Using OpenEXR and the Color Transformation Language in Digital Motion Picture Production

August 2, 2007

Florian Kainz
Industrial Light & Magic
on behalf of
The Academy of Motion Picture Arts & Sciences'
Science and Technology Council Image Interchange Framework Committee

Abstract

We present a framework for the interchange of digital images among multiple companies working on a motion picture production. The framework makes it possible to unambiguously describe what the digital images are intended to look like on the screen, without baking this information into the images' pixels early in the production pipeline. This makes it possible to make decisions about a movie's "look" as early as on the set or on location, while retaining the capability to revise this look extensively later, for example, in the Digital Intermediate process. Visual effects work can be performed on "scene-linear" images, ensuring that 2D compositing and 3D rendering calculations are performed on physically meaningful data. This framework requires that companies exchange two kinds of data: images, typically in the form of OpenEXR files, and color transforms, represented as programs written in a Color Transformation Language. The framework described here is one possible implementation of concepts that have been developed by the Academy of Motion Picture Arts and Sciences' Image Interchange Framework committee.

Introduction

In late 2004 the Science and Technology Council of the Academy of Motion Picture Arts and Sciences formed an Image Interchange Framework committee. The goal of the committee is to define a conceptual framework, file formats, and recommended practices related to color management and the exchange of digital images during motion picture production and archiving. This work is ongoing, and the committee has not yet made any official recommendations.

In addition, the Science and Technology Council has recently published an open-source interpreter for the Color Transformation Language (CTL), a programming language that has been designed for describing and implementing the operations that occur when images are moved from one color space to another during digital motion picture production. A reference manual for CTL has been submitted to the Society of Motion Picture and Television Engineers (SMPTE) for publication as a Registered Disclosure Document.

In the present document we give an overview of the conceptual framework that has so far been developed by the Image Interchange Framework committee, and we outline how these concepts might be implemented using the OpenEXR file format to exchange images, and CTL programs to exchange color transforms.

At this time the committee's work is not complete, and several aspects of what is described below are not fully resolved. This document does not attempt to define a standard. The goal of this document is to encourage interested parties to experiment with CTL and its use in actual production workflows, and to provide feedback to the author. However, the text below should not be regarded as a draft of what the Image Interchange Framework committee may recommend in the future. The committee's final recommendations may differ significantly from what is described here.

Source code for OpenEXR file I/O libraries, an interpreter for CTL, and CTL-enabled image viewing programs, as well as documentation for OpenEXR and CTL, are available online. For details see the URLs at the end of this document.

Scope

The production of a motion picture typically involves multiple companies and individuals who exchange film and digital images with each other. A good image interchange framework should at least define a common digital image file format, methods for converting images between film and digital representations, and a color management scheme to ensure that the meaning of the pixels in image files is well-defined and that when an image is displayed it looks the same, no matter where it is displayed. In addition, the participants in a movie production have other requirements that are specific to the tasks they perform:

On the Set and on Location

While shooting a motion picture with a digital camera, the director of photography must be able to establish the look of the movie, and to preview this look on a monitor. At the same time, the director of photography must not be forced to make final decisions about the look on the set. It must be possible to revise and change the look, perhaps drastically, during the later stages of the movie production pipeline.

Digital Intermediate Process

During the Digital Intermediate (DI) process images are color-corrected, scaled, and cropped. Simple compositing may also occur. The output of the DI process is the final movie; it looks as desired when viewed on the right display and with proper color rendering. The DI process should be as simple as possible. The pixels in the images are adjusted until they look right, resulting in the finished product.

Visual Effects

Visual effects production combines multiple image elements into composite images. Some elements are computer-generated; others are “practical,” that is, they have been shot with a film or digital camera. A workable interchange framework must support the production and exchange of computer-generated elements as well as combining elements from multiple sources into believable composites.

Visual effects production typically happens before the DI process. The final look of a movie is known only approximately, and it may still change. Destructive color corrections that would restrict what kind of look can be achieved later in the DI process must be avoided.

Animation

An animated motion picture may be entirely generated via 3D computer graphics. The output of the 3D rendering software may be output directly to film or to a Digital Cinema Distribution Master (DCDM), without a separate DI process. This workflow must be supported.

Film Input and Output

A digital motion picture production process must be able to exchange images with a traditional film-based production process. Conversion between film and digital images must be supported.

Video Input and Output

A digital motion picture production process must be able to import and export video images as well as various existing digital image file formats, for example, JPEG.

Digital Archival

When the final digital images for a motion picture are archived, it must be possible to reconstruct their look from the pixel data alone, without referring to color management metadata. This means that there must be a fixed agreed-upon relationship between pixel values in image files and colors on a well-known reference display. (The rationale for this requirement is that archive maintenance procedures such as conversion to a different format may not preserve metadata related to color management. It should be possible to display the images as originally intended even if metadata are lost.)

