Immersive Virtual Production with SMPTE-2110

Author: Ramiro Montes De Oca* (Immersive Dimension, SMPTE, IEEE, SBE),
Contributor: Chris Barnett° (Immersive Dimension)

Thanks!:
Paul Briscoe•(SMPTE-ST-2110-41 Board), Uday Mathur•(RED CTO), David Light•(NEP), Kenta Honjo•(RED), Ken Terry•(Dolby), Thomas True•(NVIDIA), Thomas Kernen•(NVIDIA), J.T. Rooney•(XR Studios), Hugh Hou•(@hughhou), Jeremy Hochman•(Megapixel CEO), Max Von Braun•(AGS Founder), Raymond Pao•(HTC Vive), Raymond Mosco•(HTC Vive), Tim Wen•(HTC Vive) , Mike Smith•(ROE), Evan Glantz•(Vū), Alvin Renz•(Vū CTO), Mark Swanson•(Vū Chariman), Tim Moore•(Vū CEO), Vitalii Boiko•(PixelaLabs), SMPTE•

*Author, °Contributor, • Acknowledgments
* Immersive Research:
* videofeedback@gmail.com
* https://www.linkedin.com/in/ramiromontesdeoca/
* https://github.com/videofeedback/

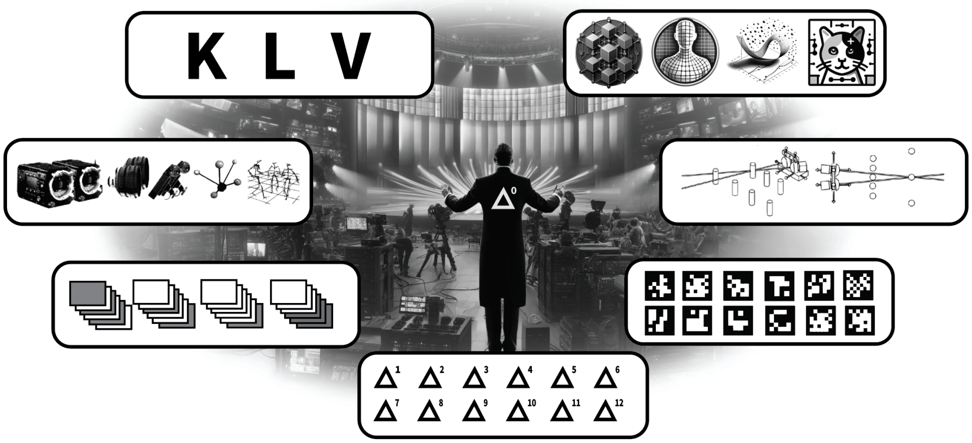

Figure 1 - Orchestrated Metadata Syncronization with SMPTE-2110-41/42

Abstract: In today's film production, Virtual Production has emerged as a pivotal tool for enhancing the creativity of the talent and directors. It enables them to bypass guesswork and immerse themselves fully in a virtual environment, promoting creative expressions. However, the adoption of these real-time tools is not without high cost and challenges with unforeseen delays. These issues often manifest across the stages of pre-production and unintentionally in production, and post-production. The arrival of Immersive Virtual Production, featuring 180°-Stereo-3D @120fps , significantly increases complexity, leaving no room for inaccuracies. To achieve flawless outcomes, a novel workflow and new methodologies are necessary. This approach must not only address new demands but also fix previously recognized obstacles. Here, we demonstrate, without new full infrastructure investment, and expanding the pallet of tools, a noble method to create Virtual Production workflows that solves old and new challenges with SMPTE-2110, which acts as primordial glue, enabling a fully tuned and synchronized Virtual Production orchestra. This coordination ensures precise real-time visualization allowing improved post-production processes with the help of new AI tools for interpolation and API integration.

Keywords: SMPTE-2110, Virtual Production, Immersive 3D, Generative AI, Interoperability, Precision Time Protocol (PTP)

1 - INTRODUCTION

1.1.1 - THE EMERGENCE OF VIRTUAL PRODUCTION:

In recent years, the film industry has increasingly embraced real-time visualization and virtual production, revolutionizing the traditional filmmaking process. This innovative approach allows directors and actors to interact with digital environments in real-time, significantly enhancing creative possibilities and immediacy. Traditional methods, reliant on blue/green screens, demand a high degree of imagination from talent, often hindering spontaneous reactions and natural performances. Virtual production, leveraging advanced technologies such as high-density LED screens and real-time CGI, enables immersive experiences that were previously unimaginable, allowing actors to perform opposite life-like digital creatures instead of placeholder props.

1.1.2 - CHALLENGES AND LIMITATIONS:

Despite its advantages, virtual production is not without challenges. High costs, technical complexities, and a shift of pressure from post-production to pre-production highlights the need for a refined approach. Current practices, including demanding requirements for Immersive 3D Virtual Production, often fall short, leading to a reconsideration of traditional methodologies. This paper acknowledges these hurdles, emphasizing the urgent need for innovative workflows that can address both existing and emerging demands within the industry.

1.1.3 - OBJECTIVE AND CONTRIBUTIONS OF THIS STUDY:

This study proposes a paradigm shift in virtual production workflows, aiming to resolve longstanding issues while accommodating the rigorous demands of modern film production. By optimizing existing hardware and introducing a SMPTE-2110 orchestrated system, we aim to enhance performance and efficiency significantly. Our research identifies critical areas for improvement, including: 1) Δ (Delta) Synchronization, 2)System Calibration, 3) Alternative Workflows, 4) Frame Remapping, 5) Deffered Rendering, 6) AI Tracking, Labeling, Classification, Interpolation and Gen-A, 7) SMPTE-2110 Metadata Management, 8) GPT-API Integration System. Through a comprehensive approach that integrates an Orchestrated Data Recording Through SMPTE-2110.

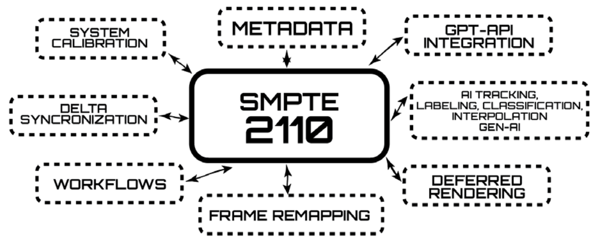

Figure 2 - Immersive Virtual Production with SMPTE-2110 System Data Integration

1.1.4 - SIGNIFICANCE AND BROADER IMPACT:

By addressing these challenges, our proposed workflows promise to streamline production processes, reduce costs, and unlock new creative potentials. This research not only contributes to the academic discourse on virtual production but also offers practical insights for industry practitioners, potentially setting new standards

1.2 - IMMERSIVE VIRTUAL PRODUCTION AS A "REALITY SIMULATION ENVIROMENT":

The distinction between Virtual Production for the Film Industry and Production for Virtual Reality applications has effectively dissolved. The array and sophistication of tools utilized in Immersive Virtual Production are comparable to, and at times exceed, those in Virtual Reality. This is why we approach everything as a "Reality Simulation" scenario rather than a theatrical solution. Once we solve how to replicate reality, we can bend those rules and physics later, in the creative process, allowing the creative mind to bend reality at will.
Below is a compilation of tools employed in Immersive Virtual Production and an explanation of why it is set to become the definitive standard for "Reality Environment Simulation Systems" (RESS). It's important to note that this perspective is rooted through System Engineering rather than being industry-specific, since the real challenge is the integration and interoperability of multiple tools for different industries. Consequently, our approach to data recording is framed through the holographic needs of multiple Research & Development areas such as Broadcasting, VR/AR/XR, Robotics, Manufacturing, Digital Twins, Machine Learning and Multimodal-Language-Models. Is the combination of all these fields of research and develepment what makes possible new alternatives to each industry.

Figure 3 - Real motion controlled scene with virtual production (Credit: Versatile Media Ltd.)[5]

1.2.1- IMMERSIVE VIRTUAL PRODUCTION WITH SMPTE-2110 POTENTIAL

In continuation, we are going to witness how the full system works. In this theoretical case study, we will exemplify how data can be collected and utilized in pre-production, and during production, and post-production, using an already motion controled production done with virtual production, and how a full production could be done instead.

1.2.2 - THE SCENE
This is a high-speed chase scene featuring a motion-controlled camera in sync with a motion-controlled car, motion-controlled lights, and a pre-rendered scene on an LED volume, covered with LED screens.

1.2.3 - PRE-PRODUCTION
Car Chase Film Plates: Part of the scenes can be shooted with real street footage on a car-plates setup. These scenes can contain stunts-coordinated scenes for realism. All these scenes records kinetic data, directly from the car's CAN-BUS[1] system and additional 5 Axis system to record 6-DOF[2] sensor information, that will help to synchronize the full scene with real life kinetics data. Cameras recording the different points of view will also include their own independent 6-DOF sensor information to have a full picture of sensorial information for each element in the recorded scene. Metadata is recorded to a SMPTE-2110-41/42 data stream, fully synchronized with the entire system.

1.2.4 - PRE-VISUALIZATION / PROGRAMMING
The full scene is animated on Previsualized CGI Sequences for cameras movements, car movements and full scene coordination. Preprogrammed camera shakes to actual motion-controlled-camera systems for realism, a digital 5 axis car chassis[3] movement recorded in real time on car plates and CGI is sequenced. All the Metadata is recorded on each digital scene with SMPTE-2110-41/42.

1.2.4 - KINETICS SIMULATION:
The scene adds real-life kinetics recorded on the Car-Chase scene to add realism to cameras and car movement. This is where fine tune happens for camera visuals and action effects. That information is feeded to a real-life Motion-Computer-Controlled car chassis. All the data is being recorded in SMPTE-2110-41/42 with synchronized Delta-Times[4] for each element in the scene.

1.2.5 - REAL LIFE CAR KINETICS INTEGRATION:
A real-life computer-controlled 6 axis car chassis is programmed with the real life kinetics recorded from previous scenes with ST-2110-41/42. Recorded kinetic recordings can simulate real G-Forces and can be attenuated or modified, but always keeping the timing associated with camera shakes and incidental movements. These kinetic recordings can simulate real G-Forces,
Real Life Computer-Control-Camera Integration: A real-life motion-controlled-camera is added to the pre-programming phase, this time fully integrated to the computer-controlled 6 axis car chassis. The scenes are coordinated with stunts and every shot is pre-programmed.

1.2.4 - PRODUCTION
LED Screen or Green Screen + Laser Projector: All the scenes are pre-rendered at high quality with already pre-defined shots. Camera movements, shake ups and cuts are part of the already scene ready from top to bottom. The talent sits on the car and the action is safetelly recorded for the talent. Improvisation is possible by switching workflows to Green / Blue or White screen workflows (see Workflows).
SMPTE-2110-41/42 recording methodology keeps all the information in one place (with the footage), without causing data loss or misplacement. The Metadata follows the footage, regardless where that footage goes. In this case, this allows spin-offs such as Park Attractions Rides, Kinetic Immersive Theaters, VR Kinetics Systems, etc.

1.2.5 - WHAT IS NEEDED:
Projects like the one described require extensive data management accuracy. Throughout this paper, we will break down each of these needs to create dynamic workflows capable of producing effortless productions like the one described.

2 - LITERATURE REVIEW

2-1 - SCOPE OF IMMERSIVE VIRTUAL PRODUCTION DATA RECORDING WITH SMPTE-2110

Here is a comprehensive list detailing what Immersive Virtual Production with Immersive Virtual Production with SMPTE-2110 will encompass in the future, including metadata and synchronized media recording; just a small sample of data collection:

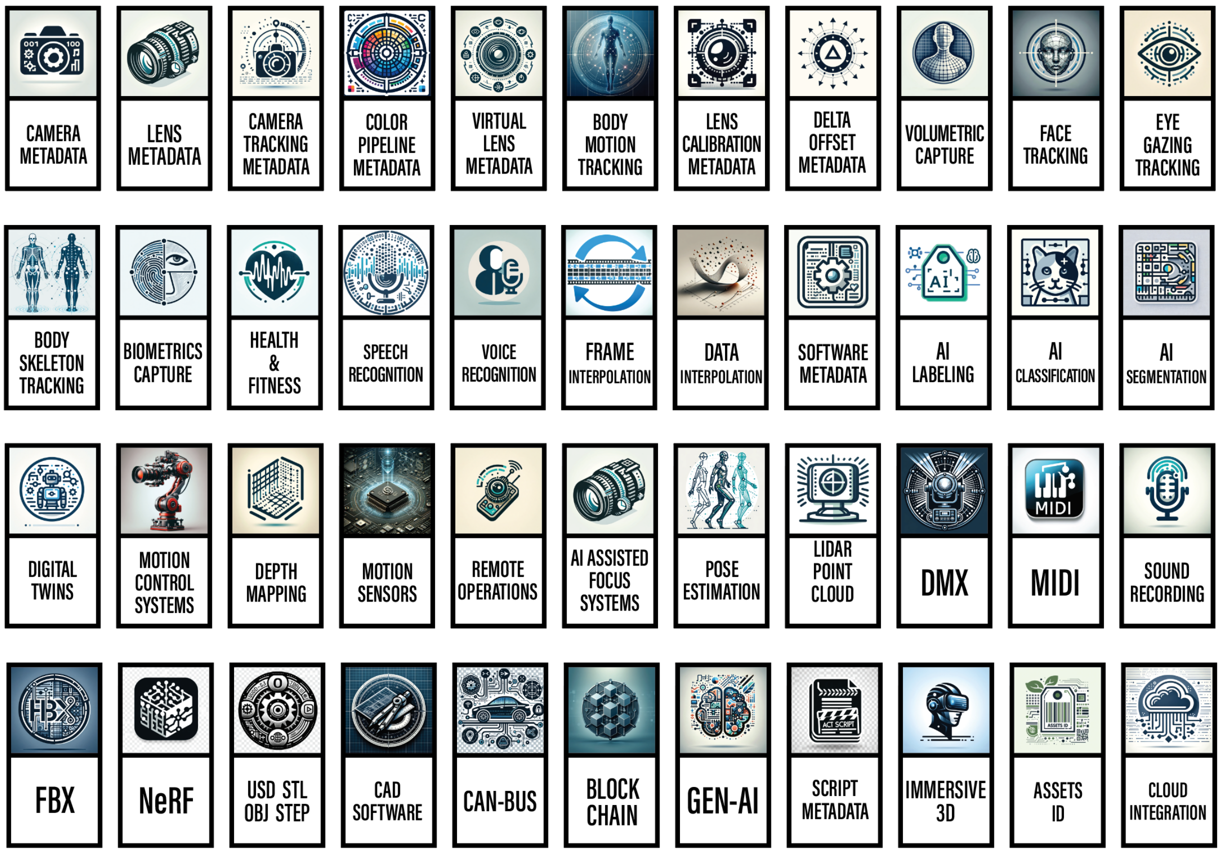

Figure 4 - Example of list of Immersive Virtual Production data recording and metadata.

2.1.1 - ENABLING INTEROPERABILITY IN IMMERSIVE VIRTUAL PRODUCTION

At the core of this Research Paper we are going to lay down the cornerstone of what will make all this possible. The Society of Motion Pictures and Television Engineers (SMPTE)[6] plays a crucial role in establishing interoperability standards that allow manufacturers, both large and small, to develop compatible systems. These industry standards enable faster development and seamless integration across different platforms and technologies.

2.1.2 - SMPTE - RIS (Rapid Industry Solutions) - OSVP:

(On Set Virtual Production): One of the key initiatives undertaken by SMPTE in recent years is the On-Set Virtual Production (OSVP) [7]program, which falls under the Rapid Industry Solutions (RIS)[8] effort. Launched in 2020, OSVP aims to develop solutions, guidelines, and best practices to address the urgent challenges faced by the industry in leveraging virtual production techniques, such as LED volume stages, camera tracking, lens information, and color calibration.
By harnessing the collective expertise of SMPTE's global membership, OSVP seeks to revolutionize the way visual content is produced, making the creation of realistic virtual environments more seamless and integrated into the physical production process.

2.1.3 - CAMDKIT (JSON METADATA):

The OSVP program has already achieved a significant milestone with the publication of CAMDKIT[9], a JSON[10] format metadata structure, on April 12th, 2023[11] . This new set of standards defines a recommended metadata structure for manufacturers to comply with, enabling better interoperability in on-set virtual production workflows. The CAMDKIT metadata structure facilitates the creation of new API[12] calls between future implementations, promoting seamless communication and data exchange across various virtual production tools and technologies.

By establishing these industry-wide standards and guidelines, SMPTE aims to address the critical need for interoperability in the rapidly evolving field of virtual production. This collaborative effort ensures that manufacturers can develop compatible solutions, enabling more efficient and streamlined workflows for content creators working with immersive virtual environments.

2.1.4 - CAMERA PARAMETERS WITH A DEFINED FORMAT IN CAMDKIT

These parameters were determined to be important for use in traditional post-VFX and having sufficiently clear definitions and practices to include in the CAMDKIT JSON format.
Most are available in the captured outputs of cameras. When available from a camera manufacturer's tools, CAMDKIT ingests and normalizes them to the listed format.
In addition CAMDKIT outputs the duration of the clip so that the sampling rate can be computed from the number of samples.

Parameter	Need vs. Like to Have	Static vs. Dynamic	Description	Units	Constraints
Active Sensor Physical Dimensions	Need	Static	Height and width of the active area of the camera sensor	micron	The height and width shall be each be an integer in the range [0..2,147,483,647].
Camera Firmware Version	Need	Static	Version identifier for the firmware of the camera	n/a	Unicode string betwee 0 and 1023 codepoints.
Camera Make	Need	Static	Make of the camera	n/a	Unicode string betwee 0 and 1023 codepoints.
Camera Model	Need	Static	Model of the camera	n/a	Unicode string betwee 0 and 1023 codepoints.
Camera Serial Number	Need	Static	Unique identifier of the camera	n/a	Unicode string betwee 0 and 1023 codepoints.
Capture Rate	Need	Static	Capture frame frate of the camera	hertz	rational number whose numerator and denominator are in the range (0..2,147,483,647].
FDL Link	Need	Static	Unique identifier of the FDL used by the camera	n/a	UUID URN as specified in IETF RFC 4122. Onlyu lowercase characters shall be used. Example: urn:uuid:f81d4fae-7dec-11d0-a765-00a0c91e6bf6
ISO Speed	Need	Static	Arithmetic ISO scale as defined in ISO 12232	unit	Integer in the range (0..2,147,483,647].
Shutter Angle	Need	Static	Shutter speed as a fraction of the capture frame rate	0.01 degrees (angular)	The parameter shall be an integer in the range (0..360000].

Figure 5 - https://github.com/SMPTE/ris-osvp-metadata/blob/main/vfx/camera-lens-tables.md

2.1.5 - LENS PARAMETERS WITH A DEFINED FORMAT IN CAMDKIT

These lens parameters were determined to be important for use in traditional post-VFX and having sufficiently clear definitions and practices to include in the CAMDKIT JSON format.
Most are available in the captured outputs of cameras ands lenses.
When available from a manufacturer's tools, CAMDKIT ingests and normalizes them to the listed format.

