THE BROADCAST BRIDGE

Live HDR Down-Conversion

Essential Guide

EG

ESSENTIAL GUIDES

Introduction from Varun Patel

Welcome to the Essential Guide covering the issues of Live HDR Down Conversion, supported by Lynx Technik AG.

By reading this guide, we hope that you will appreciate that doing HDR processing without careful consideration can result in poor customer experience.

This guide explains the challenges and best possible solution for a Live HDR down conversion.

We at Lynx Technik AG, in close corporation with RheinMain University of Applied Sciences in Wiesbaden, Germany, have researched and developed an approach to handle the issues of Live HDR Down conversion from a better and more scientific angle.

Best regards,

Varun Patel, Product Manager
Lynx Technik AG

Varun Patel, Product Manager, Lynx Technik AG

Live HDR Down-Conversion

By Tony Orme, Editor at The Broadcast Bridge

Moving to HDR is arguably one of the most innovative advances for the viewer experience in recent years. In this Essential Guide, we investigate why and how we automate the HDR to SDR conversion process. It's not as easy as it may first appear.

It's worth reminding ourselves of the problem HDR solves; it is to provide an image that more accurately represents the scene being captured. This results in greater detail in the highlights and shadows, better tonal range, and more vibrant colors, than traditional SDR video.

HDR down conversion has been available in the computer industry for many years as the digital images generated, with 16bit resolution, has far exceeded the resolution we've been working with in SDR in recent years. Consequently, a great deal of research has been conducted and many scientific papers have been written to support the research. The major difference for broadcast television is that these processes didn't work in real-time.

Modern digital cameras have a much higher dynamic range than most domestic television screens. This silent HDR camera innovation has been creeping in over the years and the camera video outputs can easily provide the HDR experience viewers are demanding.

As with most advances in broadcasting, the innovation is being driven by live sports. Not only is this where pay-per-view excels but is where the most demanding viewers are seeking the highest quality possible. The increased dynamic range and color space improves pictures enormously, especially where an event is played out on a grass field as the BT.2020 color space extends the greens to deliver beautiful rendition and saturation.

SD to HD Migration

When we moved from SD to HD back in the 1990's there was still a proliferation of cathode ray tube televisions in viewers homes so the decision to keep the existing color space made sense. Down-converting HD pictures by using aspect ratio converters made integration and simulcast transmission of HD and SD streams much easier. This encouraged the installation of HD workflows as the down-conversion to SD provided SD delivery to viewers homes that matched, or even surpassed, their existing broadcasts.

However, we are not afforded the same luxury with UHD-HDR. New flat screen technology has provided the opportunity to extend the color space to meet BT.2020 by designing new color filters in the flat screens. Combined with the increase in dynamic range and resolution, there are now three new dimensions in relation to UHD and HDR; increased vertical lines and horizontal pixels, expanded color space, and increased dynamic range.

It would be wrong to say that UHD-HDR and HD-SDR are mutually exclusive, but it's certainly difficult to derive SDR from HDR, and vice versa. The extended dynamic range of HDR cameras gives the outside broadcast vision engineer much greater latitude for balancing the extremes of luminance. In a sports stadium, where part of the pitch is in shade and the other is in bright sunlight, the vision engineers are constantly riding the iris to keep the area of interest at the correct exposure. The average picture level (APL) changes and the vision engineer must make sure the players and the ball can be easily seen.

Crushed Shadows and Highlights

The extra f-stop's of latitude will allow more of the scene to be correctly represented as there will be more detail in both the shadows and highlights. This means the vision engineer, if anything, doesn't have to ride the iris controls as much. But this is only the case when balancing for HDR workflows. As the SDR image has a reduced dynamic range and bit depth, the shadows will be crushed and highlights blown out, thus reducing the detail in both.

If the HDR to SDR solution could be achieved by direct linear mapping the luminance from the HDR to SDR, then we would just need to apply a simple scalar function to the HDR image. But the HVS responds to contrast in a non-linear fashion, and localized areas as well as global areas affect how the contrast is perceived.

