

# Prospects for Live Higher Resolution Video Streaming to Mobile Devices: Achievable Quality across Wireless Links

A.O Adeyemi-Ejeye<sup>1</sup>, M. Alreshoodi<sup>2</sup>, L. Al-Jobouri<sup>3</sup>, M. Fleury<sup>3\*</sup>

<sup>1</sup>*Univ. of Surrey, London, U.K.*

<sup>2</sup>*Qassim Univ., Buraydah, Saudi Arabia*

<sup>3</sup>*University of Essex, Colchester, United Kingdom*

\*Communications should be addressed to M. Fleury: email fleum@essex.ac.uk, tel: +44 1206 873 333, fax: +44 1206 872 900

From a review of the literature and a range of experiments, this paper demonstrates that live video streaming to mobile devices with pixel resolutions from Standard Definition up to 4k Ultra High Definition (UHD) is now becoming feasible by means of high-throughput IEEE 802.11ad at 60 GHz or 802.11ac at 5 GHz, and 4kUHD streaming is even possible with 802.11n operating at 5 GHz. The paper, by a customized implementation, also shows that real-time compression, assisted by Graphical Processing Units (GPUs) at 4kUHD, is also becoming feasible. The paper further considers the impact of packet loss on H.264/AVC and HEVC codec compressed video streams in terms of Structural Similarity (SSIM) index video quality. It additionally gives an indication of wireless network latencies and currently feasible frame rates. Findings suggest that, for medium-range transmission, the video quality may be acceptable at low packet loss rates. For hardware-accelerated 4kUHD encoding, standard frame rates may be possible but appropriate higher frame rates are only just being reached in hardware implementations. The target bitrate was found to be important in determining the display quality, which depends on the coding complexity of the video content. Higher compressed bitrates are recommended, as video quality may improve disproportionately as a result.

**Keywords** 4kUHD; emerging wireless technology; video streaming; real-time compression

## 1 Introduction

Users' expectations of video display on mobile devices are growing [7]. In fact, they may expect similar quality video to that experienced on home televisions (TVs), as they become more aware of the advanced technology deployed in tablets and smartphones [49]. TV displays are moving away from Standard Definition (SD) (720 (horizontal)  $\times$  480 (vertical) or  $\times$  576 pixels/full frame) towards higher resolutions such as High Definition (HD) (1280  $\times$  720 or full 1920  $\times$  1080 pixels/frame). With each increase in resolution, the work demanded of hardware engines for compression and display scales up in a super-linear fashion.

Nevertheless, System-on-Chips (SoCs) have been developed for HD video codecs [52] and are now under development for 4k Ultra High Definition (UHD) resolution ( $3840 \times 2160$  pixels/frame,  $16 \times 9$  aspect ratio) at 30 frames/s (fps) [45]. In fact, not only HD and 4kUHD video but also 3D in the 'video plus depth' format are of interest to mobile viewers [4].

8kUHD ( $7680 \times 4320$  pixels/frame) video satellite broadcast has been developed by the state broadcaster of Japan, NHK [70]. In addition, there is a more general need for 8kUHD so as, through zooming, cropping, and down-sampling, to better produce 4kUHD sequences for digital cinema. The results of these developments may eventually enable mobile device reception, given the incremental improvements upon existing transmission technologies, namely dual-polarized  $2 \times 2$  (or higher) Multiple Input Multiple Output (MIMO) antennas, with higher-order 1024- and 4096-Quadrature Amplitude Modulation (QAM), combined with Low Density Parity Check (LDPC) channel coding (at a cost in latency). Though, this paper confines itself to the potential for higher resolution video up to 4kUHD resolution for mobile devices, this could be a step towards 8kUHD video, which increases the realism still further because of its  $100^\circ$  Field of View (FoV) from  $60^\circ$  for 4kUHD.

The ideal viewing distance for SD video is 7.1 multiplied by the picture height [67] or equivalently when the angular resolution is 30 cycles/degree, resulting in a score of five on a five-point subjective quality scale [80]. Therefore, SD displays (and still lower resolutions such as Common Interchange Format (CIF) at  $352 \times 240$  pixels/frame) may result in unsatisfactory viewing because users of mobile devices normally position themselves closer to their screens and, thus, can see scan lines. However, for HD video the distance at which scan lines become invisible is only 3.1 multiplied by the picture height [67]. Therefore, HD resolutions are more suited to the closer viewing distances of mobile devices.

Further, in the 2014 Broadcast Asia conference 4kUHD TV was actually broadcast to mobile devices equipped with embedded Digital Video Broadcast terrestrial second-generation (DVB-T2) receivers (DVB, 2014). Video compression was through an High Efficiency Video Coding (HEVC) standard codec, which can achieve up to 50% bitrate savings over a codec of its predecessor [58], the H.264/Advanced Video Coding (AVC) standard [85]. HEVC is targeted towards higher resolutions, as evidenced by the availability of the large  $64 \times 64$  Coding Tree Unit (CTU) (a restriction to a smaller-sized CTU results in a reduction in compression performance [58]).

In early subjective tests with expert viewers [34] for three 4kUHD sequences, it was found that the average rate-distortion (RD) gain over H.264/AVC was 66.5% (measured as Bjøntegaard-Delta (BD) [11] -Mean Opinion Score (MOS)) and objectively the RD gain was

44.4% BD-PSNR. In [31], the BD-PSNR gain over the emerging (v.1 2016) Alliance for Open Media (AOM)/AVI encoder (in two-pass mode) was found to be about 38.4% using the HEVC Test Model (HM) codec implementation. However, as with its predecessor, that is the VP9 codec [53], AOM/AVI should be seen as part of the WebM open software project and may well find its way into web browsers on mobile devices, given savings in licensing fees over HM.

In our paper, when a software HEVC is employed, it is the open-source x265 encoder (in one-pass mode). Tests were conducted by the authors of [31] over a set of video sequences satisfying the HM test conditions [14] (though excluding screen content sequences). The x265 encoder (in two-pass mode) was reported [31] as having a 13.9% BD-PSNR loss relative to the HM codec. However, the commercial version of the HM codec, the HHI-HEVC codec, was reported [31] as having a 300% speedup over the HM codec. By the same speedup measure, HHI-HM's speedup was 3.4% relative to the x265 codec. Therefore, for software or hardware accelerated implementations of HEVC, the x265 codec implementation represents a good compromise in terms of RD performance and speed (especially if in one-pass mode) and, hence, is appropriate for operation on mobile devices.

However, while an HEVC codec is suited to the storage of higher resolution video, an HEVC encoder is also significantly more complex [13] than an H.264/AVC encoder. Therefore, without a suitable hardware implementation, transmitted frame rates for HD and UHD video may be restricted to 30 fps or below. Particularly for mobile devices, energy consumption at 24 fps or 30 fps is also an obstacle. On the other hand, frame rates of 60 fps or more may eventually be required [25] [80] to reduce flicker and motion blur at 4kUHD's wider FoVs (60° for 4kUHD rather than the 30° of HD). However, the prospect of higher frame rates may be some way off in time because currently 4kUHD hardware designers and implementers of HEVC encoders for mobile devices need to make rapid coding mode decisions to reduce the coding overhead, even at 30 fps. Rapid coding mode decisions in turn result in 34.6% higher bitrates [45] for the equivalent quality achieved in the HM-13.0 codec's Low Delay P mode.

One of the contributions of this current paper is to show how hardware-assisted implementations show some prospects of approaching video rate encoding. The authors of [87] ported four major H264/AVC encoding processes to the Compute Unified Device Architecture (CUDA) for parallel processing on Graphical Processing Units (GPUs). They then employed data localization to enhance thread performance on a GPU. Nevertheless, the approach was unsuitable for resolutions beyond HD and for GPUs with constrained local memory. The implementation also suffered from excessive latency when applied to real-

time/live encoding/decoding because data had to be transferred from and to the CPU main memory before accessed by other processes. The main way the 4kUHD CUDA implementation described in this paper offers an improvement is by its support for zero-copy memory, though other changes were also made.

An important consideration, apart from that of increasing the speed of the codec encoders so that they approach real-time/live encoding rates, is the need to be able to effectively stream compressed video over wireless links, even though these are error prone. Unfortunately, owing to the predictive nature of video coding video streams are sensitive to packet loss, when transported by the UDP protocol for live or real-time streaming. In [6], it was actually shown that increased HEVC compression of 4kUHD video causes greater loss of video quality compared to H.264/AVC, once coding gain is adjusted for. An earlier paper [66] made the same point when comparing MPEG-2 to H.264/AVC. The relative effect occurs because the increased compression of the later codecs makes them more vulnerable to packet loss, as the loss of a packet has more of an impact on the video quality. In general, though transmitted quality can be improved by application-layer Forward Error Control (FEC) or other forms of error control, these lead to increased bandwidth usage and/or increased latency. In fact, the long delays of large-block rateless coding was pointed out in [86]. Consequently, the authors of [86] introduced a framework for HD video frame transmission to mobile devices with adaptive low-block size FEC.

Therefore, in this paper, we examine at low error rates (less than 1%) the resulting video quality of streaming high resolution video over short-range wireless links, without error protection at the application layer so as to reduce latency. However, standard ‘previous frame’ error concealment is still applied at the video decoder. To also facilitate real-time/live streaming we have employed the IP/UDP/Real-time Transport Protocol (RTP) suite of protocols [64], either with direct packetization of the compressed video content or indirect packetization of an MPEG2-Transport Stream (TS) [10], as may be preferred for IPTV. TCP-based pseudo-streaming via some form of HTTP Adaptive Streaming (HAS) [79] can be applied to wireless communication. However, when an error occurs packets are retransmitted, which may lead to significant delays due to buffer underflow, resulting in freeze frames and user discomfort [73]. All the same, a demonstration of live 4kUHD wireless video streaming over HTTP Live Streaming (HLS), the Apple version of HAS, has been provided in [5]. The popularity of the HAS adaptive streaming process has led to proposals such as that in [39] to stabilize the end-buffer length to ensure smooth delivery while at the same time avoiding too frequent changes in the bitrate, which can have a disconcerting effect on the viewer. In

addition, there is the issue of arbitrating between the bandwidth allocation between multi-users streaming from the same server, which [89] proposes to resolve through game theory, which can be employed in a non-cooperative manner [88], as well as a cooperative manner.