Parameter	Need vs. Like to Have	Static vs. Dynamic	Description	Units	Constraints
Anamorphic Squeeze	Need	Static	Nominal ratio of height to width of the image of an axis-aligned square captured by the camera sensor	0.01 unit	Integer in the range (0..2,147,483,647].
Entrance Pupil Position	Need	Dynamic	Entrance pupil position of the lens	millimeter	rational number whose numerator and denominator are in the range (0..2,147,483,647].
F Stop	Need	Dynamic	The linear f-number of the lens, equal to the focal length divided by the diameter of the entrance pupil	0.001 unit	Integer in the range (0..2,147,483,647].
Nominal Focal Length	Need	Dynamic	Nominal focal length of the lens	millimeter	Integer in the range (0..2,147,483,647].
Focus Position	Need	Dynamic	Focus distance/position of the lens	millimeter	Integer in the range (0..2,147,483,647].
Lens Firmware Version	Need	Static	Version identifier for the firmware of the lens	n/a	Unicode string betwee 0 and 1023 codepoints.
Lens Make	Need	Static	Make of the lens	n/a	Unicode string betwee 0 and 1023 codepoints.
Lens Model	Need	Static	Model of the lens	n/a	Unicode string betwee 0 and 1023 codepoints.
Lens Serial Number	Need	Static	Unique identifier of the lens	n/a	Unicode string betwee 0 and 1023 codepoints.
T Stop	Need	Dynamic	The linear t-number of the lens, equal to the F-number of the lens divided by the square root of the transmittance of the lens	0.001 unit	Integer in the range (0..2,147,483,647].

Figure 6 - https://github.com/SMPTE/ris-osvp-metadata/blob/main/vfx/camera-lens-tables.md

2.1.6 - PARAMETERS WITHOUT A DEFINED FORMAT

Parameters without a defined format

These parameters were deemed important but with varying meanings or mechanisms for extracting them. They are not implemented in CAMDKIT.

Parameter	Need vs. Like to Have	Static vs. Dynamic	Notes		
istortion Characteristics	Like	Dynamic	Usually computed by forumla or by software based on other parameters		
Exposure Falloff Characteristics	Like	Dynamic	Based on iris position		
Tint Value	Like	Static			

Figure 7 - https://github.com/SMPTE/ris-osvp-metadata/blob/main/vfx/camera-lens-tables.md

2.1.7 - PARAMETERS REQUIRING FURTHER STUDY INCLUDING USE CASES

These parameters were deemed potentially useful, but more clarity is needed on their definitions or the use cases for them.

Parameter	Need vs. Like to Have	Static vs. Dynamic	Notes		
Active Sensor Dimensions in Photosites	TBD	Static	Useful to estimate Moire in ICVFX. Need post-VFX use case.		
Capture Color Space	TBD	Static	Log, Lin, ACES, 709, 2020, etc.		
Circle of Confusion	TBD		Used to compute depth of field based on entrance pupil position, iris position, focus, and focal length. Multiple definitions, e.g., optical vs. VFX (pixel depth-based). Need clear definition.		
Depth of Field	TBD	Dynamic	Near and far plane distance. Essential or derived from other parameters?		
Debayered Output Resolution	TBD	Static	Not always known at capture time. Present in image characteristics?		
Monitoring Encoding Curve	TBD	Static	Need use case		
Pinhole Focal Length	TBD	Dynamic	Need use case. Derivable from other data?		
LUT	TBD	Static	Which LUTs matter?		
CDL Data	TBD		Need use case.		
White Balance	TBD	Static	Usually expressed as color temperature. Different curves. Effect on image varies by maker.		

Figure 8 - https://github.com/SMPTE/ris-osvp-metadata/blob/main/vfx/camera-lens-tables.md

2.1.8 - IMAGE CHARACTERISTICS ALREADY PRESENT IN ESSENCE FORMATS

These are metadata items that may be needed that are commonly present in essence formats (images or video) and are therefore not covered in this work.

Parameter	Notes
Image Frame Dimensions	Height and width of image in pixels
Colorimetry	Mapping of component signals to red, green and blue tristimulus values
Quantization	Quantization of component signals, including the bit depth
Sampling	Spatial sampling/sub-sampling of components

Figure 9 - https://github.com/SMPTE/ris-osvp-metadata/blob/main/vfx/camera-lens-tables.md

2.1.9 - GITHUB: SMPTE/ris-osvp-metadata

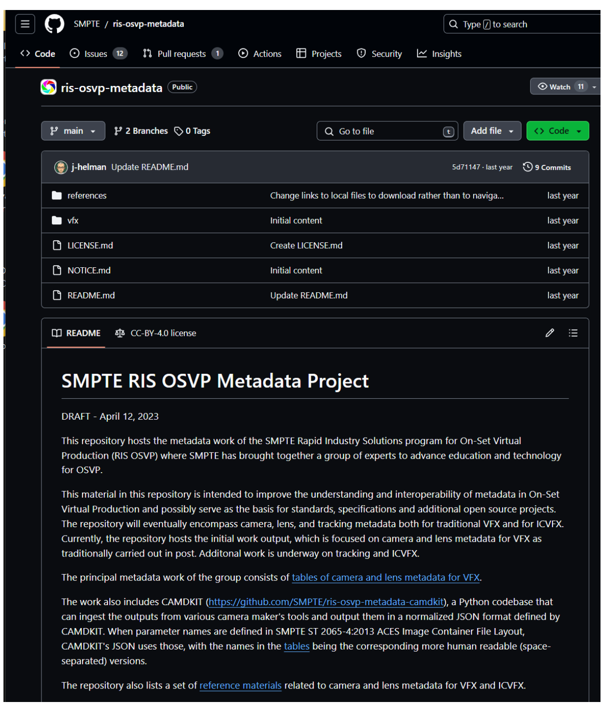

Figure 10 - https://github.com/SMPTE/ris-osvp-metadata

2.1.10 - GITHUB: SMPTE/ris-osvp-metadata/camdkit

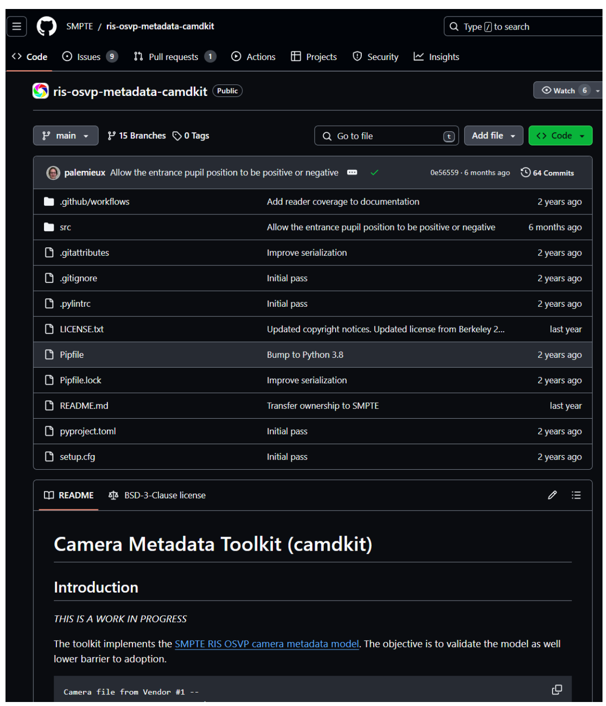

Figure 11 - https://github.com/SMPTE/ris-osvp-metadata-camdkit

2.1.11 - OUT OF THE SCOPE OF SMPTE-RIS-OSVP:

While SMPTE-RIS-OSVP focuses on specific aspects of virtual production, numerous other technologies and methodologies are currently considered outside its scope. These include:

Capture Technologies: Volumetric Capture[13]*, Biometrics Capture, Depth Mapping.
Data and Interpolation: Delta Offset[14]*, Data Interpolation[15], RAW Motion Sensors Data.
Artificial Intelligence: AI Labeling[16], AI Classification[17], AI Segmentation[18].
Advanced Technologies: Digital Twins[19], Motion Control Systems[20], Lidar Point Cloud[21].
Communication and Integration: Extended DMX[22] Data*, Cloud Integration, CAN-BUS[23].
Software and Formats: FBX[24], NeRF[25], USD[26], STL[27], OBJ, STEP, CAD Software.
Emerging Technologies: Blockchain, Generative AI (GEN-AI), Immersive 3D.
These areas, although not currently covered by SMPTE-RIS-OSVP, represent a broad spectrum of technologies that are integral to the development of a fully integrated immersive virtual production system.

2.1.12 - RATIONALE FOR EXCLUSIONS AND FUTURE INTEGRATION

The areas listed as out of scope are not included in the current SMPTE-RIS-OSVP standards due to various reasons such as the rapid pace of technological advancements, the need for specialized expertise, and the broadening scope of virtual production applications. However, the integration of these technologies is essential for the evolution of immersive experiences in entertainment, suggesting potential areas for future standardization and collaboration.

By addressing these aspects, SMPTE-RIS-OSVP and the broader community can continue to push the boundaries of virtual production, enhancing the creation and consumption of entertainment in novel and immersive ways.

2.2.0 - METADATA AT SMPTE STANDARDS:

Metadata has long been a crucial component of various SMPTE standards, including SMPTE ST 377-1 (MXF), ST 2052 (subtitles), ST 2067 (IMF) , ST 2057 (ECMS) , ST 2065-4 (ACES) , and ST 336 (KLV). However, none of these existing formats fully address the unique metadata requirements of the ST-2110 suite, which necessitates the encapsulation of precise PTP timestamps for synchronized media transport over IP networks.

2.2.1 - LEVERAGING EXISTING STANDARDS AND TECHNOLOGIES:

While the ST-2110-40 standard covers metadata in the ST-2110 domain, it is a retrofitted solution based on the legacy ST-291-1 standard, which was designed for ancillary data and subtitles. For more flexible and extensible metadata handling, the ST-336 standard's Key-Length-Value (KLV) structure[28] offers a simpler and versatile approach.

2.2.2 - PROPOSED ST-2110-41/42 (KLV):

The proposed ST-2110-41 (KLV) structire and ST-2110-42 (Package with format information), was initially proposed in 2019 by Paul Briscoe[29] (SMPTE Member) and Kent Terry [30](Dolby Laboratories) joined forces to bring "Fast Metadata" ST-2110-41 to the finish line, while ST-2110-42 is on hold until ST-2110-41 becomes a standard.
The KLV[31] structure separates metadata into three components: the Key identifies the type of data, the Length specifies the data's size, and the Value contains the actual data. This simple yet powerful structure allows for the dynamic encoding of various data formats, including JSON, FBX, Serial Data, FreeD, DMX, CAN-BUS, MIDI , Binary, hexadecima; and theorecitally also provides a way to encode custom or current formats such as JPG, PNG, USD, Python scripts, and encrypted data among others formats. However, the existing ST-336[32] standard does not natively support the crucial PTP timestamps required for ST-2110 workflows. The ST-2110-41Sub-Group is contributing to the ongoing efforts to establish a robust and interoperable ecosystem for IP-based media transport, enabling more efficient and seamless virtual production workflows, as well as other broadcast applications.

K	L	V
Key	Length	Value

Figure 12 - "Fast Metadat" ST-2110-41; (K,L,V) Data Format (Key, Length, Value)

2.2.3 - ST-2110-42 EXTRA EFFORT:
The combination of ST-2110-41 and ST-2110-42 aims to create a comprehensive and versatile metadata solution for SMPTE-2110 workflows. Similar to the flexibility of the CAN-BUS protocol used in the automotive industry and the elegance of the MIDI[33] protocol's System Exclusive (SysEX) messages, which allowed MIDI to evolve and adapt over its nearly 40-year existence, the proposed standards offer a robust foundation for fast research and development, enabling the creation of diverse and extensible data structures tailored to specific requirements. It will be the implementation of ST-2110-42 what will really bring ST-2110-41 to the next level of metadaga exchange.

2.2.4 - ST-2110-41/42 AS PROTOCOL
Unlike traditional methods that may require the development of entirely new protocols, ST-2110-41/42 builds upon the established foundations of ST-2110's RTP (Real-Time Protocol)[34] packets, transmitted over UDP[35]/IP and Multicast networks. This approach ensures that while the protocol's framework is predefined, it grants applications the flexibility to interpret, manage, or even disregard the metadata based on their specific requirements, guided by the structure provided in ST-2110-42.

2.2.5 EXPERIMENTAL IMPLEMENTATION BY DOLBY LABS OF 2110-41:
As mentioned before, although the proposed ST-2110-41 and ST-2110-42 standards have not yet been officially adopted by SMPTE, Dolby Laboratories has been utilizing 2110-41 as an experimental implementation. Kent Terry (Dolby Labs) is one of the major forces and part of the Sub-Group that is moving forward for 2110-41 to become part of the "ST" (Standarized) familty.
The proposed ST-2110-42 on the other side is being delayed until 2110-41 becomes a standard.
Dolby's experimental implementation serves not only as a precedent but also validates the need for ST-2110-41 and ST-2110-42 to become official SMPTE standards. This real-world application highlights the practical value and potential impact of these proposals on enhancing interoperability and streamlining metadata handling in SMPTE-2110 workflows.

2.2.6 - ST-2110-41/42: IP RANGES AND PORTS
Will be part of the SMPTE-2110-41/42 subgrup to determine the best way to assign recommended IP addresses ranges and Ports designation. Is always recommended to keep organized IP configurations and Ports without interferring with the rest of the ST-2110 ecosystem.

```
[F0] [MANUFACTURER ID] [MESSAGE TYPE] [LENGTH] [DATA PAYLOAD] [F7]
```

Figure 13 - MIDI SYSEX Data Format (ID, M-TYPE, LENGTH, DATA)

2.3 - MIDI AND SMPTE; A FULL CIRCLE:

The inception of the Musical Instrument Digital Interface (MIDI) organization in the early 1980s was a pivotal moment in music technology.
Over 40 years of experience handling digital media only and driven by the need for a universal standard that would enable electronic musical instruments from different manufacturers to communicate seamlessly. Pushed to the industry by visionary figures like Dave Smith of Sequential Circuits[36], MIDI was developed through unprecedented collaboration among industry competitors. The influence of SMPTE (Society of Motion Picture and Television Engineers), particularly its timecode standards, was instrumental in ensuring MIDI's compatibility with audio-visual synchronization, broadening its application beyond music to encompass multimedia production. This collaborative effort culminated in the unveiling of the MIDI standard in 1983, revolutionizing music production, performance, and digital creativity across the globe. The foundational principles of interoperability and innovation laid down by MIDI and its adaptability, as demonstrated by the integration of SMPTE timecodes, have ensured its enduring legacy and relevance in the digital age.
MIDI and SMPTE influenced each other not only on the interoperability between manufacturers, but also with the need to develop new technologies together with competitors, instead of creating their own standards.

2.3.1 - MIDI SYSEX MESSAGES:
System Exclusive (SysEx[37]) messages are a type of MIDI message used for sending data that is specific to a particular manufacturer's equipment. These messages can carry a wide range of information, from preset data and function settings to firmware updates. The structure of a SysEx message is designed to ensure compatibility and to prevent interference with the data of other manufacturers.

2.3.2 - MIDI'S PROTOCOL:
Delta time (Δ-t)[38] precise offsets is the secret sauce of MIDI implementation. The simplicity of this protocol, created in a time where 8 bits was compressed in two sets of 7 bits to allow two branch of 128 bytes messages (instead of one 256 bytes message), with the MSB (most significant bit, the first bit), marking the two different kind of 128 bits messages. MIDI created an elegant set of instructions that survived for 40 years without the need of biggest changes, since the protocol was so flexible that allowed a natural expansion of instructions through a series of simple special instructions that survived the times of 16, 32 and 64 bits instructions eras. Advanced-by-simplicity is the best way to describe the MIDI's protocol. Born at the beginning of the truly digital era, the MIDI protocol solved all the problems to be solved and future-proof in a 100% digital environment.

2.3.3 - EMBRACING AN OPEN (KLV) STRUCTURE:
The choice to integrate a KLV structure (TYPE OF DATA / LENGTH / DATA) with the precision timing offered by PTP reflects a strategic move towards creating a metadata transport mechanism that is both intuitive and self-contained. By leveraging this design, we are positioning the protocol to handle the evolving complexities of SMPTE metadata with agility and without necessitating immediate re-standardization for every new metadata type or use case encountered.

In embracing the ethos of adaptability inspired by Future-proof MIDI SysEx, ST-2110-41/42 aspires to foster innovation in metadata transport. This approach recognizes the balance between the need for a solid, standardized foundation and the demand for flexibility to address unforeseen challenges and opportunities in the dynamic landscape of media production. By doing so, ST-2110-41/42 is poised to support the next generation of media production technologies, ensuring that the standard evolves in concert with industry needs and technological advancements.

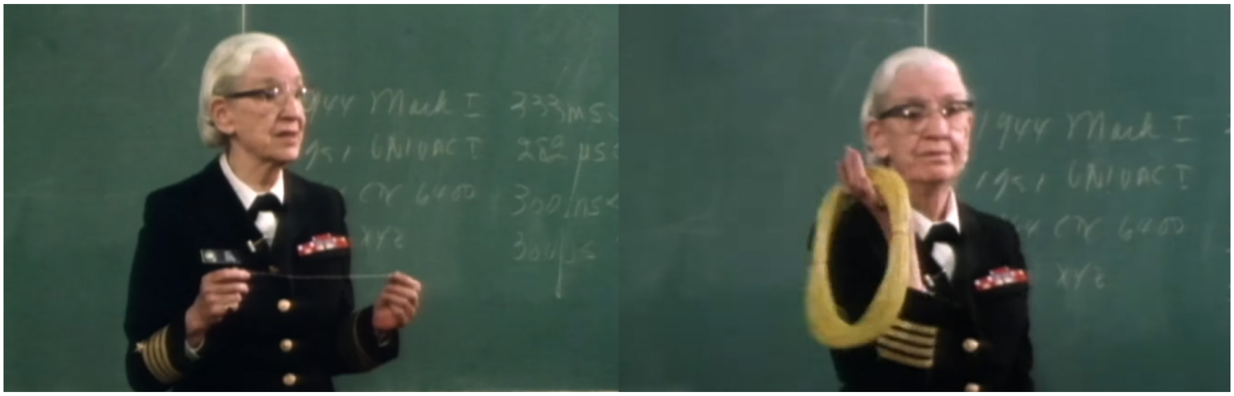

Figure 14 - Admiral Grace Hopper [39] (on the left), demonstrated the concept of a nanosecond by illustrating that it represents the maximum distance electricity can travel in one-billionth of a second (One nanosecond), approximately 11.8 inches long.On the right, demonstrating the length of a Microsecond (1μs = 984 feets).

2.4.0 - IEEE 1588, THE INNER-POWER OF ST-2110:
Also known as Precision Time Protocol (PTP)[40], is a protocol used to synchronize clocks throughout a computer network. In a local area network (LAN), it can achieve clock accuracy in the sub-microsecond range, making it significantly more precise than older protocols such as Network Time Protocol (NTP). Originally published in 2002, IEEE 1588 has undergone several revisions to enhance its accuracy and reliability. IEEE 1588 is being used by ST-2110 to carry the most precious piece of information on it's package, which is Delta-Time (Δ-t), allowing to keep track of the exact point in time that the event was created for later Time-Synchronized accuracy playback.

2.4.1 - ST-2110 Δ-t (DELTA TIME) OFFSETS
Being based on PTP time and delta time (Δ-t) precise offsets, is important to understand who PTP is the best precission tool and the reason why ST-2110 utilizes this standard for calculating offsets.

