An Experiment in Digital Intermediate Color Management Using ICC Profiles

Joseph Goldstone[*]
Industrial Light + Magic
San Rafael, California

Abstract

Traditional color management for visual effects feature-film production relied on output-referred techniques derived from film laboratory process control. The introduction of HDTV-originated content and a new goal of transferring the HDTV 'look' to film caused us to try and move beyond process control, and to experiment with graphic arts color management technology and techniques in our application domain. A feature-length motion picture, shot with HDTV cameras, was successfully recorded to film using a new workflow based on spectral measurement, CIE colorimetry and ICC profiles.

The path to delivery was a challenging one. In retrospect it is unclear whether off-the-shelf tools for the ICC-profile-based workflows of the graphic arts industry can be made to naturally encompass video encodings with legacy broadcast encodings, or films with extremely nonlinear reproduction characteristics in regions of shadow or highlight. There are also unique challenges in the metrology of projected film.

This paper describes some of the metrology and processing innovations we developed to support the abovementioned transfer. Although the client response was positive, the internal experience of production using an ICC-profile-based workflow was not uniformly so. It seems likely that in the near future, a hybrid approach combining laboratory process control techniques with colorimetric measurement and modest color appearance modeling may succeed where a pure ICC-profile-based system did not.

Pre-'Color Management' Color Management

In 2002 we delivered our first feature-length all-digital production (*Episode II: Attack Of The Clones*). Although we were pleased with the delivered work, the process of color correction prior to filmout was made more challenging by discrepancies between the color appearance of the HDTV grading monitor and the projected film. If a better match could be obtained between the two devices, more of the colorist's time could be spent on artistic (often image-region-specific) corrections and less on overall color predistortion to produce a pleasing film product. Two parallel efforts were made to improve the reproduction of image appearance from HDTV monitor to projected film.

The first of these approaches was an empirical attempt to find the inverse of the relationship between the colors produced by the *Episode II* HDTV-to-film conversion and the 'desired' colors as seen on a calibrated HDTV monitor. We created a synthetic equivalent of the well-known GretagMacbeth *ColorChecker* chart, basically four rows of six columns of black-separated color chips. This synthetic chart was recorded to film, developed, printed and projected. The same commercial color grading system that had been used to color-correct *Episode II* was brought out of the grading suite and into the screening room where it was used to manipulated the HDTV color space until the frame displayed on the grading monitor approximated the 'incorrect' projected film image. This transform (a combination of color matrixing operations, scalings and power functions) was inverted as best the system would allow, and the inverted transform was then applied to the HDTV frames prior to filmout. The hope was that prepending the inverse of the observed reproduction error to the reproduction path would produce an appearance match.

The resulting print, seen the next day, was a better approximation but not a match. We then developed a program that would transform the grading system's output prior to filmout, using a cylindrical model of a color space something like HSV. At each of three luminance levels, along six rays out from the neutral axis, three control points were available as 'distortion handles' for the color space. Handles could be locked into place to protect that area of the HSV solid from distortion, and new handles could be added. Essentially it was a 3-D LUT (a tabular function of three inputs yielding an output triplet) in HSV space with manual controls. We iterated for several weeks trying to find the set of control adjustments that would yield a projected film frame of our synthetic *ColorChecker* chart which matched the reference HDTV monitor.

The second approach was an attempt to transfer the measurement and reproduction technologies and practices of the graphic arts industry to our environment of visual effects production. Our task seemed very similar to the comparable efforts, seemingly successful, in introducing color management to the feature animation production workflow[1,2] Since no scene-referred operations were required -- the task was color correction and only subtitles were being added to the source material -- an output-

---

[*] Authors present address: Lilliputian Pictures
3553 20th St., San Francisco, CA 94110; joseph@lp.com

referred to output-referred gamut mapping problem seemed tailor-made for an ICC-profile-based system.

With early versions of the metrology tools described below in hand, it was possible to cobble together ICC profiles for the HDTV monitor and for projected film, and produce test frames (including the synthetic *ColorChecker* chart) with Photoshop. The ICC-profile-based approach was judged to give more pleasing results and was thereafter the focus of our energies.

ICC Profile Creation Methodologies

The creation of the profiles for projected film and for the HDTV monitor required device calibration, device measurement, measurement postprocessing and package-specific option selection when directing the profile-making software. We here discuss each step and point out problems we needed to solve, whether or not we actually found a solution.

