Live Demonstration: Kvazzup 4K HEVC Video Call

Joni Räsänen, Marko Viitanen, Jarno Vanne, Timo D. Hämäläinen
Laboratory of Pervasive Computing
Tampere University of Technology, Finland
{joni.rasanen, marko.viitanen, jarno.vanne, timo.d.hamalainen}@tut.fi

***Abstract*—This paper describes a demonstration setup for an end-to-end 4K video call with Kvazzup open-source HEVC video call application. The Kvazzup clients are installed on a desktop and a laptop computer powered by Intel 22-core Xeon and Intel 4-core i7 processors, respectively. The proposed two-way peer-to-peer video call setup is shown to support 2160p30 video stream from the desktop to the laptop and 720p30 stream in the reverse direction.**

Keywords— peer-to-peer video call, High Efficiency Video Coding (HEVC), 4K, Kvazzup video call application, Real-time Transport Protocol (RTP)

I. Introduction

Global IP video traffic is estimated to increase threefold between 2016 and 2021 [1]. Video communication is strongly fostering this growth through more advanced multimedia devices, faster IP networks, and widely adopted Internet video telephony services such as Skype [2] with 300 million active users. Two-party video calls and multi-party video conferencing are particularly gaining traction in the business sector where the relative growth is expected to be as fast as that of the global IP traffic [1].

The foreseen growth in video traffic can be mitigated by taking the latest video coding standard, *High Efficiency Video Coding* (*HEVC*) [3], into use in video communication. Even though there exist a large variety of open-source applications for video calling [4]-[10] only two of them support HEVC: BareSIP [9] and our Kvazzup [10]. However, BareSIP is only a *Session Initiation Protocol (SIP)* [11] User-Agent with audio and video support. Hence, Kvazzup is currently the only fully-fledged open-source application that supports HEVC video calls [12]. This work upgrades Kvazzup calls to 4K resolution, which is not provided, e.g., by Skype for Business [13].

The remainder of this paper is organized as follows. Section II presents the optimizations needed to make Kvazzup compatible with 2160p30 video calls. Section III describes how these new features are validated in practice as a part of a two-way peer-to-peer video call demonstrator.

II. Optimized Video Processing in Kvazzup

Kvazzup is an HEVC-based video telephony system written in C++ and built on Qt framework, version 5.11 [14]. It makes use of four open-source tools: Kvazaar [15] for HEVC encoding, OpenHEVC [16] for HEVC decoding, Opus [17] for audio coding, and Live555 [18] for managing *Real-Time Transport Protocol* (*RTP*) traffic.

The main components of Kvazzup are: 1) GUI; 2) Call Control; 3) Call Initiation; 4) Media Delivery; and 5) Media Processing [12]. The latter two components are optimized in this work to enable 2160p30 HEVC video processing on existing high-end processors.

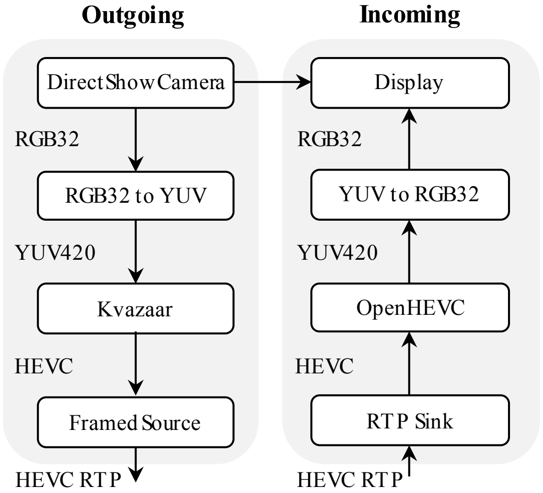

Figure 1. Incoming and outgoing video filter graphs in Kvazzup client.

In Kvazzup, the Media Processing component uses a filter architecture for video. Fig. 1 shows the implemented video filter graph of a Kvazzup client. The other Kvazzup client acts as a recipient of the outgoing HEVC RTP stream and a sender of the incoming HEVC RTP stream.