2.4.2 - NANOSECOND(ns) VS MICROSECONDS
To illustrate the difference between a PTP card with Nanosecond (ns) precission vs a PTP card with Microsecond(μs) precission, lst's use Admiral's Grace Hopper visual representation of both; whereas a Nanosecond(ns) traveling at the speed of light in a vacuum represents approcimateley 11.8 inches while a Microsecond (μs) can be represented with 984 feets:

- 1 microsecond (μs) = 1,000 nanoseconds (ns)
- 1 millisecond (ms) = 1,000 microseconds (μs) = 1,000,000 nanoseconds (ns)
- 1 second (s) = 1,000 milliseconds (ms) = 1,000,000 microseconds (μs) = 1,000,000,000 nanoseconds (ns)

2.4.3 - PTP vs MASTER CLOCK PULSE
Traditional Master Clocks can only provide up to 60Hz or **16.67 milliseconds (ms)** per pulses. Although masterclocks can be syncronized to PTP clocks, they don't provide pulses beyond their max specs. They are able to offset time, but not refference subframes.
To illustrate the difference between a PTP card vs a Masterclock, let's do the math:

$$\text{Time per pixel} = \frac{\text{Total time for one frame (in nanoseconds)}}{\text{Total number of pixels}}$$

For a 4K video at 60 fps, each pixel out of the total 8,294,400 pixels is allocated approximately **2.01 nanoseconds** within the time frame of one frame's duration, which is 16,666,667 nanoseconds. This is a theoretical calculation to understand the distribution of time across pixels in a single frame at this resolution and frame rate. And while Master Clocks can offset at nanoseconds precission, and provide Nanosecond PTP syncronization, they will not provide Nanosecond pulses that PTP cards can offer.
To summarize, that is 16,666,667 (ns) vs 2.01 (ns).
During the course of this paper, we'll learn how critical those nanoseconds are for sub-frame precission, that Master Clocks can achieve, but still limited at the Su-Frame level.

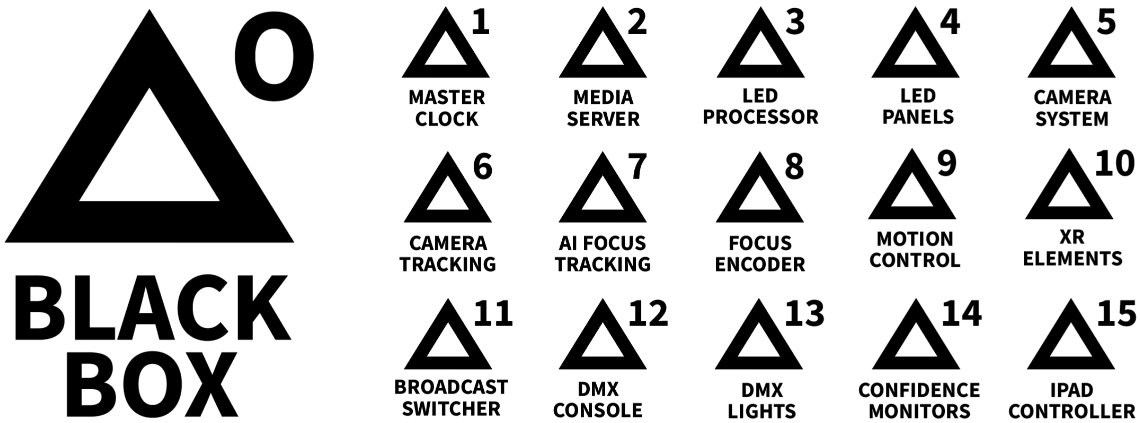

Figure 15 - Delta-offsets for each virtual production element through ST-2110's Nanosecond(ns) precission via PTP time-stamp

3 - METHODS

This section presents a series of recommendations for achieving high frame rates in shootings. It includes Immersive Stereo-3D-Virtual Production and considers techniques designed to improve upon previous virtual production workflows. The adoption of SMPTE-2110 aims to unify and simplify the key elements of this multifaceted combination of tools. Our approach begins with identifying the challenges, then developing new workflows and establishing a simplified and best operating practices for:

3.1.0 - Δ-t (DELTA-TIME) SYNCHRONIZATION:

Delta Synchronization plays a pivotal role in the integration of various components within a virtual production workflow. At its core, delta synchronization involves managing the time delays known as Delta-Offsets between these components. This is especially critical in high-speed FPS shootings, where immersive virtual productions often operate at speeds exceeding 120 frames per second (fps) per eye, culminating in a total of 240fps for stereoscopic viewing. Understanding and accurately calculating these offsets are indispensable for ensuring the synchronization of visual elements, audio, and interactive components, thereby providing a cohesive production output.

3.1.1 - Understanding Delta-Offsets:

Let's define a Delta-Offset as the time difference that each piece of equipment or system introduces into the production workflow. These offsets can significantly affect the final output if not properly managed, particularly in scenarios demanding high frame rates for creating immersive experiences. The calibration of these offsets is essential for maintaining the integrity and synchronicity of the virtual production.

3.1.2 - Illustrating Delta-Offset:
Although a comprehensive discussion of the calibration processes for each Delta-Offset is beyond the scope of this paper, examining the calibration process for a typical component, such as a camera, can shed light on the complexity and importance of this procedure. For instance, calibrating the camera's Delta-Offset involves adjusting its timing to match the virtual environment's rendering speed, ensuring that live-action footage aligns perfectly with virtual elements.

3.1.3 - THE SCOPE OF DELTA SYNCHRONIZATION:
It's important to acknowledge that every piece of equipment within the virtual production workflow, ranging from lighting systems (DMX), LED walls, motion tracking systems, lenses and lens encoders, to network infrastructure components like cloud services and network switches, plays a significant role in the overall Delta Synchronization process. Each component introduces its own Delta-Offset (e.g., CAMERA-Δt, LED-WALL-Δt, MOTION-CONTROL-Δt, MONITORS-Δt, etc.) that must be calibrated to ensure seamless integration.

3.1.4 - THE IMPORTANCE OF HIGH-SPEED SHOOTING:
High-speed shooting poses unique challenges for Delta Synchronization. As frame rates increase, the margin for error in synchronization decreases, necessitating precise calibration of Delta-Offsets to avoid visual sync aberrations with the rest of the system, specially LED walls, which can detract from the immersive experience.

3.1.5 - EXAMPLES OF NECESSARY DELTA CALCULATIONS:
Delta Synchronization is a complex, yet crucial aspect of virtual production, requiring meticulous attention to the calibration of time delays across all components. Here is a list of necessary Delta Offsets

3.1.6 - HOW SMPTE-2110 CAN HELP DELTA SYNCRONIZATION:
Automatic self-calibration systems can be achieved simply communicating their Delta Times and allowing self calibration and self-offset calculations through a centralized "ST-2110-Black Box" **(See "3.8.2 - AI ASSISTED API INTEGRATION" on page 30)**

3.2.0 - SYSTEM CALIBRATION

Calibrating complex equipment systems manually is a daunting task due to the need for constant recalibration. This section outlines the various equipment requiring frequent calibration, detailing the role of SMPTE ST-2110 in streamlining and automating these processes to enhance efficiency and accuracy.

3.2.1 - Master Clock:
- Full Setup and ST-2110 Synchronization
- LTC + SMPTE Synchronization
- Camera Synchronization
- LED Processor Synchronization
- Media Servers Synchronization
- Tracking System Synchronization
- Motion Control System Synchronization
- Sound System Synchronization
- DMX System Synchronization
- AUXILIAR Systems Synchronization

3.2.1 - Camera:
- Master Clock Synchronization
- ST-2110 Setup
- LTC + SMPTE Synchronization
- Sensor Color Calibration (Requires constant attention, even during the day)
- Color Profiles (Varies Project by Project)
- OpenColorIO (Varies Project by Project)
- Shutter Angle Refresh Verifications (Every time you use a different LED Setup)
- Firmware (Affecting The Full System Calibration)
- Δt (Delta-Time) calculation.

3.2.3 - Lenses:
- Full Distortion Calibration Profile Needed.
- Color Profile.
- Zoom Lenses: Requires re-calibration.
- Focus Assistant: Requires re-calibration.
- Zoom Assistant: Requires re-calibration.
- Manual Focus Follower: Requires re-calibration.
- AI Focus Follower: Requires re-calibration.
- Robotic Focus Follower: Lens and Robot System re-calibration.
- Δt (Delta-Time) calculation for each of these systems.

3.2.4 - LED Panels:
- FULL Color Profile Recalibration every two months (Recommended).
- Constant LED Tiles replacement and Tracking those new replacements for calibration.
- Tracking Data corruption and signal integrity.
- OpenColorIO profiles.
- Δt (Delta-Time) calculation.

3.2.5 - LED Processors:
- Master Clock Synchronization
- ST-2110 Setup
- Color Profiles Synchronization
- LED Tiles Color Management

3.2.6 - LASER Projectors:
- Master Clock Synchronization
- ST-2110 Setup
- Color Profiles Synchronization
- Beam Hot Spots Mapping

3.2.7 - DMX Lighting:
- Master Clock Synchronization
- LTC + SMPTE Synchronization
- ST-2110 Setup
- Color Calibration of each Unit (Serialized).
- Console Color Calibration.
- Δt (Delta-Time) calculation.

3.2.8 - Media Servers:
- Master Clock Synchronization
- LTC + SMPTE Synchronization
- ST-2110 Setup
- Color Profiles Management
- OpenColorIO Profiles consistency with projects.
- Δt (Delta-Time) calculation.

3.2.9 - Live Production Switchers:
- Master Clock Synchronization
- LTC + SMPTE Synchronization
- ST-2110 Setup
- Color Profiles Management
- OpenColorIO Profiles consistency with projects.
- Δt (Delta-Time) calculation.

3.2.10 - On-set Visualization and Previsualization Systems:
- Master Clock Synchronization
- LTC + SMPTE Synchronization
- ST-2110 Setup
- Color Profiles Management
- OpenColorIO Profiles consistency with projects.
- Δt (Delta-Time) calculation.

3.2.11 - Tracking Systems:
- Master Clock Synchronization
- ST-2110 Setup
- Δt (Delta-Time) calculation.
- Change of Lense Re-setup.
- World Positioning Reset.
- Camera Position Setup,
- Reflective Markers Setups.
- Outside-In Systems Setups
- Inside-Out Systems Setups
- Tracking System + Lens calibration
- Tracking System + Unreal Engine System
- Tracking System + Disguise / Axymmetry / Pixera / LiveFX /+++

3.2.12 - Motion Control (Including Robots / Wired-Suspended / Gimbal / Crane) Systems:
- Master Clock Synchronization
- ST-2110 Setup
- Δt (Delta-Time) calculation.
- New Setup: Full Robot+Camera+Lens+Lens Control Calibration.
- New Camera: Full Robot+Camera+Lens+Lens Control Re-Calibration.
- New Lens: Full Robot+Camera+Lens+Lens Control Re-Calibration.

3.2.13 - Ultra-Wideband (UWB) Tracking Systems:
- Master Clock Synchronization
- ST-2110 Setup
- Δt (Delta-Time) calculation.
- Rovers Calibration
- Receivers Calibration
- Inertial Systems Calibration
- GPS Calibration

3.2.14 - Motion Tracking Suits (Inertia / Markers / AI):
- Master Clock Synchronization
- ST-2110 Setup
- Δt (Delta-Time) calculation.
- Global Positioning
- Inertia / Markers / AI Calibration
- Camera Calibration

3.2.15 - Steady Cams Systems:
- Master Clock Synchronization
- ST-2110 Setup
- Δt (Delta-Time) calculation.
- Camera+Lens+Lens Control Calibration.
- Camera+Lens+Lens Control Re-Calibration.
- Camera+Lens+Lens Control Re-Calibration.

3.2.16 - Real-time Character Animation Tools:
- Master Clock Synchronization
- ST-2110 Setup
- Δt (Delta-Time) calculation.

3.2.17 - Volumetric Capture Systems:
- Full System Calibration
- Master Clock Synchronization
- Camera Calibration
- ST-2110 Setup
- Δt (Delta-Time) calculation.

3.2.18 - VP B-Bar and Color Pipeline:
- ST-2110 Setup
- Full Color Profiles Recalibration
- Monitors Calibration
- Δt (Delta-Time) calculation.

3.2.19 - Audio System:
- Master Clock Synchronization
- LTC + SMPTE Synchronization
- ST-2110 Setup
- Wireless Setup
- Frequencies Setup
- Δt (Delta-Time) calculation.

3.2.20- Cloud Computing and Rendering Services:
- Network Configuration
- ST-2110 Setup
- Δt (Delta-Time) calculation.
- Cloud Computing Setup
- Storage Setup

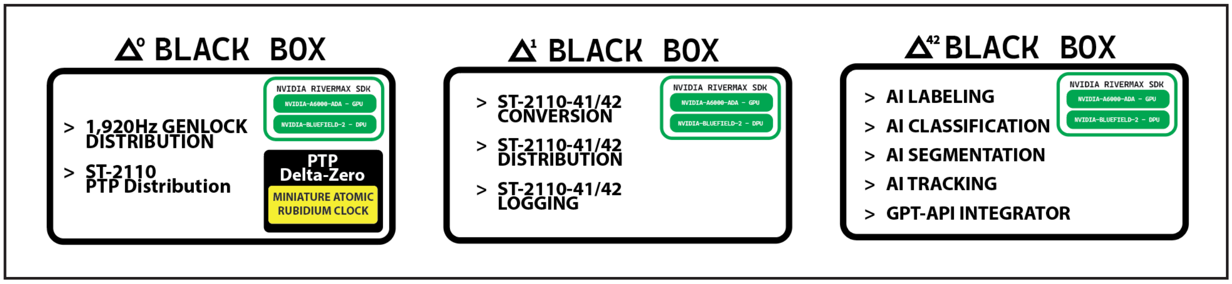

Figure 16 - Visualizing an Immersive ST-2110-41/42 Setup Block Diagram fully integrated system.

3.2.21 - ADDRESSING THE CALIBRATION CONUNDRUM WITH SMPTE ST-2110:

The complexity of manual calibration across numerous systems is a significant challenge. SMPTE ST-2110-41/42 can revolutionizes this process by centralizing settings and calibration metadata. This standard can enable automated calibration through an API-integrated GPT[41] system, significantly reducing manual calibration and enhancing system reliability and efficiency.

3.2.22 - HOW SMPTE ST-2110 ENHANCES CALIBRATION PROCESSES:

SMPTE ST-2110-41/42 will play a pivotal role by providing a framework for automated calibration. It facilitates:
- Centralized management of calibration data,
- Automated recalibration processes, reducing manual errors,
- Enhanced system interoperability and synchronization.
- This standard, coupled with advanced GPT API integration, allows for real-time, dynamic calibration adjustments, ensuring that all systems remain precisely calibrated without constant manual oversight.
[See”3.8 - API GPT SYSTEM” on page 30].

3.3.23 - RUBIDIUM vs OCXO:

The time accuracy between a Rubidium[42] atomic based clock vs an OCXO[43] (Oven Controller Crystal Oscillator) can be crucial at high frame rates.

Oscillator	200 ns	400 ns	1.1 µs	1.5 µs	5 µs	10 µs
OCXO	4 hrs	8 hrs	13 hrs	15 hrs	1.2 days	1.7 days
Rubidium	15 hrs	1 day	2 days	2.6 days	8 days	12 days

Figure 17 - This table lists typical (1 sigma confidence) values. Assume a benign temperature environment.

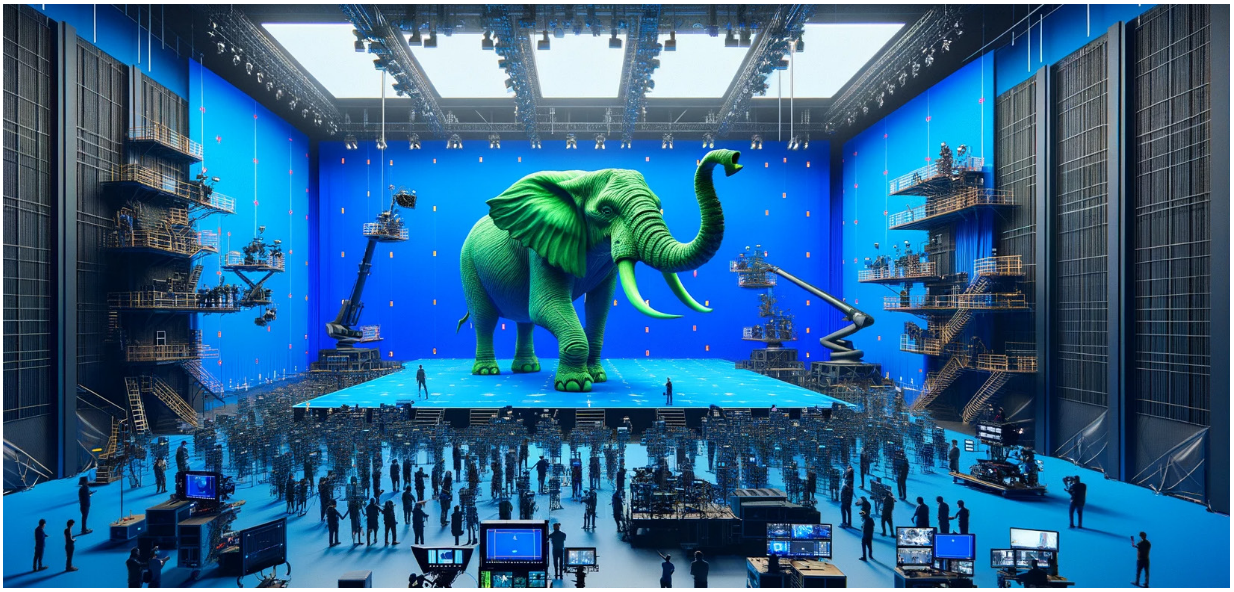

Figure 19 - A Generative-AI image of a "Green Elephant in the room".

3.3 - ALTERNATIVE WORKFLOWS

Virtual production has guides Post-Production challenges into the Pre-production stage, fundamentally altering both the creative and production landscapes of filmmaking. While virtual production promises cost savings by reimagining Pre-production constraints, this notion remains theoretical. Often, the creative process inherently progresses linearly. This paper does not aim to critique meticulously planned production schedules or budgets. Instead, it highlights the undue technical burdens placed on virtual production teams, potentially overshadowing the true value virtual production could offer. By shifting the focus towards the previsualization capabilities of virtual production, rather than the elusive "Final Pixel" ideal, we uncover the tool's optimal utility. Advocating for the positioning of virtual production as both a post-production aid and a tool for real-time, off-camera visualization, we propose an approach that not only expedites the production process but enhances post-production efficiency, focusing on compositing rather than rectifying in post.

3.3.1 - "LED SCREEN FOR FINAL PIXEL" WORKFLOW:

When it comes to solve Immersive Stereo-180-3D at 240 FPS, the technology is not ready for "Final Pixel" real-time renders.

3.3.2 - THE BIG GREEN ELEPHANT IN THE ROOM:

Which is not "Gen-AI", (See "6.3.0 - ADDRESSING THE COLOSSAL DINOSAUR-LIKE MONSTER IN THE ROOM." on page 38), but traditional green/blue screens, they offer a cost-effective solution for visual effects (VFX) but lack the immediacy and interactivity of real-time previsualization. On the other hand, LED screens, while providing real-time feedback and immersive environments for actors, come with a high financial burden. This section explores how we can bridge this gap, leveraging technology to provide cost-effective, real-time previz solutions.