An Image Interchange Framework

Output-Referred Images

Image interchange would be as simple as possible if the pixels in a digital image file directly represented the colors that are shown on a reference display, for example, on a theater screen. Image files would be self-descriptive, and no metadata related to color reproduction would have to be exchanged. However, two requirements make such an "output-referred" image representation undesirable for motion picture production:

- While it should be possible to make decisions about the "look" of a motion picture early, for example, while recording original images on the set, it should also be possible to revise those decisions later, for example, during the Digital Intermediate process.
- Many movies include visual effects whose production requires 3D computer graphics and 2D compositing. Those operations should be performed on an image representation that allows the mathematical operations employed by computer graphics and compositing to produce correct results.

Motion pictures have distinctive "film look". The images that are shown on a theater screen are not faithful representations of the depicted scenes. An on-screen image typically has a higher contrast and more saturated colors than the original scene. At the same time the dynamic range and the overall brightness of a projected image are usually much lower than in the original scene, and the projected image may have a smaller color gamut. The dynamic range, maximum brightness and color gamut of projected images are limited by the projection technology, but contrast, color saturation, and various other aspects of tone and color reproduction are subject to artistic control, and differ between movies or even between scenes in a single movie.

Processing an image to impart the desired look to it generally loses information. Pixels whose brightness levels or hues are outside the dynamic range or color gamut of the display system are replaced with values that can be displayed. For example, very dark or very bright pixels are replaced with values that correspond to the display's white or black. Information that has been lost cannot be recovered later. In a system where the pixels directly represent the on-screen colors this loss of information makes it difficult to revise the look of an image after an initial look has been baked into the pixels. Such a change tends to significantly degrade the quality of the image.

Another problem with directly storing on-screen colors in the pixels of digital images arises with visual effects: the mathematical operations implemented in 3D computer graphics rendering and 2D compositing packages assume that the pixel values in the images are closely related to light values in the depicted scene. If this assumption turns out to be incorrect, the 3D rendering and 2D compositing calculations do not work correctly, and producing believable visual effects shots becomes more difficult than necessary.

Scene-Referred Images

Instead of an output-referred image representation, one might choose a "scene-referred" representation, where the values stored in the pixels of an image directly represent the amount of light the camera receives from the corresponding objects in the depicted scene. With such a system, the desired on-screen look of an image is stored not in the pixels, but in separate "color rendering" metadata that describe how colors in the image files map to colors on the screen.

In such a system, the look of an image would be determined entirely by the color rendering metadata. Decisions about the look of the images could be revised at any time during the production pipeline, without any loss in image quality, simply by changing the color rendering metadata. In addition, 3D computer graphics and 2D compositing math would work very well because the built-in assumptions about relationships between pixel values and light in the depicted scene are correct.

A system where all images must be scene-referred would not work very well for motion-picture production. The system would limit pixel modifications to operations that maintained the scene-referred state of the images. Discussions in the Image Interchange Framework committee have shown that this restriction quickly leads to conceptual difficulties. Resizing or cropping would clearly be allowed, since those operations maintain the relationship between scene colors and pixel values. Color-balancing an image by multiplying each image channel with a constant value would probably be allowed, too, since the resulting image still depicts a physically plausible scene, although not the one that was actually present on the set or on location. However, other common and necessary operations, such as selectively altering the color of an individual object in the scene, stretch the definition of a scene-referred image to the point where that definition becomes meaningless.

ACES and Indirect-Output-Referred Images

An image interchange framework becomes much more manageable if we back off from the demand that image files are strictly scene-referred. The following describes a system where images live in an "Academy Color Encoding Space" or ACES.

The definition of the ACES is somewhat indirect: ACES is a color encoding space whose dynamic range and color gamut are conceptually unlimited. Pixels can be arbitrarily bright or dark, and they can represent any color. ACES images may or may not be scene-referred. In order to display an ACES image, a "rendering transform" (RT) and then an "output device transform" (ODT) are applied to the pixels. The combined effect of the RT and the ODT is to apply an overall "look" to an image: colors become more saturated, contrast increases, the dynamic range and color gamut are limited. The RT and the ODT are chosen so that applying the transforms to a typical scene-referred image and displaying the results produces an on-screen image that looks pleasing and fairly natural (but as mentioned above, ACES images are not required to be scene-referred).

Essentially, an ACES image is one that looks right after an RT and an ODT have been applied to it. The Image Interchange Framework Committee distinguishes this type of image from scene-referred and output-referred images by calling it “indirect-output-referred.”

With ACES pixel data, the dynamic range and color gamut limiting operations that are part of producing the look of an image still lose information; applying the RT and ODT is not a reversible operation, but the resulting "color-rendered" image exists only temporarily, for display on a screen. This color-rendered image is not saved in an image file.

Since color-rendering is performed on the fly every time an ACES image is displayed, it is typically not necessary to apply information-losing operations to the ACES pixels in order to achieve the desired look for an image. Since the RT and ODT are chosen such that a typical image already looks pleasing, achieving exactly the desired look requires only minor, tweaking of the pixel values. Tweaking is usually non-destructive in the sense that, when the pixels are modified, there exists an inverse modification that, except for rounding errors, restores the original image. Because they are non-destructive, look adjustments made early on can be revised late in a movie's production pipeline, for example, during the DI process.

ACES pixel data can also work well for visual effects production. For visual effects, images should ideally be in a state we will call "scene-linear." (We avoid terms such as "scene-referred" and "focal-plane-referred", whose usual definitions are too narrow for our purposes.)