Film Recorder Calibration

This was a straightforward calibration of our ARRILASER recorders, following the procedures outlined by the equipment manufacturer. The calibration is a straight-line relationship between device value (in the Cinéon printing density[3] color space) and measured density of the developed negative. Had we a densitometer that directly produced printing-density values (as specified by SMPTE RP 180[4]), we would expect that all of the red, green and blue device-to-density curves would have the same slope; instead we measured Status M densitometry and calibrated for a slightly-depressed red density, as per Kodak.[5] Iteratively measuring and adjusting the lookup tables loaded into the ARRILASER recorder usually lets us get to within ±0.01 (Status M) of these ideal curves, a level which produces no noticeable visual anomalies on a greyscale.[2]

Monitor Calibration

The HDTV monitor was calibrated by our video engineering group to our internal standard, derived from broadcast practice. White luminance was set to 92.5 nits, and 24% black to 2.75 nits. Setting these two points and calibrating an intermediate point to set a gamma avoided having to make a measurement of extremely dark material, where instrument precision and accuracy might be marginal.

Film Recorder Characterization

The presence of interchannel effects, even on an intermediate stock such as Kodak 2242, prohibited the use of the ICC profile format's simpler matrix/TRC model for projected film. The more elaborate multidimensional lookup table model used in the A2B*x* (device to PCS) and B2A*x* (PCS to device) tags requires sampling the output at regular intervals in device space.

The measured results form the contents of the multidimensional lookup table in the A2B*x* tag. The inverse of that 3-tuple to 3-tuple mapping is the desired content of the B2A*x* tag, augmented by gamut mapping to handle requests for PCS colors the device cannot produce.

For the measurement to correspond to the color stimuli presented to a moviegoer, the obvious methodology was to project the stimulus on an otherwise black screen and measure the CIE XYZ values reflected to the nominal ‘best seat in the house’, usually defined as just off theater center, about 30 feet back from the screen. This straightforward approach has several fatal flaws. First, one must realize that a single frame cannot be projected; held in the projector film gate, it would burn up in a matter of seconds. Second, one should also realize that when using our Minolta CS-1000 spectroradiometer for such a measurement, anything below about 40% of peak white luminance takes more than two minutes to produce a reading.

For an 11x11x11 sampling of the device space, assuming an average measuring time of one minute and 24FPS projection, one would need just under 120,000 feet of film to contain the sample set. Since this is rather impractical, one could consider exposing a short section of film (12 feet is a typical length), splicing the tail to the head, and projecting the resulting loop. Since it takes a projectionist about 30 seconds at best to unthread one loop and thread the next, an average measurement time for our 11x11x11 sample would take a minimum of 33 hours of continuous labor, again clearly outside the realm of possibility.

Our solution was to adapt the transport previously developed for an autoadvancing densitometer so as to move the print between an extremely bright light source on the one side and a GretagMacbeth Eye-One spectrophotometer on the other. The light source is an integrating sphere from a late-1980s Eikonix scanner with a film aperture plate over the exit port. Its two Solux daylight simulator bulbs (with a CRI of 99.35) are driven by a current-regulated power supply. We machined a mount that places the spectrophotometer aperture in the exact center of the film frame, suspended within 1mm of the film surface.

We use an experimental driver for the Eye-One that adapts the integration time to the stimulus luminance, but this is not strictly necessary; prior to this driver being made available to us, we achieved good results partitioning the patches by visual density (with a densitometer) and, in successive passes, reading the patches in each partition through a glass ND filter that brought those patches into the viable range of the fixed-integration-time driver. Supplying more than 25,000 nits from the lamphouse means even the most dense measured patch on Kodak Premier print stock can still be read in under five seconds; most patches are read in under two seconds.

The autoadvancing Eye-One spectrophotometer

The snouted Minolta CS-1000 in the theater

This system typically reads the patches corresponding to an 11x11x11 sample of the film recorder in slightly less than an hour. A single 4th-row Minolta CS-1000 sample of a looped hold frame of device white (*not* the reflected 'open gate' white of the projector but the reflected projection of the lowest-density neutral the film recorder can achieve for a given emulsion batch, laboratory, print stock and development process) tells us what color stimulus a moviegoer would have received from its projection. The Minolta-measured spectral power distribution of this white and the Eye-One-measured spectral power distribution of the corresponding square of the 11x11x11 sample cube together yield a wavelength-specific 'scaling spectrum' that is used to remap all the Eye-One patch measurements to their projected equivalents. These remapped spectral measurements are converted to CIE XYZ space and thence to CIE LAB space.