3.3.3 - "DEFERRED RENDER" WORKFLOW:

The big problems with LED walls are the following:
Color Mismatch (Fixing-in-post with rotoscoping and color correction). LED Wall Dropped Frames (Fixing-in-post with rotoscoping, re-tracking and re-renders)
Bad Resolution Background (Fixing-in-post with rotoscoping, re-tracking and re-renders)
Full Background Replacement (Fixing-in-post with rotoscoping, re-tracking and re-renders).
These are just few scenarios where we have to go back into post-production fixing, instead of just color correction work. Embedding all the metadata, including camera tracking into the footage with SMPTE-2110, will allow a more streamlined post-production workflow since the metadata (including tracking metadata) is going to live with the footage during the full post-production pipeline life.

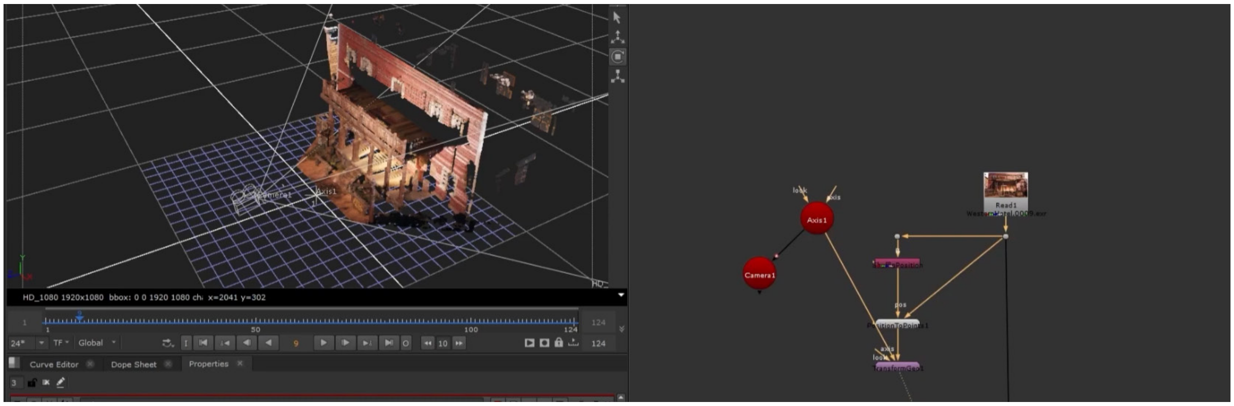

Figure 20 - Nuke's UnrealReader is a powerful tool that allows VFX artists to import real-time 3D assets and environments from Unreal Engine into Nuke's compositing pipeline

Instead of using a LED wall for Final Pixel, we can use it as a previsualization tool and use sub-frames to shoot the actual green screen instead for rapid compositing.
The motion tracked background can be rendered at full resolution without hiccups and be ready for compositing with perfect frame tracking matching and no frame offsets. Deferred render doesn't means that we are pushing the full production to Post-Production again, but allowing the render to be perfect at high resolution avoiding dropped frames and digital artifacts related to real-time render. It also gives the chance to make another render for final pixel through different render workflows such as Unreal Engine to Nuke final render. This will be the pfreffered workflow along this paper, since can generate both best features from both worlds (Real-Time Previsualization and perfect Final Pixel).

3.3.4 - "LED WALL WITH DEFERRED RENDER WORKFLOW":

Integrating Advanced Visualization Techniques: This workflow innovatively employs LED screens, traditionally used as dynamic backdrops, in a dual role that also includes blue or green screen functionalities with embedded markers for precise camera tracking interpolation. This dual functionality facilitates seamless post-production compositing by allowing the actual 3D environment to be visualized in the background during non-captured frames, thus supporting both previsualization and the capture of all necessary tracking information for a deferred rendering process. This enhanced approach does not necessitate specialized LED processors, making it adaptable to both contemporary and legacy system configurations, thereby broadening its applicability across various production scales.
The method is equally effective for traditional 2D narratives and convers immersive 3D experiences too, making it a versatile tool for a Immersive filmmaker's arsenal.

3.3.5 - DEFERRED RENDERING ENHANCING PRODUCTION QUALITY:

Deferred rendering is a cornerstone of our approach, enabling detailed real-time visualizations that do not compromise the ongoing production process. This method facilitates the generation of high-detail renders on the same day of production, surpassing the quality of conventional real-time processes without extending post-production schedules. Enhancing this approach with SMPTE-2110 Metadata recording, will secure a strong workflow stream without compromising shots.

3.3.6 - LOCAL OR CLOUD-BASED RENDERING SOLUTIONS CLOUD-BASED RENDERING:

Leveraging cloud services like AWS in conjunction with cutting-edge rendering software (e.g., Mo-Sys[44] and Disguise[45]) and camera systems (such as RED cameras[46]), we can achieve near-real-time renders without sacrificing quality. Although not instantaneous, this approach ensures a continuous workflow devoid of dropped frames, particularly beneficial for immersive stereo 3D content. HD content benefits from swift rendering times, enabling daily pre-editing reviews, while final shots can be rendered at even higher resolutions and better post-process render qualities.This progress not only alleviates the pressures of real-time content generation but also enhances the quality of previsualization and final renders, thereby benefiting the entire production team.

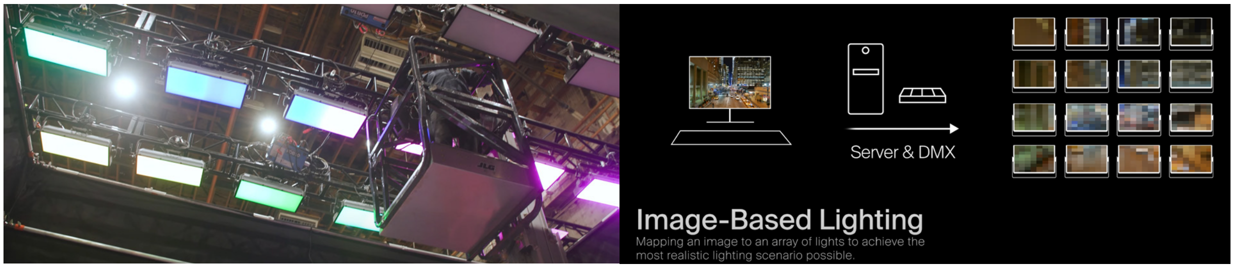

Figure 21 - Demostration at Cinegear 2023 for Aputure showcased a full array of DMX RGBWW lights for dynamic ambient illumination.

3.3.7 - REFLECTIONS AND COLOR AMBIENT COMPENSATION:
To address the need for accurate reflections in scenes, feature that LED panels excel at, a series of solutions are part of the arsenal of a Virtual Production Supervisor:

3.3.7.1 - AMBIENT LIGHT RE-INFORCEMENT:
Gaffers have now the option to control dynamic lighting situations grabbing accurate light sources from the syntetic scene through a pixel mapping array from Unreal Engine or other real time sources providing brighter lighintg sources. Creating a controlled enviroment gives gaffer the flexibility to have full control of the scene.

3.3.7.2 - GREEN/BLUE IN-CAMERA FRUSTUM:
In-between these two worlds allows deferred renders without losing secondary ambient ilumination through the led wall light source. It enables the Inner Frustum to create a better compositing scenario while leaving the rest of the screen as a more accurate ambient source.

3.3.7.3 - GREEN/BLUE IN-CAMERA SEGMENTATION:
With the help of fast AI segmentation, we can selectively isolate areas with a dynamic mask, allowing more of the LED screen to create light sources for reflections and incidental illumination.

3.3.7.4 - REFLECTION SPOTS:
A set secondary LED panels can be strategically placed for car scenes where reflections of cars or similar scenes are the main subject. This workflow, regardless of the use of LEDs for final pixel, is always one of the best approaches for these scenarios and is always in the toolbox of virtual production teams. These panels, capable of displaying low-resolution digital environments in real-time, add depth and realism through simulated reflections. For immersive content, which demands higher resolution, additional post-production may be required to achieve the desired fidelity.

3.3.7.5 - SHAPING THE FUTURE OF FILM PRODUCTION:
The deferred rendering workflow represents a significant leap forward in film production technology. By offering a versatile, efficient, and cost-effective method for integrating advanced visual effects and real-time 3D environments into filmmaking, this workflow is set to revolutionize the industry. Through the proposed SMPTE-2110 metadata workflow, we will be able to achieve full dynamic integration.

Figure 22 - Illustrated on this R&D test done by Uday Mathur (RED CTO) + David Light (NEP Sweetwater) and Ramiro Montes De Oca (Immersive Dimension), on August 16th 2023 at NEP Sweetwater in Panorama City, California;, with a production prototype of the new RED V-Raptor-XL, MegaPixel Helios, ROE LED panels (Driver@115,360hz, 1/8 Scan Line) and AGS' Ghost Frame feature; we are demonstrating Frame Remapping requested as a feature to RED cameras, what later RED Cameras is calling it "Phantom Frame" feature. We did this with Megapixel Helios running at 240 FPS, the RED V-RAPTOR-XL running at 120 FPS with a shutter angle of 180 degrees, to hide half of the 240 FPS from the LED wall. It was necessary to have a precise Delta Offset on the full workflow. Note that delay between the LED screen and the monitor is 33.36 milliseconds delay (1/240 = 0.00417) x 8 = 0.03336 Seconds = 33.36 ms. This is an example of Delta From LED Wall to Camera's Monitor, not from LED wall to Camera.A smaller Delta-Offset can be achieved if used the new SMPTE-2110 fiber output module from RED Cameras to a monitor or switcher for live XR applications.

3.4.0 - FRAME REMAPPING:

To comprehend the concept of Intra-Frames, or sub-frames, it's essential to first understand the fundamentals of high-refresh rate LED panels. These panels offer significant benefits for visual displays, thanks to their capability to refresh images at a high rate. Each panel consists of two main specifications that reveal techniques for manipulating sub-frames. The two main specifications to keep track are: #1 Refresh rate (15,360hz, 7,680hz, 3,840hz, 1,920hz), the higher the better; and #2 Scan Lines (1/8 , 1/16, 1/32, 1/64), the lower the better.NOTE: Manufacturers tend to hide the Scan Lines characteristics (1/8 , 1/16, 1/32, 1/64), which compensates negatively to the performance of final Shutter Angle options. Knowing those two Specs are the most important specs for filming purposes.

3.4.1 - THEORETICAL SHUTTER ANGLES CHARACTERISTICS:

The mechanism that governs exposure time and thus affects image capture significantly is known as the shutter. Two primary types of shutters are prevalent in modern cameras: Rolling Shutters[47] and Global Shutters[48] . Rolling Shutters expose the sensor progressively, from the top of the sensor to the bottom of the sensor (this varies from manufacturer to manufacturer), potentially leading to distortions in fast-moving scenes. Conversely, Global Shutters expose (in theory) the entire sensor simultaneously, capturing the scene without the skewing effect but possibly introducing a negligible delay. But it's crucial to acknowledge that not all Global Shutter cameras perform equally. Factors such as the camera's internal architecture and the firmware version can influence the shutter's responsiveness. Unexpectedly, even cameras equipped with faster Global Shutters may exhibit slight delays, which can affect the capture of high-speed events or interactions with fast-flashing lights, such as LED refresh rates.

3.4.2 - Variability in Global Shutter Performance:

To ascertain the precise shutter angle characteristics of your camera, we recommend conducting a comprehensive testing regimen. This involves:

Testing Against High-Speed LED Refresh Settings: Utilize LED panels with known refresh rates to simulate the conditions your camera will face.

Recording and Analyzing Results: Capture sequences at the maximum theoretical speed to fine tune your camera's sensor to your LED wall.

Comparing Firmware Versions: If possible, test your camera with different firmware versions to observe any changes in shutter responsiveness.

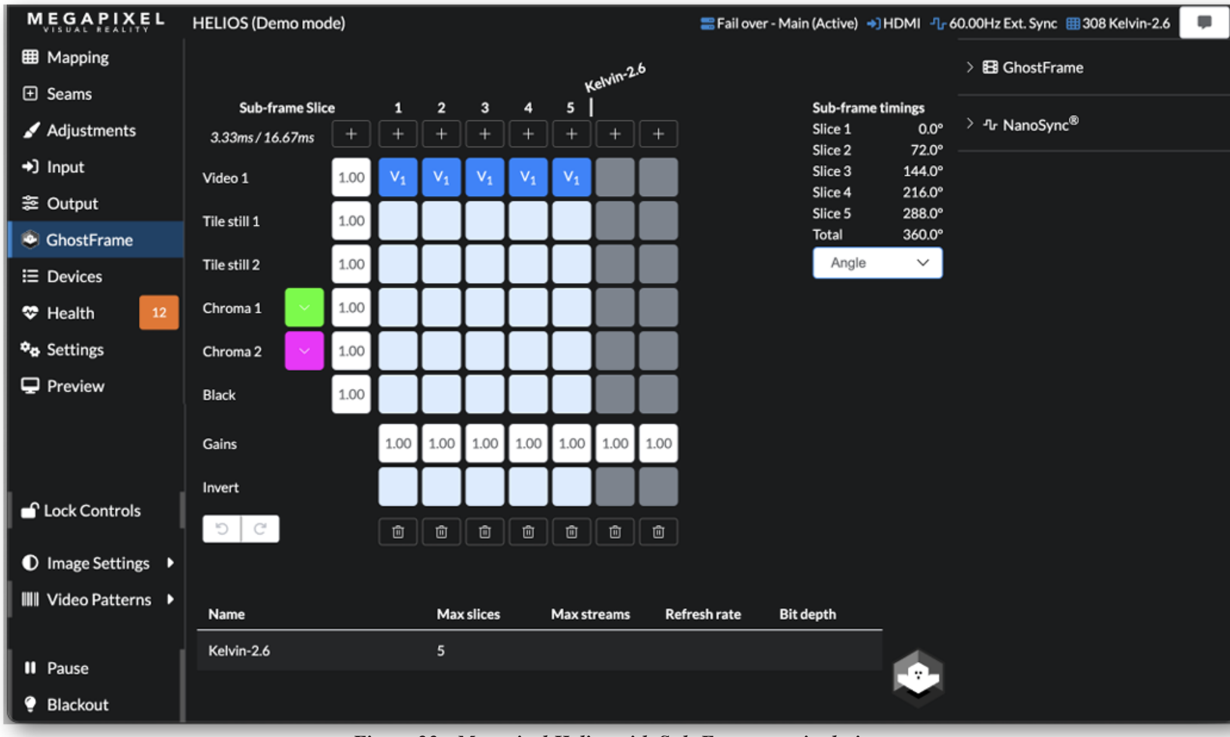

Figure 23 - Megapixel Helios with Sub-Frame manipulation.

3.4.3 - SUB-FRAMES WORKFLOWS

The use of Sub-Frames with LED Processors and Media Servers can improve your workflows options expanding the palette of possibilities on virtual production workflows.

3.4.4 - SHUTTER-ANGLE COMPENSATION:

Because we have to compensate for Slower Shutter angles that (even in the best scenario), can't reach their theoretical maximum potential, we usually end up without one of the provided sub-frames, to compensate for the camera's shutter angle delay. Different LED Processors handles this with different names, but in practicality we remove Sub-Frames until we can tach up with the camera sensor characteristic, leaving one or more sub-frames in black.

3.4.5 - GhostFrame™

GhostFrame™ [49] is a Sub-Frame manipulation patent that allows the inversion of frames or Sub-Frames to hide to the eye flickering and secondary images to the eye.

This approach proves to be an excellent solution to avoid eye fatigue. This also allows a wide variety of workflows for Immersive Virtual Production that other LED manufactuerers (other than MegaPixel) are not able to do it due the Patent restriction. Note that we are going to explore different workflows, but is important to understand that constrain for other LED Processors is this patent, not the ability to generate different Sub-Frames. As example, is that Shutter Angle compensation solutions offered by Brompton ™ and Novastar ™ are Sub-Frames solutions. GhostFrame™ ability to program Sub-Frames at will is what makes GhostFrame ™ a key tool for this workflow.

3.4.6 - GhostFrame ™ LED WALL TEST SETUP

- For this workflow setup we are using the following equipment:
- • Master Clock: Evertz 5475601-MSC
- • Media Server: x4 Pixera Nodes + x1 Pixera Director with NVIDIA
- A6000's (each)
- • LED Processor: MegaPixel Helios using x2 Display Ports inputs.
- • Switcher: MegaPixel MSM4214X NetworkSwitch
- • Led Panels: ROE - BLACK PEARL P.2.6
- • Receiving Card: MegaPixel PX1
- • Led Tiles: ROE LED Tiles: Refresh Rate 15,360Hz / Scan Lines: 1/8

3.4.7 - Scenario 1: (A-B Frames | 7,680 hz | 1/8 | @ 120 FPS | 8 SUB-FRAMES)
Stereoscopic scene with A-B frames (corresponding to Left and Right eyes) with 7,680hz 1/8 LED Tiles

SUB-FRAME 1	SUB-FRAME 2	SUB-FRAME 3	SUB-FRAME 4	SUB-FRAME 5	SUB-FRAME 6	SUB-FRAME 7	SUB-FRAME 8
LEFT EYE CGI FOR CAM-A	LEFT EYE CGI FOR CAM-A	LEFT EYE CGI FOR CAM-A	COMPENSATION BLACK OUT CAM-A	RIGHT EYE CGI FOR CAM-B	RIGHT EYE CGI FOR CAM-B	RIGHT EYE CGI FOR CAM-B	COMPENSATION BLACK OUT CAM-B

LED Tile @ 7,680 hz | 1/8 | @ 120 FPS = 8 SUB-FRAMES | 45° Sh∠ x Sub-Frame
- INPUT A: REAL-TIME RENDER (LEFT EYE)
- INPUT B: REAL-TIME RENDER (RIGHT EYE)
- In this example, the LED Processor is set at 120 FPS with 8 Sub-Frames per Frame
- Input A of the LED Processor has a 120 FPS feed for the Left Eye
- Input B of the LED Processor has a 120 FPS feed for the Right Eye
- Camera-A and Camera-B are set at 120 FPS
- Camera-A and Camera-B have a Shutter Angle of 180°
- Camera-A has Zero Delta Offset + Camera-A calculated Delta Offset.
- Camera-B has +4.1666666666 ms Offset (Half Frame) + Camera 2 Delta Offset.
- Camera offset can be done on a Master Clock or camera. (We preferred all the Offsets to be done on the Master Clock for consistency).

NOTES ABOUT THIS WORKFLOW:
"Blackout Frames" Compensation positioning depends from camera to camera characteristics.
Theoretical Shutter Angle Options: (45°, 90°,135° and 180°).
Depending the Shutter Angle Black out frame compensation, there will be a max of 3 grading options.
This workflow gives them a max of 3/8 = 37.5% Light Exposure.