An image is scene-linear if the values in the pixels are proportional to the amount of light that the camera receives from the corresponding objects in the scene. The pixels in a scene-linear image may represent the actual colors in the scene, but they don't have to. Linear operations such as scaling an image channel by a constant number (as in color timing, white balancing and exposure correction) change the image but maintain the linear relationship between the pixel values and light in the depicted scene.

When digital or film images are imported into the ACES, it is usually desirable that they are initially made as close to scene-linear as possible. Scene-linear images provide the widest latitude for later tweaking, and they are ideal for visual effects work. Color corrections during the DI process tend to make images non-linear and thus less suitable for visual effects. Whenever possible, visual effects should make use of scene-linear images that have not been color-corrected, except for white-balancing and exposure adjustments.

In many cases producing scene-linear ACES images is relatively straightforward:

The tonal response curves of the sensors in digital cameras are inherently linear. Digital cameras should be able to produce ACES images that are scene-linear with fairly good accuracy, although exact scene colorimetry may not necessarily be maintained.

Scene-linear ACES images produced by scanning film negatives are usually less accurate than those produced by digital cameras. Film stocks have known nominal response curves, but the actual curves vary depending on which laboratory develops the film, and on what day. In addition, ACES images are often derived from film scans with the assumption that the film response curves are linear over their entire usable range, ignoring the film's toe and shoulder as well as non-linear interactions between different color layers in the film.

3D computer graphics rendering software internally works in a linear color space, and images generated by 3D rendering, such as scenes for computer-animated movies or elements for visual effects shots, are exactly scene-linear.

In some cases one may not have enough information to produce scene-linear pixel data when an image is imported into the ACES. For example, it may be necessary to scan film negatives that have not been sufficiently characterized, or to import output-referred digital images such as video, JPEG files or scanned film prints.

Overall Pipeline

In a nutshell, we present an image interchange framework that works as follows:

- Original images from digital cameras, scanned film negatives, and 3D computer graphics are imported into the ACES.
- All digital image files contain ACES pixel data. Image processing under user control, such as compositing, color correction or geometric transformation, operates on ACES images and generates new ACES image files.
- Displaying an ACES image on an output device requires color rendering to achieve the desired look.

The following sections describe the framework in more detail. The text introduces a number of new names for color encoding space and for transforms that convert images between those spaces. For a schematic diagram that shows the color spaces and transforms, see Figure 1.

It should be pointed out that implementing the proposed framework is not entirely trivial. However, once implemented, the system should be easy to use. Details, such as what color transforms must be applied to an image and when, can largely be hidden from artists who work with images every day. Most artists will never have to write a color transform or look a transform's source code.

Input

An input device, such as a digital camera, produces image data in a device-specific format, for example, raw samples from a Bayer-mosaic sensor. The data are not necessarily calibrated; two units of the same camera model may produce slightly different output.

An Input Device Calibration Transform, or IDCT, converts the raw image data into a representation called Input Device Color Encoding Space or IDCES. The IDCT is specific to an individual device.

The IDCES is specific to the type of input device. For example, IDCES pixels from a film scanner may contain printing density values, but IDCES pixels from a digital camera will typically contain RGB values related to the spectral sensitivities for the camera model.

IDCES data are calibrated. For each type of input device there is a defined relationship between IDCES and the spectral power associated with the pixel values of the captured image. This relationship is the same for all units of a particular model of input device. Two different units of the same digital camera model may vary in their sensor output, but the IDCTs for the two units compensate for the variation.

An Input Device Transform, or IDT, converts the IDCES pixels to scene-linear ACES pixels. The IDT is specific to a particular input device model, but it is the same for all units of this model.

Color Rendering

A Rendering Transform, or RT, performs color rendering by transforming ACES pixels into an Output Color Encoding space, or OCES. OCES pixel data represent CIE 1931 XYZ colors on the screen of a reference display. Like the ACES, the OCES has an unlimited gamut and dynamic range.

The reference display is the display on which the final movie will be presented to viewers. The reference display may be a screen in a dark theater, illuminated by a digital projector or by a projected film print, or it may be a video monitor. The reference display is indicated by the choice of the Output Device Transform, or ODT (see below).

The RT converts ACES images into OCES images that look pleasing when reproduced on the reference display.

The Reference Rendering Transform

Developing a good rendering transform is not trivial, and most movie productions do not require project-specific rendering transforms. The Image Interchange Framework committee intends to provide a special Reference Rendering Transform, or RRT. The RRT has been carefully designed to make ACES images that are approximately scene-linear look pleasing and “film-like” when displayed on a screen.

In addition to producing a pleasing look, the RRT has been designed to be invertible. Inverting the rendering transform is necessary for importing output-referred material, such as video and JPEG images into the ACES. Invertibility is also required in some cases to create images that are suitable for long-term archiving.

It is expected that most movie production projects will use the RRT for color rendering. However, for some projects, the RRT will be replaced by an alternative RT. For example, 3D computer animation is sometimes done such that displaying the resulting ACES images directly on the reference display yields the intended result. In this case color rendering is not needed and the RT is an identity transform.