The Minolta CS-1000 can be instructed to read either emulating the 1931 2º observer or the 1964 10º observer. As most of the discussion in the ICC profile specification concerns the 2º observer, we took our measurements in that mode. We also found substantial flare proportional to the intensity of the test patch, which we attribute to bounce light from the theater walls and ceiling entering the Minolta optics from outside the angle of interest. Our solution was to add a long, black-velvet-lined PVC tube (the "snout") with a small aperture at the far end, fit snugly around the spectroradiometer lens, to screen out flare.

Monitor Characterization

Although a TRC/Matrix model could have been used for the HDTV monitor profile, we opted to use the A2B*x* and B2A*x* tags, with their 1-D shaper LUTs before and after a 3-D LUT. Measurement was carried out with the Minolta CS-1000, in a fashion similar to that done for film, but with a much smaller sample set: a 6x6x6 regular sample of the device space, and a 6x6x6 subsample of the darkest corner of the initial sampling. In addition to the previously-mentioned snout, we added an opaque plastic sheet covering those parts of the monitor face not already covered by the PVC tube, and carried out the measurements with room light brought to an absolute minimum, considerably below the dark surround normally used by the colorist.

Film Recorder Data Preprocessing

We linearly interpolated the 0-1023 recorder device space coordinates into the 0.0-255.0 range required by ProfileMaker Pro. The corresponding CIE LAB coordinates derived from remapped spectral data were *not* modified to account for flare, surround, or other elements of the viewing environment.

Flare models found in the literature did not seem to account for both the baseline flare (from, e.g., theater exit signs) and for content-dependent flare. We hoped instead that the use of the snout would give us the equivalent of flare-free readings.

We did not attempt to compensate for differences in surround. For the film image, surround conceivably could have been defined as the black velvet curtains bordering the projected area, but for the HDTV monitor the definition was considerably more difficult. As we could not decide whether monitor surround would be the top and bottom areas of the monitor (set to reference black so as to letterbox the image), or the monitor's dark bezel, or the grey carpet which surrounded the monitor cart from the viewer's standpoint. There was no obviously correct way to factor out lateral-brightness adaptation. Not being able to compensate measurements of the monitor for surround effects, we left the film measurements similarly unperturbed.

HDTV Monitor Data Preprocessing

The Rec. 709[6] 8-bit standard reserves the all-zeros (0) bit pattern and the all-ones (255) bit pattern as timing reference information. Conventional *black level* is encoded at as 16, and *nominal peak* at 235. That black is not encoded at 1, and that white is not encoded at 254, comes from the standard's analog heritage: provision for filter undershoot and overshoot as video passes through the studio. The 10-bit Rec. 709 variant reserves 0-3 and 1020-1023 for control information, 4-63 and 941-1019 for undershoot and overshoot, and 64 and 940 for reference black and white. Some (but not all) 10-bit RGB video devices will display values darker than reference black if fed RGB values below 64, and some (but not all) will display values greater than nominal peak white if fed RGB values above 940. We have long opted to use the code values between 941 and 1019 inclusive to hold highlight detail normally lost in video processing.

The unspoken assumption in the ICC profile format is that the data in the CGATS file describing the device codes that index the profile will have 0 as a valid device value and that this 0 will produce the device's darkest possible output value; and similarly, that a 255 value will produce the brightest white. A related assumption seems to be that the curve is uniquely invertible. We linearly interpolated the [64 .. 1019] range of our observations into the [0.0 .. 255.0] range expected by ProfileMaker Pro. This led us to the unpleasant position of conforming to the standard, while having a black level significantly above the darkest color the device could produce.

Film Recorder Profile Creation Options

Color crosstalk in the film and its nonlinear handling of shadow and highlight detail precluded modeling the transformation from device values to color stimulus with a matrix/TRC model. The closest match to our device in ProfileMaker Pro seemed to be a printer. The set of profile creation options for printers is extensive, but for our application few seemed appropriate. The choice of a target-device native neutral or a source-device native neutral (our preference, since the director was very much trying preserve the original HDTV look, and many theaters are not dark enough for full white-point adaptation) was of less importance when we discovered such a choice did not act on 'lighter areas'. While this constraint might make sense for paper white, it does not make much sense for projected white. We were willing to accept a lessening in projected luminance, if need be, so as to force a white point shift, but this commercial package did not give us the option. This is not surprising, considering its target user base and their applications.

The sets and locations for this production were full of extraordinarily saturated color. ProfileMaker Pro offers two styles of gamut mapping for printer profiles, one of which favors lightness reproduction, the other chromatic fidelity. We chose the latter. This was probably the best choice for us, but had consequences in the reproduction of achromatic blacks that will be described below.