In the ideal world, the sports production would have two signal paths, one for SDR, and the other for HDR. In practice, this is virtually impossible to achieve as we would need two sets of everything, including cameras. And the cameras would need to be physically in the same place to obtain the required angles.

Therefore, there is the option for the vision engineers to set the exposure levels for SDR and HDR using two different strategies due to the extended, or compressed dynamic range, depending on whether you're looking at it from the SDR or HDR perspective.

HDR Challenge

And this is the basis of the challenge broadcast facilities face, especially in real time – do they expose for SDR or HDR? There are more SDR television sets in the world, but nobody will buy HDR televisions if there are no HDR transmissions. And if nobody has HDR televisions, or the uptake is low, then the pay-per-view models that rely on them will create insufficient revenue for the investment in UHD-HDR workflows.

One solution is to acquire, process, mix, and broadcast in UHD-HDR. Then just before transmission, down-convert from UHD-HDR to HD-SDR. However, as well as changing pixel count, color space, and dynamic range, we must also change bit depth resolution from 14bits to 10bits. The analogue signal from the cameras image gathering device is processed by A/D converters to provide 14bit resolution processing within the camera. After this, the 14bit resolution is converted to 10bit for SDI distribution as both HDR and SDR use 10bits in the SDI distribution infrastructure. It is possible to use 12bits but this is rarely used in practice. This further complicates the process as the number of bits available to represent the image give rise to limited tonal variance.

The challenge of down-converting is much more difficult in the HDR world. The principal reason is due to the relative contrast within the image and the human visual response to it. Furthermore, most of the real-time computationally intensive calculations must be done in software.

Color Perception

For example, simultaneous contrast is the way two adjacent colors or shades affect each other. A grey square surrounded by a white frame is perceived as darker than the same shade grey surrounded by a dark frame.

And adjacent colors in different context affect the way the HVS interprets and perceives them. See figure 3.

The tone mapping operator (TMO) is the method of converting a pixel or image from a high dynamic range to one that has a more limited dynamic range. In this case, from HDR to SDR. The main criteria are to maintain as much dynamic range and detail as possible within the limited dynamic range.

The concept of TMO has its roots in the celluloid film industry. Capturing the massive amount of light contrast in a typical scene on a chemically limited negative proved difficult. To address this, the film manufacturers developed a method of compressing the shadows and highlights to give the familiar s-curve.

It is possible to pull the shadows and highlights from the film image in grading. Although digital cameras used for shooting feature films have a linear density to exposure relationship, the s-curve is needed to compress the shadows and highlights in the image to store as much information as possible in the limited delivery mechanism for later grading. This results in the camera providing outputs such as the s-log.

S-Curves

The analogy here is that film is given a higher dynamic range with the s-curves than the paper it is printed on or movie screen it was projected to. The point is, the graders and colorist could extract details in the shadows and highlights to provide as much detail and contrast as possible.