Making the RRT the default for most movie productions reduces the chances of displaying images incorrectly if metadata are lost or not handled correctly. In the absence of metadata that indicate an alternative rendering transform, image display software and hardware will color-render images using the RRT, and in most cases this will be the correct choice.

Output

When an image is displayed, OCES pixel colors must be converted into pixels for the actual output device that is used to display the image.

An Output Device Transform, or ODT, transforms OCES pixel data into an Output Device Color Encoding Space, or ODCES. The ODT and the ODCES are suitable for a particular type of output device.

The ODT limits the color gamut by clamping out-of-range ODCES data. The ODT may perform gamut mapping to reduce artifacts that are introduced by limiting the gamut. Taking the capabilities and the viewing environment of the output device into account, the ODT may also try to achieve a closer color appearance match between the output device and the OCES colorimetry.

An Output Device Calibration Transform, or ODCT, transforms the ODCES pixel data into values that can be sent to a specific output device unit. The ODCT performs minor adjustments of the ODCES pixel values in order to compensate for differences between different units. The ODCT is specific to a particular unit.

For example, the ODT for an sRGB monitor produces data suitable for display on an ideal monitor that conforms exactly to the sRGB specification. The ODCT for an actual monitor compensates for the differences between the actual and the ideal monitor.

Specific Input and Output Devices

Digital Camera

This section describes recording images for a motion picture with a hypothetical digital camera that directly supports the image interchange framework described in this document. (See Figure 2) At this time there exists no camera with built-in support for our framework, but such a camera could reasonably be constructed. With existing digital cameras our framework can be supported by switching the camera to "raw" or "data" mode and implementing the functions described below in software that runs on a computer that is connected to the camera.

Our hypothetical camera captures images using a Bayer-pattern mosaic image sensor. The camera's built-in processor applies the IDCT to the raw image data. The IDCT includes demosaicing, conversion to a camera-model-specific scene-linear IDCES, as well as any corrections that may be needed to compensate for deviations of the sensor and analog electronics from the camera's specification.

An IDT, which may be applied by the camera or by supporting software, converts IDCES data to ACES data. The IDT will typically be specific to the scene illumination and any white-balancing filters used on the camera lens. The IDT may be selectable to allow tweaking of the camera color analysis characteristics, although in all cases the output of the IDT must be scene-linear.

The camera or software outputs scene-linear ACES images, which are stored on a recording medium such as tape or disk.

The director of photography (or anyone else who uses the camera) wants to preview the recorded images, and most likely he or she wants to tweak the look of the images, for example, by increasing contrast and color saturation.

Either the digital camera has a built-in preview monitor or it can be connected to an external preview monitor that allows the user to see a color-rendered version of the recorded ACES images. The camera also has a number of user-adjustable controls or "knobs" that, in combination with the RT, affect the overall color rendering or look of the image. (Of course the user-adjustable controls are not necessarily implemented as actual knobs and switches on the camera; they may be user interface elements on an external camera controller unit or in supporting software running on a computer that is connected to the camera.)

The effect of the camera's knobs is implemented in a Look Modification Transform, or LMT, which operates on the ACES images that are output by the IDT. The LMT is parameterized; its behavior depends on the knob settings. The LMT outputs a new ACES image, which is then color-rendered by the RRT. The resulting OCES data are displayed on the preview monitor, using an ODT and an ODCT as described above.

The LMT, RRT, ODT and ODCT are built into the camera or supporting software, and they process the scene-linear ACES images for display on the preview monitor. The preview images are not recorded. The settings of the camera's knobs and an identifier for the LMT are, however, recorded as metadata in the scene-linear ACES images.

The LMT is specific to a particular camera model, but independent of the knob settings with which it is used. The LMT changes only when the camera receives a hardware or software upgrade. Therefore it is not necessary to store the actual LMT along with the ACES images; a unique identifier is enough. The camera identifies the LMT by storing a string such as "Acme_Inc_Model_5_Version_2_20" in the image files, and the camera manufacturer makes a portable representation of the LMT available to users.

Recording scene-linear ACES images together with the knob settings and an LMT identifier allows the intended look of the images to be re-created. At the same time recording images this way directly supports visual effects and other operations that are best performed before a look has been baked into the images.

Since the look has not been baked into the ACES pixels, it can be changed later without compromising image quality. Once a final look has been decided upon and all scene-linear operations have been performed, the look can be baked in, and the LMT can be discarded.

(Note: the hypothetical digital camera described above is similar to the camera-raw based workflow that was described in a presentation at the 2007 HPA Technology Retreat: "Production to Presentation: CineForm Online RAW Post-Production Workflow" by David Newman, CineForm and Mark Pedersen, Pedersen Media Group.)

Film Input

This section describes scanning and importing images that were shot on motion picture negative film.

Film images are digitized by a film scanner, which records raw negative density samples. The relationship between the raw samples and the actual density of the negative depends on the type of film scanner and may vary between different units of the same type of scanner.

The film scanner's IDCT transforms the raw density samples into the IDCES. In the IDCES for film scanners, pixels are represented as calibrated red, green blue printing density triples. When the same negative is scanned on different film scanners, the scanners' IDCTs ensure that nearly identical IDCES data are produced.