Monitor Profile Creation Options

The characterization of a monitor in ProfileMaker Pro is much simpler than the corresponding process for a printer profile (as we used for the film recorder). Essentially one chooses between a matrix/TRC device model or a multidimensional LUT model for the device. The TRC stored is not a true TRC but a single number indicating a power function exponent (a per-channel 'gamma'); this led us to choose the multidimensional LUT.

We also chose the observed white point of the monitor as our profile's white point. This seemed consistent with our mapping the reading of the 1019, 1019, 1019 patch to LAB (100.0, 0.0, 0.0).

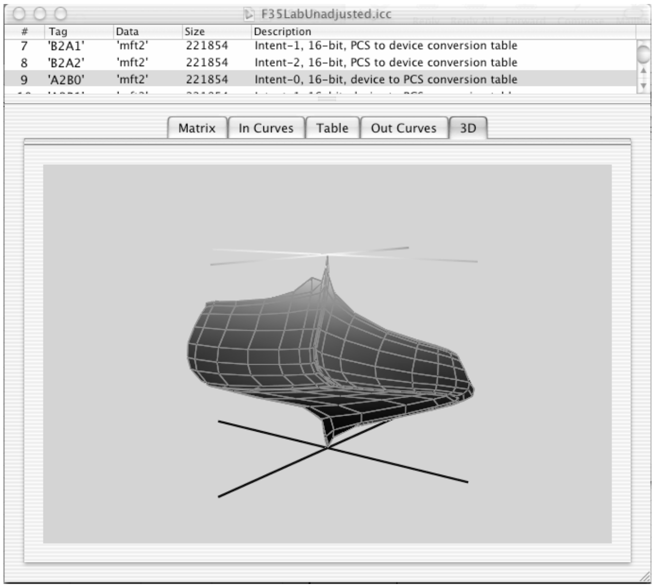
Results

Some characteristics of the projected film images were identified as needing enhancement: the darkest few stops of the image seemed crushed together, especially along the neutral axis; the black level seemed elevated above the customary Kodak Vision black; and the white point was warmer than that of the monitor. At the time, we imagined that these problems would be readily remedied in a few weeks of data adjustment. This proved optimistic.

ICC profiles are binary encodings of a structure whose format is highly variable. Although a variety of profile-inspection tools are on the market, they invariably stopped short of the level of detail we desired in investigating the abovementioned problems. It was a good month before we had a rudimentary library for profile header and tag manipulation, on which we then built extended profile inspection and editing tools.

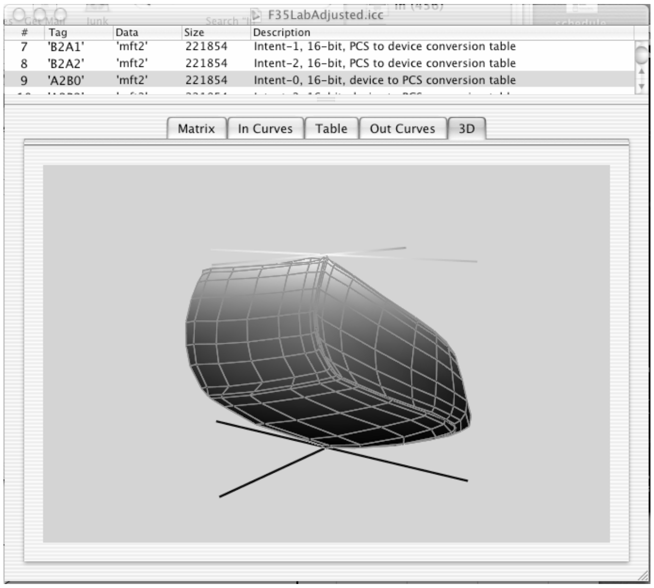
Printer profiles created by ProfileMaker Pro use CIE LAB as their PCS. A calibrated recorder will have a linear relationship between device code value and negative density, but the printing process introduces a toe and a shoulder. The transformation of CIE XYZ to CIE LAB values further compounds this compression. The figure below shows the mapping of the device code values to their projected film counterpart values in LAB space.

The achromatic white and black points give a certain luminance range to the device's projection into LAB space. These points seemed extremely sensitive to the presence of chroma (thus the 'sharp' appearance of the gamut solid at top and bottom). The slightest introduction of chroma into a shadow lifted it up in luminance and produced a 'milky' black. We hypothesized this behavior to be a gamut-mapping artifact. Although a comparable process was probably taking place with the highlights, this was not as visually apparent.



