High-Speed Content Capture for Maximum Slow-Motion Impact – Set Free from Cables

David Edwards
IMT Vislink
Hemel Hempstead, UK

Bart van Dijk
Grass Valley
Breda, The Netherlands

William Norris
NEP Group
Bracknell, UK

Abstract - *Advances in increased visual clarity have played out before our eyes on 4k and HDR capable screens. It has been demonstrated that better pixels can translate to greater visual impact as they mimic the capabilities of the human visual system. However, it is also noticeable that in these latest systems, while the sharpness of static objects is impressive, the clarity delivered from moving objects has not kept pace. This joint paper from Grass Valley, NEP and IMT Vislink will describe the impact of balancing spatial and temporal clarity alongside the rollout of high dynamic range (HDR) while also considering the limitations of data bandwidth. It will propose the employment of frame rates of up to 150fps across time-interleaved UHD infrastructure, while also exploring the deployment and operational affordability of implementing a high frame rate, slow motion system that enables full slow-motion effects in conjunction with real-time video presentation.*

INTRODUCTION

Television audiences choose to view live Sports programming for many reasons. They may be an avid supporter of a particular sporting team who subscribes to every game, they may be caught up in the excitement of an international competition and feel compelled to participate in the talking point, the viewer's content discovery/EPG system may have recommended the event, or they may have simply stumbled across a piece of programming and been drawn-in to continue watching. Whatever the human nature pull towards an event, Program producers can significantly affect the viewer engagement by exercising skilled use of the resources available to them. Generating and holding viewer engagement in Sports programming comes, in-part from creating an innovative mix of visual impact, expert analysis, storytelling and getting the viewer to the heart of the action.

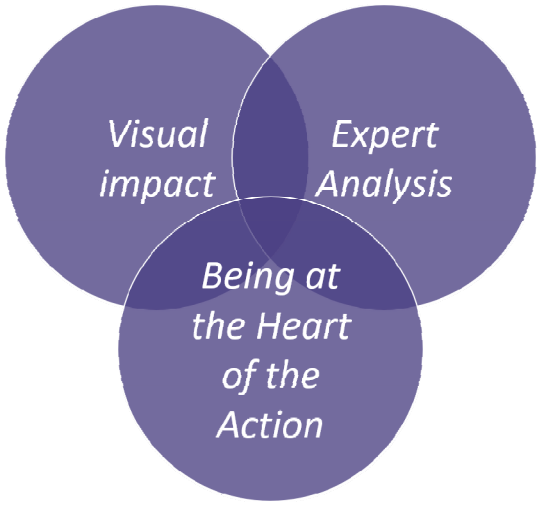

FIGURE. 1: CORE COMPONENTS IN GENERATING VIEWER ENGAGEMENT IN SPORTS PROGRAMMING.

There have been big advances in improving visual impact in recent years with many viewing platforms offering UHD tiers. Take-up of these higher resolution services is having different levels of impact across different regions with household penetration figures currently standing at 25% in North America, 17% in Western Europe and 9% in Japan [1]. High Dynamic Range as an additional tool is now in the first phases of rollout to the consumer. Both technologies create a visual impact on the consumer in different ways, each aiming to create a visual experience that more accurately reflects the capabilities of the human visual system.

As is so often the case, a complicated picture is emerging of which platform operators, content providers, and content types favor the emerging technologies. Economics of cost to produce and cost to distribute versus content subscription value will pay a key part in format selection. Consequently, the jury is still out on the relative success of UHD + HDR and HD + HDR.

The accelerating pace of change borne out by the migration to new video formats reflects the wider transformation taking place in the video industry. Organizations are willing to invest in technology – but within a controlled investment landscape. There is a desire to invest in technological change to better engage with the viewer with the goal of increasing market share, to invest to increase responsiveness to market demands, to invest to reduce costs and drive up efficiency, to invest to be seen to innovate and demonstrate thought leadership – All under a financial umbrella of doing more for less.

The visual impact improvements (UHD and HDR) that are now in the market have demonstrated that they not only support the visual impact drive for greater view engagement, they also underpin the viewer engagement that comes from Expert Analysis. Was a ball "in" or "out"? Was a foul committed? Who crossed the line first? The greater clarity gives greater certainty in the talking points for debate during the event, post-match and in homes and bars throughout the country. The enhanced spatial resolution that UHD or HDR provides cannot provide a complete picture to give absolute certainty on the unfolding action. Temporal resolution is also critical.