The IDT transforms the IDCES data into approximately scene-linear ACES data. Usually a fairly simple IDT is used; it assumes that there is no non-linear interaction between the negative's red-, green- and blue-sensitive layers and that the relationship between density and exposure is a linear function. In some cases one may want to employ an IDT that is based on a more realistic model of the relationship between exposure and density. Such an IDT may produce more accurate scene-linear data by accounting for the film's toe and shoulder and for interactions between layers. However, using a highly accurate IDT is often difficult because it tends to be very sensitive to a film processing lab's day-to-day negative processing variations.

Film Output

Recording ACES images on film can be done two ways:

Traditional Mode

In this scenario images shot on camera original film are scanned and imported into the ACES. After processing the ACES images, for example, to perform visual effects work or to insert subtitles, the images are recorded on intermediate stock. The intermediate stock will be used together with camera original stock in a traditional film-based workflow that includes cutting and splicing, and optical color timing.

Film recording in traditional mode is the reverse of film scanning: The RT is an identity transform, as the rendering is provided by the film print; the ODT is the exact inverse of the IDT; and the film recorder's ODCT converts the printing density ODCES values into data that drive the film recorder.

If an image is imported into ACES and then exported in the above fashion, without first altering the ACES data, then the resulting negative is essentially a duplicate of the camera original.

Previewing ACES images in traditional mode simulates printing and projecting the negative. This requires a helper transform that models the relationship between red, green, blue printing density triples and the on-screen color that results from printing the negative and projecting the print in a theater. This helper transform is called a Print Film Emulation Transform, or PFET. A given PFET is accurate only for a specific print film stock, developed at a particular laboratory.

The RT for previewing ACES images in traditional mode is formed by concatenating the inverse of the IDT with the PFET. This RT may produce colors that cannot be shown correctly because they fall outside the gamut of the preview display, but as the ACES data are not changed by previewing, the recorded film images will still be correct.

Reference Display Mode

In this scenario, images shot on camera original film are scanned and imported into the ACES. While artists work on them, the images are displayed using a video monitor or a digital projector. Color rendering is performed by the RRT or by another rendering transform that does not make use of any particular PFET.

In reference display mode, work on the ACES images is complete when the images look right on the reference display, that is, when they look right with the chosen RT applied. Recording ACES images on film, and projecting a print should result in on-screen images that look close to what is seen on the reference display, subject only to differences between the gamut and dynamic range of the film system versus the reference display.

In reference display mode, the RT for film recording is the same as for previewing, and the ODT for the film recorder is essentially the inverse of the PFET for the print film stock and laboratory that will be used to make release prints. The ODT may not be an exact inverse of the PFET because the color gamut of the reference display typically differs from the gamut of the projected film print, and the ODT may adjust the images to account for the differences.

Video and other Output-Referred Material

From time to time it will be necessary to import output-referred images into the ACES, for example, video, JPEG files or scanned film prints. In this case, it may be neither possible nor desirable to create scene-linear pixel data. Importing output-referred material should result in ACES images that look, when color-rendered with the RT, just like the original video, JPEG image or film print. This can be achieved by using an IDCT that converts from the material's native representation to CIE XYZ, followed by an IDT that is the inverse of the RT. Note that this implies that the RT for any project that uses output-referred material must be invertible.

Digital Projector and DCDM Output

DCDM-encoded images for a digital projector are created by starting with ACES values and applying an RT, followed by an ODT that converts from OCES to the DCDM encoding.

Long-Term Archiving

Metadata related to color rendering should not be part of archival copies of digital motion pictures. In order to achieve this, the rendering transform for archived movies must always be the RRT. The ACES files for any motion picture project that uses the RRT during production can be archived directly. When a motion picture project is produced using a different, project-specific RT, then a separate set of archival ACES files must be produced. This is done by applying first the project-specific RT and then the inverse of the RRT to the original ACES images.

ACES Pixel Data

ACES image files contain pixel data that represent CIE 1931 XYZ tristimulus values. Depending on the on-disk encoding, the files may either contain actual XYZ triples, or any other representation that can be converted to XYZ. (But see the note at the end of the File Formats section, below.)

File Formats

The workflow described above can be implemented using CTL for most of the transforms and OpenEXR as the ACES image file format.

Some of the transforms required for this workflow will be built directly into input and output devices, and no portable CTL representation of those transforms is needed. (In the case of a digital camera with a Bayer-mosaic sensor, the IDCT includes de-mosaicing and thus cannot be implemented in CTL anyway because CTL processes each pixel in complete isolation from all others.)

If multiple companies work on a motion picture project they need to share certain transforms, and those transforms should have CTL representations:

Transforms that are not shared	Transforms that are shared
IDCT for any input device	IDT for film scanner
ODCT for any output device	ODT for film recorder

IDT for digital camera	ODT for RGB display
	RT, including the standardized RRT
	LMT

The image interchange framework described here assumes that only ACES images are saved in image files. IDCES, OCES and ODCES images exist only temporarily and are never saved in files. Thus, the entire image interchange framework needs only two file formats, one for ACES images and one for color transforms.

