THEO

ENHANCEplayer

INTRODUCTION

The volume of online streaming is rapidly increasing, in terms of number of hours viewed and in terms of average bitrate. This puts a severe load on communication networks, both for the operators of these networks and for video service suppliers that need to pay for the capacity.

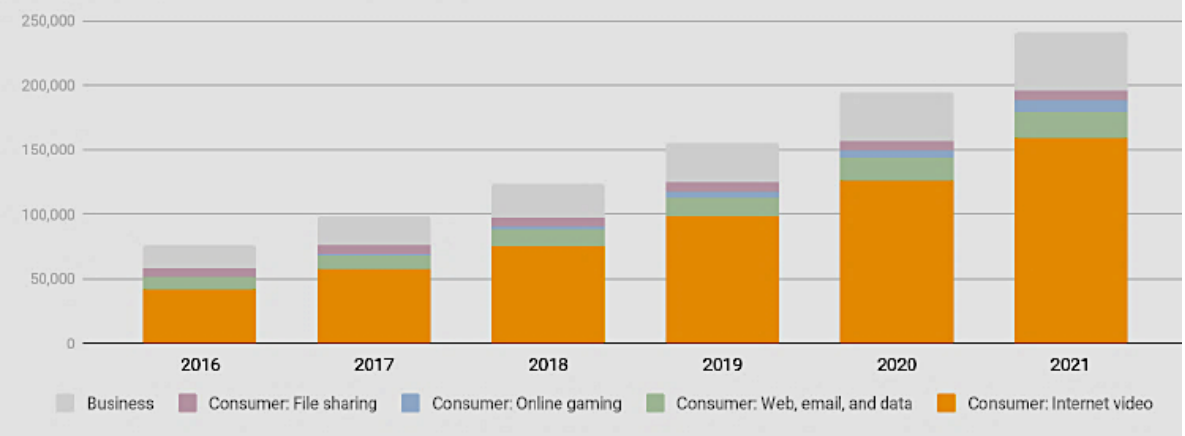
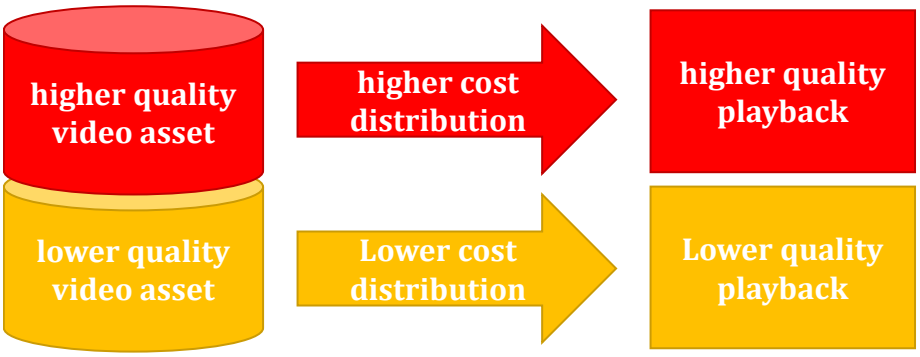


Figure 1: Evolution of the IP Traffic (Petabytes / month) since 2016.

Viewers ideally want the highest possible quality. Besides optimizing the encoding parameters, higher high quality is achieved by higher resolutions and higher bitrates. Higher bitrates mean higher costs for the video service suppliers.

Unfortunately, higher bitrates are not always realistic, especially in crowded locations or areas with reduced wireless coverage. In that case viewers fall back to a lower bitrate and associated lower quality.