A. Outgoing Video Stream

The outgoing video stream is captured by our new DirectShow Camera filter that selects the camera resolution through DirectShow API. Kvazzup accepts RGB32 and YUV420 video formats from the user camera. If the input format is RGB32, Kvazzup converts it to YUV420 for the subsequent filter stage. The conversion is made by a dedicated RGB32 to YUV filter that is SSE4 optimized for 2160p30 processing.

Kvazzup uses Kvazaar to encode YUV420 format to HEVC video. The Kvazaar filter outputs either one or multiple HEVC slices per picture and Live555 sends the slices across the network. A single-threaded Live555 is not able to copy 4K video in real time. Therefore, pictures are encoded into multiple slices, which are copied by the Framed Source filter to Live555 buffer. In Live555, Event Triggers manage the communication with the Framed Source filter and H265VideoStreamDiscreteFramer encapsulates the slices into packets (≤ 1498 bytes) for RTP transmission.

B. Incoming Video Stream

Live555 receives an incoming RTP stream from another Kvazzup client and RTP Sink filter stores the HEVC slices into memory. The OpenHEVC filter combines the slices to pictures and decodes them to YUV420 format that is

This work was supported in part by the European Celtic-Plus Project Virtuose and the Academy of Finland (decision no. 301820).

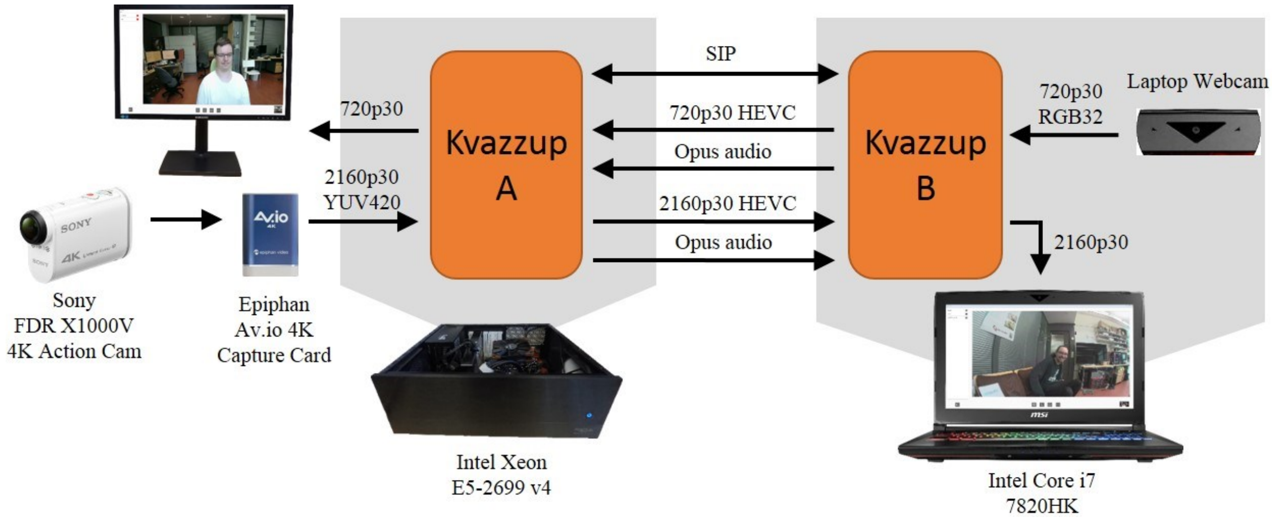

Figure 2. Demonstration setup for Kvazzup peer-to-peer 4K HEVC video call.

TABLE I. Platform Features and Client Settings

Feature	Desktop (Kvazzup A)	Laptop (Kvazzup B)
Processor	Intel Xeon E5-2699 v4	Intel Core i7 7820HK
Cores	22	4
Frequency	2.20 GHz	2.90 GHz
Memory	64 GB	32 GB
Compiler	64-bit MinGW	64-bit MinGW
System	64-bit Windows 10	64-bit Windows 10
Camera resolution	2160p30	720p30
Profile	ultrafast	ultrafast
QP	37	37
Intra period	64	64
Kvazaar	16 threads	3 threads
YUV to RGB32	3 threads	2 threads
OpenHEVC	4 threads	2 threads
RGB32 to YUV	1 thread	1 thread

converted to RGB32 for video rendering with Qt. The conversion is made by a dedicated YUV to RGB32 filter that is AVX2 optimized and multi-threaded for 2160p30 processing. The Display filter supports rendering the video using software rasterization with Qt QWidget or OpenGL [19] with Qt QOpenGLWidget. The video can be drawn as a part of the UI or in fullscreen.

