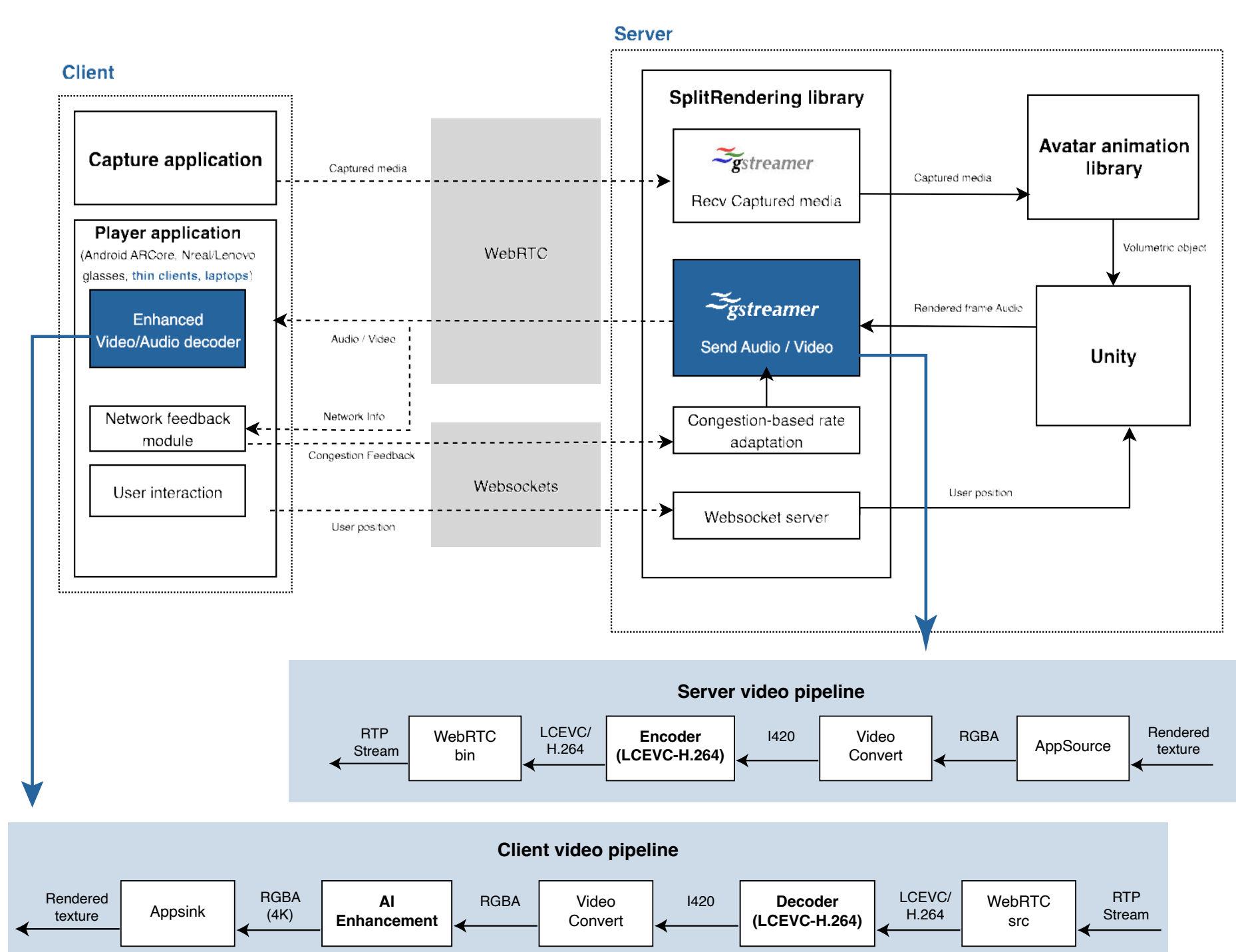


Fig.: Conceptual architecture of AIVATAR integrated in the SPIRIT platform and the real-time animation and streaming of realistic avatars use case.

Key results

↑ **30%** better compression

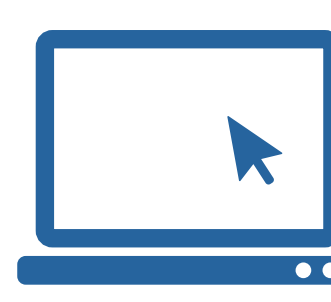
Reduces H.264 bandwidth needs while **keeping video quality high**, matching H.265, without changing codecs.

↓ **38%** lower CPU usage

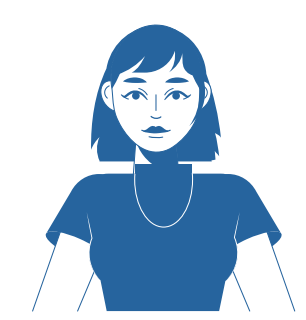
Compared to H.265, delivers efficiency with far less processing power, **cutting operational costs** for large-scale deployments.



<200 ms latency with 4K output for real-time XR.



Client-side AI Superresolution, optimized for low-power devices like laptops.