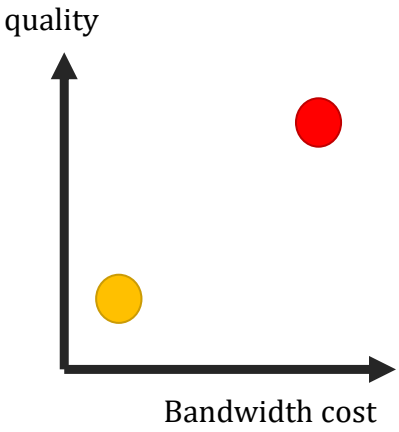


Figure 2: Today viewers must select either a high cost / high quality solution, or a low (bandwidth) cost / los quality solution.

The ENHANCEPlayer project targets to overcome this by offering a solution whereby a lower bitrate (and hence lower quality) version of the videos is transferred to the client, followed by a quality enhancement step on the client. This approach significantly reduces the network load, while still being able to offer a high quality to the viewers.

Several approaches exist to increase the quality of video streams on the client, including

- Approaches to increase the resolution of the video, also called upscaling
- Approaches to increase the framerate of the video
- Approaches to improve the dynamic range and color gamut of the video.

This leads to a high(er) quality for a low bandwidth cost.

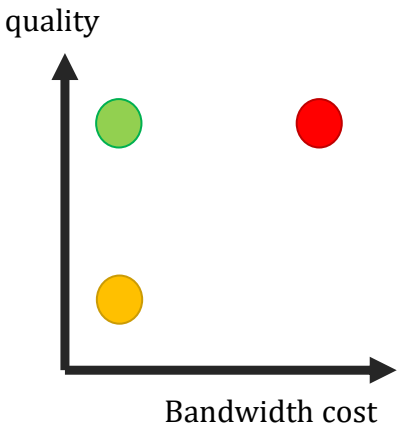


Figure 3: Upscaling of the video on the client allows to combine a low bandwidth cost with a higher quality.

ENHANCEplayer will focus on the first approach. The ENHANCEPlayer solution will upscale incoming videos on the client device. Several solutions exist, from interpolation techniques to artificial intelligence (machine learning) based solutions. Since machine learning solutions give better results, we adopt these techniques in the ENHANCEplayer project.

ENHANCEPLAYER BENEFITS

The ENHANCEPlayer system allows to use lower bitrate streams for a quality comparable to higher bitrates. The obtained savings can be important. A lower resolution immediately reduces the bitrate. As follows from figure 4, selection a lower resolution significantly lowers the required network capacity.

16:9 aspect ratio	H.264/AVC
416 x 234	145
640 x 360	365
768 x 432	730
768 x 432	1100
960 x 540	2000
1280 x 720	3000
1280 x 720	4500
1920 x 1080	6000
1920 x 1080	7800

Figure 4: Apple's authoring guidelines for the bitrate per quality (resolution).

Consequently, the ENHANCEPlayer leads to cost savings for the video service providers and leads to higher visual qualities for the viewers. Video service suppliers need to pay CDN costs, proportional to the amount of data transferred. For large video suppliers this implies a cost saving of several millions. The quality improvements moreover lead to more customer loyalty, boosting revenues.

ENHANCEPLAYER SYSTEM

The ENHANCEPlayer system uses a lower resolution video to be sent to the clients. At the client, the player decodes the low-resolution video. The decoded low resolution video is then processed by an upscaler module. The output of the upscaler module is then rendered.

