Lens Metadata Standards:
Experience and Recommendations

Ron Fischer
Technical Director
Universal Virtual Stage
Universal Studios
ronald.fischer@nbcuni.com
+1-818-370-6117

9/27/2012

This paper responds to a SMPTE technical committee request for the Universal Virtual Stage's experience with lens metadata and how standards might be applied to advance the state of the art.

Why?

Lens metadata is used in special effects pipelines for simulating camera optical systems. In feature film effects work camera simulation matches rendered CG images to plates of takes on a set or location. For on-set visualization, lens metadata is part of real-time imaging solutions, whose quality and adoption has improved rapidly over the last five years.

What?

Lens metadata provides some of the parameters for mathematical models of a lens & camera optical system. Related metadata distinct from the optical system would be, e.g. camera position and orientation, stereoscopic rig pose and light field array configuration.

This metadata has a crucial relationship to the image and it's time of exposure, e.g. through a rotating shutter or sensor scanning technique. The metadata also concerns multisample exposure techniques such as Red's HDX.

Some parameters of a camera system remain relatively fixed, while others are dynamic. The imaged area on a sensor is fixed due to geometry of pixels, however focal length, focus distance, entrance pupil, optical center and distortion values change during operation of a lens.

Dynamic variation occurs in response to, e.g. rotating the focus and zoom rings.

How?

Metadata parameters can be derived through optical or mechanical means. Optical means include laser sighting on a bench, and analysis of projected, prepared or free landmarks in images. Mechanical means include encoders within a lens barrel. These techniques can be used individually, together, before or after shooting a take.

Recording of lens metadata has varied broadly, from serial or network protocols captured directly by application software, to metadata channels in HD-SDI video streams. The relationship of the captured lens metadata to the exposure time of its related image is often not clear, in some cases only accurate to within +/-1 frame time.

It is also the case that, largely because current lens metadata formats are proprietary, many applications discard it when transforming images or passing HD-SDI video between devices.

Users and Vendors

Ours being a diverse industry of creative professionals, general opinion of the value of lens metadata varies.

In Production and Post these broadly fall into 1) metadata of all sorts is too often lost, and robust pipelines can only depend on approximate information about the camera, plus the captured take and reference images (calibration boards), or 2) the value of on-set visualisation makes real-time data useful, and it can also be used to assist post-production solutions if the data can be preserved during transport into post if a commitment is made to some set of proprietary protocols and capture mechanisms.

Lens and camera system vendors see clients using lens metadata to “remote read” lens barrel position. An emerging group of new clients uses lens information for on-set visualisation and virtual sets. The key difference in these two groups is that on-set visualisation crucially requires that lens metadata be directly related to the exposure time of the imaged frame, and that it be provided in real-time.

Position

Universal Virtual stage’s experience using systems such as Lightcraft’s Previzion, Autodesk Motionbuilder and Brainstorm, has been that streamed real time metadata is both immediately useful on set and also in post-production.

“Through the lens” techniques have improved but still produce “visually plausible” solutions that are are mathematically incorrect. In some cases this does not matter, but in others it can cause shot production cost to scale up disproportionately.

Hybrid tracking solutions which combine metadata from several sources, e.g. mapping tables generated beforehand, mechanical lens barrel encoders, and through the lens refinement, significantly narrow the large parameter space in which high precision post production systems derive an accurate camera simulation.

Robust solutions which scale incrementally with shot difficulty are more likely to arise from hybrid techniques rather than a single metadata source technique.

Any standard recommended to industry professionals should strongly support vendor solutions which are:

- robust
- incremental scaling of effort
- low intrusion
- broad application
- supportive of innovation and experimentation

While systems like Previzion, Brainstorm, etc. make good use of “raw lens metadata” capture, lens metadata vendors have been slow to provide access, and reluctant to invest in means to relate the data to the camera’s exposure shutter (e.g. genlocking).

This has resulted in a “chicken and egg” problem where, because metadata is often not available at all, discarded after capture, or is provided at greatly varying quality, the potential professional consumers of this data cannot take advantage without significant risk.

Scoping

A lens metadata standard relates to the inputs of a lens+camera moving picture optical system model. These models have measurement and rendering processes, either of which may take metadata as inputs.

Primary input data includes lens identity, camera imaged area, encoder values, scene identity, images, visual landmark positions, aperture and exposure timing information. Depending on the model chosen, this data may be considered static or dynamic.