High-throughput wireless technologies are, as is shown in this paper, a ‘bright spot’ for transmission of high resolution video, including 4kUHD. As reviewed in Section 2, since the acceptance of the draft standard IEEE 802.11 part n in 2007 [62], suitable high-throughput technologies have proliferated, including IEEE 802.11ac, building on IEEE 802.11n and operating at 5 GHz, and IEEE 802.11ad operating at 60 GHz at short ranges with directional antennas [27]. IEEE 802.11ac and IEEE 802.11ad are also briefly reviewed in Section 2.

In regard to assessing video quality, studies have shown [38] that the widely used Peak Signal-to-Noise Ratio (PSNR) and Mean-Squared Error (MSE) are flawed in differentiating structural contents of video frames because different types of impairments can occur and still have the same MSE value. This paper, therefore, employs the efficiently-computed Structural Similarity (SSIM) index [84], especially since structural impairments will be easily noticed [54] at higher-resolutions, diminishing the end-user’s Quality of Experience (QoE).

In summary, the contributions of this paper are as follows:

- The reader is guided across the different wireless LAN (WLAN) technologies that can now enable mobile streaming access to higher resolution video, including 4kUHD.
- Experiments and implementations of uncompressed and compressed video are analysed, including limitations of range and throughput.
- For live streaming, real-time encoding is required. The paper introduces practical work in hardware acceleration of encoding, including that of the authors that may eventually complement the first chipsets recently reported or announced for 4kUHD encoding.
- The paper also summarizes experiments across high-throughput WLAN links to determine network latencies, video quality in response to packet loss, and relative codec response, from either H.264/AVC or HEVC codecs.

The remainder of this paper is organized as follows. Section 2 concisely reviews contemporary high-throughput indoor wireless standards and their appropriateness for streaming HD and 4kUHD video. The following Section 3 examines related research on wireless broadcast and streaming both of uncompressed and compressed video at high resolutions. Section 4 considers the feasibility of hardware acceleration of 4kUHD video with reference to a preliminary investigation by the authors. Section 5 then reports on streaming experiments for 4kUHD across each of the IEEE high-throughput standards, from 802.11n to

802.11ac and 802.11ad. Finally, Section 6 draws some conclusions about the prospects for higher-resolution streaming to mobile devices.

## 2 Context

There is a number of high-throughput indoor wireless standards, as summarized in Table 1. Both IEEE 802.11n and IEEE 802.11ac employ the 802.11 Medium Access Control (MAC) in the 5 GHz band, as does IEEE 802.11ad at 60 GHz, and they can, therefore, be characterized as WiFi extensions. The IEEE 802.15.3c MAC [8] is organized as a piconet and does not employ the negotiable, distributed access of WiFi. Though the latter may be an advantage for IEEE 802.15.3c, it does not appear to have met the approval of the market place, as the task force ‘went into hibernation’ in 2009. Consequently, it is not considered further in this paper. As will be seen from Table 1, these wireless technologies have benefited from the widespread move to Orthogonal Frequency Division Multiplexing (OFDM) [72] (bringing increased resilience in the face of multipath interference) together with higher rates of modulation such as 256-QAM. Streaming over multiple channels through MIMO antennas along with space-time coding has also led to very significant rises in throughput.

The unlicensed 60 GHz band is open to the interference that occurs in the crowded unlicensed 2.4 GHz. However, more importantly, oxygen absorption at 10 dB/km peaks at around 60 GHz [68] and remedial amplification is restricted both by practical considerations and national standards to around 10 dBm. Therefore, it is usually assumed that reliable indoor propagation without beam-forming is restricted to 10 m, though depending on the extent of shadowing, the range might be extended to 20 m. A 5×5 (or 6×6) antenna array can compensate by about 25 dB [59], leaving aside material absorption, such as by the concrete in walls. Fortunately, the antenna size is much reduced at 60 GHz because of the millimetre waves, allowing patch antenna chips to be incorporated into high-end laptops.

Table 2, offers an interesting insight on the prospects of transmitting higher resolution video over high-throughput wireless. Notice that, though the majority of this paper is concerned with compressed video, uncompressed video in Table 2 avoids the additional latency arising at an encoder. Depending also on the expected compression ratio (approximately 100:1 for MPEG-2 video), the estimates in Table 2 can be scaled by the reader accordingly. Notice also that three 8-bit RGB color channels are assumed at the start of the Table but High Dynamic Range (HDR) of luminosity with a bit depth of 10-pixels per color channel is under active investigation [12] and, in fact, is specified in HEVC’s Main 10 profile, as are the higher

frame rates for higher resolution video discussed in Section 1. (In fact, 8-bit sampling was dropped from the UHD Rec. ITU-R BT.2020 standard.) 4kUHD and HDR are combined in [41] and, as previously mentioned, HDR is integrated within HEVC [28]. Those proposals are for broadcast TV, while for mobile devices such prospects seem long in the future.

**Table 1** Features of indoor high-throughput IEEE wireless standards (other standards include ECMA 387 and WIGWAM)

Standard	Features	Modulation	Spectrum band	Max. gross datarate
IEEE 802.11n (WiFi)	MIMO up to 4 streams, 40 MHz channels	OFDM	2.4 GHz, 5 GHz	288.9/600 Mbps (20/40 MHz channels)
IEEE 802.11ac (WiFi)	80-160 MHz channels, up to 256-QAM, MIMO up to 8 streams	OFDM	5 GHz	1 Gbps
IEEE 802.11ad (WiGigTM)	Beam-forming, up to 4 channels	BPSK/QPSK or OFDM	60 GHz	Approaching 7 Gbps
IEEE 802.15.3c (WirelessHDTM for AV OFDM)	Beacon with superframe MAC	OFDM (HSi and AV), SC-SDE	60 GHz	5.28 Gbps (SC-FDE), 25 Gbps (WirelessHD™ for < 1m)

AV=Audio Video, HSI = High Speed Interface, MIMO = Multiple Input Multiple Output antennas, OFDM = Orthogonal Frequency Division Multiplexing, QAM = Quadrature Amplitude Multiplexing, SC-FDE = Single-Carrier Frequency Division Equalization

Then in Table 3 (after [69]) estimates are provided for the required compressed data-rates, including next generation 8kUHD video. The Table simply takes into account the approximate 50% increase of the compression ratio with the emergence of each new standardized codec, which on historical evidence seem to appear every ten years [29]. This Table implies that either an advanced WiFi solution will be needed, as experimented with by the authors, or a converged network coordinating two wireless technologies will be required to meet the datarate needs of mobile devices. Advanced WiFi solutions are now briefly reviewed.

**Table 2** Data-rates for emerging and projected uncompressed HDTV formats according to frame rate and bits per channel per pixel for a progressive display

Pixels per line	Active lines per picture	Frame rate per second	No. of bits per pixel	Data rate (Gbps)
1280	720	24	24	0.531
1280	720	30	24	0.664
1440	480	60	24	0.995
1280	720	50	24	1.106
1280	720	60	24	1.327
1920	1080	50	24	2.488
1920	1080	60	24	2.986
1920	1080	60	30	3.732
1920	1080	60	36	4.479
1920	1080	60	42	5.225
1920	1080	90	24	4.479
1920	1080	90	30	5.599

**Table 3** Estimated data-rates for compressed HD and UHD formats according to frame rate and codec type (without HDR)

Codec	HD 30 fps (1920 × 1080) (Mbps)	4kUHD 30 fps (Mbps)	4kUHD 60 fps (Mbps)	8kUHD 60 fps (7680 × 4380) (Mbps)
MPEG-2	16-20	64-80	90-112	360-448
H.264/AVC	8-10	32-40	45-56	180-224
HEVC	4-5	16-20	22-28	80-112

## 2.1 IEEE 802.11n

IEEE 802.11n [61] operates both at 2.4 GHz and 5 GHz. It is an improvement on previous IEEE 802.11 standards such as parts a/b/g mainly by the addition of MIMO. When operating at 2.4 GHz or 5 GHz it can employ 20 MHz-wide or 40 MHz-wide channels, the latter reducing the latency of the narrower channels. The introduction of frame aggregation allows the full exploitation of the available data rates arising from these advancements at the physical layer. Two types of frame aggregation standards are available, namely MAC Protocol Data Unit (MPDU) and the MAC Service Data Unit (MSDU). These frame aggregation standards group several data frames into one large frame [78]. However, the advantages of jumbo frames may be dependent on support within a mobile device for gigabit Ethernet.

## 2.2 IEEE 802.11ac

Of feasible wireless technologies, IEEE 802.11ac standard of 2013 [9] provides a high-throughput wireless local area network (LAN) in the unlicensed 5 GHz and relative to 2.5 GHz uncrowded band. In theory, a single spatial channel has a maximum throughput of 867 Mbps. This is principally due to an increase in mandatory channel width to 80 MHz and up to 256-QAM. However, when 256-QAM is selected, the impact of noise increases significantly. The increased channel width that restricts 802.11ac operation to the 5 GHz band also results in shorter wavelengths at this frequency. Prior study reported in [21] considered the likely IEEE 802.11ac data-rates (in the earlier Wave 1 version) in an indoor environment. A number of smartphones support IEEE 802.11ac, such as Samsung's Galaxy S7 of 2016 which supports Wave 2 802.11ac. Wave 2 of 802.11ac includes Multi-User (MU)-MIMO antennas ( $2 \times 2$  configuration). MU-MIMO allows high datarate downloads even when several such devices are connected to the same access point.



### **2.3 IEEE 802.11ad**

The IEEE 802.11ad amendment to the IEEE 802.11 standard was ratified in 2012 [55], with the WiGig industry-supported standard integrated into it. The need for directional beam-forming at 60 GHz was mentioned earlier in this Section. IEEE 802.11ad introduced the concept of virtual antenna sectors to formalize the selection of antennas to focus a beam, with an antenna array varying in size between classes of device types. Phase-weighted arrays have been implemented [81] as patch antennas within radio transceiver chipsets.

Modulation over a single carrier can be relatively simple, given the high data rates anyway achievable at 60 GHz, Binary Phase-Shift Keying (BPSK) or Quadrature Phase Shift Keying (QPSK). Using Reed-Solomon (RS) channel coding distinguishes a low-power, single-carrier, mode over using LDPC codes. Multi-carrier OFDM is a higher-energy alternative, not suited to the mobile devices of this paper. Because directional beam-forming results in ‘deaf’ spots outside the beam, it is necessary to modify the IEEE contention-based MAC, which IEEE 802.11ad does by offering a choice of three solutions [18]. A polling-based solution is similar to IEEE 802.11’s Point Coordination Function (PCF), hitherto defunct, but now adapted to directional beams. A time-scheduled allocation of access, likewise is similar to IEEE 802.11’s Hybrid Coordination Function (HCF). Finally, 802.11ad also offers the usual CSMA/CA contention, provided a pseudo-omnidirectional beam pattern is employed. By 2016, notebooks such as the Acer TMP648-MG-789T had incorporated 802.11ad wireless interfaces as part of a ‘triband’ offering (with 2.4 and 5 GHz).

## **3 Related work**

Uncompressed 4kUHD video over optical networks has been evaluated in [74] and [32], with an application in digital cinema. The minimum requirement for uncompressed UHD video starts at 2.39 Gbps for 8-bit 4:2:0 chroma subsampling at 24 fps (with 3.98 Gbps for 8-bit 4:2:2 subsampling at 30 fps). All the same, uncompressed 4kUHD wireless transmission has been implemented [2] (with four wireless 60-GHz parallel HD channels, though the range was very short to reduce interference between the channels). An earlier transmission of uncompressed HD was reported in [76], which reduced delay from retransmissions by a combination of unequal error protection between MSBs and LSBs, multiple Cyclic Redundancy Checks (CRCs) and error concealment. Uncompressed video transmission avoids coding delay when streaming live video such as for sports. It is also attractive for applications such as streaming to a wireless monitor, TV, and projector [44], as delay when

streaming compressed video to such displays is noticeable, being as much as an average of 170 ms [17].

However, particularly for uncompressed video, an issue is how to deliver the video stream to the wireless access point. The research in [17] provided a proof-of-concept of transmission of HD video formatted for High Definition Multimedia Interface (HDMI) across a single-mode optical fiber, known as Radio over Fiber (RoF). The wireless video transmission was by horn antennas over no more than a third of a metre at 60 GHz or 100 GHz. It is also worth noting that 4 Gbps transmission of uncompressed video at 60 GHz was demonstrated in [51] using Orbital Angular Momentum (OAM) antennas, though, as with horn antennas, these may not be applicable to mobile devices. All the same, without streaming support, the storage requirements for uncompressed video are considerable. Thus, an uncompressed 4kUHD 30-minute video clip will require 537.75 GB of storage. Therefore, compressed formats are currently preferred, especially for mobile devices.

Compressed UHD video transmission over optical networks has also been extensively explored over the years. For example, a JPEG 2000 codec was used in [75] to compress/send or receive/decompress 4kUHD video in real time with visually-lossless quality at bit rates of 200-500 Mbps over a 1-Gbps IP network. Turning to wireless transmission, H.264/AVC compressed 4kUHD over IEEE 802.11n wireless operating at 5 GHz has been experimented with [3]. The chroma subsampling was varied between 4:2:0, 4:2:2, and 4:4:4 (the modes in the UHD Rec. ITU-R BT.2020 standard) at 20 Mbps with a Group of Pictures (GOP) length of 40 and frame structure of IPPP....

In [33], 4kUHD was split into a base layer (BL) and an enhancement layer (EL) to allow transmission of broadcast video as two streams, with a reported rate-distortion loss of 10% to 30%. In preliminary tests, scalable 4kUHD video allowed either full HD to be decoded or 4kUHD to be decoded at an average rate of 38 fps. It also allowed the 4kUHD scalable streams to be transmitted over separate wireless media (DVB-T2 and satellite are planned). However, only HEVC decoder rates were reported in [33], as a broadcast service was assumed.

This scalable coding scheme for transmission has recently been extended up to 8kUHD (7680 × 4320 pixels/frame) in [69] with two ELs for 4kUHD and three ELs for 8kUHD. Transmission is split between the BL and lowest EL through DVB-T2 broadcast and the higher two ELs by Long Term Evolution (LTE). Thus, without transmission over high-throughput IEEE 802.11ac or 802.11ad then scalable coding and network convergence may well become a necessity.

As mentioned in Section 2, 60 GHz transmission is vulnerable to oxygen absorption, severely limiting its effective range to below 10 m, which is why the feasibility of multi-hop transmission has been explored and shown by the authors of [1] to be achievable with low latency. However, streaming HD and 4kUHD streaming was at only 24 fps. The use of relays to increase the 60 GHz transmission range for uncompressed HD video broadcast at 1.5 Gbps is treated in a theoretical fashion in [42], though it is assumed that high-gain antenna are already used between camera and relay to increase the 60 GHz range up to 100 m.

Adaptive transmission of H.264/AVC HD video according to the Modulation and Coding Scheme (MCS) of IEEE 802.11ad was shown by simulation in [90]. In terms of choosing the MCS to maximize the video quality, the decision is based on an estimate from the built-in Channel State Information (CSI). There are ten non-OFDM MCSs in IEEE 802.11ad, one of which is selected in [90], along with a quantization parameter (QP), and a delivery deadline.

## **4 Accelerating video encoding**

Hardware acceleration will be necessary for any HD and particularly 4kUHD encoder implementation if real-time streaming is to be supported. As discussed below, H.264/AVC acceleration via GPUs for 4kUHD video is within reach, albeit reaching beyond 25 fps may be problematic, despite the desirability for display rates of 60 fps and higher, discussed in Section 1. In respect, to HEVC, Nvidia have released a number of GPU implementations in support of the Main 10 (8- to 10-bit with support for 4:2:0 chroma subsampling) and Main 12 (with additional support for 12-bit pixel depths) profiles. In addition, the GP1 chip from the GoPro Hero6 camera in 2017 is said to be able to encode 4kUHD at 60 fps.

After examining hardware acceleration, this Section then rounds off by considering optimization of HEVC codec software, through pruning of intra-coding modes and increased accuracy of rate control.

### **4.1 CUDA programming**

CUDA programming has proven effective as a way to transfer compute intensive components to a GPU. A source function is compiled, becoming a ‘kernel’. One or more kernels are subsequently downloaded to a GPU, which acts as a coprocessor to the CPU. Within the GPU, ‘threads’ are the mechanism for parallel processing, with the threads executing instances of the kernel code in parallel. Threads within a thread block can co-work with each other through the shared memory and can synchronize their execution to coordinate memory

access, though there is a limit to the number of threads within any one block. However, though in one CUDA example, [87], major H.264/AVC processing components were parallelized, memory transfer latency for resolutions beyond HD was not considered. Data localization allows threads to work efficiently on a GPU. However, in [87] data must pass via GPU memory before it can be accessed. The implementation by the authors of the current paper supports zero-copy memory, taking data from CPU memory and passing it directly to the GPU threads executing the kernels.

In addition, the GPU architecture of the Streaming Multiprocessor (SMX) architecture with CUDA compute capability enables dynamic parallelism. In dynamic parallelism, a kernel spawns new threads to receive CPU instructions. Consequently, a CPU no longer needs to issue new instructions when (say) a macroblock is to be divided into smaller units. Dynamic parallelism is now employed in the authors' implementation for: inter- and intra-prediction; entropy coding; and de-blocking filter components. Additionally, the number of intra-prediction modes has been drastically pruned. This and the other two main implementation contributions are now described in more detail.

## **4.2 CUDA implementation details**

### *Zero-copy memory mapping*

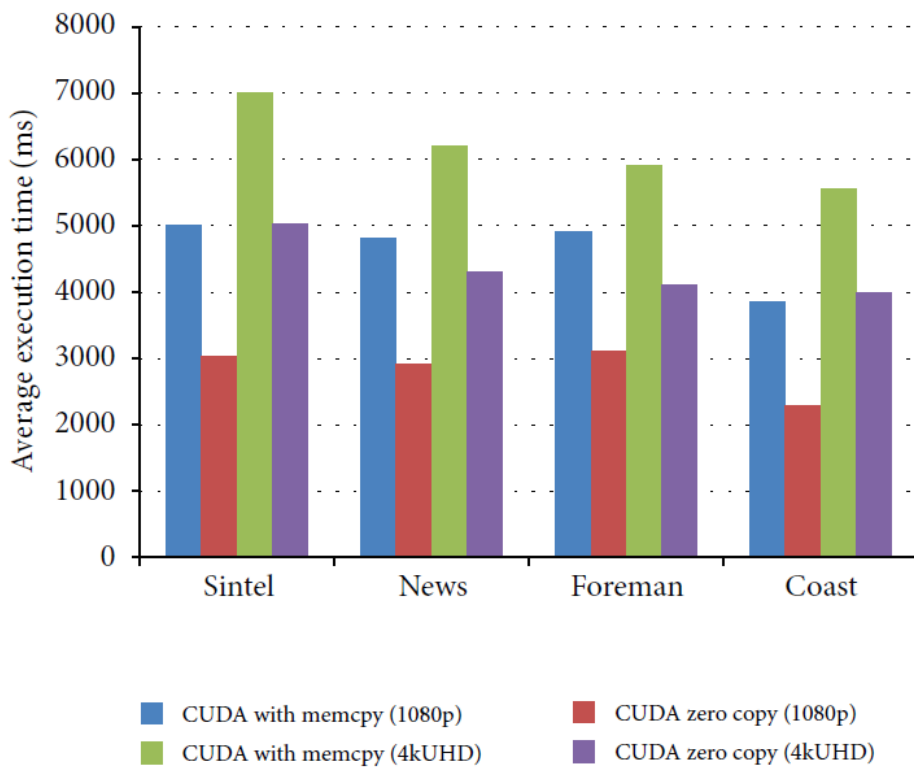
Irregular memory access patterns can be successfully handled by a conventional CPU due to its extensive memory hierarchy, which reduces access latency by caching. Unfortunately, the same patterns may hinder the efficient utilization of GPU memory bandwidth because of restrictions on access patterns if efficient memory performance is to be achieved.

In a normal CUDA application, memory is allocated as pageable. Consequently, memory is only allocated when needed. However, pageable memory results in an increase in memory access latency, as this memory will eventually page out and it will only be reallocated when needed. In the current implementation, to counteract the increase in memory latency, an independent CPU (or host in CUDA parlance) Memory Management allocation Unit (MMU) was implemented. In this way, the implementation caters for live video streaming (or real-time streaming for interactive video applications such as video conferencing). It was assumed that the maximum memory size for the application was 2 GB, based on practical experiments with x264 video streaming.

At start-up time, the implementation enables host mapping and the MMU allocates CPU pinned memory for input data. (Pinned memory is memory that cannot be swapped out, thus improving access time.) CUDA kernel pointers were set to allow access from the GPU. Then

the kernel pointers were allocated addresses from the host memory, as if it were the GPU's global memory. This memory allocation technique enables the overlapping of encoding and packetization for video transport, as data are accessed via Direct Memory Access (DMA) (residing on the GPU) from the CPU, without any explicit data transfer to the GPU memory. At runtime, CUDA kernels are normally asynchronous in respect to the CPU; therefore, each block also created an atomic counter for synchronization purposes. All blocks were executed with non-divergent branching [35] and data could also be read from previous threads. Notice that the CPU does wait until the encoding process is complete before releasing the memory buffers for packetization and buffer refill of unprocessed frames.

The disadvantage of direct GPU access of CPU memory is that such transactions were not as fast as might have been expected because the bandwidth of the Peripheral Component Interconnect express (PCIe) expansion bus could not be fully exploited. Because the paged memory could be swapped or reallocated, the PCIe driver needed to access each page, copy it to a buffer, and then pass it on to DMA. Nevertheless, memory access overhead was substantially reduced, resulting in an overall saving in execution time, Fig. 1 illustrates. The test video sequences in Fig. 1 were obtained from [24] and [77].



**Fig. 1** Average execution time comparison between CUDA memcpy and CUDA zero copy

### *Dynamic parallelism*

To obtain accurate inter-prediction values, H.264/AVC standard allows partitioning of a standard-sized  $16 \times 16$  macroblock (MB). Each MB may be split into  $4 \times 4$ ,  $4 \times 8$ ,  $8 \times 4$ ,  $8 \times 8$ ,  $16 \times 8$ ,  $8 \times 16$  sub-MBs or blocks. In the current implementation, using dynamic parallelism in the CUDA implementation, the  $16 \times 16$  macroblock acting as a parent kernel spawned thread blocks for the sub-MB blocks, without needing any extra instructions from the CPU, which instructions would otherwise add considerably to the computation time (see frame rate comparison in Section 4.3).

### *Pruning of intra-prediction modes*

During intra-prediction, previously reconstructed blocks of pixels act as reference pixels. Strong correlations exist across neighboring blocks resulting in multiple prediction modes, with nine such modes in H.264/AVC intra-prediction. This makes intra-prediction an obstacle to parallel execution, particularly when an MB is decomposed into  $4 \times 4$  blocks. In the latter, by default the reconstruction of one block cannot begin before another has finished and follows a zig-zag pattern across the blocks in order to synchronize the decoder with the encoder. In [46], the zig-zag processing order of the blocks at the encoder was modified to allow processing at the decoder of two blocks at a time. Thus, a block could be predicted while a previously predicted block was being reconstructed in a two-stage pipeline. In [40], the order of block encoding was also changed but in this case so as to prune the number of prediction modes in the most judicious manner. In fact, only three prediction modes, DC mode, horizontal and vertical mode were used for nine of the blocks, allowing pipelining, while compression of the remaining seven blocks was unaffected by pruning because of the prediction reordering. According to [40], compared to the reference JM algorithm without pipelining, there was a 41% reduction in processing time, with virtually no loss of compression efficiency.

However, the current implementation increases the degree of parallelism further still by reducing the direction of predictions to just the horizontal intra mode. This enables four-way parallelism across the four rows of each MB when decomposed into  $4 \times 4$  blocks. Thus, the main criterion for the drastic pruning was to optimize parallelism for live- and real-time streaming. There is consequently a loss of compression efficiency with its relative assessment being part of future work.

### 4.3 Implemented frame rates

Table 4 reports the H.264/AVC codec settings (with HEVC settings reported for a later experiment) for a performance text of the authors' implementation of live streaming. The relative frame rates for HD and 4kUHD video Constant Bitrate (CBR) encoding are reported in Fig. 2. Inspection of the 4kUHD output frame rates of the hardware accelerated encoder showed that coding complexity reduced Sintel's output to 15 fps, whereas Coast's output approached 20 fps, with the other two clip's rate between that in order of motion activity, as further discussed in Section 5.

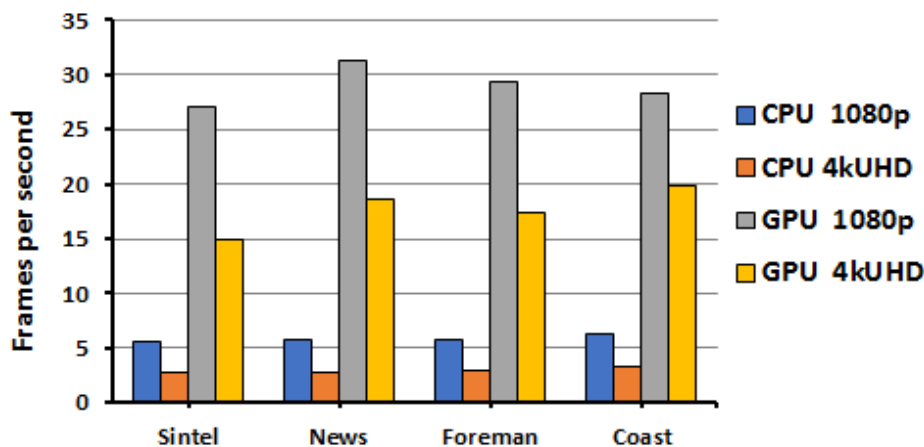


Fig. 2 Mean frame rates either by H.264/AVC CPU encoding or by GPU accelerated encoding

Table 4 Codec parameters for test

Parameter	H.264/AVC	HEVC
Profile	High (5.1)	Main 10
Processing unit	Macroblock (MB)	CTU
Processing unit size	16×16	64×64
GoP size	40	25
GoP frame structure	IPPPP...	IBBPBBP...
CBR bitrates	20 Mbps	13.5 Mbps

Prior implementation of GPU acceleration for H.264/AVC concentrated on variable block size motion estimation (ME), not surprisingly as ME can contribute to up to 70% of the computational complexity on a CPU [50]. For example, the authors of [15] contributed a CUDA thread scheduler and additionally organized the variable blocks so that there was a more even flow of computation (by splitting variable sized blocks into  $4 \times 4$  blocks). A speed-up of 12 with a GPU over a purely CPU executed codec. Then to further improve the speed-up, a task scheduler was developed, as described in [65], so that ME computations

could take place across multiple GPUs. By those means, it was reported in [65] that four GPUs can achieve real-time, full ME (without using a fast, approximate search) with a  $32 \times 32$  search window for an HD ( $1280 \times 720$  pixels/frame) video.

#### **4.4 HEVC acceleration**

Turning to HEVC inter-prediction, compared to the seven ME block sizes of H.264/AVC, there are 12 different HEVC block sizes (with blocks now arranged as Prediction Units (PUs) within each Coding Unit (CU), a CU being a subdivision of a CTU). When Rate-Distortion analysis is turned on, up to 425 different ME calculations may be needed before a prediction mode is selected (if a fast mode decision algorithm is not deployed). In [83], variable block size ME along with fractional pixel interpolation was performed on a GPU in parallel with the other processing tasks performed on the CPU. For a  $2560 \times 1600$  pixels/frame video sequence, a speed-up of 113 was reported with a frame rate of 24. However, because in [83] ME was performed on a CTU line-by-line basis there was a cost of 0.7 dB degradation in video quality. To further reduce the loss in video quality, in [50], processing was reorganized in the Low Delay P profile so that a CTU pipeline was created, with adaptive ME range based on the motion found in co-located blocks.

To achieve 60 fps for HEVC encoding of 4kUHD in the Main 10 profile, further computational resources were needed in [36], namely 2 PCs each with 32 cores. ME was again GPU assisted but additionally Sum of Absolute Difference (SAD) calculations were delegated to the Single Instruction Multiple Data (SIMD) hardware assistance present on x86 processors. The authors of [36] report a speedup of 13 over the well-known, open source x265 codec, with 0.03 dB video quality loss (0.5 dB loss measured against the HM reference codec).

The need to deploy multiple levels of parallelism and several forms of hardware support indicates that software-based real-time 4kUHD at appropriate frame rates may be some way off in time for mobile devices. The aforementioned GP1 chip for the GoPro Hero6 appears to deliver this performance, though video quality and configuration details appear to be unavailable. Nvidia GPUs have built-in support for HEVC encoding, with some said to support resolutions up to 8kUHD. It should be noted that Nvidia does not provide encoding support for the VP8 and VP9 codecs, though it does provide decoding support for those codecs. Unlike the one-pass HEVC reference software (though not the later open-source x265 implementation) VP9 is a two-pass codec by default, which may explain the Nvidia policy.



Two-pass features, e.g. adjusting the bit-allocation according to the video content, especially the motion characteristics found in the first pass, are likely to result in coding gains, allowing reduced bitrates for the same video quality. However, in [30], comparing low-delay settings in High profiles, it was found that one-pass HEVC reference software achieved a reduction in bitrate for equal video-conference quality (assessed by PSNR) relative to two-pass mode VP9 of 30.6% (in units of BD-bitrate (BR)). On the other hand, HEVC reference software was slower than two-pass VP9 by a factor of 6.12. However, considerable research has now been undertaken to improve the encoding time of HEVC codecs.

In terms of HEVC intra-prediction, some important contributions have been made, of which a few highlights are now reviewed. There are up to 35 intra modes available in HEVC. However, early tests showed that when HEVC is confined to intra frames, HEVC's coding efficiency gain over H.264/AVC is limited to about 22% [48], providing a case for intra-mode pruning. In a full implementation of the standard, the rate-distortion (RD) of each candidate mode needs to be evaluated after each Large Coding Unit (LCU) is split into CUs, which in turn are split into Prediction Units (PUs). The size and shape of a PU determines the number of intra-modes that can be applied. In [19], the edges within each PU were first categorized into five directional types. The dominant edge type of the five then guided the selection of the intra-modes to be evaluated. For each type, a set of just nine angular intra-modes along with the DC and planar modes were evaluated. [19] reports an average 20% reduction in encoding time compared to version 4.0 of the reference software, with negligible reduction in video quality (assessed by PSNR). Then in [71], for lossless HEVC encoding used in screen-content compression on mobile devices, selection of intra mode was confined to just three modes designed by the author for up to 54% reduction in encoding time for a 7% reduction in bitrate (without loss). For 'lossy' compression, in [91], a speed-up over the default reference software method of approximately 2.5 is reported. This is achieved by performing a progressive rough search beforehand by means of a fast Hadamard transform in a way that returns fewer candidates for RD evaluation. CU splitting into sub-CUs is also terminated at an early stage to reduce the number of evaluations.

In terms of improving HEVC rate control, i.e. optimizing the bitrate according to the distortion, QP is no longer the dominant parameter determining RD characteristics. This is because in HEVC there is no longer a fixed-size transform block, which the prior  $\rho$ -model relied upon when modeling RD. Instead, [14] advocates the  $\lambda$ -domain, where  $\lambda$ , the Lagrangian multiplier, is the slope of the RD curve, from which the optimal RD point is

selected. As a result, HEVC now incorporates R- $\lambda$  modeling in rate control, whereby a set of coding parameters are employed to select for  $\lambda$ . However, when HDR is implemented, more bits may be needed in some regions relative to others, especially in very bright, high-contrast areas that may occur within CUs. Thus, in [63] three different R- $\lambda$  models are selected from according to regional contrast, with encouraging results in more accurate modeling of the RD.

## 5 4kUHD transmission over WLANs

Experiments in 4kUHD resolution video streaming across the IEEE 802.11 high-throughput standard wirelesses are considered in this Section.

### 5.1 IEEE 802.11n streaming

This Section demonstrates that 4kUHD video streaming is even possible with H.264/AVC compression across an IEEE 802.11n WLAN operating at 5 GHz [62]. The video was transmitted indoors over 20m from a PC 802.11n dongle to an Access Point (AP) and onwards over 20m to another PC 802.11n dongle. To reduce external interference within the laboratory, experiments were conducted at night. At the time of the experiments a mobile display at the requisite resolution was not available and, therefore, four HD (1080p) displays were mosaicked by means of a Nvidia NVS 450 GPU with support software (Nvidia, 2008). The test video was Sintel (see Fig. 1) in YUV format and compressed with the x264 codec implementation according to the settings for H.264/AVC of Table 4. However, the CBR bitrate was varied according to chroma subsampling mode and target frame rate, as recorded in Table 5, with a maximum average compression ratio of 160. Owing to the network transmission software configuration, MPEG2-TS encapsulation was employed at the application layer, rather than direct IP/UDP/RTP resulting in a maximum of seven 188 B MPEG2-TS packets within each IP/UDP/RTP packet.

From Table 6's results, 4kUHD video quality was reasonable for a wireless link, given the SSIM range from 0 to 1. One-way network latency of 100 ms is typically aimed at [26], allowing time for IPTV channel changes or Video-on-Demand (VoD) responses of 1 to 2 s. In Table 6, the average network latency over two 20 m hops was found to be 71 ms, with the majority of that time likely to have been traversing the AP. At 30 fps, a delay of 71 ms is just over two frames at 30 fps. That is to say, for real-time video such as of sports events, two frames delay between capture and display if other components of streaming delay were not

present. Given the average PLRs of Table 6, as further discussed in Section 5.2, one might expect, if the HEVC codec had been used, lower video quality. However, intuitively using the less efficient H.264/AVC codec actually results in less compressed data per packet than had an HEVC stream been packetized.

**Table 5** Streaming configuration of 1000 frames of Sintel

Subsampling mode	Frame rate (fps)	Uncompressed bitrate (Gbps)	Avg. compressed bitrate (Mbps)
4:2:0	24	2.39	20
	30	2.99	30
4:2:2	24	3.19	20
	30	3.99	30

**Table 6** Experimental results from streaming 4kUHD over IEEE 802.11n

Subsampling mode	Frame rate (fps)	Avg. video quality (SSIM)	Avg. network latency (s)	Avg. packet loss rate (%)
4:2:0	24	0.87	0.040	0.80
	30	0.80	0.065	1.65
4:2:2	24	0.82	0.045	0.90
	30	0.70	0.071	2.85

## 5.2 IEEE 802.11ac streaming

This Section reports high-resolution video streaming experiments over IEEE 802.11ac wireless. The same test sequences as used for Fig. 1’s results were employed, with Table 7 recording their characteristics in terms of recommendation ITU-T P.910’s Spatial Index (SI) and Temporal Index (TI). (The much higher temporal complexity of Sintel explains its lower frame rates in Fig. 2). To compress 500 frames of each of the sequences prior to packet loss visibility (PLV) assessment, the x265 implementation of HEVC was employed. HEVC was configured in its Main 10 profile, with settings as in Table 4 with 8-bit depth and 4:2:0 chroma subsampling. The configuration of IEEE 802.11ac was similar to that in [21], using the high-throughput features present in the Broadcom BCM4360 chipset. The settings are given in Table 8.

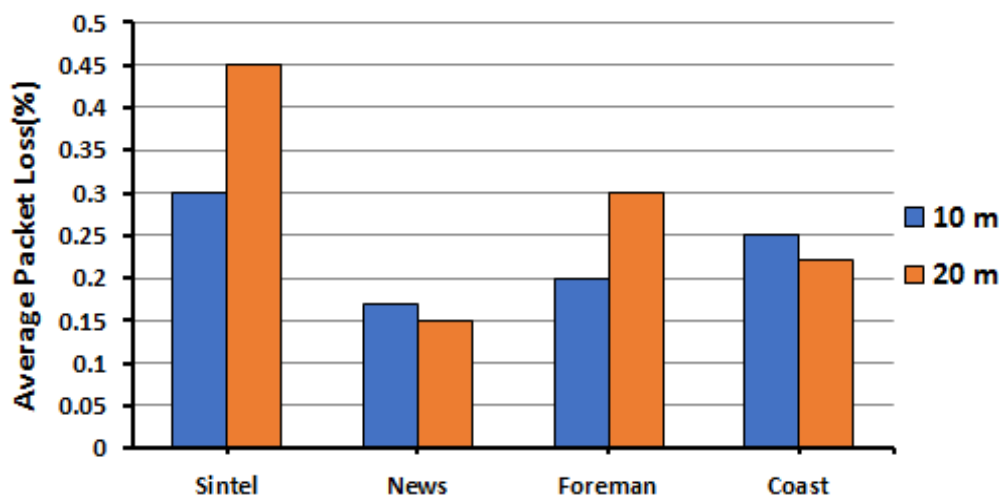
**Table 7** Test video sequences content type

Video sequence	SI	TI	Motion classification
Coast	10.84	16.92	Moderate
News	17.52	21.24	Moderate
Foreman	19.71	38.29	High
Sintel	16.39	72.26	High

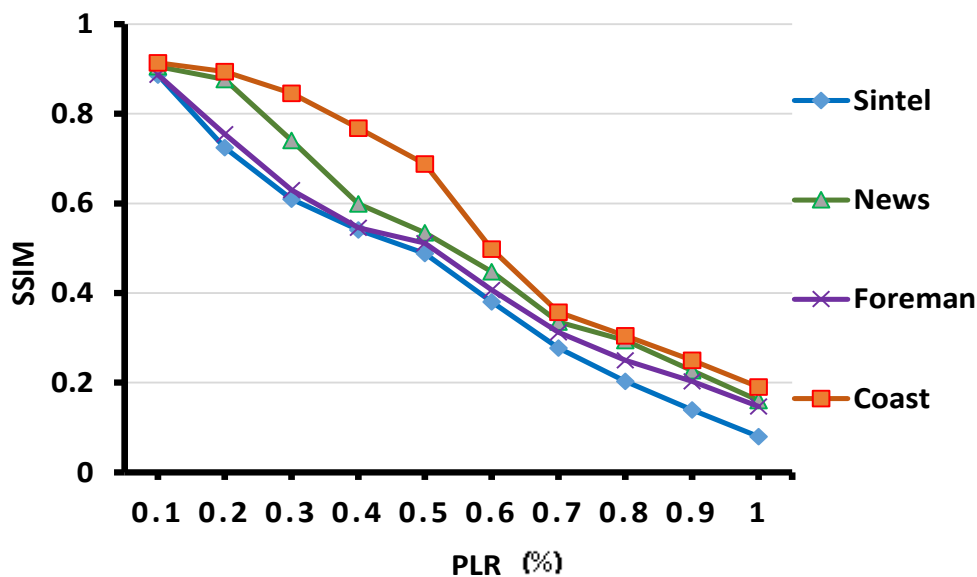
**Table 8** Settings for IEEE 802.11ac measurements

Setting	Value
Channel bandwidth	80 MHz
Channel number	36 (5.180 GHz)
No. of spatial channels	3 (3 × 3 MIMO)
Transport protocol	UDP
Datagram size	1472 B
Modulation	256-QAM

Fig. 3 considers packet losses at 10 m and 20 m (average of 20 tests). For 10 m transmission was unimpeded but for 20 m standard office furniture was present. Unreported tests showed that any intervening partition walls led to a sharp fall in datarates. From Fig. 3, it may appear that up to 20 m Packet Loss Rates (PLRs) are not distinguished by distance. However, the PLRs for higher-activity Sintel in particular are certainly higher. We, therefore, hypothesize that the total time during which a video was transmitted (according to the frame rates reported in Table 1 as output from the in-house CUDA encoder) exposed the video to more packet loss events. This was despite the fact that channel 36 (Table 8) was chosen and transmission was at night, both to reduce external interference.

**Fig. 3** PLRs at two distances during IEEE 802.11ac transmission of 4kUHD video sequences

Video quality in Fig. 4 with similar PLRs to those recorded in Fig. 3 was assessed by the SSIM index (see Section 1). From Fig. 4's plots, especially for PLRs between 0.2% and 0.5%, it is apparent that motion activity of a test video sequence (refer back to Table 5) strongly influences of the packet loss impact. Coast in particular benefits in that way. It would be unwise to stream 4kUHD video without error protection if the PLR was over 0.5%. On the other hand, Fig. 3 gives an indication that a PLR of 0.5% over a short to medium distance may be a rarity for 802.11ac streaming when reduced external interference is present. However, SSIM does not assess temporal quality and a frame rate of 50 to 60 fps rather than 25 fps may be preferable. Further comparing SSIMs for Sintel with those for the same video under H.264/AVC compression in Table 6, it is apparent that the video quality is generally lower under similar PLRs. As remarked in Section 5.1, though the comparison is an approximate one owing to the changed experimental circumstances, the main reason that the video quality is degraded is likely to be due to the greater compression efficiency of HEVC. The latter leads to a greater impact upon video quality from the loss of a packet.



**Fig. 4** SSIM video quality assessment for a range of PLR percentages with HEVC codec encoding and 4kUHD resolution

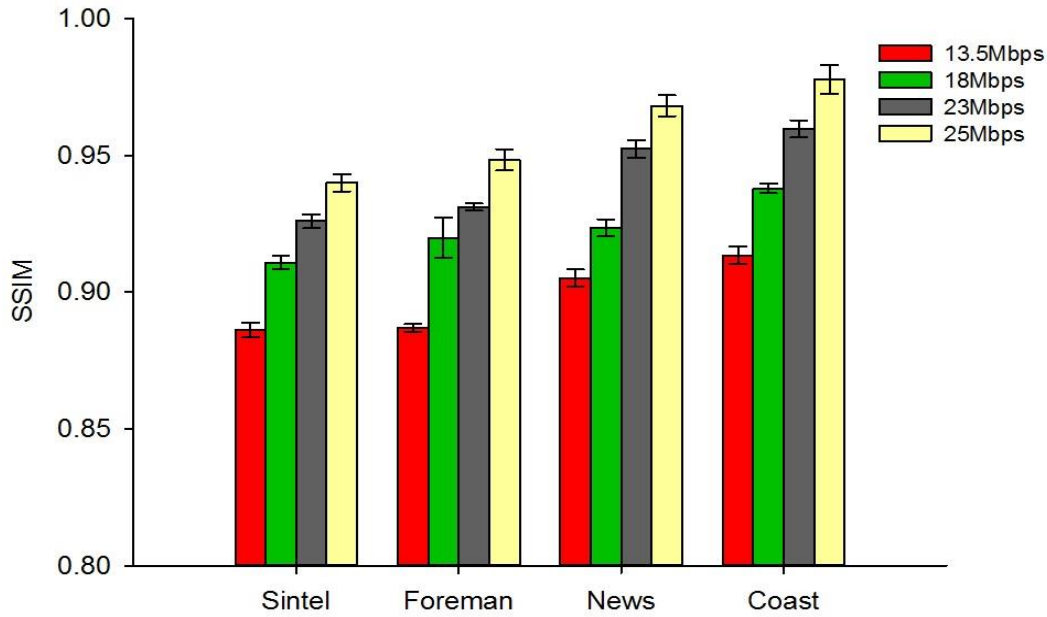
### 5.3 IEEE 802.11ad streaming

In this Section, an IEEE 802.11ad 60 GHz transmitter took the place of the IEEE 802.11ac transmitter in the previous Section. The video configuration was similar to that for the same video sequences as in the previous Section. However, the target CBR bitrate was incrementally increased. The PLRs were recorded as approximately 0.1%. The results of

these experiments are shown in Fig. 5, which includes standard error bars (one standard deviation of the mean).

It can be seen that, as the compressed bitrate is increased, the relative effect of the packet losses seems to reduce. This can be attributed to the amount of coded information that is distributed amongst the packets, each of the same size, except the final packet, i.e. larger packets did not lead to increased PLR [43]. For example, at a CBR rate of 13.5 Mbps video will have a fewer number of packets but this means that the amount of coded information per packet is higher compared to a 25 Mbps stream, where the distribution of compressed data over multiple packets reduces the sensitivity to packet loss. Notice that the PLR, that is not the number of lost packets, remained approximately the same, 0.1%, as a result of stable wireless channel conditions during the experiments. This is the same effect that was previously remarked upon when going from H.264/AVC to HEVC compression. The general point is that, despite the coding efficiency of HEVC, for high throughput 60 GHz transmission, it is better to select as high a bitrate as is available to fully exploit the available bandwidth.

For comparison purposes CBR encoding was used in the above experiments. In practice, for live streaming, Constant Rate Factor (CRF) with a Video Buffer Verifier (VRF) limit may be preferable for streaming. This is because CBR encoding with a hard bitrate runs the risk of additional bits being added by the encoder to the compressed stream simply to meet the hard bitrate. CRF is a form of Variable Bitrate (VBR) but takes motion across frames into account rather than simply imposing a hard QP limit. For live streaming, one-pass encoding was preferred to reduce latency. Two-pass encoding is available in the open-source x265 codec implementation of HEVC and for VoD applications allows file storage limits to be matched more accurately. However, two-pass encoding is not appropriate for live streaming because of potential latencies involved and in the above experiments two-pass encoding was not used. The gain in x265 codec video quality from two-pass encoding relative to the extra latency involved in gathering content statistics in a first pass are reserved for future research.



**Fig. 5** Video quality after transmission over IEEE 802.11ad link for 4kUHD video compressed by HEVC at various bitrates (PLR approx. 0.1%)

#### 5.4 Discussion: Internet latency

This paper has considered delay over a wireless network, which in the experimental environment was low. However, in the journey over the wired Internet to the wireless access network, levels of end-to-end delay and jitter, variation in delay, may well be higher. Consequently, a corrective to the findings of this paper in that respect may be needed. As remarked in Section 1, for HAS-type streaming, start-up-delay, which combines end-to-end delay and buffer filling, as well as jitter has an important impact on QoE [73]. It has long been known that the length of stalls and their frequency, arising from jitter-induced congestion, has a considerable impact on QoE [60 [82]. Initial delay is largely preferred by end users rather than stalls [37]. Given that in HAS streaming, the QoE is already impacted by bitrate switches, infrequent though they may be, studies such as [22] identify the buffering ratio, i.e. in HAS when the video chunks are being stored rather than displayed, as the most important factor affecting QoE. In general, buffer size can be increased or more sensibly dynamically adapted according to network conditions [26]. However, live video streaming, especially sports events, is generally not streamed through HAS. This is because buffer size is limited to a few seconds and all viewers are approximately synchronized. Thus, sports viewers are particularly sensitive to the buffering ratio, as they are to video quality [22], which implies a dedicated optical path in the Internet optical core to reduce this issue. There is also variation in the subjective response according to the type of VoD, either long (over 35 mins.) or a short clip, and the level of end user engagement, which can be forgiving of

Internet jitter if the content is of interest. In summary, the behavior of the wireless access network needs correction, in all but home cinema applications, to the impact of network jitter in the Internet. QoE is dependent on the behavior of the network core, the type of content, the user engagement, and, for HAS-type streaming, on the buffering size and adaptability.

## 6 Conclusion

This paper has provided an overview of research activity towards real-time transmission of high resolution video over wireless links. Some current smart phones such as the Sony Xperia Z5 Premium, iPhone 6s, and Nokia Lumia 1520, to select a few at random, can capture 4kUHD video. Nevertheless, they do not currently display such a high resolution, even though the ideal viewing distance of HD already matches how handheld devices are viewed. at close range. However, given that a good number of smartphones and tablets already support IEEE 802.11ac, it may not be long before these phones not only capture and transmit video but may also be used for live streaming and display. Dell's Wireless Dock D5000 is a means of transmitting video at 60 GHz based upon available triband chips, with a maximum throughput of 4 Gbps. Together with the Dell Latitude 6430u Ultrabook with a 1601 WiGig network card, this wireless dock or AP was used in multi-hop streaming experiments of 4kUHD video. Thus, uncompressed or compressed video may already be streamed with commercial products at a maximum of 1 Gbps because of the WiGig Ethernet interface.

The paper has presented experiments at 4kUHD resolutions showing the likely video quality at short range over low packet loss channels. Findings show that though motion complexity strongly influences the attainable quality at a given CBR. However, by using a less efficient codec, e.g. H.264/AVC rather than HEVC, it is possible to reduce the impact of any packet loss. Given the available bandwidths, especially over an IEEE 802.11ad link, increasing the target bitrate is sensible, as this results in a disproportionate gain in video quality whenever packet loss does occur. Encoding latency is important for live streaming and in these circumstances there is a tradeoff between latency and resilience to packet loss. Application layer channel coding and its impact is a subject for future work. It too can have an impact on latency during live streaming. Energy consumption and battery longevity on mobile devices is also an important factor and a subject of future research in respect to the arrival of high resolutions on such devices.