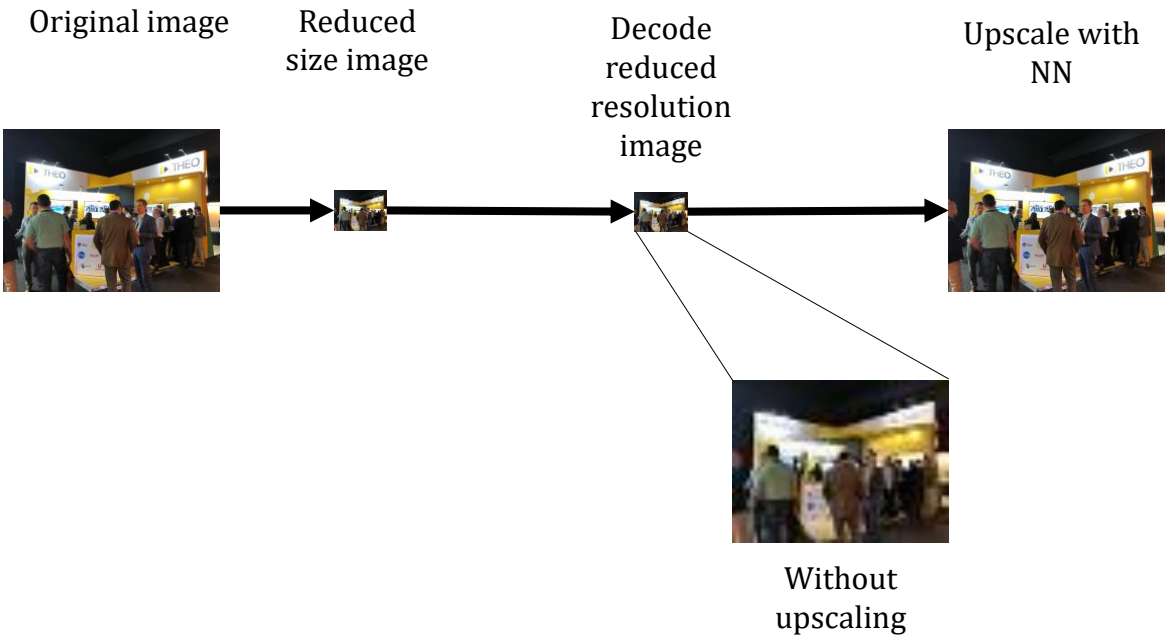


Figure 5: The ENHANCEPlayer system sends a lower resolution image to the client. At the client, this image is upscaled to a higher resolution.

UPSCALING TECHNIQUES

TRADITIONAL UPSCALING

Traditional upscaling starts with a low-resolution image and tries to improve its visual quality at higher resolutions. This improvement is done by

- Repeating pixels
- Interpolating pixels (e.g., bilinear interpolation or a nearest neighbor interpolation)
- Applying image processing filters to reduce artefacts

All these techniques apply mathematical formulae without looking at the content of the images.

NEURAL NETWORKS

Upscaling through neural networks take a different approach. The neural network uses the low-resolution image and interprets that image to predict a higher resolution image. Because the neural networks "understand" and internally represent the content of the image, these networks can better upscale the image and improve the upscaling.

> A neural network is a mathematical system that consists of so-called neurons that are connected to each other and that together perform a complex function such as image classification. The exact operation of the network is defined by its architecture and by the parameters of the network (also called weights). These parameters are not explicitly programmed but are "trained" by presenting a large range of examples, together with the ideal solution, to the neural network. The training phase of the neural network will define all the parameters (weights) such that the network displays the desired behavior, which is to mimic / repeat the training set. As of that moment, the network will independently work on new inputs as well.

The result of upscaling through neural networks typically looks sharper, has more detail and clarity.

CHALLENGES

Streaming video is often consumed on constrained devices, such as mobile phones and in web browsers.

For mass distribution to (mobile) devices, high compression ratios are applied. As a result, artefacts are introduced. The neural network can be tuned to cope with these artefacts and even remove the artefacts.

IMPROVEMENTS

Several methods were employed in parallel to achieve a significant reduction in network memory and number of mathematical operations (speedup). Specifically, three strategies were employed, (1) Network pruning, (2) Quantization (3) Depthwise Separable Convolution, and (4) Optical flow.

Network pruning

The memory size and execution speed of a neural network can be optimized by reducing the number of neurons and the number of connections between the neurons. This has been done by removing neurons and connections that are hardly relevant for the final result by carefully ensuring that the removal has no noticeable impact.

Quantization

Each operation of the neural network can be executed at normal precision (32 or 64 bits) but can also be executed at lower precision such as 8 or 16 bits.