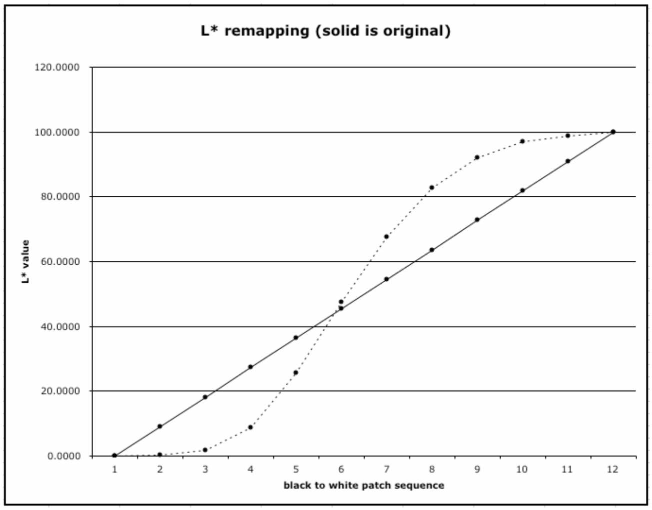
#	Tag	Data	Size	Description
7	'B2A1'	'mft2'	221854	Intent-1, 16-bit, PCS to device conversion table
8	'B2A2'	'mft2'	221854	Intent-2, 16-bit, PCS to device conversion table
9	'A2B0'	'mft2'	221854	Intent-0, 16-bit, device to PCS conversion table

*The A2B*x *gamut, prior to L* adjustment*



#	Tag	Data	Size	Description
7	'B2A1'	'mft2'	221854	Intent-1, 16-bit, PCS to device conversion table
8	'B2A2'	'mft2'	221854	Intent-2, 16-bit, PCS to device conversion table
9	'A2B0'	'mft2'	221854	Intent-0, 16-bit, device to PCS conversion table

*The A2B*x *gamut, after L* adjustment*



The L remapping curve*

ICC A2B*x* and B2A*x* multidimensional lookup tables are a concatenation of 3 1-D input shaper LUTs, the 3-D lookup table, and 3 1-D output shaper LUTs. Our remedy for the observed compression was to derive a shaper LUT through which we passed the L* values of the LAB measurements being fed to ProfileMaker Pro. We then took the resulting profile and rewrote its B2A*x* tags to compose this shaper LUT on the existing contents of the tags' input shaper LUT curves. To support soft proofing the inverse of this shaper LUT was applied to the existing contents of the comparable A2B*x* tags' output shaper LUT.

The result was a considerably more spread-out projection of the device gamut into LAB space:

Although the pure blacks became blacker and image shadow detail was restored to what we were seeing on the monitor, we began to see new problems. Most noticeably actors with naturally dark red skin tones would, if they were lit in such a way as to increase their normal saturation, evince a shift from red towards magenta.

We revisited our profile creation path and, unable to immediately find an explanation for the problem and with production deadlines imminent, attempted to remedy the problem by a succession of programmatic edits to the recorder profile. The most useful profile-editing tool was probably one that would take a kidney-shaped section of LAB space, with the projected center on the neutral axis, and adjust the data in the B2A*x* and A2B*x* tags to reflect values that were rotated to a degree specified by the user. Radial and rotational fade-on and fade-off of the effect were similarly adjustable. (Note that this tool is disturbingly close to the manually-operated lattice distorter we earlier rejected.) In practice we found that (a) most of the problems were near a very narrow wedge in the a,b plane roughly centered around the positive a=b line, and (b) adjustments to that narrow wedge tended to cause blotchiness in skin tones.

Ultimately, we ended up capturing the relationship between the RGB space of the monitor and the RGB space of the recorder in a proprietary lattice format and then programmatically editing this RGB-to-RGB transform to eliminate observed blotchiness. There were perhaps 20 shots for which even this captured, edited lattice did not adequately solve the problem of overly-red complexions. These were handled by manually pre-distorting those parts of the imagery prior to filmout. While this may sound like an indictment of the overall effort, one should compare the 20 shots required such pre-correction with the more than 2,000 shots needing similar correction in *Episode II*.

Observations

The ICC 4.0 profile specification offered a reassuring level of detail that gave us the confidence to try using tools from another industry. In practice, however, it became something of a Procrustean bed as we tried to adapt our devices to a model with unspoken assumptions, e.g. that all device values are intended to represent color information, and that the highest valid device value represents a 100% white diffuse reflector. Not all output-referred color spaces are so well behaved, Rec. 709 most especially.