III. Kvazzup 4K30p Video Call Demonstrator

Fig. 2 depicts the individual components of the demonstrator with two Kvazzup clients: Kvazzup A on the desktop and Kvazzup B on the laptop. The features of the applied platforms and client settings are detailed in Table I.

Kvazzup A receives raw 2160p30 YUV420 video from Sony FDR x1000V 4K Action Cam via Epiphan AV.io HDMI capture card. It encodes the raw video and sends 2160p30 HEVC output over the network to Kvazzup B that decodes the video for playback on a 4K laptop screen. Correspondingly, the laptop camera provides raw 720p30 RGB32 video to Kvazzup B that encodes and sends it to Kvazzup A for decoding and playback in a separate display. The computers are connected with an Ethernet cable.

The demonstration allows users to interact with Kvazzup clients. A user can start and end a call and disable audio or video during a call. The audio is played through headphones and the video can be seen as a part of the UI or at full resolution in a fullscreen mode by double clicking the view.

In this demonstrator, the end-to-end latency from Sony Action Cam to the laptop screen is shown to be around 500 ms, and the respective latency from the laptop camera to the separate display around 700 ms. The bit rates for 2160p30 and 720p30 HEVC videos are around 0.5 - 0.8 Mbit/s and 0.1 - 0.2 Mbit/s under typical call conditions, i.e., a single person with a static background. The low bit rate is the result of 1) relatively little motion in the video call; and 2) the usage of Kvazaar HEVC encoder with appropriate coding parameters, e.g., having intra period of 64 and the Quantization parameter (QP) value of 37.

References

[1] Cisco, *Cisco Visual Networking Index: Forecast and Methodology*, 2016-2021, Sep. 2017.

[2] *Skype* [Online]. Available: http://www.skype.com/

[3] *High Efficiency Video Coding*, document ITU-T Rec. H.265 and ISO/IEC 23008-2 (HEVC), ITU-T and ISO/IEC, Apr. 2013.

[4] *Linphone* open-source voip software [Online]. Available: http://www.linphone.org/

[5] *Jitsi* [Online]. Available: https://jitsi.org/

[6] *Ring* [Online]. Available: https://ring.cx/

[7] *Tox* [Online]. Available: https://tox.chat/

[8] *Wire* [Online]. Available: https://wire.com/

[9] *BareSIP* [Online]. Available: http://creytiv.com/baresip.html

[10] *Kvazzup* [Online]. Available: https://github.com/ultravideo/kvazzup

[11] IETF RFC 3261 SIP: Session Initiation Protocol [Online]. Available: https://www.ietf.org/rfc/rfc3261.txt

[12] J. Räsänen, M. Viitanen, J. Vanne, and T. D. Hämäläinen."Kvazzup: open software for HEVC video calls," *in Proc. IEEE Int. Symp. Multimedia*, Taichung, Taiwan, Dec. 2017.

[13] Skype for Business client video resolutions. [Online]. Available: https://docs.microsoft.com/en-us/skypeforbusiness/plan-your-deployment/clients-and-devices/video-resolutions

[14] *Qt* [Online]. Available: https://www.qt.io/

[15] *Kvazaar HEVC encoder* [Online]. Available: https://github.com/ultravideo/kvazaar

[16] *OpenHEVC HEVC decoder* [Online]. Available: https://github.com/OpenHEVC/openHEVC

[17] *Opus Codec* [Online]. Available: http://opus-codec.org/

[18] *LIVE555* [Online]. Available: http://www.live555.com/

[19] *OpenGL* [Online]. Available: https://www.opengl.org/