Depthwise Separable Convolution

Additionally, we chose to employ Depthwise Separable Convolution to further reduce the computation burden of the network by breaking a single 3D convolution up into multiple 2D convolutions followed by a 1D convolution.

Optical flow

Optical flow is a motion detection concept that describes the relative motion of objects, edges and surfaces to an observer in a scene. The use of optical flows reduces the amount of computation needed

to increase resolution of a video. Instead of processing video frame by frame, adapting optical flow allows only increasing resolution of "key" frame for a sequence of frames which are in the same scene.

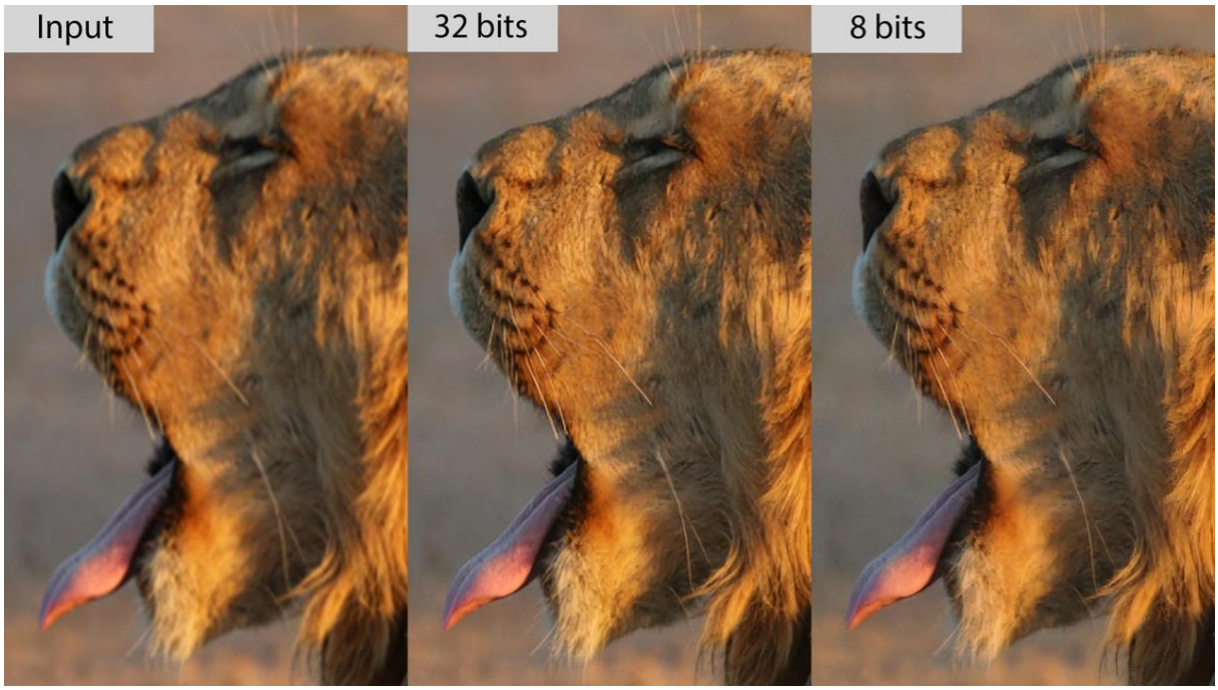


Figure 6: Left: Input image. Middle, 4x Upscaling using 32bit floating point precision. Right, 4x Upscaling using 8 bits quantization with only a minor loss in visual quality.

