ACES Color Management
an industry op-ed by Steve Wright

In the last several years the subject of color management has become a colossal issue in the visual effects community, and for two distinctly different reasons. The first is that a modern visual effects project will get images from a wide variety of cameras and sources, so how do we get all of these different images to play nice together in a shot? Secondly, how can you render photorealistic CGI for a movie when the live action clips you need to match to bounce around in different color spaces? The answer is ACES, the *Academy Color Encoding Specification* that promises to unify all images into a grand unified field theory of color.

So what's the problem here? So what if some scenes are shot with a RED, others with an Arri Alexa, a few with a Sony CineAlta, and let's throw in a few shots with classic 35mm film? Can't we just cut them all together then color correct them to make them all look like they were shot by the same camera? No. And even if you could, what about the CGI? What colorspace should it be in to match the various live action sources?

The Academy of Motion Picture Arts and Sciences has at last tackled this problem in a comprehensive manner to produce ACES (Academy Color Encoding System), a color management system created specifically for filmmaking and visual effects that also, by the way, is perfectly suitable to video and television production. The short form of the story is that it takes all captured footage regardless of camera type, converts it to a master all-encompassing "color space" - perfect for editing, image processing, compositing, CGI, and color correction – then outputs it to whatever display colorspace you want – digital projectors, 35mm film, HD TV, iPad, iPod, iPhone, etc. It even future-proofs the master for any conceivable new display system into the far future.

Color Space

Since our problem is trying to get all these different color spaces to play nice together, let's take a minute to define what a color space actually is. You can think of a colorspace as a 3-dimensional cube like the one shown here. Instead of XYZ to locate any point in 3D space, the axis are RGB to locate any pixel in color space. The problem here is that the actual color of a given pixel depends on the exact color your red, green, and blue is. And these vary a great deal between different types of equipment. There are some other fiddly issues, but this is the main one.
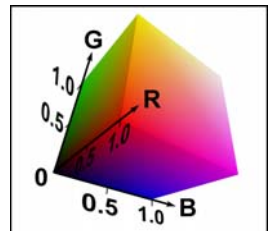
Capture Referred data

The problem with cameras (digital or film) is that they "bake" their particular color space into the images they capture, hence Capture Referred data. If several different cameras were used to capture the same scene you end up with several different versions of that scene – and this pretends that you could somehow get all of them to match the same exposure. Capture Referred data is not conducive to mixing multiple cameras on the same project and certainly does not help the CGI folks know what to render to.

Scene Referred data

What is needed is to somehow represent the original scene in front of the lens in a <u>linear</u> light space, meaning if you double the code value of a pixel it doubles the amount of light that it represents. No log

images or gamma corrected video, but a one-to-one relationship between the brightness of the scene and the digital data that represents it. Since this data refers to the original scene it is called Scene Referred data. This is the ACES color space. The CGI folks love it since linear light space is exactly where all CGI is rendered internally anyway because CGI is a mathematical model of the real world and the real world is linear light space. It is upon output that the CGI is converted to whatever color space the job is set up for thus knocking it right out of the original linear light space.

So Scene Referred liner light space is perfect for the CGI folks, but what about mixing all those different cameras together? That was the other problem. ACES addresses this too. While the camera has baked in its personal color bias into its captured images, it is possible to measure this and mathematically back out the camera's bias to restore the original scene illumination creating the much sought after Scene Referred data.

Now, in principle, you can capture the same scene with 10 different cameras and correctly convert them all to Scene Referred data, all 10 images would be essentially identical. They can now be mixed and matched in the same shot or cut between them with no problem. And any CGI can be mixed with them too because it is linear light space to begin with. Things are starting to look up here.

The ACES gamut

A key part of the ACES color space is its enormous gamut. Gamut is the range of color that a given color system can handle and that is yet another variable in these different cameras. CGI has an unlimited gamut because all of its calculations are purely mathematical so there are no physical limits, but all cameras have physical limits. Problem is, their limits are different. ACES solves this problem neatly by defining its gamut to be so huge that no physical device could ever max it out. It even far exceeds that of human vision, the gold standard of imaging systems.

So we have captured our various scenes with various cameras and convert them all to ACES Scene Referred linear light space so they all play nice together. We can now do any image processing, compositing, transforming, look development, or any other processing we wish knowing that it will all match. And by utilizing the 16 bit floating point precision of the OpenEXR file format we can crunch the pictures all we want without introducing artifacts. In fact, the ACES standard is implemented in a special "constrained" format of EXR images so everybody follows the rules and you can trust an ACES image no matter who creates it.

Display Referred data

The last issue is the output for display devices. Here we have the same problem we had with the cameras, only in reverse. The cameras capture different data for the same scene, and the display devices (projectors, TV sets, etc.) produce different appearing images with the same data. You've seen this a million times. The same image looks different on your workstation compared to a TV set compared to a digital projector – even on different brands of projectors.

To address this ACES incorporates a third and final stage of the color pipeline to convert the Scene Referred data (which is universal and omnipotent) to Display Referred data where the display device characteristics are "baked" into the image data so it looks nice on a particular display. Of course, you need one set of Display Referred data for a theatrical digital projector, another for shooting out to 35mm film, another for HDTV, etc. But at least we can all meet in the middle with the ACES Scene Referred color space then go from there to any display device we want knowing that all the images from all the different sources are going to play nice together and look great.

Steve Wright