Temporal Capabilities of the Human Visual System

In understanding an event, positioning unfolding action in time is equally as critical as positioning the action in space. But just how quickly can the Human Visual System (HVS) comprehend the unfolding action?

A number of psycho-physical studies have been undertaken to gain insights into the temporal capabilities of the human visual system. The studies have investigated differing aspects of the time taken to comprehend key components related to image understanding.

The 2014 study in the Attention, Perception & Psychophysics journal [2] suggests that some humans are able to rapidly detect certain image types within a short time frame as low as 13ms – as part of a Rapid Serial Visual Presentation of images – If they have been prompted as to which image element they should search for amongst a series of images. The study used image examples such as a smiling couple as the target image for recognition.

FIGURE 2. FASTEST HVS IMAGE TYPE RECOGNITION.

Whilst the study into understanding the meaning of an image indicates that we are able to process images remarkably quickly it does not indicate that a human can obtain critical detail from an image within such short time durations.

A further study in the Journal of Neuroscience [3] examined the time frames involved in the HVS' ability to comprehend detail in an image. The research suggests that some people have the capability to recognize the orientation of an object in as low as 9ms. Color recognition takes longer – 25ms. Critically, recognition of both color and orientation takes even longer – within a 30ms time window – for simplistic images.

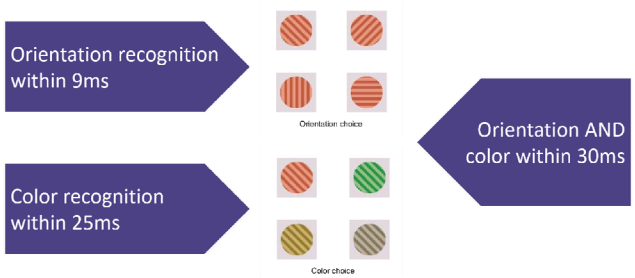

FIGURE 3. HUMAN VISUAL SYSTEM TAKES LONGER TO RECOGNIZE IMAGE DETAILS.

Aligning the results of these studies on the Human Visual System with the need to analyze unfolding events within a fast-moving sporting event to create an in-depth comprehension of the action reinforces the understanding that it is not possible for a human to accurately and reliably process all critical action within a real-time time frame. To accurately analyze a critical piece of sporting action there is now good evidence to back the long-held practice in slowing down time as part of Sports event analysis – The need for Slow Motion.

Slow Motion Solutions

In fulfilling the need to enhance the analysis of a sporting event and create the discussion points as part of the event story-telling, content producers need to reflect on how best to deploy the slow-motion need. Slow Motion solutions sit into two distinctive categories – Slow motion camera systems and synthetic slow-motion processing.

Synthetic slow-motion systems have found their place in the broadcast work flow. They operate by processing video frames to interpolate the differences in object position, relying on vector and object transformation modelling to predict frame content for any time point between real video frames. The computational power involved in the video processing now provides a good fit with the transition to a COTS, IT and software centric video

environment that the broadcast industry is becoming. Use of synthetic slow motion can lead to smooth action in slower than real time frame presentation. However, the challenge of generating perfect, artefact-free content is great and questions will exist in the Expert game analysis if the slow-motion content viewed was real or the result of imperfect image generation.

Slow motion camera systems provide an alternative method of acquiring slow motion content. They operate at a higher frame rates than standard video, which, when played back at normal frame rates provide slower than real time video presentation. By directly capturing every video frame, slow motion cameras provide certainty and accuracy in their image capture as all content is real and not interpolated.

When utilizing slow motion camera systems to provide high frame rate content capture there is a balance to be struck in the camera to achieve the required image quality. The ambient lighting levels, camera sensitivity and frame rate need to be managed with greater care than at lower (normal) frame rates. Similarly, the video quality threshold needs to be balanced through selection of frame rate and the resulting data payload. However, since the rollout of UHD, high data rate infrastructure has been developed and practical experience gained - driving down the cost of equipment and deployment.

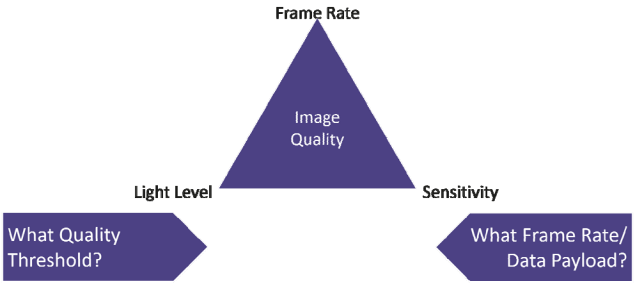

FIGURE 4. SLOW MOTION CAMERA SYSTEM TRADEOFFS.