Measurement processes take the primary input data and derive secondary data such as FOV or focal length, entrance pupil, depth of field, optical center, and distortion. Measured data may be considered dynamic or static, e.g. FOV might be derived per frame, or related to encoder positions through a two dimensional map of encoder positions for focus and zoom.

Once measurement is completed the optical system model can render matching images with a smaller subset of dynamic metadata.

The full scope of potential optical system models used by the industry ranges far beyond this paper, including the potential for stereoscopic, lightmaps, depth maps, motion estimation and non-imaging camera data to be used in populating optical models.

Recommendation: Concept

Similar to the revolution in camera imaging created by “raw” image formats and EXIF metadata, it is proposed that the only primary or “raw” lens metadata be standardized. This is a conservative approach which is likely to bring the greatest benefit while avoiding the problems of prematurely locking in concepts related to optical models and derived data. Vendors remain free to derive secondary parameters as they see fit for their optical models and processes, and gain the significant benefit of a well defined and transportable primary lens metadata stream upon which to build.

Recommendation: Primary Data

As an example, systems certified to the standard might provide the following:

- Unique ID for lens
- Unique ID for camera body
- Mount time (either actual or as detected)
- Imaged area on sensor
- Mechanical lens encoder positions (focus, zoom, aperture)
- Master Frame ID (e.g. timecode allowing for speed change)
- Frame subsample type (color, HDx, depth, etc.)
- Frame subsample ID (e.g. subframe time)
- Exposure duration
- Exposure type (rotating shutter or sensor scan type)
- Extended name/value data (optional)

The unique ID is based on the existing concept of serializing lenses and camera bodies. Most digital devices have a unique ID such as a MAC network address. These are meted out by a standards organization. One possible method of assigning unique IDs would be through TCP/IPv6 addresses or RFID numbers.

The mount time allows for detection of lens changes and suggests rechecking parameters such as optical center. Lens mount times can also be related to a len’s last alignment on an optical bench, or a post shoot test of parameters.

The imaged area of a sensor is crucial to making derived parameters transportable. There are technically inaccurate “marketing oriented” definitions of sensor size to be avoided. Suggested units are sensor pixel width and height in mm.

Dynamic lens operation data provided by mechanical encoders, while imperfect, can in some applications be entirely adequate for the completion of a shot. For large formats, lens encoder data can be used to dramatically reduce the search space for through the lens tracking solutions. Suggested units would be a continuous, 16 bit value from lock to lock of barrel travel, or start to stop of internal lens element movement.

Being dynamic, lens metadata must be accurately related to frame identity.

It’s proposed that frame identity be provided at two levels, the exposure start/stop time of the image, and the exposure start/stop time of subsamples, e.g. HDRI and other multisampling).

Likewise, it may be useful for the standard to include more camera-related data such as the shutter phase and angle, or sensor scan technique.

An optional area for extension data (name/value pairs) is recommended. The contents of this area would be unspecified, but might carry derived parameters from measurement processing, and allow for experimentation with new forms of primary data.

Recommendation: Control

Devices certified to the standard would be required to allow manual control of lens values for measurement purposes, i.e. focus distance and zoom (and any other relevant values) must be settable for measurement processes.

Recommendation: Transport Preservation

A standardized embedding format for this data in serial, network, HD-SDI and single frame image files would define its transportability through various media. Software and devices would be certified on the basis of their ability to preserve this data during transmission and format conversions.

Conclusion

While there are a number of published and proprietary standards for capture of primary lens metadata the industry would significantly benefit if it were standardized, along with how it’s embedded in transport media. A standard would address the current “chicken and egg” problem, wherein applications do not develop because lens metadata is neither standardized nor typically preserved by transport layers. By making lens metadata more reliable and broadly available, it can be rationally depended on by real-time and post-production clients without overt risk.

While higher level optical models exist, it is felt that a minimum set of “raw” lens-related metadata has the greatest immediate value, and allows the professional community to create added value specific to their solutions.

This standardized basis will allow also vendors to explore solutions for generating derived camera model parameters, whatever those camera models might be.

It is hoped that this standard will encourage solution vendors to focus on providing better real-time and post-production camera/lens simulation systems which meet goals of robustness and scalability.

As lens and imaging system models evolve in the next years additional standardization could take place via examination of the application “ecosystem” which develops atop this “raw” lens metadata standard.