In current practice, film scanning and recording houses deliver scanned images as DPX files and accept DPX files for recording. At present most scanning and recording houses tend to avoid dealing with image file formats other than DPX. In this framework, they would have to handle ACES files with corresponding transforms to and from density. The assumptions listed in the previous paragraph could be relaxed such that images from a film scanner and OCES images for film recording are saved as DPX files with a standardized and well-defined relationship between printing densities and DPX code values. In this case film scanning and recording houses would continue to handle only DPX files.

The existing OpenEXR image file format allows images with arbitrary RGB color encoding spaces; the encoding space of a given file is specified by a `chromaticities` attribute in the header. Software that reads OpenEXR files is expected to pay attention to this attribute. In some cases it may be difficult to modify existing software or hardware to handle arbitrary primaries correctly. A fairly simple extension to OpenEXR's existing RGB interface could implement "virtual primaries": when an image is opened for reading the application specifies the desired RGB space, and the OpenEXR file I/O library automatically converts the pixels into this space by performing a matrix multiplication. When an image file is opened for writing, the application specifies the RGB space of the pixels; the library stores the pixels and the RGB specification in the file. At this time virtual primaries are not implemented in OpenEXR.

(Note: the Image Interchange Framework committee has not yet decided how pixels in ACES files should be encoded. The possibilities under consideration include parametric RGB spaces, virtual primaries as described in the previous paragraph, as well as a single, fixed ACES RGB space.)

Color Rendering Metadata

ACES image files may be tagged with metadata attributes related to color rendering. The table below gives the names and types of attributes as they appear in OpenEXR file headers. For other image file formats the attributes must be translated to an appropriate representation. The attributes are optional, that is, not every file is required to specify them. If a particular attribute is not present in a given file then software that reads the file should use an appropriate default value, as specified below.

Type and name — description

`Chromaticities chromaticities` — Defines the CIE (x,y) coordinates of the primaries and white point of the image file. The C/C++ definition of type `Chromaticities` is as follows:

```
struct V2f
{
    float x;
    float y;
};

struct Chromaticities
{
    V2f red;
    V2f green;
    V2f blue;
    V2f white;
};
```

The red, green, blue and white fields define the (x,y) coordinates of the RGB triples (v,0,0), (0,v,0), (0,0,v) and (v,v,v), where v is an arbitrary value greater than zero.

If the chromaticities attribute is missing, then the primaries and white point are as specified in Rec. ITU-R BT.709.

V2f adoptedNeutral

Specifies the CIE (x,y) coordinates that should be considered "neutral" during color rendering. Pixels in the image file whose (x,y) value matches the adoptedNeutral value should be mapped to neutral values on the display.

If the adoptedNeutral attribute is missing, then the white value specified in the chromaticies attribute should be used. If both the adoptedNeutral and the chromaticities attribute are missing, then the Rec. ITU-R BT.709 white point should be used.

ImageState imageState

Specifies the image state of the file. The C/C++ enumeration type ImageState has three possible values:

```
enum ImageState
{
    SCENE_LINEAR = 0,
    INDIRECT_OUTPUT_REFERRED = 1,
    OUTPUT_REFERRED = 2
};
```

SCENE_LINEAR means that the pixels in the file are scene-linear. The file contains an ACES image, and displaying it requires color rendering.

INDIRECT_OUTPUT_REFERRED means that the file contains an ACES image whose pixels have been adjusted so that the image looks as intended when displayed with color rendering.

OUTPUT_REFERRED means that the file contains an OCES image whose pixels have already been color-rendered. When the image is displayed, it is necessary to apply an ODT and an ODCT, but the RT must be skipped. (The image interchange framework described here does not make use of OCES image files; OCES images exist only temporarily, as intermediate results when an image is sent to an output device. However, saving an OCES image file may be useful for debugging.)

If the imageState attribute is missing, software that processes the image should assume that the image state is INDIRECT_OUTPUT_REFERRED.

string renderingTransform

Specifies the name of the CTL function that implements the intended RT for this image. If the renderingTransform attribute is missing, the name "transform_RRT" should be used.

string lookModTransform

Specifies the name of the CTL function that implements the intended LMT for this image. If the lookModTransform attribute is missing, then the image has no LMT.

Setting the Image State Attribute

During most stages in the production of a motion picture it is not important to know whether a given ACES image is scene-linear or not. Compositing and 3D computer graphics for visual effects are an exception; visual effects work is ideally done with scene-linear images. The `imageState` attribute listed in the table above distinguishes between scene-linear and other ACES images.

Original ACES images from a digital camera or a 3D renderer, and images produced by scanning film stock with known response curves, are at least approximately scene-linear. The digital camera, rendering software or film scanner should tag the images by setting the `imageState` attribute to `SCENE_LINEAR`.

Images imported from video, JPEG files and other output-referred sources, and images produced by scanning film with unknown response curves, are most likely not scene-linear. The `imageState` attribute in the corresponding ACES files should be set to `INDIRECT_OUTPUT_REFERRED`.

Simple color-timing of an image, that is, adjusting its white-balance and exposure by linearly scaling each channel, as is typically done prior to visual effects work, does not affect scene-linearity. Color-timing software or hardware should preserve the value of the `imageState` attribute.