COMBINED STANDARD FRAME RATE AND SLOW-MOTION CAMERA SYSTEMS

Technological developments have now brought together the ability for real-time standard frame rate live content capture and high frame rate capture to utilize readily available, affordable, broadcast infrastructure. Critically, the developments enable the live video and slow-motion video to originate from a combined workflow.

These combined solutions implement high frame rate capture at 3x normal frame rate. They utilize UHD connectivity to carry the high frame rate playout across multiple synchronized 3G-SDI interfaces. The production system then selects every 3rd frame for the HD, standard frame-rate production flow and ingests all video frames for the slow-motion frame store and playback production flow.

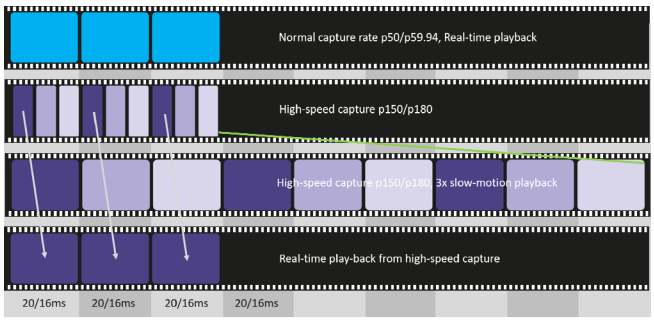

FIGURE 5. HIGH FRAME RATE CAPTURE FOR SLOW MOTION CAPTURE AND REAL-TIME PLAYBACK.

THE POWER OF "3" – WHY 3X SLOW MOTION?

The combined workflow, slow motion camera systems utilizes 3x normal frame rate – but why?
3x normal frame rate has been selected for bridging the best compromise across creative demands, technical limitations and the characteristics of the Human Visual System.

As examined earlier in this paper, the Human Visual System, whilst very fast in identifying some key elements of object recognition (sub 20ms) requires significantly longer to garner full details from an image – The research points to at least 30ms for simple images. The scientific research points towards, that for Events that are critical to the story telling or analysis, fast occurring action lasting just a few video frames may need to be slowed to at least 2x frame duration, preferably longer.

Examining the technological limitations, we have seen that utilizing higher frame rate content capture requires greater camera sensitivity, larger lenses or brighter lighting. Higher frame rates of course also lead to higher bit-rate video payloads and to higher costs associated with managing these payloads. Widely available and affordable in-camera video infrastructure is currently limited to solutions that have grown out of UHD – providing 4x synchronized 3G-SDI interfaces – in theory creating the opportunity to implement 4x slow motion.

Much higher frame rate technological solutions also exist under the Ultra Slow-Motion banner – However these systems require very specialized workflows available at a very different cost point.

Creatively, examining the fastest/slowest rate limits for audience perception of slow-motion content suggests that audiences struggle to identify 2x slow motion as having a slow-motion effect. At the other extreme, real time action sequences that undergo slow motion analysis as part of the storytelling of an event still need to convey a sense of pace and excitement. Slow motion story-telling sequences cannot last too long. It appears that the time taken to show 4x slow motion playback of a critical event may start to lose the

audience focus. In creativity of event production, 3x slow motion seems about right.

Bringing together the findings of what is desirable and possible across the spectrum of technical capabilities, needs of the Human Visual System to comprehend action sequences and the art of creating engaging productions, 3x slow motion appears to be the best fit across all the contributing factors.

GENERATING COMPELLING CONTENT BY GETTING TO THE HEART OF THE ACTION

In producing engaging programming that generates viewer interest, Event directors are continually exploring ways to evoke a sense of drama and to convey the unfolding story surrounding the action. Innovative camera angles are sought that provide a "really there" and personal point of view (PoV) experience.

Deploying equipment into shooting locations that deliver innovative camera angles can prove a challenge. Cable runs to line cameras may be lengthy. This may require complex preparation or create health and safety issues. Having multiple cameras locked-off to provide short duration PoV shots may prove difficult to justify on cost grounds as an addition to the core camera locations. In creating engaging and innovative camera content, flexibility is key – The camera must be free to be re-purposed to shoot from any location, free from cabling – Freedom to roam. Wireless camera systems deliver these benefits.

Wireless camera systems have existed for some years with the capability to deliver HD content and more recently UHD content. Wireless camera solutions fall into two technological categories: – Solutions that use bonded cellular technology – depending on widely deployed mobile phone infrastructure, and solutions that utilize a dedicated, private transmission network deployed especially for the event. For major, high profile, events and sporting occasions the private network solutions are preferred as they deliver a more reliable quality of service without the unknown risk of poor network contention that comes from a network share with mobile phone use of the venue spectators.

Private network wireless camera systems, typically built around variants of OFDM technology, have been designed to deliver a good balance of guaranteed high video quality and low latency within a constrained RF bandwidth. Such wireless camera transmission systems operate within licensed bandwidths of between 6MHz to 10Mhz channels. Clearly these bandwidths have data payloads significantly lower than the 3Gbit/s necessary to carry raw 1080p SDI video. Like many other components of the broadcast industry they utilize video (and audio) compression to reduce the required transmitted bandwidth to manageable rates. Latest state of-the-art systems use HEVC compression to achieve UHD video transmission within the bandwidth limitations of available frequency allocations.

These systems have the capability to provide RF transmissions with an operational range of sports stadiums or larger horse racing venues and provide a level of robustness and resilience to RF interference from multi-path reflections that comes from transmitting within an urban environment.

IMPLEMENTING A TRIAL WIRELESS SLOW-MOTION CAMERA SYSTEM

NEP Broadcast Services, Grass Valley and IMT Vislink recently partnered to trial and develop the ability to combine the enhanced story-telling and expert analysis that comes from slow motion with the flexibility, camera resource efficiency and innovative production values that comes from the use of wireless cameras to attempt to deliver a new tool to get the viewer to the heart of the action.

First trialled in the UK at Cheltenham race-course for a horse racing meet, the trial brought together a high-speed 1080p resolution broadcast camera capable of delivering 3x normal frame rate for content capture. The video connectivity from the camera utilized standard 3G-SDI interfaces with the high frame rate video synchronized across time-interleaved 3x SDI paths in a similar way to how quad-split UHD video is transmitted and synchronized across 4x 3G-SDI interfaces.

The high-speed camera was provided with wireless capability through the use of a wireless camera encoder/transmitter operating in "UHD mode" - Ensuring that each 3G-SDI signal was synchronized and presented through the transmission system without any risk of frame re-ordering.

The video feed from the high-speed camera was received at the production suite with the real-time, normal frame rate content and high frame rate content originating from the same camera, presented for use to the event director in both real-time and slow-motion workflows.

Having proved a success in delivering more engaging action during the first trial, the solution was then rolled out to a higher profile event at Aintree, Liverpool (again in the UK) for the premier National Hunt jump race - The Grand National. The solution was able to deliver content as part of the live video presentation as well as slow motion replay sequences – from one camera system, from one production flow. With Directors gaining confidence in the abilities of the system and exploring the creative possibilities it can deliver, new sequences have been put to air - live, such as a

time-slip shot of a jockey dismount - smoothly transitioning from real time to slow motion.

KEY FINDINGS AND FURTHER WORK

Creatively, born out by the content going to air, the ability to deploy slow motion and real time camera shots through a combined workflow appears to have led to more engaging programming and more immersive storytelling. Commercially, the benefits of deploying a wireless camera system that matches the capability of line camera equipment delivers an attractive rate of return by being able to be dynamically repurposed during the course of the event.

There is still work to investigate further the linked video compression trade-offs within the wireless transmission system between picture quality, compressed video bit rate and latency. Whilst some Event productions require seamless switching between line and wireless cameras with no through latency differential – now achievable with the latest wireless camera technology, under certain circumstances very complex and demanding images can stress the video compression algorithms. Greater learning of how to manage the event and camera direction can help build production guidance on how to get the best from the deployed solutions and deliver the story-telling, expert analysis and immersive event production that TV audiences seek.

EQUIPMENT AND SOLUTIONS UTILIZED

- Grass Valley LDX-86 Camera
- IMT Vislink HCAM and UltraReceiver wireless camera system
- Grass Valley LiveTouch slow-motion & instant replay system
- NEP Broadcast Services content production

REFERENCES

[1] Hisakazu Torii, Paul Gray, “Is there anybody out there?” IHS Markit Media & Technology Digest January 2019, Issue 568

[2] Potter, Mary, C., Wyble, Brad, Hagmann, Carl, Eric, McCourt, Emily, S. “Detecting meaning in RSVP at 13 ms per picture” *Attention Perception & Psychophysics,* February 2014

[3] Bodelon C, Fallah M, Rynolds JH. “Temporal Resolution for the Perception of Features and Conjunctions” *The Journal of Neuroscience,* 24 January 2007