3.4.8 - Scenario 2 - (A-B Frames | 15,360 hz | 1/8 | @ 120 FPS | 16 SUB-FRAMES):
Stereoscopic scene with A-B frames (corresponding to Left and Right eyes) with 15,360hz 1/8 LED Tiles

SUB-FRAME 1	SUB-FRAME 2	SUB-FRAME 3	SUB-FRAME 4	SUB-FRAME 5	SUB-FRAME 6	SUB-FRAME 7	SUB-FRAME 8	SUB-FRAME 9	SUB-FRAME 10	SUB-FRAME 11	SUB-FRAME 12	SUB-FRAME 13	SUB-FRAME 14	SUB-FRAME 15	SUB-FRAME 16
LEFT CAM-A	LEFT CAM-A	LEFT CAM-A	LEFT CAM-A	LEFT CAM-A	LEFT CAM-A	LEFT CAM-A	BLACK OUT CAM-A	RIGHT CAM-B	RIGHT CAM-B	RIGHT CAM-B	RIGHT CAM-B	RIGHT CAM-B	RIGHT CAM-B	RIGHT CAM-B	BLACK OUT CAM-B

LED Tile @ 15,360 hz | 1/8 | @ 120 FPS = 16 SUB-FRAMES | 22.5°x Step
- INPUT A: REAL-TIME RENDER (LEFT EYE)
- INPUT B: REAL-TIME RENDER (RIGHT EYE)
- In this example, the LED Processor is set at 120 FPS with 16 Sub-Frames per Frame
- Input A of the LED Processor has a 120 FPS feed for the Left Eye
- Input B of the LED Processor has a 120 FPS feed for the Right Eye
- Camera-A and Camera-B are set at 120 FPS
- Camera-A and Camera-B have a Shutter Angle of 180°
- Camera-A has Zero Delta Offset + Camera-A calculated Delta Offset.
- Camera-B has +4.1666666666 ms Offset (Half Frame) + Camera 2 Delta Offset.
- Camera offset can be done on a Master Clock or camera. (We preferred all the Offsets to be done on the Master Clock for consistency).

NOTES ABOUT THIS WORKFLOW:
"Blackout Frames" Compensation positioning depends from camera to camera characteristics.
Theoretical Shutter Angle Options: (22.5°, 45°, 67.5°, 90°,112°,135°, 157.5° and 180°).
Depending the Shutter Angle Black out frame compensation, there will be a max of 7 grading options.
This workflow gives them a max of 7/16 = 43.75% of Light Exposure.

3.4.9 - Scenario 3 - (A-B Frames | 7,680 hz | 1/8 | @ 120 FPS | 8 SUB-FRAMES):

Stereoscopic scene with GREEN+TRACKER MARKERS frames (Both Camers on Sync) with 7,680hz 1/8 LED Tiles

SUB-FRAME 1	SUB-FRAME 2	SUB-FRAME 3	SUB-FRAME 4	SUB-FRAME 5	SUB-FRAME 6	SUB-FRAME 7	SUB-FRAME 8
GREEN + MARKERS CAM-A+B	GREEN + MARKERS CAM-A+B	GREEN + MARKERS CAM-A+B	GREEN + MARKERS CAM-A+B	GREEN + MARKERS CAM-A+B	GREEN + MARKERS CAM-A+B	BLACK OUT CAM-A+B	2D PREVIZ FOR EYE AND REF

LED Tile @ 7,680 hz | 1/8 | @ 120 FPS = 8 SUB-FRAMES | 45° Sh ∠ x Sub-Frame

- INPUT A: GREEN/BLUE SCREEN WITH TRACKING MARKERS (LEFT+RIGHT EYE)
- INPUT B: REAL-TIME PREVIZ RENDER (LOW RES)
- In this example, the LED Processor is set at 120 FPS with 8 Sub-Frames per Frame
- Input A of the LED Processor has a 120 FPS feed for the Left Eye
- Input B of the LED Processor has a 120 FPS feed for the Right Eye
- Camera-A and Camera-B are set at 120 FPS
- Camera-A and Camera-B have a Shutter Angle of 180°
- Camera-A and Camera-B have Zero Delta Offset + Camera-A calculated Delta Offset.

NOTES ABOUT THIS WORKFLOW:

"Blackout Frames" Compensation positioning depends from camera to camera characteristics.

Theoretical Shutter Angle Options: (45°, 90°,135°, 180°, 225°, 270° and 315°).

Depending the Shutter Angle Black out frame compensation, there will be a max of 6 grading options.

This workflow gives a max of 3/4 = 75% Light Exposure.

On this workflow, depending on the chosen Shutter Angle, both (Black out frame and Previz Frames can be moved around for optimal eye viewing).

3.4.10 - Scenario 4 - (A-B Frames | 15,360 hz | 1/8 | @ 120 FPS | 16 SUB-FRAMES):

Stereoscopic scene with GREEN+TRACKER MARKERS frames (Both Camers on Sync) with 15,360hz 1/8 LED Tiles

SUB-F RAME 1	SUB-F RAME 2	SUB-F RAME 3	SUB-F RAME 4	SUB-F RAME 5	SUB-F RAME 6	SUB-F RAME 7	SUB-F RAME 8	SUB-F RAME 9	SUB-F RAME 10	SUB-F RAME 11	SUB-F RAME 12	SUB-F RAME 13	SUB-F RAME 14	SUB-F RAME 15	SUB-F RAME 16
GREEN + MARKERS CAM-A+B	GREEN + MARKERS CAM-A+B	GREEN + MARKERS CAM-A+B	GREEN + MARKERS CAM-A+B	GREEN + MARKERS CAM-A+B	GREEN + MARKERS CAM-A+B	GREEN + MARKERS CAM-A+B	GREEN + MARKERS CAM-A+B	GREEN + MARKERS CAM-A+B	GREEN + MARKERS CAM-A+B	GREEN + MARKERS CAM-A+B	GREEN + MARKERS CAM-A+B	GREEN + MARKERS CAM-A+B	BLACK OUT FRAME	EVE PREVIZ	EYE PREVIZ

LED Tile @ 15,360 hz | 1/8 | @ 120 FPS = 16 SUB-FRAMES | 22.5°x Step

- INPUT A: GREEN/BLUE SCREEN WITH TRACKING MARKERS (LEFT+RIGHT EYE)
- INPUT B: REAL-TIME PREVIZ RENDER (LOW RES)
- Similar to (3.8.1), we are using a 15,360hz Refresh Rate LED tile this time, increasing +9.75% LED light exposure from 37.5% to 43.75%.
- In this example, the LED Processor is set at 120 FPS with 16 Sub-Frames per Frame
- Input A of the LED Processor has a 120 FPS feed for the Left Eye
- Input B of the LED Processor has a 120 FPS feed for the Right Eye
- Camera-A and Camera-B are set at 120 FPS
- Camera-A and Camera-B have a max Shutter Angle of 180°
- Camera-A and Camera-B have Zero Delta Offset + Camera-A calculated Delta Offset.

NOTES ABOUT THIS WORKFLOW:

"Blackout Frames" Compensation positioning depends from camera to camera characteristics.

Theoretical Shutter Angle Options: (22.5°, 45°, 67.5°, 90°,112°,135°, 157.5°, 180°, 202,5°, 225°, 247.5°, 270° and 292.5°).

Depending the Shutter Angle Black out frame compensation, there will be a max of 13 grading options.

This workflow gives them a max of 7/16 = 43.75% Light Exposure.

On this workflow, depending on the chosen Shutter Angle, both (Black out frame and Previz Frames can be moved around for optimal eye viewing).

Figure 24 - Virtual Production with Laser Projectors for Previsualization (no Final Pixel on Camera).

3.4.11 - LASERS PROJECTORS

By integrating laser-projected previz frames within the intervals not captured by the camera, we offer a solution that bypasses the need for LED walls. This approach not only enhances spatial awareness for the crew and actors but also simplifies post-production by providing a 'clean canvas' for compositing, further discussed in the section on Deferred Rendering. In other words, laser projection is used for visualization only, not for Final Pixel on camera. In this way, you can project any resolution, since the only goal is to provide a visual reference of the simulated environment.

3.4.12 - PREVISUALIZATION ONLY

We propose the use of laser projectors as an innovative alternative for stage previz. By integrating laser-projected previz frames within the intervals not captured by the camera, we offer a solution that bypasses the need for expensive LED walls. This approach not only enhances spatial awareness for the crew and actors but also simplifies post-production by providing a 'clean canvas' for compositing, further discussed in the section on Deferred Rendering. In other words, laser projection is used for visualization only, not for Final Pixel on camera. In this way, you can project any resolution, since the only goal is to provide a reference.

3.4.13 - SYNCHRONIZED LED LIGHTING SOLUTIONS:

At 230.720Hz or 115,360Hz and building on the concept of synchronized lighting, this subsection introduces a high-speed LED lighting solution capable of operating at refresh rates similar to those proposed for the LED panels. Unlike existing products, these lights focus on powerful beam projection and high-speed black frame generation, offering new possibilities for scene illumination and backdrop generation in sync with laser projectors. By comparing with existing technologies such as KINOFLO MIMIK, we highlight the unique features this solution, emphasizing sub-frame synchronization and the ability to provide a wider range of shutter angle options for cinematographers.

3.4.14 - ***Note: We are aware that some of these products might exist in the lab, we are just pointing out the need of it.***

3.4.15 - PROPOSED LASER PROJECTION SETUP

For this workflow setup we are using the following equipment:

- Master Clock: Evertz 5475601-MSC
- Media Server: x2 Pixera Nodes + x1 Pixera Dirwith NVIDIA A6000/ea
- LED Processor: MegaPixel Helios using x2 Display Ports inputs.
- Switcher: MegaPixel MSM4214X NetworkSwitch
- Led Panel: Kinoflo MIMIK 120
- Receiving Card: MegaPixel PX1
- LASER PROJECTOR: Christie Griffyn Series
- Camera: RED KOMODO 6K with SDI output
- White Screen background.

3.4.16 - KINOFLO MIMIK 120 WORKFLOW

NewLED-Image-Base panels, equipped with video input capabilities, have the potential to drive an entire sound stage with these lights, creating a fully immersive environment by synchronizing with LASER projectors over a “White Screen” background. Operating at the speed of 30,720Hz with 1/3 of the scan lines, they offer an astonishing 10,240 FPS of freedom.

3.4.17 - Scenario 5 - (GREEN LIGHT +LASER Frames | 30,720Hz| 1/3 | @ 120 FPS | 85.3 SUB-FRAMES):

LASER PROJECTORS in combination of GREEN or BLUE LED lighting + LED Processor.

FRAME 1	FRAME 2	FRAME 3	FRAME 4	FRAME 5	FRAME 6	FRAME 7	FRAME 8
GREEN LIGHT + MARKERS CAM-A+B	LASER PREVIZ	GREEN LIGHT + MARKERS CAM-A+B	LASER PREVIZ	GREEN LIGHT + MARKERS CAM-A+B	LASER PREVIZ	GREEN LIGHT + MARKERS CAM-A+B	LASER PREVIZ

LED Laser Previz @ 240 FPS + GREEN LED Lights = (1 skipped frame)180° Sh ∠ x Frame

- WHITE SCREEN WITH GREEN/BLUE LED LIGHTS
- REAL-TIME PREVIZ RENDER (LOW RES)
- In this example, the LASER PROJECT is set at 240 FPS (Playing 60 FPS content at 120 FPS Media server for Real-Time PREVIZ)
- A white screen is used as backdrop with GREEN or BLUE LED lights.
- Camera-A and Camera-B are set at 120 FPS
- Camera-A and Camera-B have a Max Shutter Angle of 180°
- Camera-A and Camera-B have Zero Delta Offset + Camera-A calculated Delta Offset.

NOTES ABOUT THIS WORKFLOW:
Because the high-refresh rate of Mimik LED lights, this workflow doesn’t needs blackout frames.
The use of a “White Screen” will offer the the versatility to cover large areas with LASER projectors covering large areas, as big as a typical sound stage.
Green LED patterns can potentially over-shoot over LASER projections (extra sub-frames for safty).

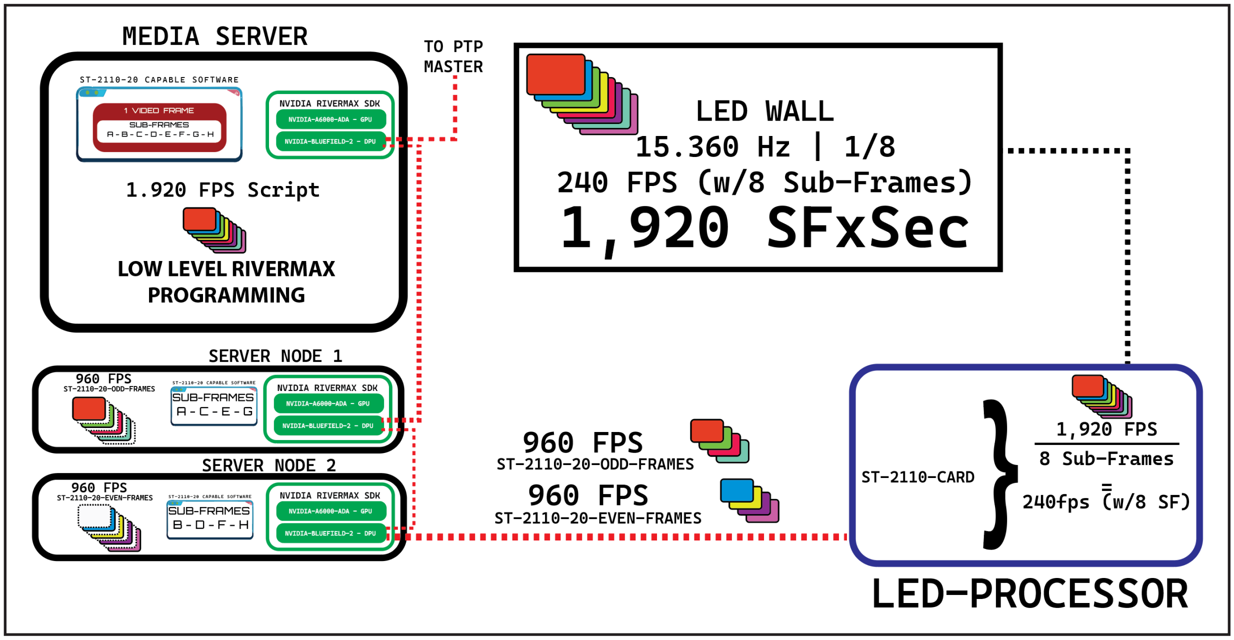

Figure 26 - Theoretical Media Server for Sub-Frame manipulation.

3.4.18 - THEORETICAL MEDIA SERVER SETUP WITH SMPTE-2110 VIDEO THROUGH NVIDIA RIVERMAX:

One of the advantages of NVIDIA Rivermax SDK is the ability to manipulate data at the lowest operating levels of SMPTE-2110 Data Streams. NVIDIA Rivermax Cards allows video vrames buffers directly through DMA (Direct Memory Access) bypassing the CPU to reduce latency and increase throughput. These cards leverage NVIDIA's GPU technology for hardware acceleration. This includes video decoding and encoding, color space conversion, scaling and other image processing tasks. Having access to DMA, allows ST-2110 software such as Disguise or Pixera to generate video output at the Sub-Frame level. The ability to create scripts on these ST-2110 softwares allows the same level of Sub-Frame manipulation done on LED-Processors, to be done at the Media Server Level.

3.4.19 -240 FPS TEST AT NAB 2023:

Almost one year ago, at NAB 2023, we setup this demo[50] for **SiliconCore**'s 1.9mm 7,680Hz 1/8 @ 240 Hz. We did it whit x2 Pixera Servers + 1 Pixera Director to an array of x8 Color-Light Processors to drive 240 FPS at 4K. The LED side of the system performed perfectly, but the two NVIDIA-A6000ADA cards started to overheat the system at the hour, so we had to dial down to 120fps. More R&D with NVIDIA shows that NVIDIA's Sync Card 200hz speed limitation, does not Sync x2 nodes, hence the need to keep it in one system.
NVIDIA Rivermax is currently the only card able to Sync 240hz on separate systems, thanks to SMPTE-2110's PTP, and the only workflow available for over 200fps setups, unless the system stays encapsulated in one server.

Figure 25 - Sucessful 240 FPS shooting test at NAB 2023 with SiliconCore's SiliconCore's 1.9mm 7,680Hz 1/8 @ 240 Hz. LED Panels, driven by Colorlight LED Processor.[51]

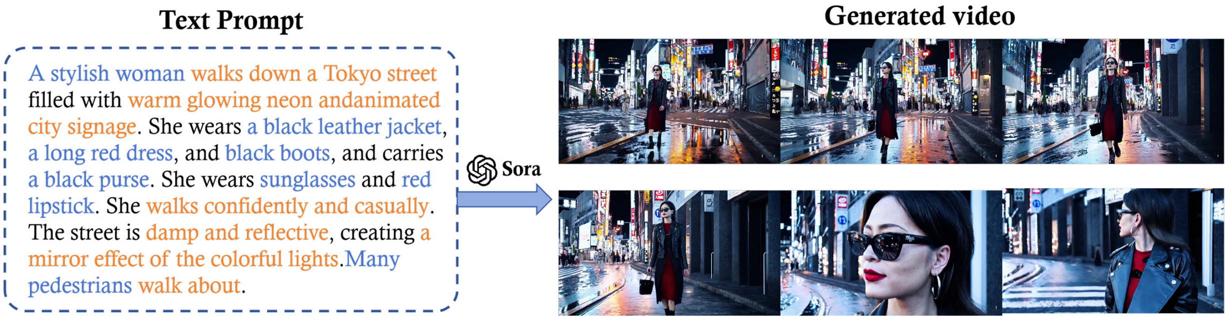

Figure 27 - The introduction of OpenAI's SORA[58] (Social Robot) raised eyebrows regarding the methods used for training. Although not disclosed by OpenAI, many speculations have been made about the use of several AI-Labeling techniques and AI-tracking metadata information (Detailed Labeling, Segmentation, and OpenPose being some of them).

3.5.0 - AI TRACKING, LABELING, INTERPOLATION AND GENERATIVE AI

The combination of multiple areas of AI converging at the same point through Multimodal Language Models (MLMs)[52], allows Machine Learning (ML)[53] training to generate more accurate models with the help of better context. It is not enough for a tracking system to just look at a tracking point and determine its position just focusing on that tracking point. New training data takes in consideration 'segmentation' (to reduce computational load), labeling of that particular tracking point that belongs to a human shoulder, allowing better calculations through constraints and possible movements. Each piece of information allowing better context allows models to perform exponentially better.

3.5.1 - LABELING = CONTEXT

The secret of Generative AI is to give context. The more context Multimodal Language Models have, the better the intended results are. When AI models are presented with more context, the results are exponentially better. This is why it is not surprising how good the new SORA model is, since it has been trained with very accurate tracking information, including Pose Estimation and segmentation labeling techniques.
Is not enough for a MLM to classify: [Human, Male, Age, Hair Color, Eye Color, Ethnicity]. New classification models include: (Motion Tracking, Posture, Activity, Emotions, Intonation, Intention, BMI, Personality, Speech Patterns, etc.). Generating metadata that provides better context, will allow AI models to train on better curated data.

3.5.2.1 - MOTION TRACKING DATA

With SMPTE-2110 embedding metadata this information can travel through the whole pipeline, allowing better post-production workflows and creating better data to be interpolated with different motion tracking sources from Post-Production software. Will be the combination of multiple tracking data metadata with the use of MLMs what will allow close to perfect tracking results.

3.5.2.2 - RAW MOTION TRACKING DATA

This is an area of R&D limited to tracking companies, but allowing also RAW metadata of different tracking systems to interact with each other allows machine learning training to learn from their particular system and generate better models than that the original manufacturer can provide thanks to localized environment training. RAW metadata needs to be processed internally by each device with their own characteristic fine tunning and through (Kalman Filters for example), but allowing a MLM to decide the correct Kalman Filters interpolating multiple motion tracking inputs (stereo-visual, kinetic, robotic encoders, aruco markers, etc); will generate a robust set of data able to generate greater confidence in tracking information. This area of computer vision is one of the most advanced areas in AI, and we know that RAW data will be crucial for feeding directly to AI rather than manually updating values without a full context.