Creatively color-correcting an image to achieve a desired look, or baking a look modification transform into the pixels of an image, destroys the image's scene-linearity. The resulting image files should be tagged by setting the `imageState` attribute to `INDIRECT_OUTPUT_REFERRED`, or equivalently, by removing the `imageState` attribute.

Color Channels, CTL Inputs and Outputs

In order to make CTL transforms portable between different kinds of application software, the signatures of some transforms in this image interchange framework must be standardized. Below is a list of standard signatures. Transforms may have input and output parameters in addition to those listed. Additional input parameters must have a default value; a transform with additional input parameters must work even if application software does not specify values for those parameters. Additional output parameters must not alter the interpretation of the standard output parameters; application software or later CTL transforms must be able to ignore the additional output parameters.

The function names in the CTL function signatures below are not to be taken literally; they are placeholders for the actual names.

Rendering Transform, ACES to OCES

```
void
transform_name
    (input varying half R,
     input varying half G,
     input varying half B,
     input uniform Chromaticities chromaticities,
     input uniform V2f adoptedNeutral,
     output varying half X_OCES,
     output varying half Y_OCES,
     output varying half Z_OCES)
```

`R`, `G` and `B` define the red, green and blue components of the input pixels. `chromaticities` defines the chromaticities and white point of the RGB color space of the input pixels. The input pixels may or may not be scene-linear.

The `R`, `G` and `B` values are scaled such that "middle gray" corresponds to a pixel with (x,y) coordinates equal to `adoptedNeutral` and a luminance of 0.18. If `adoptedNeutral` is equal to `chromaticities.white`, middle gray corresponds to the RGB triple (0.18, 0.18, 0.18).

`X_OCES`, `Y_OCES` and `Z_OCES` define the CIE X, Y and Z components of the output pixels. The output values are linear with respect to the amount of light on the display. The `X_OCES`, `Y_OCES` and `Z_OCES` values are scaled such that middle gray corresponds to a pixel with Y_OCES equal to 0.18. The range of the output values is unlimited.

Look Modification Transform, ACES to ACES

```
void
transform_name
    (input varying half R,
     input varying half G,
     input varying half B,
     input uniform Chromaticities chromaticities,
     output varying half ROut,
     output varying half ROut,
     output varying half ROut)
```

R, G and B define the red, green and blue components of the input pixels. ROut, GOut and BOut define the red, green and blue components of the output pixels. chromaticities defines the chromaticities and white point of the RGB color space of the input and output pixels.

Typically the LMT will have a number of parameters in addition to those listed above. The additional parameters are specific to a particular digital camera model; their values correspond to the settings of the camera's "knobs."

Output Device Transform, OCES to ODCES for an RGB display

```
void
transform_name
    (input varying half X_OCES,
     input varying half Y_OCES,
     input varying half Z_OCES,
     input uniform Chromaticities displayChromaticities,
     output varying half R_display,
     output varying half G_display,
     output varying half B_display)
```

X_OCES, Y_OCES and Z_OCES define the CIE X, Y and Z components of the OCES input pixels.

R_display, G_display and B_display define the red, green and blue components of the output pixels. displayChromaticities defines the chromaticities and white point of the RGB color space of the display device. The RGB values are linear with respect to the amount of light on the display. Output values must between 0.0 and 1.0, corresponding to the display's minimum and maximum light output.

The gamut of the display is implicitly defined by the displayChromaticities and by the limited range of the output values.

Output Device Transform, OCES to ODCES for negative film

```
void
transform_name
    (input varying half X_OCES,
     input varying half Y_OCES,
     input varying half Z_OCES,
     output varying half DR_film,
     output varying half DG_film,
     output varying half DB_film)
```

X_OCES, Y_OCES and Z_OCES define the CIE X, Y and Z components of the OCES input pixels.

DR_film, DG_film and DB_film define the desired red, green and blue printing density values on the negative. The values are linear with respect to density (negative log of transmittance). They must be between 0.0 and 1023.0, corresponding to densities of 0.0 and 2.046 respectively.

Input Device Transform, IDCES for negative film to ACES

```
void
transform_name
    (input varying half DR_film,
     input varying half DG_film,
     input varying half DB_film,
     input uniform Chromaticities chromaticities,
     output varying half R,
     output varying half G,
     output varying half B)
```

DR_film, DG_film and DB_film define the red, green and blue printing density values on the negative. The values are linear with respect to density. The values are between 0.0 and 1023.0, corresponding to densities of 0.0 and 2.046 respectively.

R, G and B define the red, green and blue components of the ACES output pixels. chromaticities defines the chromaticities and white point of the RGB color space of the output pixels.

Other Transforms

The IDTs for digital cameras, as well as IDCTs and ODCTs, are not typically shared. Transforms that are not shared must interface properly with their predecessors or successors. For example the IDT for a digital camera must output ACES RGB pixels as expected by the input side of the RT. Beyond that, standardization of transforms that are not shared is not necessary, and no CTL implementations of those transforms are required.