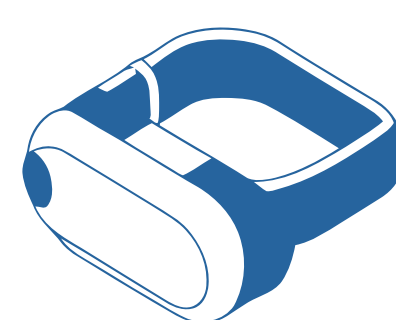
Sharper avatars with reduced artifacts and enhanced facial detail.



AI-enhanced 4K streaming for realistic avatar presence.



Deployable on consumer devices via GStreamer and WebRTC.



Integrated with SPIRIT for scalable immersive telepresence.



Enhance the quality of experience by delivering high-resolution and quality video over any network, regardless of bandwidth constraints.



Carbon reduction by streaming and encoding with an energy-efficient, greener codec.



Legal and technical compliance thanks to a business-ready solution with patent reporting and licensing, fully compatible with your tech stack and infrastructure.

References (selection)

MPEG-5 Part 2: Low Complexity Enhancement Video Coding (LCEVC). ISO/IEC 23094-2.
Netflix VMAF (Video Multi-Method Assessment Fusion). <https://github.com/Netflix/vmaf>
Fluendo Codec-Pack. 2025 <https://fluendo.com/products/fluendo-codec-pack/>
Fluendo AI plugins. 2025 <https://fluendo.com/products/fluendo-ai-plugins/>

Input 720p



AI-super-resolution (4K)

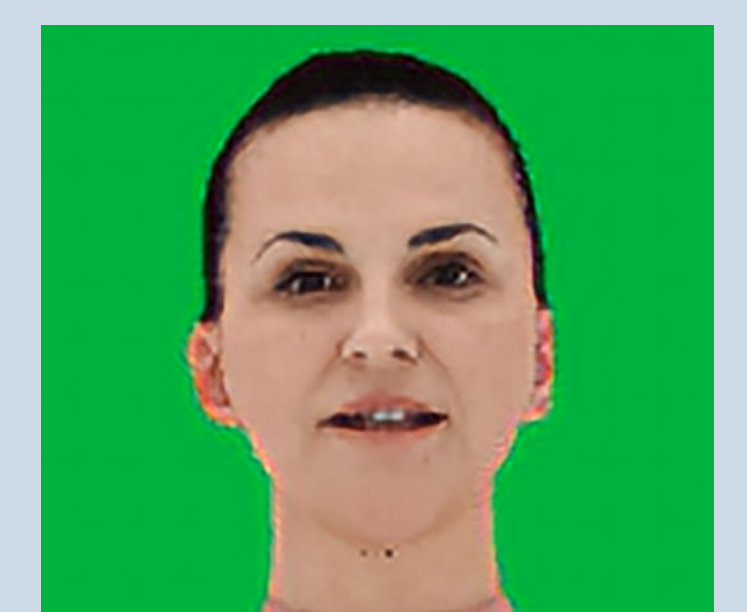
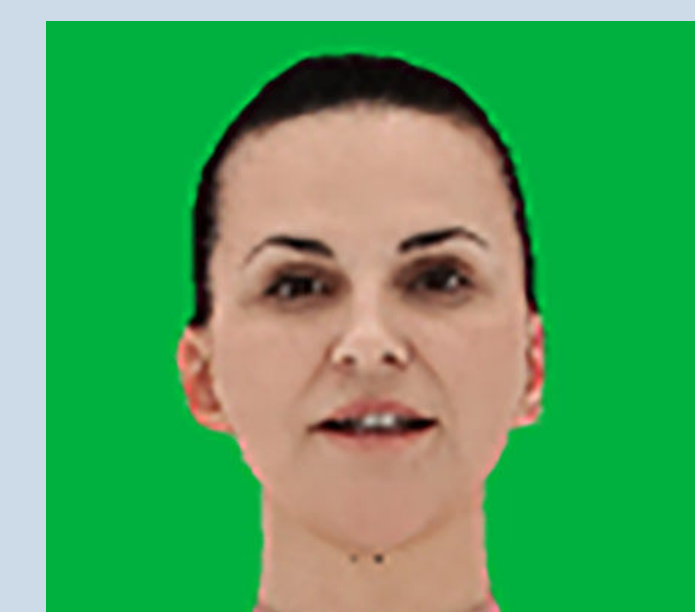


Fig.: Preliminary results comparing original 720p avatar video frames from the SPIRIT use case with the AI-based super-resolution upscaling from the original 720p to 4K.

Added value and impact

AIVATAR efficiently balances compression and processing demands, **enabling affordable, high-quality XR communication** even on mainstream hardware, empowering scalable deployment in enterprise, education, and healthcare. This work complements **Fluendo's LYNX and Raven innovation projects**, as well as the **Fluendo Codec Pack and Fluendo AI Plugins** products, advancing codec efficiency, edge AI, and sustainability in immersive communications.



Contact:
Sergio Sánchez-Carballido
Fluendo
ssanchez@fluendo.com



Funded by the European Union



Project funded by
Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra
Federal Department of Economic Affairs,
Education and Research EARE
State Secretariat for Education,
Research and Innovation SERI



UK Research and Innovation

Learn more:

