THE BENEFITS OF APPLYING AI TO COMPRESSION

J.L. Diascorn

Harmonic, France

ABSTRACT

Artificial intelligence (AI) is a popular subject today. Currently used across various verticals, from medicine to autonomous vehicles and finance, it is projected to have a significant impact. Today, AI is used for video compression, not just to provide bitrate savings but also to improve the quality of experience (QoE) and savings in processing power.

This paper will present three applications of AI for video compression, explaining how each helps with the delivery of video content over broadcast and OTT networks. The applications that will be examined include Dynamic Encoding Style (DES), which enables a better trade-off between video quality and bitrate; Dynamic Resolution Encoding (DRE), which enables a superior QoE and density; and Dynamic Frame Rate Encoding (DFE), which allows for improved density and QoE.

After a brief presentation of the methods, the paper will then present the results of implementing these technologies in the real world.

INTRODUCTION

Video compression for broadcast TV services started more than 20 years ago. Over time, several key improvements, such as dual-pass, statistical multiplexing, and software migration, were made to compression technology in order to boost performance. Artificial Intelligence (AI) is driving the next frontier of video compression enhancements.

AI is effective at detecting objects and at surveillance. Machines are capable of detecting cancer cells with excellent accuracy, which can be a great help for medical doctors (1, 2).

AI algorithms can also be useful at processing a lot of data. Some companies use it to clean large data sets, an activity called data wrangling.

More and more, AI can be used for decision-making. The autonomous vehicle collapses many of these uses. Indeed, detection is important in an autonomous car, as other vehicles, persons, objects, and signs on the road need to be clearly identified along with their motion. Together with the internals of the car, it becomes a lot of data to process. The autonomous car has to constantly make decisions about the speed, direction, signaling, and more.

In other terms, AI is very effective at predictions (3).

More details on the evolution from human-designed algorithm to using AI for live video compression can be found in (10).

In the VOD encoding domain, Netflix has been the pioneer in developing an AI-based system to assist file encoding, known as per-title or per-chunk encoding (4).

Those techniques only apply to offline encoding and cannot be used for live video.

This paper presents three examples of AI applied to live video encoding to optimize broadcast and OTT content delivery. The first three sections present the three examples. For each example, the paper presents a brief presentation of the methods followed by the results, including real-life effects.

In this paper, both "AI" and "machine learning" expressions are used, knowing that machine learning is, in fact, a part of AI.

DYNAMIC ENCODING STYLE (DES) OR CONTENT-AWARE ENCODING (CAE) FOR BITRATE SAVINGS

In this first application, the video compression algorithm itself has improved thanks to machine learning technology. The goal is to improve the video quality/bitrate trade-off, meaning reducing the bitrate while maintaining the video quality or keeping a bitrate and improving the video quality.

This is done by the means of encoding styles. Encoding styles are compression algorithm configurations well-suited for particular content.

Results

DES has been thoroughly tested across a lot of material, and it has shown a bitrate reduction vs. deployed system from 20% up to 30% on VBR content in broadcast applications, and 35% on average up to 50% compared with CBR for streaming applications.

Table 1 shows the comparison of the AI-based algorithm with the deployed solution for a customer's use case. The AI-based algorithm is run at different lower bitrates compared with the deployed solution, from 10% to 30% lower. At 10% the AI-based algorithm is better, at 20% it is equal and at 30% it is worse. The last two columns provide a comparison of lowering the bitrate for both algorithms for verification purposes.

The conclusion is that the AI-based algorithm provides a 20% gain.

Prog	Channel	AI version Pool bitrate -10%	AI version Pool bitrate -20%	AI version Pool bitrate -30%	Both versions Pool bitrate -10%	Both versions Pool bitrate -20%
1	Documentary	=	AI slightly lower	AI lower	AI better	AI better
2	Cartoon	=	=	AI lower	=	=
3	General Entertainment	=	=	AI lower	=	AI slightly better
4	Movie	=	=	AI slightly lower	=	AI slightly better
5	Sport	AI better	=	AI lower	AI better	AI better
6	High action shows	AI better	AI slightly better	=	AI better	AI better

Table 1 – Video quality comparison on different channels between deployed and AI-based algorithm

DES and CAE have been deployed in many streaming situations, with some examples and results shown below.

The first example is a large streaming service with more than 1 million subscribers and more than 50 channels. This service supports live, VOD, cloud DVR, time-shift and server-side dynamic ad insertion. Due to the COVID-19 global health crisis, the service provider observed a dramatic increase in the bandwidth use and needed a solution to relieve the pressure without changing its infrastructure. By turning on DES and CAE the service provider saw significant improvements on their network. The backbone traffic was reduced by 50%, and the CDN peak usage was reduced by 30%.

Figure 1 shows the backbone traffic reduction after DES/CAE was activated.

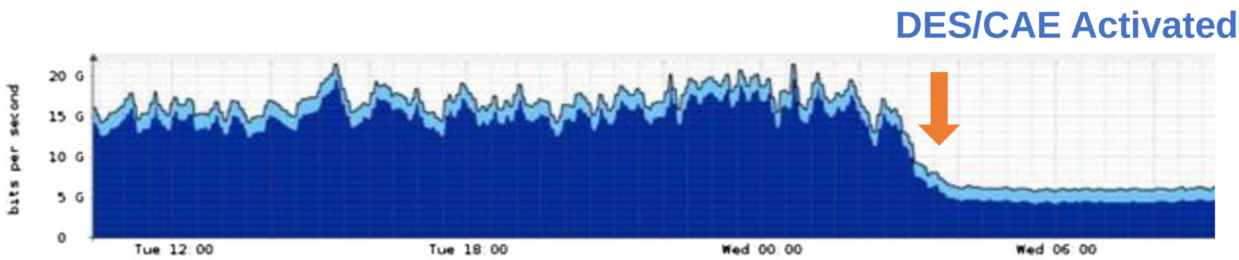

Figure 1 - Backbone traffic reduction thanks to DES and CAE

The second example involves a large European streaming provider. The measurements were also made during the lockdown period due to COVID-19. In this example we show the average bitrate variation between normal compression and with DES/CAE turned on.

For sports content, a bitrate reduction of 30% was measured, and for studio content a bitrate reduction of 40% was observed. Studio content includes television programs, such as talk shows and games shows.

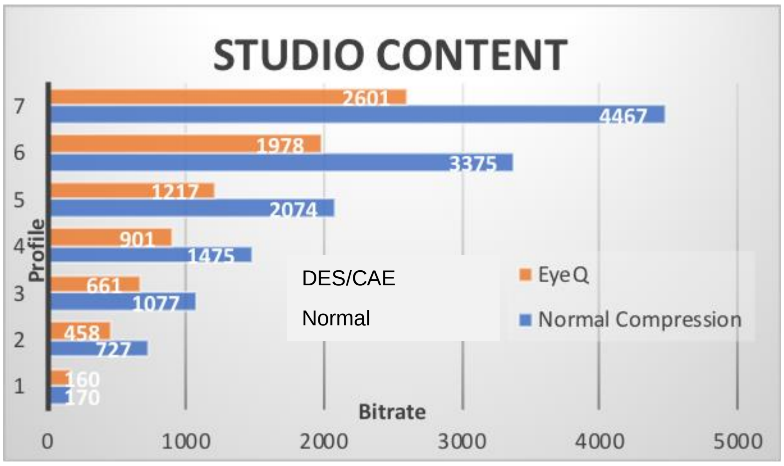

Figure 2 - Studio content average bitrate reduction thanks to DES/CAE

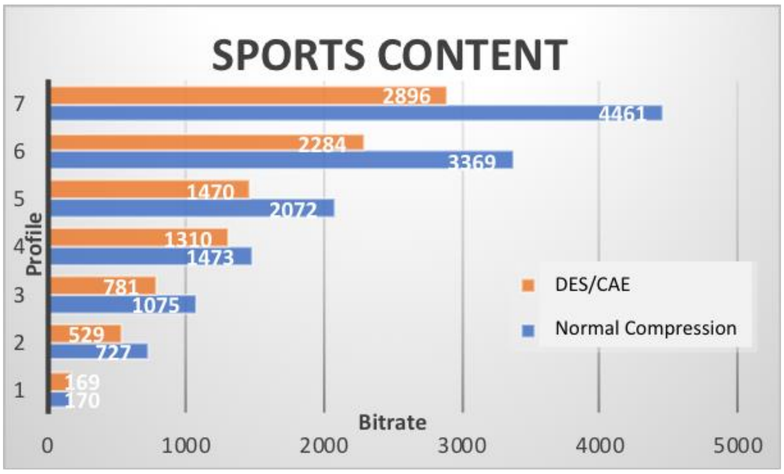

Figure 1 - Sports content average bitrate reduction thanks to DES/CAE

In the example above the streaming service is using seven profiles (See Table 2). It is interesting to note that the higher bitrate savings is obtained on the highest resolution profiles, which are also using the most bits.

	Resolution		
Profile 1	384*216	Profile 4	960*540
Profile 2	480*270	Profile 5	960*540
Profile 3	640*360	Profile 6	1280*720
		Profile 7	1920*1080

Table 2 - Profiles used in the example in Figure 2 and 3

Interoperability

We tested the streams with a variety of set-top boxes (STBs) in TS and HTTP modes and did not find any problems on the tests done in AVC and HEVC. This was expected since the dynamic configuration of encoding decisions does not impact the compliance of video streams to the standards.

DYNAMIC RESOLUTION ENCODING ENABLES A SUPERIOR QOE AND DENSITY

DRE consists of always providing the resolution that will deliver the best QoE. The appropriate resolution is selected by a machine learning algorithm.

It is well known that a high-resolution picture with little or no artifact looks better compared to a picture with lower resolution. On the contrary, if the picture encoding reaches its limits at a given bitrate and produces artifacts, then a lower resolution picture will not show these artifacts and will look better than the high-resolution video with many artifacts.

Depending on the video, one would sometimes want to deliver a low-resolution video and sometimes a high-resolution video in order to always achieve the best possible quality for end users.

Additionally, DRE allows operators to use fewer CPU cycles to process video. One server will be able to process more services in a given hardware or cloud instance and will cost less.

Results

We performed an experiment in the lab and also at a customer site and found that the results are always better with DRE than without using objective and subjective evaluations.

QoE Improvements

The end-user QoE improvements are impressive. Visible gains are seen on many scenes at several operating points for many resolutions. One important aspect to mention is that no visual discomfort was experienced when the resolution changed, even within a continuous scene.

For a low-resolution reference (at low bitrate) DRE achieved around 80% higher resolution most of the time.

For a high-resolution reference (at a higher rate), DRE maintained high resolution 30% to 40% of the time. It means that 60% to 70% of the time the video quality improved since DRE chose a lower resolution that looked better than the nominal high resolution.

Figures 6 and 7 show the statistics for the low-resolution case, 480p, at 1.5 Mbps, and the statistics for the high-resolution case, 1080p, at 3.9 Mbps.

Two sets of tests were performed, as shown below. In the first test the change of resolution is allowed as soon as the predicted video quality is even slightly increased. For the second test, the change is allowed only if there is a visible improvement in video quality. This is indicated by the delta0.3 VQ below, which is the improvement visibility threshold, on a five-step MOS scale.

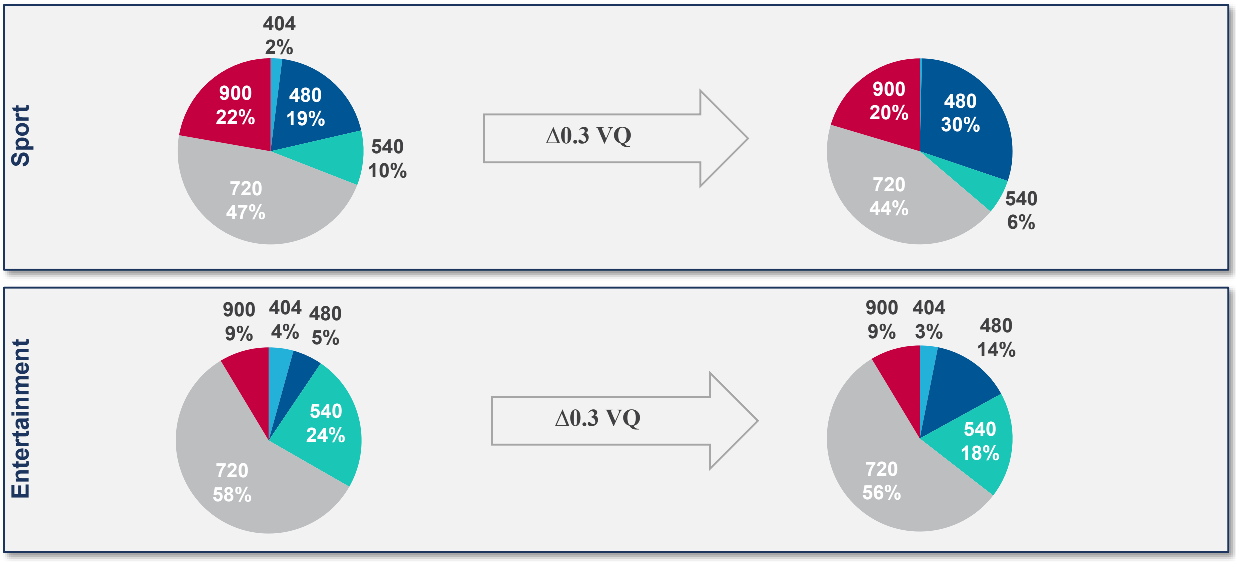

Figure 6 - Percent of duration using a different resolution than 480p at 1.5 Mbps nominal service profile

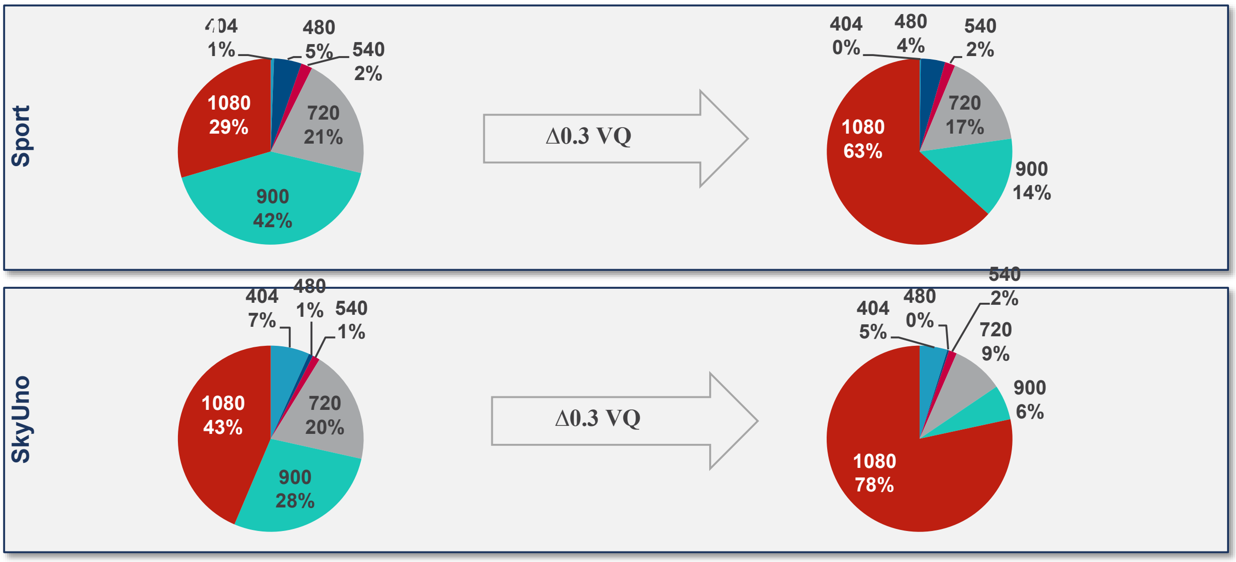

Figure 7 - Percent of duration using a different resolution than 1080p at 3.9 Mbps nominal service profile

Bitrate saving

Dynamic Resolution Encoding has been tested across many types of content and using different standards. About 50% bitrate savings has been observed.

For example, in HEVC, 4K quality can be encoded at 8 Mbps instead of 16 Mbps, 2.5K can be encoded at 5 Mbps instead of 10 Mbps, and 1080p at 3 Mbps instead of 6 Mbps.

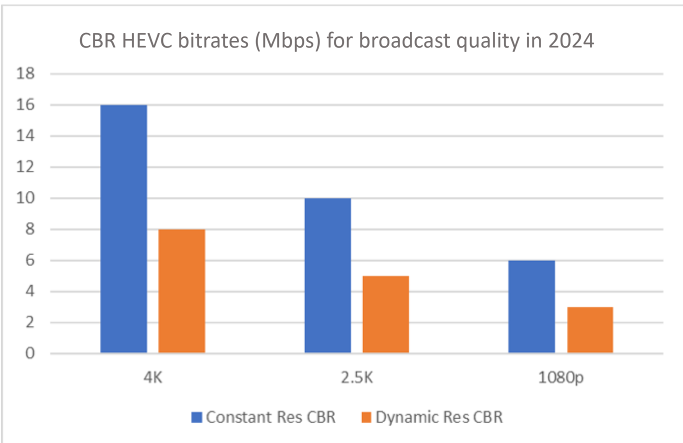

Figure 8 - CBR HEVC bitrates (Mbps) for broadcast quality

DRE technology can be applied to all encoding standards. We started testing with VVC, and Figure 9 and Figure 10 show the resolution selection. The content was encoded at 5 Mbps.

For a lot of content, the highest resolution is often selected, and for sport content, which is the most complex to encode UHD resolutions (2160p and 1440p) are still selected 60% of the time.

Figure 9 - Clips used in Figure 10

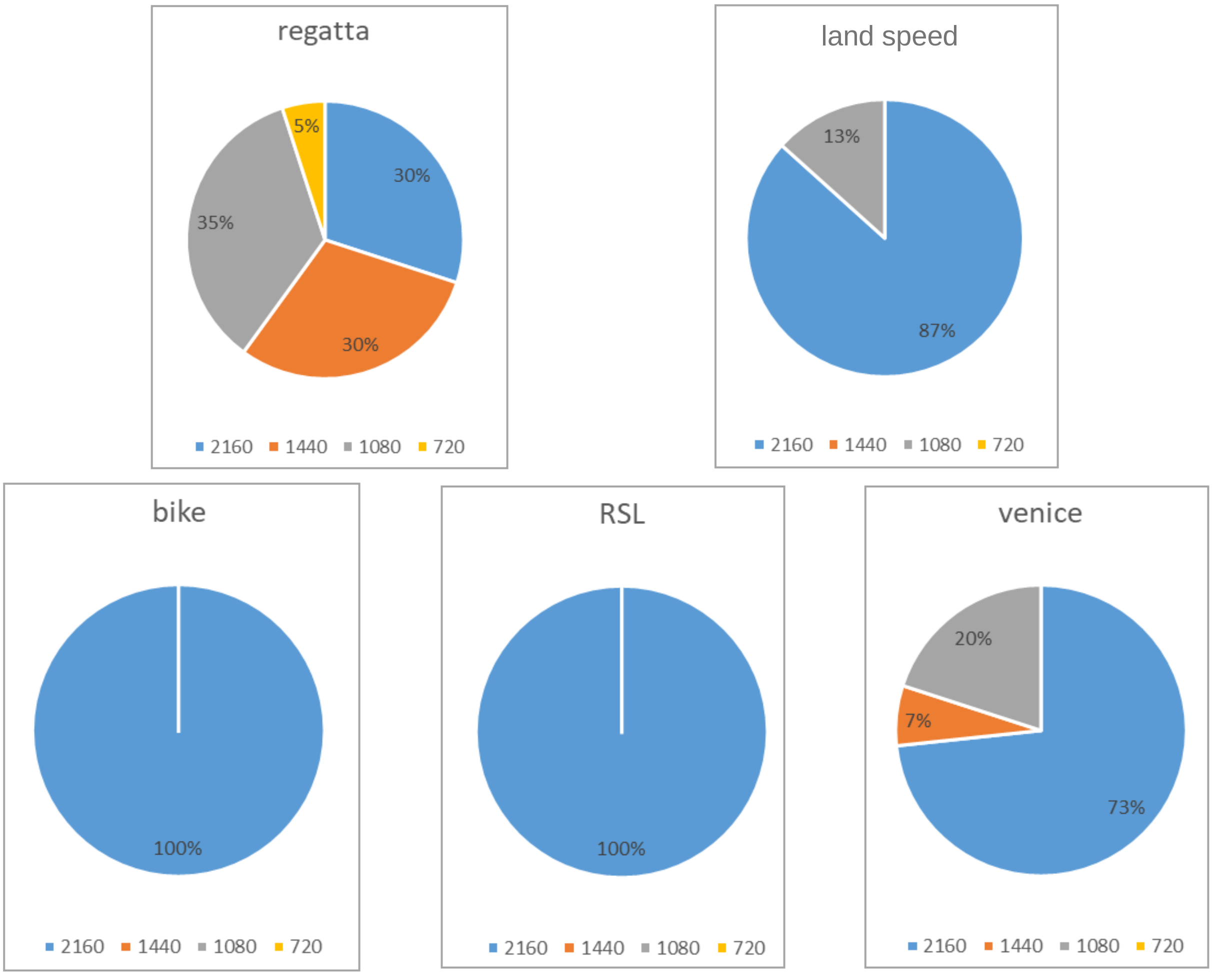

Figure 10 - Resolution selections on five clips

Video quality scores

For the five clips above, we looked at the quality, both with objective measurement and subjective viewing. Table 4 shows the VMAF scores.

clip	VMAF score	comment
regatta	77.95	+2.51 avg vs pure 4K, up to +10.5
land speed	92.05	+0.01 avg vs pure 4K, up to +0.2
bike	92.97	
RSL	93.32	
venice	87.56	+0.05 avg vs pure 4K, up to +0.4

Table 4 - VMAF scores for the five clips at 5 Mbps with DRE

For the subjective assessment, we can conclude that the quality is good, even for the regatta content, which has a lower VMAF score. We checked the quality of the streams using a dynamic resolution (i.e., regatta, land speed, Venice) in comparison with pure 4K quality. There is no perception of resolution change. Encoding artifacts are not visible on the most complex scenes, and the lower resolutions do not exhibit a loss of details.

Interoperability

We tested the streams with a variety of OTT STBs in HTTP mode and did not find any problems on the tests done in AVC and HEVC.

Density Improvements

The density has been measured and shows significant improvements. It is interesting to note that those improvements depend on the type of content.

For complex content, using a 720p resolution instead of 1080p provides about 50% CPU savings. Indeed, the number of pixels from 1080p*1920 to 720p*1280 is 56%, but the selection of the lower resolution does not happen all the time, so on average, the gain is 50%. Going down to 540p resolution can save around 65% of CPU.

For simple content, together with CAE, the encoding of one single 1080p representation instead of three representations can save ~40 to 50% CPU.

DYNAMIC FRAME RATE ENCODING ALLOWS FOR IMPROVED DENSITY AND QOE

DFE consists of optimizing the frame rate used for encoding depending on the video to encode. This topic has been researched in various papers (5, 6).

The appropriate frame rate is selected by a machine learning algorithm. It is well known that video with fast movement requires a fast frame rate, otherwise the video will look jittery and unnatural.

However, if the movement is slow or if there is little movement at all, like in a freeze frame, then encoding at a high frame rate will not produce an improved video. In fact, encoding at full frame rate will cost CPU to encode all the pictures, even if a lot of blocks will be simply copied from previously encoded pictures at the end. The encoding at full frame rate also generates a higher bitrate because the prediction algorithm will not produce a null bitrate for pictures, which have not been skipped. In fact, the encoded frame rate will vary depending on the movement in the video.

DFE provides density and bitrate savings.

Results

Density improvements

The density savings are dependent on the video source content. Let's start with average results. For HD content and a mix of movies, sports, and general entertainment the algorithm eliminates up to 40% of images. When measuring the CPU usage, it gives a gain of up to 30%. When looking at specific content, we find that the gain will be different for sports with constant movement.

The gain can be higher for documentaries and movies, where motion is generally not very high, while for sports it will be lower. For many sports, there is fast action but also many shots with less or no movement, like a global view of a field, portraits, and displays. In that case, the CPU gain is lower but quite close to the average gain.

The test sequences used to demonstrate the technology at the 2019 NAB Show showed a savings of 36% in frames using a wide variety of sequences (i.e., documentaries, sports, news, movies).

One important point to mention is that compared with the classical approach where all frame rates will have to be computed (p15, p30 or p60), the DFE technique will save encoding by a factor of three.

Table 5 shows the average frame drop over a large number of clips.

1080p5994 input clips	% frame drop
Global all clips (2702 clips)	**38.8%**
260 clips short clips 1080p59	33.0%
2 long clips (from 1080i) 1080P59	35.2%
1 long clip movie1 1080p59	59.5%
3 long clips sport 1080p50	8.0%
6 long clips movie2 1080p50	49.6%

Table 5 - Dynamic Frame Rate Encoding average frame drop on progressive content

1080i2997 input clips	% frame drop
Global all clips (272 clips)	**32.4%**
260 clips short clips 1080i29	30.1%
2 long clips 1080i	29.7%
1 long clip movie1 1080p29	40.0%
3 long clips sport 1080i25	16.0%
6 long clips movie2 1080i25	44.4%

Table 6 - Dynamic Frame Rate Encoding average frame drop on interlaced content

Bitrate saving

The bitrate savings also depend on the source content. With a mix of movies, sports, and general entertainment, the average gain is around 10% bitrate savings when using the MPEG-4 AVC codec. With the HEVC codec the bitrate gain is around 5% to 10%. As expected, the gain is less in HEVC than AVC as HEVC has higher performance to remove redundancy between pictures.

Interoperability

We have tested the streams with a variety of STBs in TS and HTTP modes and did not find any problems on the tests done in AVC and HEVC.

CONCLUSIONS

Table 7 summarizes (1) the different benefits of the AI-based compression developed in this paper.

	Bitrate saving	QoE improvement	Native density improvement	Density improvement vs. legacy techniques	Interoperability
DES	20-40%	NA	NA	NA	TS/HTTP
DRE	up to 50%	Yes	40% CPU saving	3-4x	HTTP
DFE	5-10%	NA	30% CPU saving	3x	TS/HTTP

Table 7 – Summary of the three examples

DES and CAE have been deployed by many of our customers with great success. We are now deploying the first customer with DFE and are working on commercially deploying the DRE technique.

The proposed methods have been designed for live applications and have demonstrated very good interoperability results on TS and HTTP (i.e., DES, DFE) and HTTP (i.e., DRE).

In this paper, we have developed three different applications of AI for video compression. Indeed, it shows once more how general and widespread machine learning technology can be.

It shall be noted that these techniques can be combined. In that case, there are collective benefits, including bitrate savings, QoE improvements, and CPU savings, at the same time. Of course, as mentioned in the paper, depending on the content or the technique, one will be more effective than the others.

Finally, the paper has shown that the potential for progress using AI for video compression is quite wide. The path is set, and the speed of advancement is increasing rapidly.

REFERENCES

1. Konstantina, K.D. 2015. Machine Learning Applications in Cancer Prognosis and Prediction. Computational and Structural Biotechnology Journal, Volume 13, pp. 8-17.
2. Liu, Y., et al. October 2018. Artificial intelligence-based breast cancer nodal metastasis detection. Arch Pathol Lab Med.
3. Agrawal, A., Gans, J., Goldfarb, A. April 2018. Prediction machines: The simple economics of artificial intelligence. Harvard Business School Press.
4. Netflix. December 2015. Per-title encode optimization. https://medium.com/netflix-techblog/per-title-encode-optimization-7e99442b62a2.
5. Katsenou, A. V., Ma, D., and Bull, D. R. 2018. Perceptually aligned frame rate selection using spatio temporal features. Picture Coding Symposium (PCS. IEEE. pp. 1–5.
6. Huang, Q., Jeong, S.Y., Yang, S., Zhang, D., Hu, S., Kim, H.Y., Choi, J. S., and Kuo, C.-C. J. 2016. Perceptual quality driven frame-rate selection (pqd-frs) for high-frame-rate video. IEEE Transactions on Broadcasting, vol. 62, no. 3, pp. 640–653.
7. Gary, M. September 2019. DeepMind's losses and the future of artificial intelligence. Wired. https://www.wired.com/story/deepminds-losses-future-artificial-intelligence/
8. Cyrus, R. July 2019. Bias in AI: A problem recognized but still unresolved. Tech Crunch. https://techcrunch.com/2019/07/25/bias-in-ai-a-problem-recognized-but-still-unresolved/
9. Ohm, J.R. June 2018. Recent developments in video compression standardization. CVPR CLIC Workshop, Salt Lake City.
10. Diascorn, J.L. April 2019. AI technology is changing the future of video compression. NAB 2019 technology paper, Broadcast Engineering and Information Technology Conference.
11. Diascorn, J.L. October 2019. How AI technology is dramatically improving video compression for broadcast and OTT content delivery. SMPTE 2019 Proceedings.

ACKNOWLEDGEMENTS

Many thanks to Patrick Dumenil, Thierry Fautier and Xavier Ducloux at Harmonic for their additions and review of this paper.