## References

1. Abe, A., Walker, S.D.: Multi-hop 802.11ad wireless H.264 video streaming. In: IEEE Int. Conf. on Telecomm. and Sig. Proc. pp. 94-99 (2016)
2. Adeyemi-Ejeye, A.O., Walker, S.D.: Uncompressed quad-1080p wireless video streaming. In: IEEE 4th Comput. Sci. and Electron. Eng. Conf., pp. 13–16 (2012)
3. Adeyemi-Ejeye, A.O., Walker, S.D.: Ultra-high definition wireless video transmission using H. 264 over 802.11n WLAN: Challenges and performance evaluation. In: 12th Int. Conf. on Telecommun., pp. 109–114 (2013)
4. Adeyemi-Ejeye, A.O., Alreshoodi, M., Al-Jobouri, L., Fleury, M., Woods, J.: Packet loss visibility across SD, HD, 3D, and UHD video streams. *J. of Video Commun. and Image Representation* **45**, 95-106 (2017)
5. Adeyemi-Ejeye, A.O., Alreshoodi, M., Walker, S.D.: Implementation of 4kUHD HEVC-content transmission. *Multimed. Tools and Apps.* **76**(17), 18099-18118 (2017)
6. Alreshoodi, M., Adeyemi-Ejeye, A.O., Al-Jobouri, L., Fleury, M., Al-Zahrani, B.: Packet loss visibility for higher resolution video on portable devices. In: IEEE Int. Conf. Consumer Electron., pp. 237 – 238 (2017)
7. Bae, S.H., Kim, J., Kim, M., Cho, S.: Assessments of subjective video quality on HEVC-encoded 4K-UHD video for Beyond-HDTV broadcasting services. *IEEE Trans. Broadcast.* **59**(2), 209-222 (2013)
8. Baykas, T. et al.: IEEE 802.15.3c: The first IEEE wireless standard for data rates over 1 Gb/s. *IEEE Commun. Mag.* **49**(7), 114-121 (2011)
9. Bejarano, O., Knightly, E.W.: IEEE 802.11ac: From channelization to multi-user MIMO. *IEEE Commun. Mag.* **51**(10), 84-90 (2013)
10. Bing, B.: 3D and HD broadband video networking. Artech House, Boston, MA (2007)
11. Bjøntegaard, G.: Calculation of average PSNR differences between RD-curves. VCEG Meeting, ITU-T SG16 Q.6 (2001)
12. Borer, T., Cotton, A.: A “Display Independent” High Dynamic Range television system. BBC Research & Develop. White Paper, WHP 309 (2015)
13. Bossen, F., Bross, B., Sühring, K., Flynn, D.: HEVC complexity and implementation analysis. *IEEE Trans Circuits Syst Video Technol.* **22**(12),1685-1696 (2012)
14. Bossen, F.: Common HM test conditions and software reference configurations. JVT-VC, document JVTVC-LI1100 (2013)
15. Chen, W.-N., Hang, H.-M.: H.264/AVC motion estimation implementation on Compute Unified Device Architecture (CUDA). In: IEEE Int. Conf. Multimed. and Expo, pp. 697-700 (2008)
16. Cheng, L., Liu, C., Dong, Z., Yu, J., Chang, G.-K.: 60-GHz and 100-GHz wireless transmission of High-Definition video services in converged Radio-over-Fiber systems. In: Conf. on Lasers and Electro-Optics., pp. 1-2 (2013)
17. Choi, M., Lee, G., Jin, S., Koo, J., Kim, B., Choi, S.: Link adaptation for high-quality uncompressed video streaming in 60-GHz wireless networks. *IEEE Trans. Multimed.* **18**(4), 627-642 (2016)

18. Cordeiro, C.: Evaluation of medium access technologies for next generation millimeter-wave WLAN and WPAN. In: IEEE Int. Conf. Commun. Workshops, pp. 1-5 (2009)
19. Da Silva, T.L., Agostini, L.V., da Silva Cruz, D.A.: Fast HEVC intra prediction mode decision based on edge direction information. In: 20<sup>th</sup> Europ. Sig. Process. Conf., pp. 1214-1218 (2012)
20. De Simone, F., Naccari, M., Tagliasacchi, M., Dufaux, F., Tubaro, S., Ebrahimi, T.: Subjective assessment of H.264/AVC video sequences transmitted over a noisy channel. In: Int. Workshop on Quality of Multimed. Experience, pp. 204-209 (2009)
21. Dianu, M.-D., Riihijärvi, J., Petrova, M.: Measurement-based study of the performance of IEEE 802.11ac in an indoor environment. In: IEEE Int.Conf. Commun., pp. 5771-5776 (2014)
22. Dobrian, F., Sekar, V., Awan, A., Stoica, I., Joseph, D., Ganjam, A., Zhan, J., Zhang, H.: Understanding the impact of video quality on user engagement. In: ACM SIGCOMM, pp. 362-373 (2011)
23. DVB: DVB Broadcast Asia 2014 exhibition & conference preview. <http://epublishbyus.com/ebook/ebook?id=10036869#/2> (2014)
24. Elemental Technologies: 4K Test sequences. <http://www.elementaltechnologies.com/resources/4k-test-sequences> (2013)
25. Emoto, M., Sugawara, M.: Critical fusion frequency for bright and wide field-of-view image display. *J. Display Technol.* **8**(7), 424-429 (2012)
26. Evans, J., Filsfils, C.: Deploying IP and MPLS QoS for multiservice networks. Morgan Kaufmann, San Francisco, CA (2007)
27. Fleury, M.: Streaming uncompressed HD over wireless channels. In: Int. Broadcasting Conf. (2012)
28. François, E., van de Kerkhof, L.: A single-layer HDR video coding framework with SDR compatibility. In: Int. Broadcast. Conf. (2016)
29. Ghanbari, M., Crawford, D., Fleury, M., Khan, E., Woods, J.: Future performance of video codecs. Research Report for Office of Commun. (Ofcom) UK (2006)
30. Grois, D., Marpe, D., Nguyen, T., Hadar, O.: Comparative assessment of H.265/MPEG-HEVC, VP9, and H.264/MPEG-AVC encoders for low-delay video applications. In: SPIE Proc., vol. 9217, Applications of Digital Image Processing XXXVII (2014)
31. Grois, D., Nguyen, T., Marpe, D.: Coding efficiency comparison of AVI/VP9, H.265/MPEG-HEVC, and H.264/MPEG-AVC encoders. In: Picture Coding Symp. (2016)
32. Halák, J., Krsek, M., Ubik, S., Žejdl, P., Nevřela, F.: Real-time long distance transfer of uncompressed 4K video for remote collaboration. *Fut. Gen. Comput. Syst.* **27**(7), 886-892 (2011)
33. Hamidouche, W., Cocherel, G., Le Feuvre, J., Raulet, M., Déforges, O.: 4k real time video streaming with SHVC decoder and GPAC player. In: IEEE Int. Conf. on Multimed. and Expo Workshops (2014)
34. Hanhart, P., Rerabek, M., De Simone, F., Ebrahimi, T.: Subjective quality evaluation of the upcoming HEVC video compression standard. EPFL Tech. Paper (2012)
35. Harris, M.: Optimizing parallel reduction in CUDA. NVIDIA Developer Technol. (2007)
36. Heng, T.K., Asano, W., Itoh, T., Tanizawa, A., Yamaguchi, J., Matsuo, T., Kodama, T.: A highly parallelized H.265/HEVC real-time UHD software encoder. In: IEEE Int. Conf. on Image Process., pp. 1213-1217 (2014)

37. Hoßfeld, T., Egger, S., Schatz, R., Fiedler, M., Masuch, K., Lorentzen, C.: Initial delay vs. interruptions: Between the devil and the deep blue sea. In: 4th Int. Workshop QoMEX, pp. 1–6 (2012)
38. Horé, A., Ziou, D.: Image quality metrics: PSNR vs. SSIM. In: IEEE 20th Int. Conf. on Pattern Recognit., pp. 2366–2369 (2010)
39. Jiang, J., Sekar, V., Zhang, H.: Improving fairness, efficiency, and stability in HTTP-based adaptive video streaming with FESTIVE. *IEEE/ACM Trans. Network.*, **22**(1), 326–340 (2014)
40. Jin, G., Lee, H.-J.: A parallel and pipelined execution of H.264/AVC intra prediction. In: IEEE Int. Conf. on Comput. and Info. Technol. (2006)
41. Kamata, H., Kikuchi, H., Sykes, P.J.: Real-world live 4K Ultra HD broadcasting with High Dynamic Range. In: Int. Broadcast. Conf. (2016)
42. Kim, J., Tian, Y., Mangold, S., Molisch, A.F.: Quality-aware coding and relaying for 60 GHz. In: IEEE Int. Conf. Commun., pp. 5148-5142 (2013)
43. Korhonen, J., Wang, Y.: Effect of packet size on loss rate and delay in wireless links. In: IEEE Wireless Comms. and Network. Conf., pp. 1608-1603 (2005)
44. Kunić, S., Šego, Z.: Beyond HDTV technology. In: 55th Int. Symp. ELMAR, pp. 83-87 (2013)
45. Lee, S., Kim, H., Eum, N.: Reduced complexity single core based HEVC video codec processor for mobile 4K-UHD applications. In: IEEE Int. Conf. on Consumer Electron., pp. 94-95 (2016)
46. Lee, W., Lee, S., Kim, J.: Pipelined intra prediction using shuffled encoding order for H.264/AVC. In: IEEE Region 10 Conf., pp. 1–4 (2006)
47. Li, B., Li, H., Li, L., Zhang, J.:  $\lambda$ -domain rate control algorithm for High Efficiency Video Coding. *IEEE Trans. Image Process.* **23**(9): 3841-3854 (2014)
48. Li, B., Sullivan, G., Xu, J.: Comparison of compression performance of HEVC Draft 6 with AVC High Profile. In: JCTVC10409 JCT-VC Meeting – Geneva (2012)
49. Li, Z., Huang, Y., Liu, G., Wang, F., Zhang, Z.L., Dai, Y.: Cloud transcoder: Bridging the format and resolution gap between Internet video and mobile devices. In: ACM Int. Workshop on Network and Operating Syst. Support for Digital Audio and Video, pp 33-38 (2012)
50. Ma, J., Luo, F., Wang, S., Ma, S.: Flexible CTU-level parallel motion estimation by CPU and GPU pipeline for HEVC. In: IEEE Visual Commun. and Image Process., pp. 282-285 (2014)
51. Mahmoudi, M.E., Walker, S.D. 4-Gbps uncompressed video transmission over a 60-GHz Orbital Angular Momentum wireless channel. *IEEE Wireless Commun. Lett.* **2**(2), 223-226 (2013)
52. Mehendale, M., Das, S., Sharma, M., Mody, M., et al.: A true multi-standard, programmable, low-power, full HD video codec engine for smartphone HD SoC. In: IEEE Int. Solid State Conf., pp. 226-228 (2012)
53. Mukherjee, D., Bankoski, J., Grange, A., Han, J., Koleszar, J., Wilkins, P., Xu, Y., Bultje, R.: The latest open-source video codec VP9 – An overview and preliminary results. In: Picture Coding Symp., pp. 390-393 (2013)
54. Murakami, T.: The development and standardization of Ultra High Definition video technology. In: Mrak G, Grgic M, Kunt M (eds) High-Quality Visual Experience. Signals and Commun. Technol. Springer Verlag, Berlin, Germany, pp. 81-135 (2010)