Glossary

Transforms

IDCT — Input Device Calibration Transform, converts raw values from an input device such as a camera or film scanner to calibrated IDCES data such that there is a known relationship between IDCES pixel values and light in the recorded scene. The IDCT compensates for differences between the specifications for a particular type of input device and an actual unit. Each unit has its own IDCT.

IDT — Input Device Transform, converts calibrated IDCES data into ACES data. If possible, the ACES data will be scene-linear. The IDT can be shared between different units of the same type of input device.

LMT — Look Modification Transform, modifies ACES data in order to impart a particular look to ACES images. The LMT is used by digital cameras. It allows a camera to record scene-linear ACES as well as the intended look of those images. Since the look is not baked into the images it can be revised later in the production pipeline by altering the LMT.

ODCT — Output Device Calibration Transform, converts calibrated ODCES data into output device values. The ODCT compensates for differences between the specifications for a particular type of device and an actual unit. Each unit has its own ODCT.

ODT — Output Device Transform, converts OCES data into ODCES data for a particular type of output device. The ODT can be shared between different units of the same type of output device.

RRT — Reference Rendering Transform. A default rendering transform that is used for the production of a movie, unless there is a specific reason to use another, project-specific rendering transform. Use of the RRT is mandatory for final ACES data that will be archived.

RT — Rendering Transform, converts ACES pixel data into OCES pixel data.

Color Encoding Spaces

ACES — Academy Color Encoding Space. A color space with conceptually unlimited gamut and dynamic range. ACES images typically start out as scene-linear, but become increasingly non-linear as they are modified by various stages of a movie production pipeline.

IDCES — Input Device Color Encoding Space. A color encoding space that is specific to a particular type of input device. Where possible, the IDCES is chosen such that it is possible to construct an IDT that converts from IDCES data to scene-linear ACES data.

ODCES	Output Device Color Encoding Space. A color encoding space that is specific to a particular type of output device. The ODT for this type of device converts OCES data to ODCES data.
OCES	Output Color Encoding Space. A color space with conceptually unlimited gamut and dynamic range. OCES pixel data describe reference display colorimetry.

Source Code

The C++ source code for the reference implementation of the CTL interpreter can be downloaded from

http://www.oscars.org/council/ctl.html

The CTL interpreter depends on a set of low-level utility libraries, called IlmBase, which can be downloaded from

http://savannah.nongnu.org/projects/openexr

The OpenEXR_Viewers package, available at

http://savannah.nongnu.org/projects/openexr

includes source code for OpenEXR still image and moving image viewers, both with CTL support. In addition to the packages listed above, building the image viewers requires the OpenEXR and OpenEXR_CTL, packages, which are available at

http://savannah.nongnu.org/projects/openexr

and

http://www.oscars.org/council/ctl.html

The source code packages mentioned contain instructions for building and installing the respective libraries and executables.

The "Color Transformation Language User Guide and Reference Manual" is available at

http://ampasctl.sourceforge.net/CtlManual.pdf

Documentation for the OpenEXR file format, the OpenEXR file I/O libraries and the OpenEXR image viewers is available at

http://www.openexr.com/documentation.html

Figures

Figure 1: Color encoding spaces and color transforms

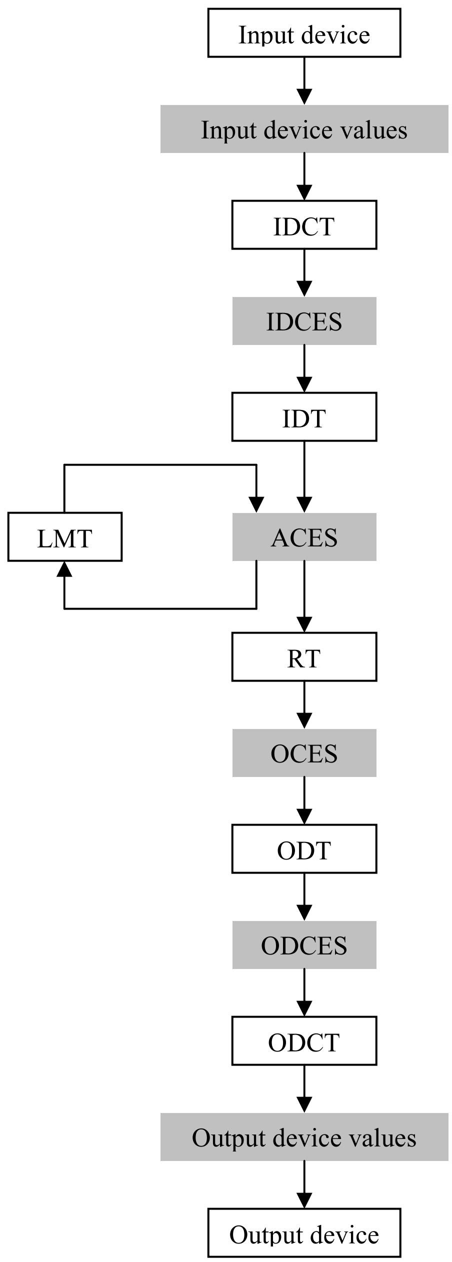

Figure 2: A digital camera that directly supports this image interchange framework