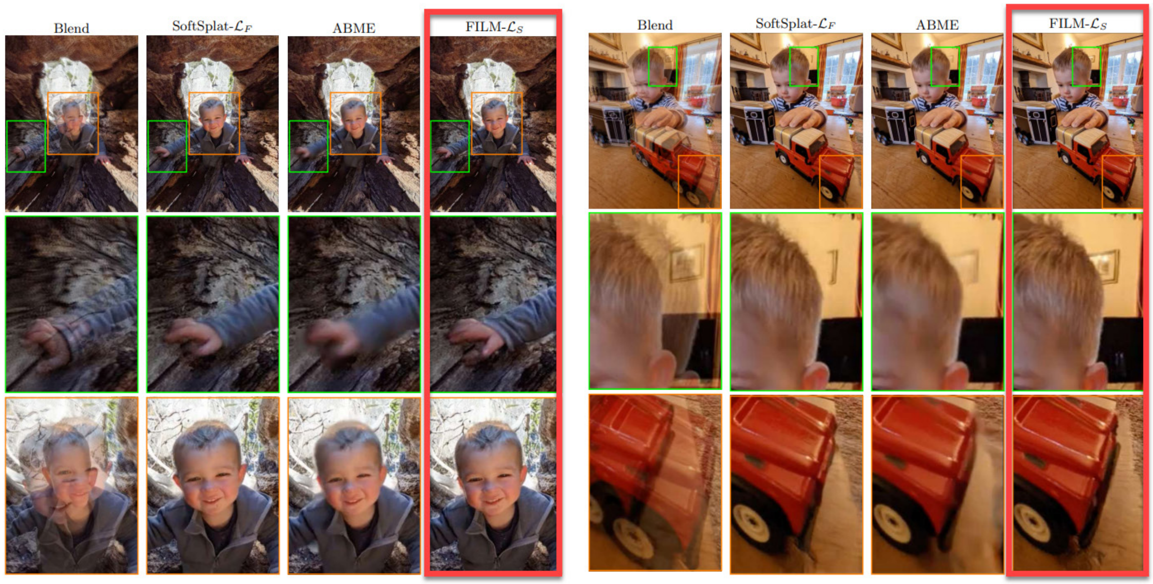

Figure 29 - Frame interpolation as demostrated in 2022 by Google Research paper "FILM: Frame Interpolation for Large Motion"

3.5.3.1 - AI INTERPOLATION:

AI interpolation refers to a set of techniques used in artificial intelligence and machine learning to estimate or predict intermediate values within a range of data points. This process is crucial in various applications, from generating high-resolution images and videos to simulating missing data in time series and enhancing the resolution of digital audio. Unlike traditional interpolation methods, which typically use mathematical formulas to estimate values based on nearby data points, AI interpolation leverages the power of neural networks and deep learning models to understand and predict the underlying patterns and relationships in the data.

3.5.3.2 - "FILM: FRAME INTERPOLATION FOR LARGE MOTION":

When Google Research released the paper "FILM: Frame Interpolation for Large Motion"[54] in 2022, a small but noticeable revolution occurred in the film industry. Companies like Topaz AI[55] began popping out with models based on this paper, and new possibilities were on the toolbox of filmmakers. Although not perfect, they lit the path to new options for slow motion and footage restoration, improving upon old mathematical interpolation approaches versus the use of AI to reconstruct missing steps between two images.

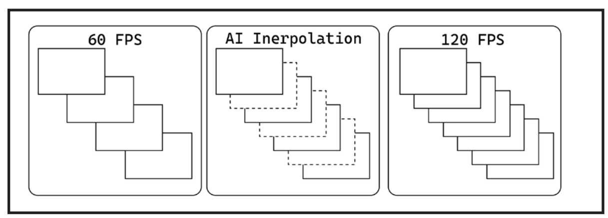

Figure 28 - Frame interpolation is bringing multiple benefits to Immersive Productions allowing retrofitting old footage to bring them to a better resolution in size and FPS.

3.5.3.3 - IMMERSIVE FRAME INTERPOLATION

In the context of Immersive Stereo-3D-180-120 Fps, this tool could become one of the most important tools in the arsenal of an Immersive film producer. Not only because solves complicated frame-remapping workflows, but also because could be a solution for alternative lower-resolution methods to be restored to better resolutions and frame rates.

3.5.3.4 - MOTION TRACKING INTERPOLATION:
As previously mentioned in ***("3.5.2.1 - MOTION TRACKING DATA" on page 25),*** the use of tracking metadata can be interpolated by AI to fill the gaps on missing information, such is the case of FPS missmatch and low cycles tracking data. This is just another point in favor to acquire RAW metadata from tracking systems and allow advanced AI models to interpolate the full spectrum of tracking metadata in conjuction with the processed metadata and other tracking systems sources.

3.5.3.2 - AI SOUND INTERPOLATION
Another area that will benefit from ST-2110-41data interpolation is the sound department. With the use of multiple microphones well tracked on the space domain, it can generate a holographic sound map as a powerful multidimensional recording device. The capabilities of a multilayered and multi-spectrum recording system can solve many sound engineering problems. This area of research has many applications and with the use of AI tools, the possibilities are endless. [Never a dull moment to recommend: "The Conversation" Francis Ford Coppola 1974[56]]

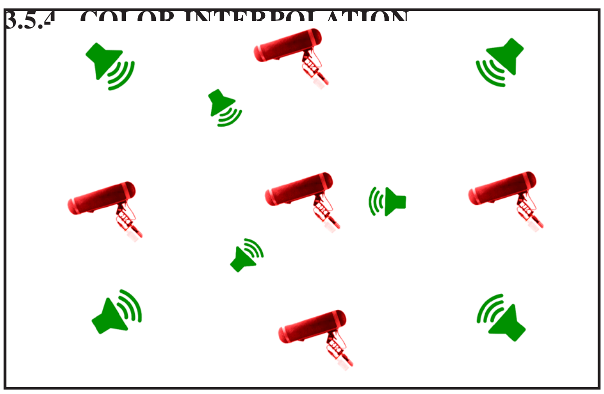

Figure 30 - Array of multiple microphones with precise location and orientation information via ST-2110-41/42

Figure 31 - "The Conversation" Francis Ford Coppola 1976, this film reveals audio survilance techniques based on frequencies filters to mask desired frequencies.

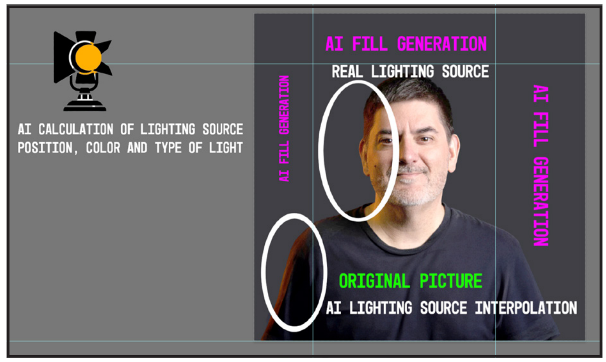

Figure 32 - Light Source AI interpolation will become a key tool in the arsenal of gaffers in film production, with the help of DMX array systems providing lighting from multiple angles.

On the area of color interpolation, AI models can be capable of recreate lighting conditions based on images. Either digital lighting on 3D scenes or practical lighting on sets. This will allow recreation of scenes from any picture as refference to be duplicated on set with arrays of DMX lighting sets connected to AI-Color models. In order to generate these models, ST-2110-41 Metadata will be a key element for future workflows.[57]

3.5.4 - COLOR API INTEGRATION
The implementation of ST-2110-41/42 will allow full color workflows to ease the full color pipeline integration with the adition of API-AI-Assisted integration.

3.5.5 - INTERPOLATION AND AI TRAINING
AI interpolation can be applied to every other area where a constant stream of data information is provided. This is also true on the color space domain, but of course any other tracking technique, including body motion tracking applications and including generative ai.
While there are more areas where AI interpolation can be applied, the one in particular of main interest to AI researchers is for training purposes. This invaluable data can train new AI models used for generative-ai, such as SORA and similar text to video models.

Figure 33 - Source: 3D model with primitives shapes or complex shapes can serve as a real-time background.

Figure 34 - Result: Generative-AI feeded with primitives through the use of ControNet can generate photo-realistic images.

3.6.0 - VIRTUAL PRODUCTION + GEN-AI

Whether it's an immersive environment or a simpler 2D format, virtual production workflows will complement each other in very dynamic situations. Here are some techniques on experimentation:

3.6.1 - PRIMITIVES TO GENERATIVE-AI

In the context of generative AI and image synthesis, the use of primitive shapes as a foundation for more complex images is a powerful concept. Here's how this approach can enhance the generation of generative AI images. Primitive shapes such as spheres, cubes, cylinders, and cones can simplify the representation of complex objects. By decomposing an object into these basic components, a generative model can learn to construct complex scenes from simpler, more fundamental elements. This simplification aids in the learning process, making it easier for the model to grasp the underlying structure of diverse objects.

3.6.2 - FROM SIMPLE OBJECTS TO INTRICATE SHAPES WITH CONSISTENCY.

In order to achieve consistency, a novel method is to generate a full 3D mesh representation of the space and allow Generative-AI to apply its respective materials. This workflow is already being implemented in commercial 3D rendering software. By reducing big problems to compartmentalized problems, we can achieve consistency with segmentation and 3D mesh generation processes.

3.6.3 - REAL-TIME AI BACKGROUND

Real-time backgrounds can be generated for reference and can be locked by choosing the corresponding "Seed" to avoid color changes. This method is still in experimentation, but the results are very promising.

3.6.4 - REAL TIME FACE REPLACEMENT

Although in its infancy, real-time face replacement is one of the most impressive features of real-time Control-Net algorithms and pre-trained models for this purpose. Although not a consistent result, with the implementation of different techniques such as segmentation, face tracking, eye gazing, and OpenPose, this area is going to improve exponentially, allowing very complex operations to be done in almost real-time.

3.6.5 - USE OF ST-2110

Generative AI will highly benefit from this workflow, allowing the flow of metadata to feed multiple processes in this area, limited only by our imagination.

3.6.6 - NOTE ON THIS SUBJECT:

This section could be 20 pages long and we want to limit the conversation around Immersive Virtual Production only, but we recognize the impact that Generative-AI is making to the entretainment industry. I will have more to say at the end of this paper on ("6.3.0 - ADDRESSING THE COLOSAL DINOSAUR-LIKE MONSTER IN THE ROOM." on page 38).

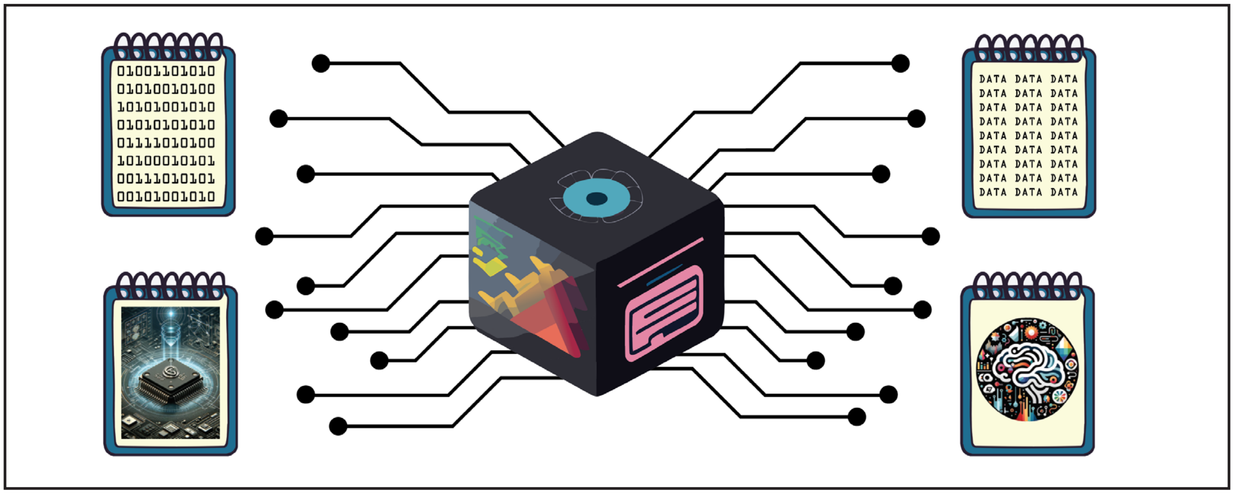

Figure 35 - A Theoretical SMPTE-2110-41/42 "BLACK BOX" Data Recorder, AI Classification, Segmentation, Tracking , Media Management and GPT-API Integration.

3.7.0 - METADATA MANAGEMENT:

Metadata management is one of the most necessary areas in film production while at the same time is one of the least used or appreciated. But when interoperability is present on the pipeline, is that metadata what makes possible to make progress on large convoluted workflows. The current state of Metadata Management is the lack of unification and standarization. The real thedious work is to keep track of atomized metadata logs, keep track of every single software updates and syncronization of multiple sources. There are few companies creating centralized metadata, but each one of those companies uses their own proprietary data format, making interoperability a very difficult task.

3.7.1 - SMPTE-RIS-OSVP METADATA

While SMPTE-2110 Metadata is not available yet, SMPTE-RIS-OSVP CAMDKIT (JSON format) metadata has been already standardized for Cameras and lenses, and is being standardized as we speak for, tracking and color management. [see CAMDKIT].
Implementation is being put in practice by companies like Mo-Sys, RED, Canon, Blackmagic Design, Sony, and all the major companies members of SMPTE-RIS-OSVP.

3.7.2 - SMPTE-RIS-OSVP Metadata Management

On the other side, companies like RED, Mo-Sys, Ncam (Zeiss) and Epic Games are already incorporating their own metadata formats to the OSVP standar, being at the forefront of the OSVP metadata movement.
But is important to explain, that although JSON metadata is not yet standardized for SMPTE-2110, and this is where we think, it will accelerate implementation and use.

3.7.3 - THEORETICAL "ST-2110 BLACK BOX"

A theoretical "ST-2110 BLACK BOX" will unify all the metadata streams in one backbone of data for better collection and real-time processing without having to access obscure menus on different equipment in the workflow. This will allow all the departments to have custom tailored solutions for their needs, instead of having a centralized system.

3.7.4 - "ST-2110 BLACK BOX" SOFTWARE DEV:

Even if we don't have a full workflow generating ST-2110-41 metadata on its own, current metadata management sofware can potentially become the big hero and convert the Metadata into this format until manufacturers joins the ST-2110-41/42 on their own. But in order to do that, this sofware has to create the best estimate of what Delta-Offset for each piece of Metadata. Although is not a small task, just converting all that Metadata into a ST-2110-41 format, could signify a big progress on this not-so-simple task.

3.7.5 - "ST-2110 BLACK BOX" TRANSMITTER:

Helping this workflow, a ST-2110 hardware module converting Metadata into ST-2110-41 can also help to accelerate the development of a ST-2110 BLACK BOX. The smaller the better.

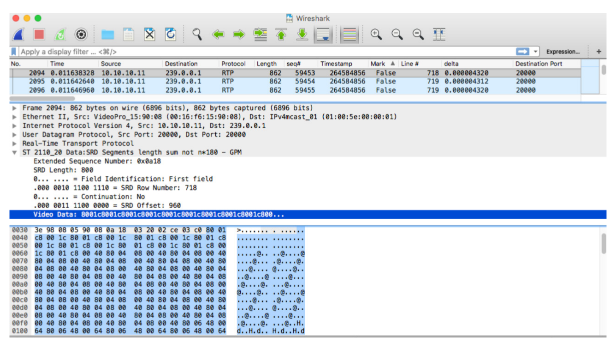

Figure 36 - For dissecting ST-2110 packages in our R&D, we use Wireshark for data analysis and even data recording.

3.8 - API GPT SYSTEM

When it comes to writing new API communication calls from different equipment in a broadcast engineering environment, nothing is more fun than digging into SMPTE Standards and sniffing what is going on and what is being communicated through those communication lines.
But with the use of GPT's (Generative Pre-Trained Transformers)[59] and feeding User Manuals and SDKs to the "Knowledge" files of a custom GPT, the use of SDK's (Software Development Kits) and their correspondent API calls are this time magnified by the power of generative ai ability to understand two systems and apply the knowledge to make them communicate without problems.

3.8.1 - API INTEGRATION

The final goal of this paper, is to create a system that will allow to train a custom designed system (custom for each user), able to communicate each piece of equipment through a comprehensive manufacturer documentation with SDK information with API calls.
The gerates the SDK information from each manufacturer, the better interoperability will be possible. API communication from manufacturer to manufacturer varies, that's why is key for manufacturers to allow users to have access to the most fundamental API calls for Setting and Remote Operations. If manufacturers are able to release SDKs for full remote operations of a system, without having to use the User-Panel interface, the list on System Clibration ***(See "3.2.0 - SYSTEM CALIBRATION" on page <?>)*** could become as easy as a click of a mouse. And although this sounds amazing and scary at the same time, this implementation could be easy as just changing settings through a simple menu.

3.8.2 - AI ASSISTED API INTEGRATION

We already generated API calls between RED Cinema Cameras to Evertz to MegaPixel Helios without major problems. Replicating this experiment to the rest of the workflow shouldn't be a big problem. GPTs are just tools and with the use of SMPTE-2110 Metadata, full integration could be as easy as flipping a switch.

Figure 37 - GPT-API integration with a Theoretical "ST-2110-41/42 BLACK BOX"

3.8.3 - NVIDIA RIVERMAX SDK

With the flexibility that RIVERMAX SDK[60] offers of ST-2110 communication and package conversion, a system with RIVERMAX SDK connected to a GPT code generator, can implement syntetic dynamic solutions able to create new communication workflows in the system.

Figure 38 - RED V-Rapter[X] and Entaniya Fisheye HAL 200 killer combo.

3.9.0 - IMMERSIVE-STEREO-180° FORMAT:
To achieve out of the box this new format, is recommended to format the media at: 4K for each eye, 170° 180° FOV, 96 FPS / 120 FPS / 240 FPS** (Multiples of 24 for transmedia compatability, and muliples of 8 for cubic frame interpolation).

3.9.1 - TRANSMEDIA CONSIDERATIONS:
To facilitate different conversion on all transmedia formats, a friendly FPS should be produced at the top of the the chain. In this case, **(Immersive-180-S-3D >IMAX-S-3D > IMAX-2D > 16:9-S-3D > 16:9-2D)** following the "x8 cubic" rule, where multiples of 8 are the only options. Being the most compatible with 24 fps in the transmedia chain: **(24, 48, 72, 96, 120, 144, 168, 192, 216, 240).**

3.9.2 - IMMERSIVE LENS EQUIPMENT:
On this area there is little to almost none competition, but Entaniya legendary HAL lense (featured on "2001: A Space Oddesy"), **[Entaniya Fisheye HAL 200**[61]**]**, it's without a doubt the way to go in order to shoot in large formats. There is a need for Speed-Buster solutions, and there is no doubt that this will come eventually.