55. Nitsche, T., Cordeiro, C., Flores, A.B., Knightly, E.W., Perahia, E., Widmer, J.C.: IEEE 802.11ad: Directional 60 GHz communication for multi-Gbps Wi-Fi. *IEEE Commun. Mag.* **52**(12), 132-141 (2014)
56. Nvidia: Nvidia quadro NVS 450. [http://www.nvidia.co.uk/object/product\\_quadro\\_nvs\\_450\\_uk.html](http://www.nvidia.co.uk/object/product_quadro_nvs_450_uk.html) (2008)
57. NVIDIA Developer: Video Encode and Decode GPU Support Matrix. <https://developer.nvidia.com/video-encode-decode-gpu-support-matrix> (2017)
58. Ohm, J., Sullivan, G., Schwarz, H., Tan, T., Wiegand, T.: Comparison of the coding efficiency of video coding standards — including High Efficiency Video Coding (HEVC). *IEEE Trans. Circuits Syst. Video Technol.* **22**(12), 1669-1684 (2012)
59. Park, M., Gopalakrishnan, P.: Analysis on spatial reuse and interference in 60-GHz wireless networks. *IEEE J. Sel. Areas Commun.* **27**(8), 1443-1452 (2009)
60. Pastrana-Vidal, R.R., Gicquel, J.C., Colomes, C., Cherifi, H.: Sporadic frame dropping impact on quality perception. In: *SPIE 5292, Human Vision and Electronic Imaging IX* (2004)
61. Paul, T.K., Ogunfunmi, T.: Wireless LAN comes of age: Understanding the IEEE 802.11n amendment. *IEEE Circuits Syst. Mag.* **8**(1), 28-54 (2008)
62. Perahia, E., Stacey, R.: Next generation wireless LANs: 802.11n and 802.11ac (2nd ed.). Cambridge Univ. Press, Cambridge, UK (2013)
63. Perez-Daniel, K.R., Sanchez, V.: Luma-aware multi-model rate-control for HDR content in HEVC. *IEEE Int. Conf. on Image Process.*, pp. 1022-1026 (2017)
64. Perkins, C.: RTP: Audio and Video for the Internet. Addison-Wesley, Boston, MA (2003)
65. Pieters, B., Hollemeersch, C.F., Lambert, P., Van de Walle, R.: Motion estimation for H.264/AVC on multiple GPUs using NVIDIA CUDA. In: *SPIE Appl. of Digital Image Process. XXII*, vol.7443 (2009)
66. Pinson, M.H., Wolf, S., Cermak, G.: HDTV subjective quality of H.264 vs. MPEG-2, with and without packet loss. *IEEE Trans. Broadcast.* **56**(1), 86–91 (2010)
67. Poynton, C.: Digital video and HDTV: Algorithms and interfaces. Morgan Kaufmann, San Francisco, CA (2003)
68. Rappaport, T.S., Heath Jr., R.W., Daniels, C.R., Murdock, N.J.: Millimeter wave wireless communications. Prentice Hall, Upper Saddle River, NJ (2014)
69. Ryu, Y., Park, K., Wee, J., Kwon, K.: An efficient 4K and 8K UHD transmission scheme on convergence networks with broadcasting and LTE by using coordinated multi-point transmission system. *KSII Trans. Internet and Info. Syst.* **11**(8), 4092-4104 (2017)
70. Saito, S., Shitomi, T., Asakura, S., Satou, A., Okano, M., Murayama, K., Tsuchida, K.: 8K terrestrial transmission field tests using dual-polarized MIMO and higher-order modulation OFDM. *IEEE Trans. Broadcast.* **62**(1), 306-315 (2016)
71. Sanchez, V.: Fast intra-prediction for lossless coding of screen content in HEVC. *IEEE Global Conf. on Sig. and Info. Process.*, pp. 1367-1371 (2015)
72. Schulze, H., Lüders, C.: Theory and applications of OFDM and CDMA. Wiley & Sons, Chichester, UK (2005)

73. Seufert, M., Egger, S., Slanina, M., Zinner, T., Hoßfeld, T., Tran-Gia, P.: A survey on Quality of Experience of HTTP Adaptive Streaming. *IEEE Commun. Surveys & Tutorials* **17**(1), 469-492 (2015)
74. Shirai, D., Kawano, T., Fujii, T., Kaneko, K., Ohta, N., Ono, S., et al.: Real time switching and streaming transmission of uncompressed 4K motion pictures. *Fut. Gen. Comput. Syst.* **25**(2), 192-197
75. Shirai, D., Yamaguchi, T., Shimizu, T., Murooka, T., Fujii, T.: 4K SHD real-time video streaming system with JPEG 2000 parallel codec. In: *IEEE Asia Pacific Conf. on Circ. and Systems*, pp. 1855-1858 (2006)
76. Singh, H., Niu, H., Xiangpin, Q., Shao, H., Kwon, C., et al.: Supporting uncompressed HD video streaming without retransmissions over 60 GHz wireless networks. *IEEE Wireless Comms. and Network. Conf.*, pp. 1939-1944 (2008)
77. Sintel: Sintel 4K. <http://www.sintel.org/news/sintel-4kversion-available/> (2011)
78. Skordoulis, D., Qiang, N., Hsiao-Hwa, C., Stephens, A.P., Changwen, L., Jamalipour, A.: IEEE 802.11n MAC frame aggregation mechanisms for next-generation high-throughput WLANs. *IEEE Wireless Commun. Mag.* **15**(1), 40-47 (2008)
79. Sodagar, I.: The MPEG-DASH standard for multimedia streaming over the Internet. *IEEE Multimed.* **18**(4), 62-67 (2011)
80. Sugawara, M., Choi, S.-Y., Woods, D.: Ultra-High-Definition Television (Rec. ITU-R BT.2020): A generational leap in the evolution of television. *IEEE Sig. Proc. Mag.* **31**(3), 170-174 (2014)
81. Valdes-Garcia, A., Reynolds, S., Natarajan, A., et al.: Single-element and phased-array transceiver chipsets for 60-GHz Gb/s communications. *IEEE Commun. Mag.* **49**(4), 120-131 (2011)
82. van Kester, S., Xiao, T., Kooij, R.E., Brunnström, K., Ahmed, O.K.: Estimating the impact of single and multiple freezes on video quality. In: *SPIE 7865, Human Vision and Electronic Imaging XVI* (2011)
83. Wang, X., Soni, L., Chen, M., Yang, J.: Paralleling variable block size motion estimation of HEVC on CPU plus GPU platform. *IEEE Int. Conf. Multimed. and Expo* pp. 1-5 (2013)
84. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.* **13**(4), 600-612 (2004)
85. Wiegand, T., Sullivan, G., Bjøntegaard, G., Luthra, A.: Overview of the H.264/AVC video coding standard. *IEEE Trans Circuits Syst. Video Technol.* **13**(7), 1685-1696 (2003)
86. Wu, J., Yuen, C., Cheung, N.-M., Chen, J.: Delay-constrained high definition video transmission in heterogeneous wireless networks with multi-homed terminals. *IEEE Trans. Mobile Computing* **15**(3): 641-655 (2016)
87. Wu, N., Wen, M., Su, H., Ren, J., Zhang, J.: A parallel H.264 encoder with CUDA: Mapping and evaluation. In: *IEEE 18th Int. Conf. on Parallel and Distrib. Syst.*, pp. 276-283
88. Yuan, H., Fu, H., Liu, J., Hou, J., Kwong, S.: Non-cooperative game theory based rate adaptation for dynamic video streaming over HTTP. *IEEE Trans. Mobile Computing*, DOI: 10.1109/TMC.2018.2800749 (2018)
89. Yuan, H., Wei, X., Yang, X., Xiao, J., Kwong, S.: Cooperative bargaining game-based multiuser bandwidth allocation for dynamic adaptive streaming over HTTP. *IEEE Trans. Multimed.* **20**(1), 183-197 (2018)

90. Zhang, D., Liu, D.: An adaptive cross-layer optimization scheme for light compressed High-definition video transmission based on 60 GHz system. In: IEEE Int. Conf. Commun. Technol., pp. 671-675 (2013)
91. Zhang, H., Ma, Z.: Fast intra mode decision for High Efficiency Video Coding (HEVC). IEEE Trans. Circuits Syst. Video Technol. **24**(4), 660-668 (2014)