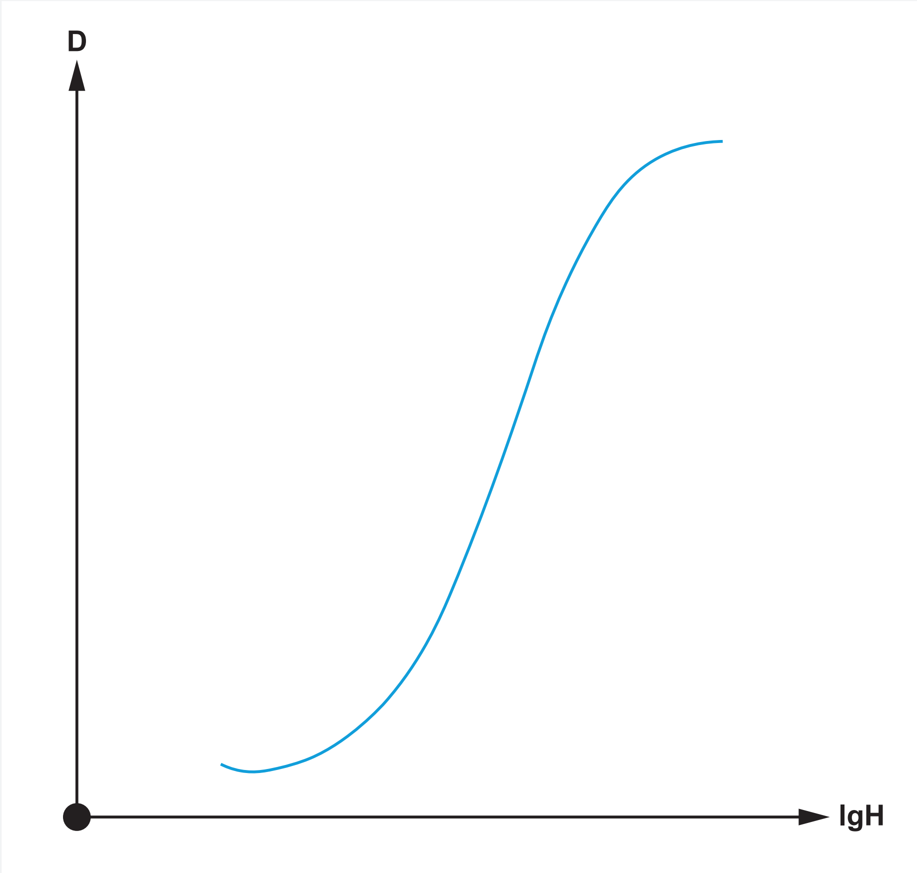


Figure 1 – In photography, each type of film stock has its own type of s-curve. The IgH (light-H) axis is the log of the exposure and the D is the image density on the film. Changing the emulsion properties moves the curve left and right. In modern digital workflows, the TMO equivalent is provided by computers and real-time processing.

And this highlights another interesting situation. In a blockbuster movie, period drama, or documentary, the colorist will work very closely with the director of photography and program creatives to make sure they are providing the correct aesthetic look. In this context, the TMO is providing an aesthetic solution, but that might not always be the case. It might be that the program makers prefer a more technically correct look, such as in a wildlife documentary.

For many years, the television camera sensor has been able to record a higher dynamic range than the television can display. But if the raw camera image, without processing, was presented to the vision path, it would show significant levels of high-level clipping (depending on how the exposure was set). The video knee is used to limit this as it extends the dynamic range of the highlights at the expense of compressing them.

When a production is shot and exposed for HDR, the mapping to SDR is not always linear and trying to apply generic transfer functions can easily result in flat pictures giving little dynamic range.

TMO's have different characteristics depending on the outcome required. It might be that the producers prefer an aesthetically pleasing image in preference to a technically accurate one. Or another creative may want to maintain as much details as possible and to maximize the contrast. One of the challenges of applying the s-curve is that it compresses the image potentially causing loss of detail in both the shadows and highlights.

Due to the proliferation of computer imaging, a great deal of research has taken place over the years and TMO's have divided into two main types; global (spatially uniform) and local (spatially varying).