3.9.3.0 - IMMERSIVE CAMERA EQUIPMENT:
We have done very extensive research with RED Cameras, specially with the new RED V-RAPTOR[62] that allowed us to create a full SMPTE-2110 Workflow thanks to their SMPTE-2110 output card. The compact size is another factor to take in consideration. When it comes to Stereo, everything doubles the size, weight and specially price.

3.9.3.1 - CAMERA OPERATION:
For the camera operation is recommended to use RED CONNECT[63] wireless software, since is impossible to access all the settings directly through the camera whith this tight setup.

3.9.3.2 - RED PHANTOM TRACK R&D :
Immersive Dimension R&D helped RED to streamline this new feature to allow simultaneous A-B recordings allowing real-time visualization and Green-Blue Screen with trackers.

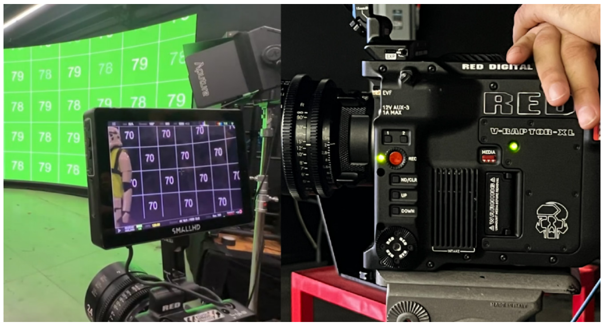

Figure 39 - Early stages of development with Immersive Dimension for RED's Phantom Track feature allowing A-B shooting options for Virtual Production. (Aug 16th 2023).

* **For 3D Immersive Virtual Production is recommended to use the A-B frame option at 240 FPS.**

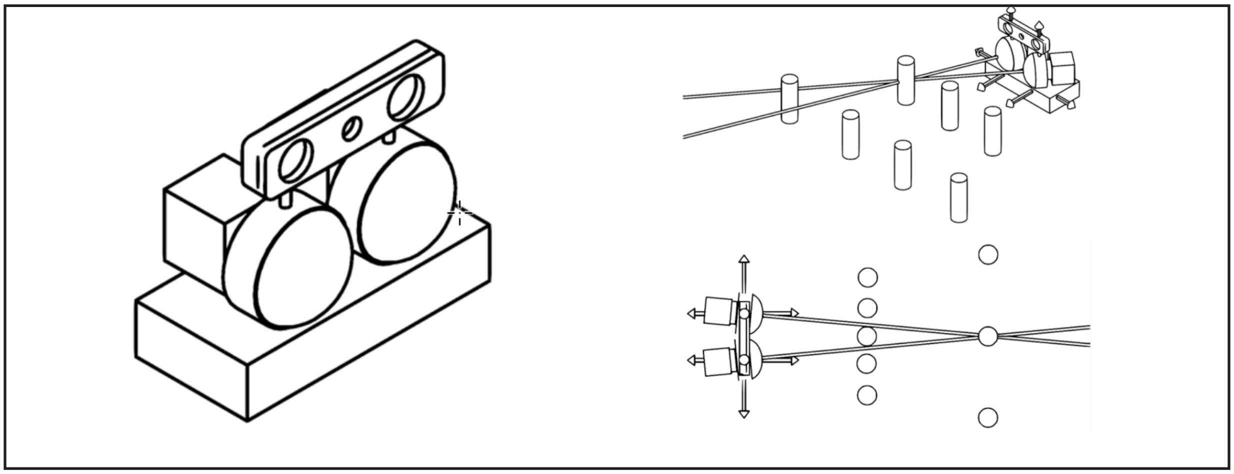

Figure 40 - Prototype for immersive visual feedback interaction with Apple Vision Pro Eye Gaze.

3.10.0 - STEREO CAMERA SETUPS

With the extensive research and development conducted by James Cameron and Vincent Pace on multiple Stereoscopic 3D workflows, it might appear that there is little left to explore. However, with the introduction of new immersive hardware, such as the Apple Vision Pro and the Quest 3, another area that merits attention is eye gaze tracking.

3.10.1 - EYE GAZE TRACKING[64]

The concept of tracking the nerves and muscles in the eyes seems almost like science fiction, yet it is precisely what gives cyberpunk characters their distinctive look for a good reason. By interpolating the data from electrodes positioned around the eyes, we can predict the direction in which the eyes will move within microseconds in advance. This is perfect for immersive environment systems. This prediction is sufficient to foresee a new point of view (POV) in front of the user. Although this may not seem like an intuitive or practical method for moving a camera, real-time systems can benefit from this workflow. Whereas the Cameron-Pace patent is limited to crossing the eyes to compensate for focus tracking on a subject, this method allows for independent movement of each eye in all axes to simulate human eye movement.

3.10.1.2 - REMOTE VR EYE GAZE TRACKING

Building on the same concept, the Apple Vision Pro already includes this feature. With the low latency of SMPTE-2110, remote VR eye gaze tracking could be a viable solution for live applications, although it has not yet been tested. This presents a theoretical approach to enhancing live interactive experiences.

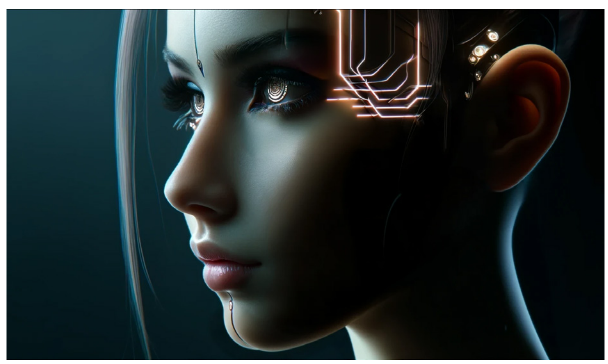

Figure 41 - Futuristic eye gazing augmentation.

3.10.1.3 - EYE GAZE WITH AI-SEGMENTATION

Another method to generate eye-gaze-like workflows is to create segmentation of the footage and create a digital plane movements based on the eye movements. This requires extra computer power to process real-time segmentation.

3.11 - REAL TIME VISUALIZATION

With 3D to VR set software like Assimilate's LiveFX and SGO Mistika VR, real time visualization on VR sets helps for full real-time visualization.

Figure 42 - DALL-E: "A 16:9 image of a futuristic virtual production set."

4 - RESULTS

4.1.0 - SYNCRONIZATION:
Integrating Precision Time Protocol (PTP) synchronization into our workflow presented significant challenges due to ongoing development and limited support for low-level API integration, especially in maintaining consistent development support across different hardware. To mitigate these challenges and avoid delays caused by troubleshooting across various cards vendors, we opted for an NVIDIA-exclusive approach. This decision was not intended as a universal solution but was a strategy to expedite problem resolution within our specific development context.

4.1.1 - MASTERCLOCK AT HIGH SPEEDS(+200Hz):
The inherent 200Hz speed limit of NVIDIA Sync cards restricts synchronization capabilities to 200fps across nodes, posing a significant limitation. To address this, we leveraged NVIDIA BlueField-2 cards, which enable synchronization of A6000-ADA cards through PTP time. While this workaround suits our current needs, we advocate for the broadcast industry to adapt to emerging standards, including 240 fps workflows for Master Clock tick source. Despite our inquiries, solutions for 240Hz master clock synchronization remain elusive. Our experimentation was limited to available hardware, underscoring the need for more scalable and cost-effective solutions.

4.2.0 - SYSTEM CALIBRATION:
As outlined in *(see ""3.2.0 - SYSTEM CALIBRATION" on page 12)*, our system calibration process is manual and time-intensive, and it does not guarantee flawless results. This highlights the importance of integrating with the "ST-2110 Black Box" to streamline and enhance the calibration process, ensuring higher accuracy and efficiency.

4.3.0 - ALTERNATIVE WORKFLOWS:
This document does not aim to serve as a comprehensive guide to workflows and setups. However, it's important to acknowledge the variety of available workflows and alternatives that complement our recommendations. Our key takeaway is the pivotal role of deferred rendering in achieving immersive virtual production. We assert that virtual production stands out as the most effective method for creating immersive content, encouraging the exploration and adoption of diverse workflows to enhance production outcomes.

4.3.1 - GREEN/BLUE LED PANELS + PREVIZ:
Combining green/blue LED panels with previsualization represents an optimal strategy in virtual production. Previsualization remains a critical feature, enabling creative teams to preview and adjust scenes before final production. This integration not only enhances the efficiency of real-time virtual production but also supports a more dynamic and intuitive creative process.

4.3.2 - LASER PROJECTORS VIRTUAL PRODUCTION:
Laser projectors offer a promising avenue for virtual production, potentially simplifying adoption compared to traditional methods. However, their successful integration requires as much coordination and planning as any other virtual production technique. Detailed workflow development and team alignment are essential to harnessing the full potential of laser projectors in this context.

4.3.3.0 - FRAME REMAPPING:
Frame remapping introduces a versatile approach to virtual

production, offering multiple solutions for diverse challenges. Through sophisticated frame manipulation techniques, enabled by SMPTE-2110 workflows, we can significantly enhance production flexibility. While this document intentionally omits the specifics of tracking markers and multi-camera frames to avoid complexity, it's worth noting that frame remapping plays a crucial role in sub-frame workflows. The advent of ST-2110-41/42 promises to further streamline these processes, reinforcing the importance of frame remapping in immersive and stereo 3D content production.

4.3.3.1 - FRAME REMAPPING CHALLENGES:
Advancements in frame interpolation technology, exemplified by tools such as Topaz AI, have significantly improved the quality of high frame rate video, particularly in transitions from 120 FPS to 240 FPS, where results are nearly indistinguishable from native high frame rate footage. While achieving 240 FPS at 4K resolution is within reach, it introduces practical challenges, notably in managing the thermal output of cameras operating at such high frame rates. Although 240 FPS is the preferred frame rate capturing stereo footage, 120 FPS currently presents a more feasible target for a wide range of production scenarios, including green screen and visualization workflows. This balance between high frame rate aspirations and practical thermal management underscores the ongoing need for innovation in both hardware and software to fully realize the potential of high frame rate production.

4.3.4 - DEFERRED RENDER:
Deferred rendering will remains a cornerstone of modern Immersive Virtual Production, valued for its enhanced control over the final output. This method significantly reduces the need for post-production fixes, offering a superior approach to managing visual quality. While applicable across various production types, its benefits become particularly crucial in the production of immersive content, where options are currently more limited.

4.3.5 - AI TRACKING, LABELING, INTERPOLATION AND GENERATIVE AI:
Advancements in AI for tracking, labeling, interpolation, and generative tasks are being pursued on multiple fronts. Our goal is to harmonize these AI capabilities into a cohesive system, enhancing interoperability within the virtual production ecosystem. A focus on standardizing metadata structures under SMPTE-2110-41/42 will facilitate seamless integration, offering a robust foundation for AI-driven production enhancements.

4.3.6 - SMPTE-2110 METADATA MANAGEMENT:
Effective management of SMPTE-2110 metadata presents ongoing challenges, particularly in parsing, organizing, and standardizing data across diverse production environments. Utilizing tools such as Wireshark has been instrumental in monitoring and classifying metadata. However, the absence of a universally accepted standard complicates these efforts. Future discussions within standardization groups will likely focus on refining and implementing ST-2110-41/42 protocols to enhance metadata. For now, we continue doing customized solutions as the rest of the industry does, but with the goal of standarization towards ST-2110 communication.

4.3.7 - API INTEGRATION WITH OPENAI'S GPTs:
Our project successfully demonstrated API integration across a diverse range of hardware, including RED Cinema Cameras, MegaPixel Helios, Evertz 5700, Pixera servers, and Unreal Engine with Vive Mars trackers data culminating in SMPTE-2110 outputs. This integration was facilitated by leveraging OpenAI's GPT for custom code generation, utilizing SDKs and API documentation provided by hardware manufacturers. The resulting system exhibited advanced capabilities, akin to what might be expected from Artificial General Intelligence (AGI), even before AGI has been officially recognized. This experience suggests that the evolving landscape of custom GPT models and AI agents is rapidly blurring the lines between current AI capabilities and the future potential of AGI, marking a significant shift in our approach to intelligent systems integration. “Focus is all you need”.

5 - DISCUSSION

5.1.0 - ST-2110-KLV AND SMPTE-RIS-OSVP:

This document serves two main purposes. First, it aims to educate readers about the emerging field of Immersive Stereo 3D with Virtual Production. Secondly, it proposes a collaborative effort between our SMPTE-RIS-OSVP team and the ST-2110-41/42 group. We suggest joining forces with the Sub-Group to integrate CAMDKIT into the ST-2110 framework and future implementation through R&D and implementation. This integration is intended to enrich the ST-2110 standard with comprehensive metadata outlined in this document, paving the way for future immersive content creation.
Recognizing that the current volume of data represents only an initial phase, we anticipate further expansions. Our goal is to develop a resilient **ST-2110-KLV** format that not only supports rapid future development and implementations but also establishes a robust communication pipeline for the immersive industry. This approach is designed to ensure scalability and adaptability, accommodating future advancements in immersive technology.

5.1.1 - INTEGRATION PROCESS:

We propose a detailed plan to incorporate CAMDKIT within the ST-2110 ecosystem. This plan will outline specific steps, required resources, and potential challenges, along with strategies to address them. Our objective is to ensure a seamless integration process that leverages the strengths of CAMDKIT and future integrations to the ST-2110 standard.

5.1.2 - METADATA CONSIDERATION:

A critical component of our proposal involves the careful integration of metadata relevant to immersive content. We will specify the types of metadata under consideration, their significance to immersive experiences, and the methodology for their integration into the ST-2110-KLV format. This detailed approach ensures that the future standard remains relevant and comprehensive.

5.1.3 - FUTURE-PROOFING STRATEGIES:

Acknowledging the rapid evolution of technology, we aim to design a format with the flexibility to adapt to future innovations achieving a future-proof standard, ensuring the longevity and relevance of the ST-2110-KLV format.

5.2.0 - ST-2110 LOW LEVEL PROGRAMMING:

In 2024, SMPTE-2110 infrastructure is affordable and scalable. With the untilization of NVIDA RIVERMAX SDK we are able to manipulate packages at the lowest processing level, but the need for alternatives requires mass education for the full industry.

5.2.1 - S-2110 EDUCATION:

Grabbing my SMPTE hat, I want to recommend SMPTE.org as a primary educational source for all the SMPTE standards, specially for ST-2100 data management. This recommendation goes beyond Broadcast Engineers, and a reminder that everyone in the industry is invited to be part of SMPTE (Society of Motion Pictures and Television Engieneers), specially if your profession is broadcast related.
(http:// https://www.smpte.org/)

5.3 - NEW PROPPOSED EQUIPMENT:

Here is a list of new propossed equipment on this paper:

5.3.1 - GREEN/BLUE LED PANEL:

A genlocked single driver RGB LED panel at 115,360hz for Blue/Green screen only. This should be diffused with white diffuser front and use the standard LED Panels frames. Lowering LED pitch requirements and reducing the full BOM, this configuration ensures many aplications where an LED matrix is not necessary. There is a wide market (including influencers) needing for this product.

5.3.2 - LIGHTING: PANELS SPOTS / PANELS:

Similar to the previous product, but for lighting equipment with genlock capabilities to a single RGBWW driver. Also at high refresh (115,360hz minimum)*.

5.3.3 - 240HZ ABLE GENLOCK

There is a real need of this product. 240hz will ensure a perfect LED calibration to sub-frames at high speeds.

* I will be surprised if is not realeased at NAB 2024.

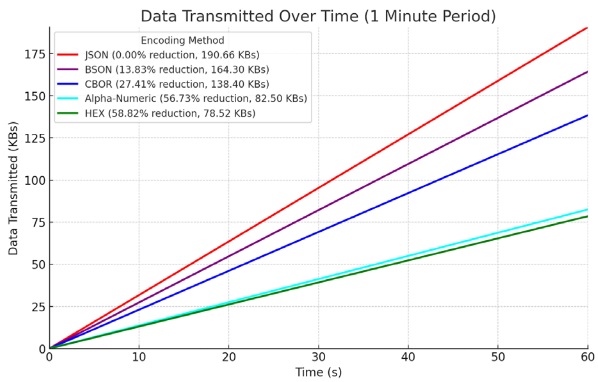

Figure 43 - Comparison of data volume over 1 minute for (JSON, BSON, CBOR, Alpha-Numeric and HexaDecimal packets).

5.4 - HIGH VOLUME METADATA

In the specialized field of virtual production, the SMPTE-RIS-OSVP Camera Metadata Group, Lens Metadata Group, and Camera Tracking Group have defined and recommended a set of metadata standards. These standards, formatted in JSON, are designed for both real-time and post-production use, as depicted in Figure X. The adoption of these SMPTE standardized formats aims to ensure compatibility with the next generation of virtual production hardware and software, facilitating a seamless integration into existing and future workflows.

5.4.1 - DATA TRANSMISSION EFFICIENCY

A pivotal aspect of selecting an optimal data format in virtual production and embedded systems is understanding the trade-offs in data transmission efficiency. A recent comparative analysis over a one-minute interval revealed significant differences in the data throughput of JSON, BSON, CBOR, Alpha-Numeric, and HEX formats. Notably, this comparison did not account for data serialization complexities, focusing solely on the raw transmission efficiency:

These findings underscore the impact of data format selection on transmission efficiency, particularly in bandwidth-constrained environments.

5.4.2 - JSON METADATA LIMITATIONS IN HIGH-PERFORMANCE ENVIRONMENTS:

Despite JSON's widespread adoption and human-readability, its suitability for high-performance data processing, especially within embedded systems, requires careful evaluation. This section delves into the implications of using JSON for metadata in scenarios demanding efficient data handling, contrasting it with alternative representations that offer enhanced performance.

5.4.3 - EFFICIENCY IN EMBEDDED SYSTEMS:

Embedded systems, often operating with limited processing capabilities (ranging from 16MHz to 96MHz) and memory constraints, pose significant challenges for data management. Continuous data streams, such as those from sensor outputs or device-to-device communications, amplify the inefficiencies associated with JSON parsing due to its textual nature and the computational overhead needed for its parsing and serialization. Consequently, serialezed binary data formats like HEX, which demonstrated a 58.82% reduction in data transmission size compared to parsed data in JSON, offers a more efficient alternative in these contexts.

5.4.4 - COST IMPLICATIONS:

In high-speed, high-volume data scenarios, the choice between JSON and more compact, binary formats has profound implications for both hardware requirements and overall system costs. Efficient JSON handling may necessitate advanced processors, such as a $13 FPGAs vs a $0.50 STM32 chip thereby increasing both the direct costs of hardware and the indirect costs associated with development and intellectual property that FPGAs brings. The empirical data showing significant efficiency gains with formats like serialized HEX data (58.82% reduction) highlights the importance of evaluating total cost of data interchange format.

5.4.5 - EXPLORING ALTERNATIVES:

Given JSON's limitations in high-performance contexts, the exploration of alternative formats is crucial. Binary JSON variants, such as BSON and CBOR, retain the structured advantages of JSON while providing greater efficiency.
But if we need extreme data reduction, serialized HEX data showcases the highest transmission efficiency in our tests.

5.4.6 - CONCLUSION ON HIGH VOLUME DATA:

This section is not intended to choose one or the other, a comparative analysis presented here illustrates the potential for significant efficiency improvements, guiding developers towards making informed decisions that optimize both performance and cost-effectiveness.