Although the ICC specification suggests "[P]rivate tags allow CMM developers to add proprietary value to their profiles", it concludes the same paragraph with "…the use of private tags should be kept to an absolute minimum." In fact, contemporary vendors of profile-making packages know the user may choose from a variety of CMMs, and so the "secret sauce" which they promote as distinguishing their system must be embedded in standard tags. The business model thus encourages a conflation of device characterization and gamut mapping in the B2A*x* tags, which is unfortunate. When we encountered problems after our reshaping of L* in the B2A*x* and A2B*x* tags, we were left wondering if ProfileMaker Pro was doing something in its gamut mapping of the perturbed LAB that was appropriate and clever for graphic arts printing and that relied on the L* values of the blacks being so close together. Our difficulties involved colors that were well within gamut. Had gamut mapping been in a separate tag, we could have done better problem isolation.

We also found that no amount of intellectual argument could dissuade production from judging projected imagery without simultaneously seeing monitor imagery. None of us had any formal training in color science, but we had read enough to be continually astonished that in an environment which effectively prevented chromatic adaptation, using an approach that ignored virtually every known color appearance factor, the projected film's match to the monitor was considered better than anything we had previously accomplished.

Future use of an ICC-profile-based workflow for output to motion-picture film would require a better resolution of the crushed-blacks problem, along with remedies for a few other problems not mentioned in this paper. In retrospect, too much of our time was spent in attempts to understand a B2A*x* tag creation algorithm optimized for another medium entirely. We would probably have done better to have gone back to the open literature and developed our own, giving us much greater understanding, predictability and control. This would have also afforded an opportunity to experiment with moving gamut mapping into an abstract profile. Similarly, writing our own TRC tags with real TRCs and not shorthand gamma values might have given us real improvement on our monitor characterization.

An alternative to an ICC-profile-based workflow would be to augment the process-control-based system of the past with colorimetric measurement of HDTV monitors and with an analytic model of the relationship between Status M densitometry, Status A densitometry, and projected film colorimetry. Such an approach is beginning to be available as a commercial product,[7] which, although still a black box, is at least a black box designed for digital intermediate, and not for graphic arts reproduction.

Acknowledgements

Natasha Leonnet, the show's colorist, gave invaluable, continual and lucid descriptions of the state of the work. Fred Meyers of ILM provided the opportunity to try this new approach. Patrick Endaya and Thomas Senn of GretagMacbeth provided advice and access to engineering talent. Finally, Rod Bogart rescued the project from dead ends in both concept and execution; without his assistance the project would have been at best a desaturated shadow of its final realization.

References

1. Arjun Ramamurthy, Lem Davis and Frank Herbert, "Achieving Color Match Between Scanner, Monitor and Film: A Color Management Implementation for Feature Animation", *SMPTE Journal,* **108**(1999) 6, p. 363-373.
2. Ado Ishii, "Color Space Conversion for the Laser Film Recorder using 3-D LUT", *SMPTE Motion Imaging*, **111**(2002) 11, p. 525-532.
3. Glenn Kennel and David Snider, "Gray Scale Transformations of Digital Film Data for Display, Conversion, and Film Recording", *SMPTE Journal*, **102**(1993) 12, p. 1109-1119.
4. SMPTE , *Spectral Conditions Defining Printing Density in Motion-Picture Negative and Intermediate Films*, SMPTE Recommended Practice RP-180, 1994.
5. Kodak, *Kodak Digital LAD Test Image – User's Guide and Digital Recorder Calibration and Aims*, Kodak publication H-387, 2001.
6. ITU, *Basic Parameter Values for the HDTV Standard for the Studio and for International Programme Exchange*, ITU-R Rec. BT.709, 1990.
7. Richard Kirk, *Truelight Software Library*, 2003 (available at http://www.filmlight.ltd.uk/truelightLibrary.pdf).

Biography

Joseph Goldstone received his B.S. degree in Electrical Engineering and Computer Science from the Massachusetts Institute of Technology in 1982. He managed the film scanning and recording department at Digital Domain in Venice, California from 1994 to 1998, and worked in the film group at Industrial Light + Magic from 1998 to late 2003. From spring of 2002 to the fall of 2003 his work was largely involved with the metrology and color management of film projectors,, digital cinema projectors and HDTV displays. He is a member of the ACM and the SMPTE.