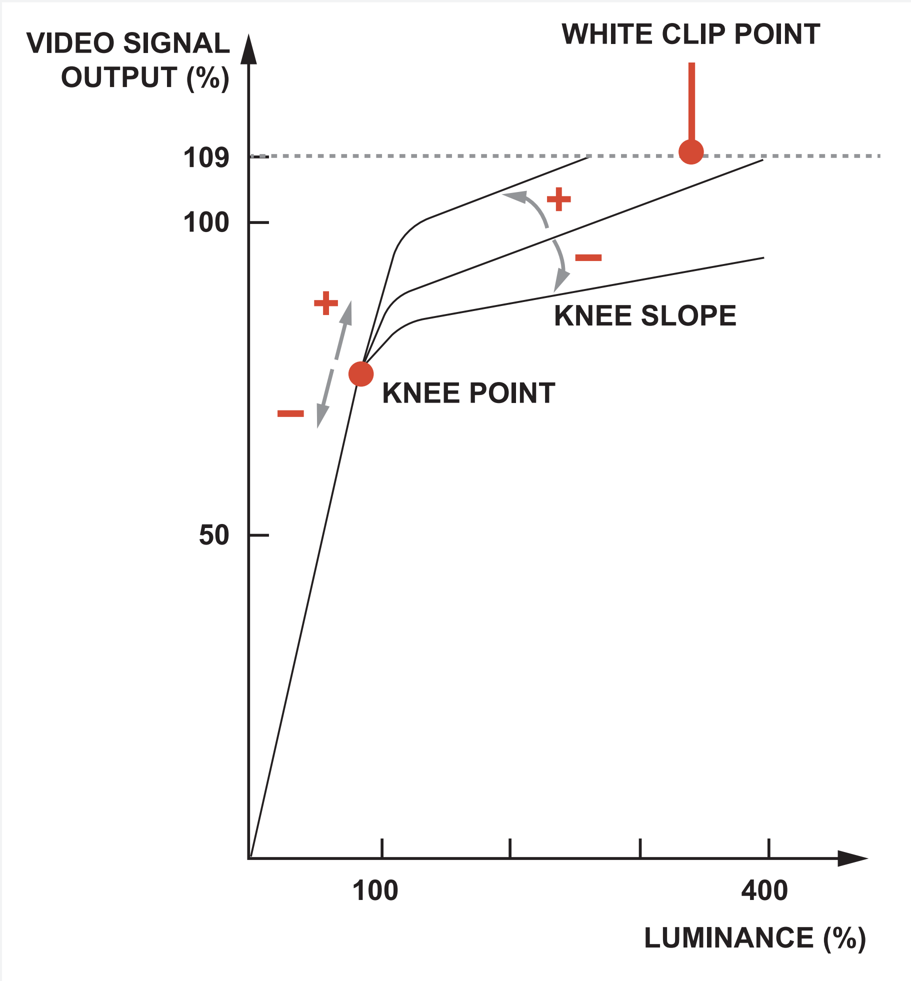


Figure 2 - Adding the knee helps extend the video dynamic range but the highlights are compressed. Strategies such as high-frequency video gain and edge enhancement are used to increase the perception of the detail.

Figure 3 – This is an example of simultaneous contrast. The two boxes on the top have the same color and the two colors on the bottom have the same colors. But we can see by looking at them that they clearly affect each other in our perception of the intensity and hue. The colors themselves haven't changed, but this phenomenon is caused by how they are laid out relative to each other and hence our perception of them.

TMO Optimization

Global operators analyze the image average luminance and optimize the TMO characteristic accordingly. Once the correct parameters have been calculated, they are applied equally to the whole picture and each pixel in it. Computationally, they are very fast as they are based on look-up tables (LUT's). However, they can cause loss of contrast and detail as they have a tendency to average out peak transients.

Local operators analyze each pixel and the immediate neighborhood of pixels around it. The non-linear function changes the value of each pixel sequentially depending on the content of the pixels around it. These algorithms are complicated and computationally expensive. However, they perform better than global operators for local areas as they mimic the HVS sensitivity to simultaneous contrast.

In effect, the global TMO's are replicating the iris and black level adjustments the vision engineer makes throughout the transmission. And the camera's AK (aperture corrector), whose function is to increase the contrast at the high frequency transients, is analogous to the local TMO.

Interesting, the TMO's do represent the s-curve in the film transfer function discussed earlier.

Furthermore, there is a tradeoff between the local and global operators. If the algorithm varies the parameters too much, relative to the neighboring pixels, this can have a negative effect on the overall global operators.

Research has continually demonstrated that six significant artifacts are known to influence the automated HDR to SDR conversion using TMOs; global flickering, local flickering, temporal noise, temporal brightness incoherency, temporal object incoherency, and temporal hue incoherency.