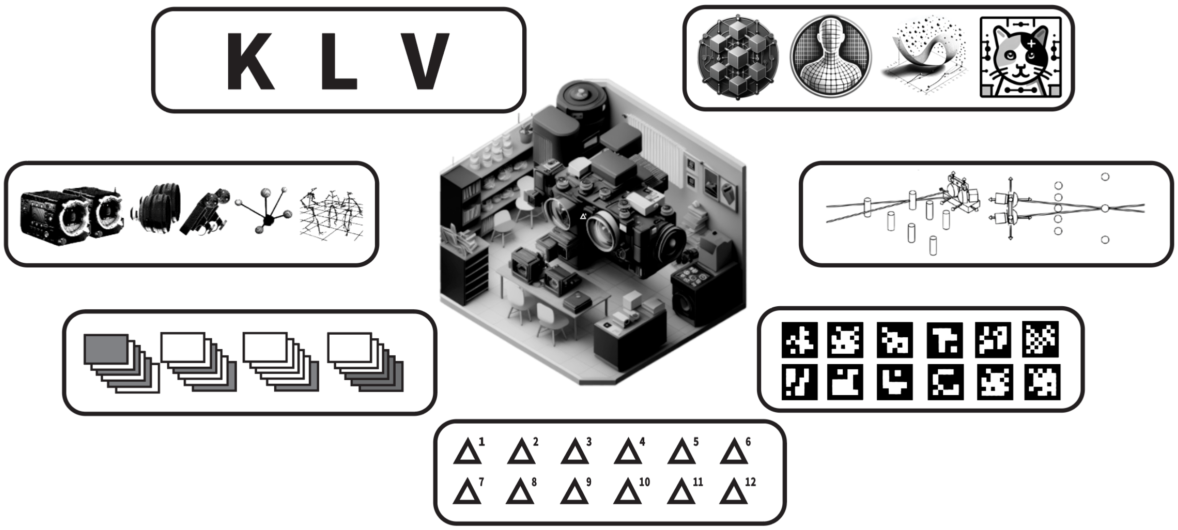

Figure 44 - "Immersive Research"

6 - CONCLUSSION

6.1 - RESEARCH & DEVELOPMENT:

The more independent research and development that comes from outside the manufacturing and software industries, the easier it will be for the entertainment industry to invest in the right tools for the job. New ideas born from passion should be encouraged by everyone in the industry. I love to share ideas that will steer development in the right direction. The reason for sharing all this information is not based on sponsorships (which we don't have) or promotions (which we are not conducting). We are sharing all this information because we see a vibrant technology capable of achieving more than what it was designed for, and it is time to move forward with immersive content. Technology is moving fast, and we are embracing those changes. I hope this paper will help in the future to create amazing immersice content.

6.2 - INNOVATION AND EXPERIENCE

Pushing new technology quickly is an effective startup method for learning and failing fast. There was a lot of failure through these new LED Screen workflows, but there was also a lot of learning. We personally learned that experience, when coupled with new technology, creates multiple workflows, and each one of those deserves a full research paper.

Virtual Production is intrisicly a TV Broadcast engineering problem to solve, and the film industry didn't fully crossed the river to the TV Broadcast industry yet, understanding its limitations and achieving perfect syncronization needed to address these challenges.

Virtual Production proved to be not only hard to adapt for the film industry but also a multidisciplinary industry where there is a lack of common communication jargon. Broadcast engineers, filmmakers, motion control engineers, color scientists, VR programmers, machine learning scientists, and artists walk into a bar...

Filmmaking is now an open universe of multidisciplinary professionals where they feed new ideas from each other in a Roman feeding feast before a final war against an imaginary dragon that theoretically will kill us all.

But beyond fearmongering, we are about to enter a new paradigm of transmedia, which, beyond AI, (ready or not), this industry (Immersive VP), will take over our production schedules, and we have to solve this technically challenging Rubik's cube as a team and working in sync.

Embracing large Green/Blue screen sets with Laser Projectors and LED screens with deferred rendering workflows idea, didn't came from a vacuum. It came from listening the crew and VFX expert complaining of A and B, and rather creating excuses, I personally took it as a mission to solve those problems instead of insisting on something that was half working or not working at all (beyond my powers), and working with what I had. Just focusing on "making it work".

Wen you multiply experience, you create something new.

Figure 45 - A 16:9 image of a virtual production film set . On the center a colossal, Japanese dinosaur-like monster.

6.3.0 - ADDRESSING THE COLOSAL DINOSAUR-LIKE MONSTER IN THE ROOM.

Here we go... we are going to talk about this.
And these are very philosophical and challenging times. Is the author of this paper opinion that we should embrace AI rather than fight the monster. AI is good finding the right path between “0” and “42”.

6.3.1 - USING AI AS A TOOL:
Althoug, a well critizised phrase, I stick with it. “Use AI as a tool”.
I personally believe in an interplanetary future for humanity rather than a Terminator timeline, both communicated by the same industry that allow us to dream impossible dreams.
“All I have to say is that, whatever you’ll dream will happen”.
I choose to dream a world where AI will help us to achieve our maximum potential. This “monster” can help us to protect us from ourself.

6.3.2 - CGI GENERATIVE-AI:
I trully believe that CGI creators are the most trained towards the use of AI as a tool. Big progress on new workflows will come from the same industry instead of programmers.

6.3.3 - AI AND THE AUTHOR
I don’t allow generative AI to think for me; instead, I use it to organize my ideas coherently for others to understand. As a neurodivergent individual, AI enables me to communicate new ideas in more creative ways. Throughout this paper, it’s evident where ChatGPT was utilized for grammar and structure, as well as where my own unedited words appear.

If you want to share your thoughts with me on this subject, don’t doubt to reach at to me directly:
videofeedback@gmail.com [Subject: White Paper],
I listen to everyone.

Figure 48 - "Where the puck is going".

6.4.0 - SKATE WHERE THE PUCK IS GOING:

Supply chain and hardware manufacturing slowed down during the last pandemic, but development derived from 2019 hardware didn't stop ([65] [66] [67]). We might slowed down four years of market addoption, but we accelerated software development with already old tech. µLED technology[68] (Micro-LED) has been developed to to a point that lightweight devices reached the desired quality and cheap price. As we experienced in LED China 2024, Micro-LEDs are ready to conquere the AR/XR market ([69]). Users walking around with Augmented Reality devices will not be a YouTube stunt anymore, but our everyday reality. Users will be already conditioned to wear an extension cable to an extended battery pack+spatial computer device (as we already are with our phones). These lightweight devices will serve also as wearable entretainment systems able to display 180° images at 60 FPS for each eye and the new generation of VR/Satial Computer headsets will be lightweight and most important, affordable.

6.4.1 - DEVELOPING 180 IMMERSIVE STEREO 3D CONTENT

Virtual Reality animations are cool, but you know what is better than that? Organic real Immersive content. Before we can't distinguish CGI characters from real life characters, Immersive Stereo 180 3D content can fill the gap with uncanny content. Creating 180 Stereo 3D libraries at maximum quality will have immediate audience, since the lack of PG content is making this format to slow down the format.

If you have an Immersive 180° Stereo 3D project and want to be technically produced by us, just let us know:
ramiro@immersivedimension.com, chris@immersivedimension.com

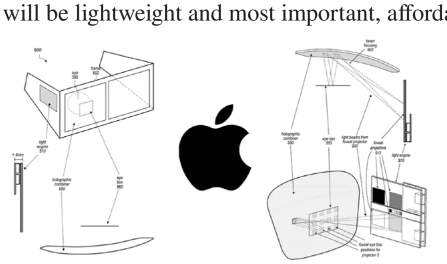

Figure 46 - This 2019 AR glasses patent filed by Apple. [70] [71] [72]

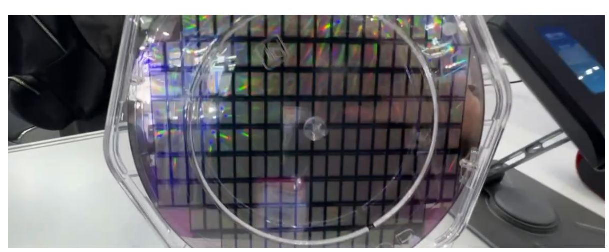

Figure 47 - LED China 2024 showcases multiple AR Micro-LED glasses, ready for prduction and affordable manufacturing cost.

REFFERENCES

1 CAN-BUS: controller area network (CAN bus) is a vehicle bus standard designed to allow microcontrollers and devices to communicate with each other.https://en.wikipedia.org/wiki/CAN_bus

2 Six degrees of freedom (6DOF), or sometimes six degrees of movement, refers to the six mechanical degrees of freedom of movement of a rigid body in three-dimensional space. https://en.wikipedia.org/wiki/Six_degrees_of_freedom

3 Car rotating rig https://inmotionfilmservices.com/machines/car-rotating-rig/

4 Delta timing: https://en.wikipedia.org/wiki/Delta_timing

5 Robots, Unreal Engine & Crashing Cars (Versatile Media Ltd): https://www.youtube.com/watch?v=jcozOywz-b3E

6 SMPTE (Society of Motion Pictures and Television Engineers) https://en.wikipedia.org/wiki/Society_of_Motion_Picture_and_Television_Engineers

7 SMPTE OSVP (On Set Virtual Production) https://www.smpte.org/rapid-industry-solutions/on-set-virtual-production

8 SMPTE-RIS (Rapid Industry Solutions) https://www.smpte.org/rapid-industry-solutions

9 CAMDKIT: https://github.com/SMPTE/ris-osvp-metadata-camdkit

10 JSON file: https://en.wikipedia.org/wiki/JSON

11 SMPTE RIS OSVP Metadata Project: https://github.com/SMPTE/ris-osvp-metadata

12 Application programming interface (API): https://en.wikipedia.org/wiki/API

13 Volumetric Capture: https://en.wikipedia.org/wiki/Volumetric_capture

14 Delta Offset on Network Timpe Protocol: https://en.wikipedia.org/wiki/Network_Time_Protocol

15 Motion Interpolation: https://en.wikipedia.org/wiki/Motion_interpolation

16 Automatic image annotation: https://en.wikipedia.org/wiki/Automatic_image_annotation

17 Computer Vision (AI Classification): https://en.wikipedia.org/wiki/Computer_vision#Recognition

18 Image Segmentation: https://en.wikipedia.org/wiki/Image_segmentation

19 Digital Twin: https://en.wikipedia.org/wiki/Digital_twin

20 Motion Control: https://en.wikipedia.org/wiki/Motion_control

21 Point Cloud: https://en.wikipedia.org/wiki/Point_cloud

22 DMX512: https://en.wikipedia.org/wiki/DMX512

23 CAN https://en.wikipedia.org/wiki/CAN_bus bus: https://en.wikipedia.org/wiki/CAN_bus

24 FBX : https://en.wikipedia.org/wiki/FBX

25 NeRF (Neural radiance field): https://en.wikipedia.org/wiki/Neural_radiance_field

26 USD (Universal Scene Description) file: https://en.wikipedia.org/wiki/Universal_Scene_Description

27 STL file: https://en.wikipedia.org/wiki/STL_(file_format)

28 KLV Data Structure: https://en.wikipedia.org/wiki/KLV

29 Paul Briscoe: ST-2110-41/42 Fast Metadata: https://www.youtube.com/watch?v=F0B2yJFXdrE

30 Kent Terry (Dolby) Advanced Audio Applications with SMPTE-ST 2110-41: https://www.youtube.com/watch?v=kz9_-g-TGCc

31 KLV Data Structure: https://en.wikipedia.org/wiki/KLV

32 ST-336 Data Encoding Protocol using Key-Length-Value: https://ieeexplore.ieee.org/document/8019807

33 MIDI Protocol (Musical Instrument Digital Interface): https://en.wikipedia.org/wiki/MIDI

34 Real-time Transport Protocol: https://en.wikipedia.org/wiki/Real-time_Transport_Protocol

35 User Diagram Protocol (UDP): https://en.wikipedia.org/wiki/User_Datagram_Protocol

36 Dave Smith (Engineer): https://en.wikipedia.org/wiki/Dave_Smith_(engineer)

37 MIDI SysEx (System Exclusive Messages) : https://blog.landr.com/midi-sysex/

38 Delta timing: https://en.wikipedia.org/wiki/Delta_timing

39 Grace Hopper: https://en.wikipedia.org/wiki/Grace_Hopper

40 Precision TIme Protocol: https://en.wikipedia.org/wiki/Precision_Time_Protocol

41 Generative pre-trained transformer: https://en.wikipedia.org/wiki/Generative_pre-trained_transformer

42 Rubidium : https://en.wikipedia.org/wiki/Rubidium

43 Crystal oven: https://en.wikipedia.org/wiki/Crystal_oven

44 Mo-Sys Cloud Rendering: https://www.mo-sys.com/product/camera-tracking/virtual-production/mo-sys-vp-pro/

45 Disguise cluster rendering: https://www.disguise.one/en/insights/blog/cluster-rendering-everything-you-need-to-know

46 RED Cloud upload: https://tinyurl.com/mrymf6m7

47 Rolling Shutter: https://en.wikipedia.org/wiki/Rolling_shutter

48 Rolling Shutter: https://en.wikipedia.org/wiki/Rolling_shutter

49 GhostFrame: https://megapixelvr.com/ghostframe/

50 NAB 2023 demo working for Vu Studios, with Evan Glantz as project manager, Ramiro Montes De Oca as Broadcast and System enginee.r

51 Silicon Core 240fps Demo: https://www.youtube.com/watch?v=D7Oj1IZzqlo

52 MLMs Survey: https://en.wikipedia.org/wiki/Large_language_model

53 Machine Learning: https://en.wikipedia.org/wiki/Machine_learning

54 Frame Interpolation: https://github.com/google-research/frame-interpolation

55 Topaz AI: https://www.topazlabs.com/

56 “The Conversation”, Francis Ford Coppola 1974: https://www.imdb.com/title/tt0071360/

57 “Immersive Research” is developing a ML model able to interpolate light from a picture and

58 OpenAI SORA: https://openai.com/research/video-generation-models-as-world-simulators

59 GPTs: https://en.wikipedia.org/wiki/Generative_pre-trained_transformer

60 NVIDIA Rivermax SDK: https://developer.nvidia.com/networking/rivermax

61 Entaniya Fisheye HAL 200 Series: https://products.entaniya.co.jp/en/list/hal-200-series/

62 RED V-Raptor [X]: https://www.red.com/v-raptor

63 RED Connect: https://www.red.com/red-connect

64 Eye Tracking: https://en.wikipedia.org/wiki/Eye_tracking

65 Apple's AR Patents: https://www.macrumors.com/2019/02/26/apple-glasses-patent/

66 Apple's AR Patents: https://kguttag.com/2019/11/20/apple-ar-patent-clutter-and-hiding-ar-patents/

67 Apple's AR Patents: https://www.slashgear.com/apples-latest-work-on-ar-glasses-19592247

68 Micro LED: https://en.wikipedia.org/wiki/MicroLED

69 Miro-LED: https://www.idtechex.com/en/research-report/micro-led-displays-2024-2034-technology-commercialization-opportunity-market-and-players/968

70 Apple's AR Patents: https://www.macrumors.com/2019/02/26/apple-glasses-patent/

71 Apple's AR Patents: https://kguttag.com/2019/11/20/apple-ar-patent-clutter-and-hiding-ar-patents/

72 Apple's AR Patents: https://www.slashgear.com/apples-latest-work-on-ar-glasses-19592247

END OF REFFERENCES

BACK PROPAGATION (AI)

A.2- BACK-PROPAGATION:
Back-propagation is a fundamental algorithm used in training artificial neural networks, serving as the backbone for a wide array of machine learning and deep learning applications. Its core purpose is to efficiently calculate the gradient of the loss function with respect to each weight in the network, which is essential for optimizing these weights through gradient descent or other optimization algorithms. The process involves two main phases: the forward pass and the backward pass.

A.3 - FORWARD PASS:
In the forward pass, input data is passed through the network, layer by layer, until it reaches the output layer. At each layer, the inputs are transformed using the current weights of the network and an activation function, producing an output that becomes the input for the next layer. The final output is then used to compute the loss (or error) by comparing it against the true values.

A-4 - BACKWARD PASS (BACK-PROPAGATION):
The backward pass is where back-propagation truly comes into play. Starting from the output layer, the algorithm computes the gradient of the loss function with respect to each weight by applying the chain rule of calculus, essentially determining how much each weight contributed to the error. This process is repeated layer by layer, propagating the error gradients backward through the network, hence the name “back-propagation.”

A-5 - GRADIENT DESCENT:
With the gradients calculated, the network’s weights can be updated in the direction that minimally reduces the loss, typically using a method called gradient descent. This step is crucial for learning: by iteratively adjusting the weights in small steps proportional to the negative of their gradients, the network gradually reduces the loss and improves its predictions.

APPENDIX INFORMATION

A.1 - Examples of Theoretical LED TILES Max FPS

REFRESHRATE (Hz)	SCAN LINES		FPS	# SHUTER OPTIONS	OPTIONS
15,360 Hz	1/8	=	1,920 FPS	1	360 (Open)
15,360 Hz	1/16	=	960 FPS	1	360 (Open)
7,680 Hz	1/8	=	960 FPS	1	360 (Open)
7,680 Hz	1/16	=	480 FPS	1	360 (Open)
7,680 Hz	1/32	=	240 FPS	1	360 (Open)
7,680 Hz	1/64	=	120 FPS	1	360 (Open)
3,840 Hz	1/8	=	480 FPS	1	360 (Open)
3,840 Hz	1/16	=	240 FPS	1	360 (Open)
3,840 Hz	1/32	=	120 FPS	1	360 (Open)
3,840 Hz	1/64	=	60 FPS	1	360 (Open)
1,920 Hz	1/8	=	240 FPS	1	360 (Open)
1,920 Hz	1/16	=	120 FPS	1	360 (Open)
1,920 Hz	1/32	=	60 FPS	1	360 (Open)
1,920 Hz	1/64	=	30 FPS	1	360 (Open)

REFRESHRATE (Hz)	SCAN LINES		SUB-FRAMES	FPS	# SHUTER OPTIONS	STEPS
15,360 Hz	1/8	=	1,920	24	80	4.5°
15,360 Hz	1/8	=	1,920	60	32	11.25°
15,360 Hz	1/8	=	1,920	120	16	22.5°
15,360 Hz	1/8	=	1,920	240	8	45°
15,360 Hz	1/8	=	1,920	480	4	90°
15,360 Hz	1/8	=	1,920	960	2	180°
15,360 Hz	1/8	=	1,920	1,920	1	360° (Open)

REFRESHRATE (Hz)	SCAN LINES		SUB-FRAMES	FPS	# SHUTER OPTIONS	STEPS
7,680 Hz	1/8	=	960	24	40	9°
7,680 Hz	1/8	=	960	60	16	22.5°
7,680 Hz	1/8	=	960	120	8	45°
7,680 Hz	1/8	=	960	240	3	90°
7,680 Hz	1/8	=	960	480	2	180°
7,680 Hz	1/8	=	960	960	1	360° (Open)

REFRESHRATE (Hz)	SCAN LINES		SUB-FRAMES	FPS	# SHUTER OPTIONS	STEPS
3,840 Hz	1/8	=	480	24	20	18°
3,840 Hz	1/8	=	480	60	8	45°
3,840 Hz	1/8	=	480	120	4	90°
3,840 Hz	1/8	=	480	240	2	180°
3,840 Hz	1/8	=	480	480	1	360° (Open)

END OF DOCUMENT