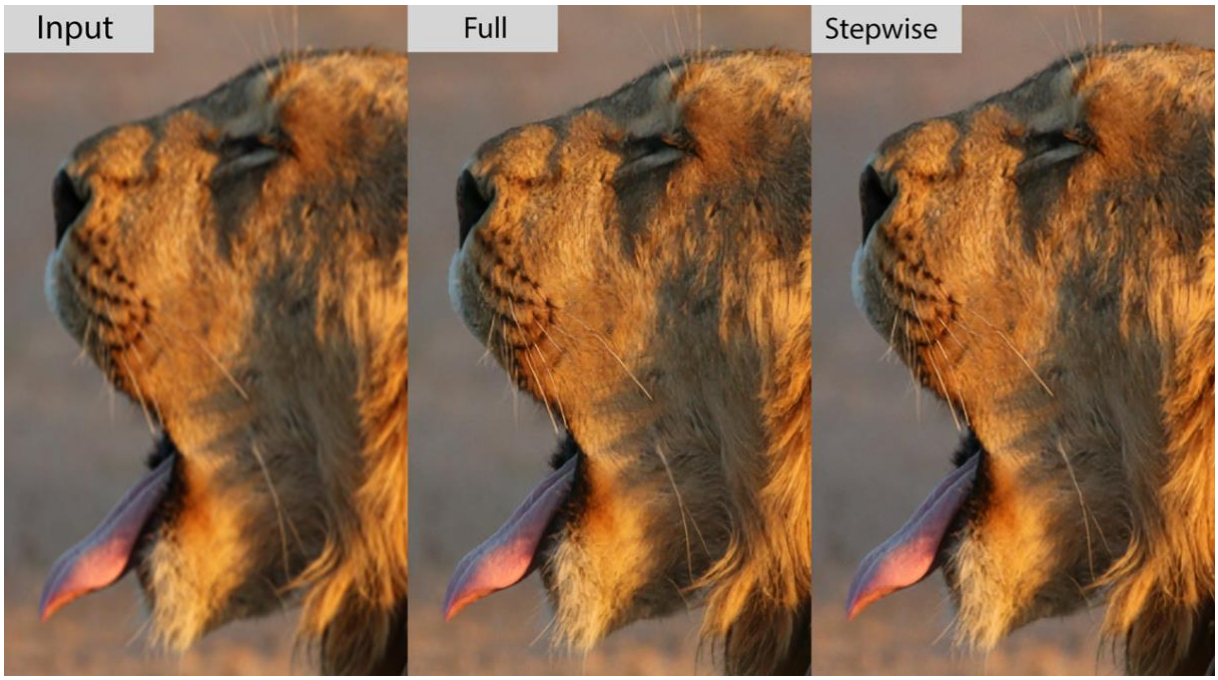


Figure 7: Left: Input image. Middle: 4x Upscaling using Full 3D convolution. Right: 4x Upscaling using Depthwise Separable Convolution.

ENHANCEPLAYER RESULTS

The ENHANCEPlayer solution outperforms traditional approaches. The traditional interpolation results in less sharp faces and backgrounds. The upscaling preserves a natural and crisp look and feel.

Bicubic interpolation

Upscaled with AI

Ground truth

Traditional upscale ENHANCEPlayer

Figure 8: ENHANCEPlayer results in crisper, sharper and more natural images.

LIMITATIONS

At this stage, the ENHANCEPlayer technology works fine for the most recent mobile chipsets with neuron coprocessors (e.g. Apple iPhone 11). Other devices do not have the required performance.

OPPORTUNITIES

The neural network models for upscaling streaming video can be tuned to the content for better results. This further improves the upscaling quality for the same (or even lower) computational requirements.

CONCLUSIONS

The ENHANCEPlayer solution is a very promising technology to bring significant bandwidth savings and higher visual qualities. The visual quality is much better than upscaling qualities obtained using traditional interpolation techniques. In blind tests, the results obtained from ENHANCEPlayer give excellent results. The ENHANCEPlayer technology pushes the state of the art when it comes to executing highly demanding neural networks on resource-constrained platforms such as mobile devices,

The ENHANCEPlayer technology has been partially funded by a EC H2020 grant

CONTACT US

Discover what THEO can do for you.

www.theoplayer.com
contact@theoplayer.com