Global Flickering

Abrupt changes in successive HDR frames can have the effect of creating flicker in the SDR image. Although the HDR image brightness frame sequence may be stable over time, the TMOs use image statistics to determine the SDR luminance level which may be unstable over time, and this may exhibit consecutive frame variance leading to the SDR video flickering. The concept is best described in statistical process control (SPC) where 99.7% of measurements appear within the 3rd standard deviation. The 0.3% that occur outside of this can lead to long term instability. Therefore, in video HDR conversion, it is the extreme events that are of concern to us and can cause the global flickering. Long term time analysis is employed to stop this occurring.

Local Flickering Artifacts

Local TMOs generally work in a small area within the image to restore the relative contrast of the image in that area. The successive changes over many frames can lead to statistical instability in a similar way to global flickering. However, the flickering only occurs in the localized region of the image at the area in question.

Temporal Noise

To provide the best viewing experience, a group of TMOs deal with reproducing small amounts of detail in the SDR image from its HDR source. However, noise can make this difficult to achieve as its often a challenge to differentiate between high frequency, low-level transients, and video noise caused by the sensor.

Temporal Brightness Incoherency

This occurs where a large change in brightness occurs in the HDR image, but it is not represented in the SDR image after down-conversion. If the video cuts from a high luminance level to a low luminance level, the APL will have changed significantly. However, the TMO may not be aware such a large transition is intended and may significantly increase the luminance level in the SDR image to match the previous frames.

Temporal Object Incoherency

If a camera pans across a pitch, the APL changes due to the shadows caused by the sunlight falling across the stadium. If the TMO adopted the strategy of just adjusting the SDR APL based on the HDR APL, it's quite likely that the players would vary significantly in luminance level making them unrecognizable. In this example, the vision engineer working in HDR would adjust the exposure of the camera so the player would be better represented. However, this may not map to the SDR image acceptably.

Temporal Hue Incoherency

This may occur due to color clipping resulting in a color shift in the SDR image. If the red, green, and blue, values are not preserved by the TMO, the hue and saturation of one or more of the colors mapped into the SDR image may become unbalanced.

These primary artifacts show that mapping directly from HDR to SDR is a complex and difficult task and automating it in a live environment is even more so. The more frames that can be analyzed leads to a more accurate SDR representation. However, analyzing too many frames leads to excessive latency, something that is unacceptable for live transmissions.

The holy grail is to provide a UHD-HDR program that can deliver a far superior viewing experience for the audiences compared to the traditional HD-SDR service. But at the same time, respecting the SDR audience and providing the best program possible. As we have seen, it's almost impossible to provide a dual HDR and SDR workflow, but those that can are financially prohibitive. The optimal solution is to produce the program in UHD-HDR and provide an automated HD-SDR down-conversion that is as good as, or if not better than, the HD-SDR programs available today.

The Sponsors Perspective

Task-Specific FPGAs Offer Cost-Effective Solution To Simultaneous HDR/SDR Production

By Michael Grotticelli

The emergence of high dynamic range (HDR) acquisition in live and studio production has brought numerous possibilities for improved picture quality and creating an immersive experience for viewers. But it has also added cost and complexity to production workflows that broadcasters and content distributors are not eager to pay.

The greenMachine multipurpose 4K/UHD or 3G/HD/SD quad channel video and audio processing platform.

Consumers are increasingly seeking HDR videos on Netflix or similar streaming video providers, thereby prompting content producers to seek less costly ways to create more HDR content.

Many of the broadcasters' OB fleets are now 4K UHD capable and have been used in a number of broadcasts, but many professionals are still looking for cost-effective and practical ways to simultaneously shoot and distribute 4K UHD and standard-definition (SDR) content with HDR attributes. Indeed, another major challenge when implementing HDR is to maintain good backward compatibility with existing SDR displays and receivers. That's what most consumers are watching TV on.

Therefore, to support the overwhelming number of SDR TVs in the world, many productions have tested deploying separate but simultaneous workflows, shooting in SDR as well as in HDR. Besides the extra cost for dedicated HDR crew and equipment, having two paths for the same content is a challenge for the technical director that operates the OB production switcher, who now has to adjust the image for both audiences using a single camera Iris control. Producers also face difficulties when they have to produce content in multiple color spaces simultaneously.

In addition, those who produce in HDR and need to deliver both HDR and SDR content have struggled with the plethora of hardware necessary to deliver the two video streams.

Native HDR Workflows

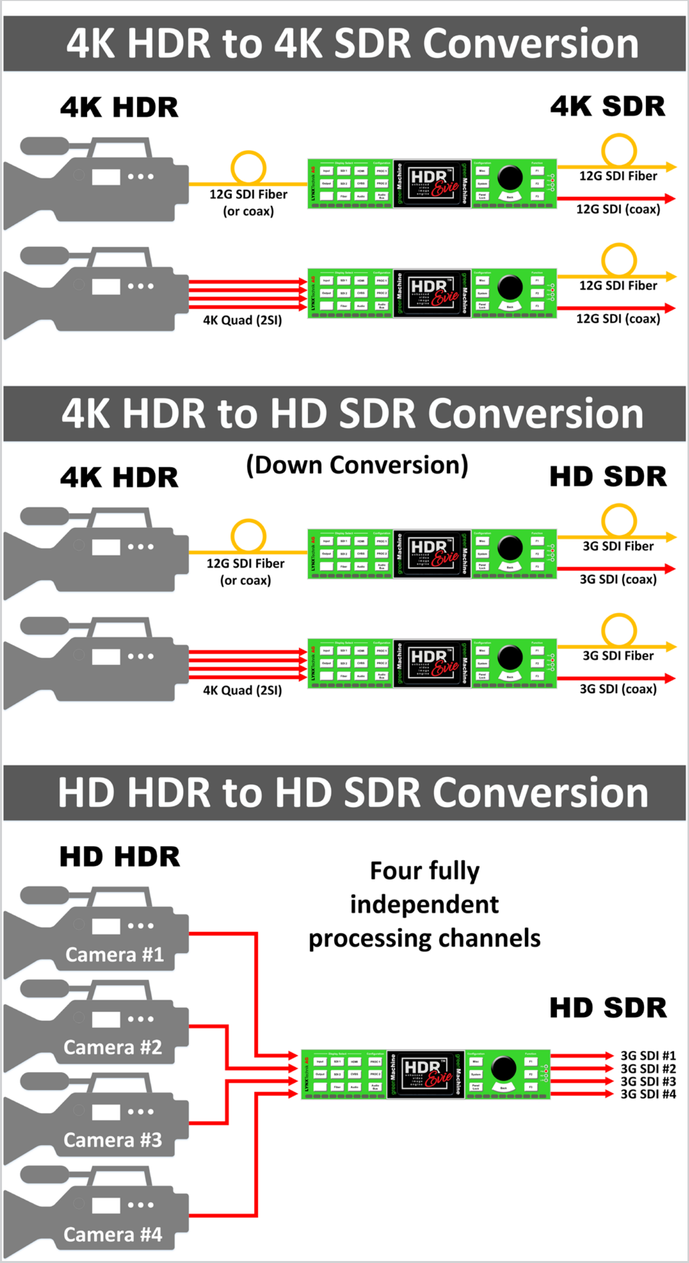
There are several production methods that can be used to deploy HDR but one workflow that is gaining traction is to produce content in a native HDR workflow, whereby all the switching accomplished in HDR and the signal is later down-converted to SDR to feed legacy displays. Most of the graphics or logos would be done in the SDR, which would then be upconverted to HDR (without any color changes) to be used in the HDR program stream. This will ensure the color of the graphics on both SDR and HDR signals remains the same.

This native approach needs to be carefully thought out. Looking to reduce the complexity and number of external devices required for two separate signal paths, Lynx Technik has developed a new FPGA hardware-based general-purpose audio and video processing platform that can be used to perform many different functions using a series of signal processing components (called “constellations”) in different configurations—depending upon the application at hand.

The greenMachine

This platform is called the greenMachine and it allows customers to configure the hardware for a specific application for as long as it is needed and then reconfigure it for a different purpose. This allows the customer to have one piece of hardware that can be used for multiple purposes and is launched with just a simple mouse click for different application. Specially designed greenGUI software, available for both Windows and Mac operating systems, can control an entire greenMachine system from a central location.

The greenMachine hardware is customer configured, using one of the pre-defined constellations. A constellation is a template or package of features. The constellation can also be defined as a factory-configured combination of processing tools such as HDR, Frame Sync, Up down cross converter, embedding-de-embedding, and color adjustments.



By installing different constellations (configuration of processing features) the greenMachine can be used for many different applications.

For example, the greenMachine titan hardware constellation processes four 3G/HD/SD-SDI video streams or a single 4K/UHD video input. It supports up to 12G processing (3840 x 2160 UHD @60Hz) and can convert between single-link 4K video (12Gps) and quad-link 4K video (2SI, 4x 3G).

Look To The Constellations

Two other greenMachine constellations are HDR static and HDR Evie (Enhanced video image engine), which are used for handling high dynamic ranges and color gamuts. This presents viewers at home with more dynamic images than previously seen, even without an up-to-date HDR display.

The HDR Static and HDR Evie constellations help streamline HDR workflows. HDR Static is recommended for video productions where lighting conditions are not dynamically changing and remains constant. It is best suited for the studio or indoor environment. HDR Evie on the other hand is as good as Static HDR in a studio environment but gives amazing results for outdoor environments where the lighting conditions are dynamically changing; such as in outside broadcasting.

The two solutions have their own use cases; with HDR static being capable of SDR to HDR conversion along with HDR to SDR, while HDR Evie is one of the world's first frame-by-frame HDR-to-SDR converters. HDR Evie applies color and contrast parameters to each frame. The two constellations also come with features like video adjustments, embedding/de-embedding, audio processing and shuffling.

Hardware Outperforms Software Processing

While many have considered software-centric systems that can process signals and perform tasks in parallel, Lynx Technik suggests that a software-centric approach is often expensive and time-consuming. In live production, the content needs to be distributed in real-time with no delay. HDR/SDR processing requires high computational speed with no latencies. Software running on GPUs offer slower processing speeds than FPGAs. And FPGAs have high cache memory that reduces the memory bottlenecks associated with external memory access.

The greenMachine hardware devices use the latest high-speed programmable Xilinx technology and dual ARM processors. The overall delay in processing HDR-SDR video is one frame in both the HDR Static and the HDR Evie constellations. This makes it an ideal solution for simultaneous SDR/HDR live event production.

Adding to the flexibility of the greenMachine for customers, the constellations are sold as licenses that can be activated or scaled back as needed. When more than one greenMachine is connected on a network, the license can be shared among several greenMachines. However, only one license can be activated on one machine at a time. This licensing model has proven to be invaluable to content distributors in other areas of the content chain and looks to have equal power for live production.

The Future Of HDR In Live Production Is Certain

There's no doubt that HDR delivers a "wow" factor when it comes to live production, such as sports. Some think an HD HDR image looks equal to a 4K UHD picture. That's why most experts expect to see more HDR video projects going forward, especially HD HDR since operators can reuse their already deployed 3G SDI infrastructures. Due to the extra cost involved, larger mobile production companies will move to UHD HDR sooner than smaller firms, but it is coming to projects of all sizes. It's only a matter of when business models are made to combine with technology strategies in a practical way.

Supported by

LYNXTechnik AG
Broadcast Television Equipment

WP
WHITE PAPERS

EG
ESSENTIAL GUIDES

MEDIA

CS
CASE STUDIES

Find Out More

For more information and access to white papers, case studies and essential guides please visit:

thebroadcastbridge.com

9/2019

Supported by

LYNXTechnik AG®
Broadcast Television Equipment