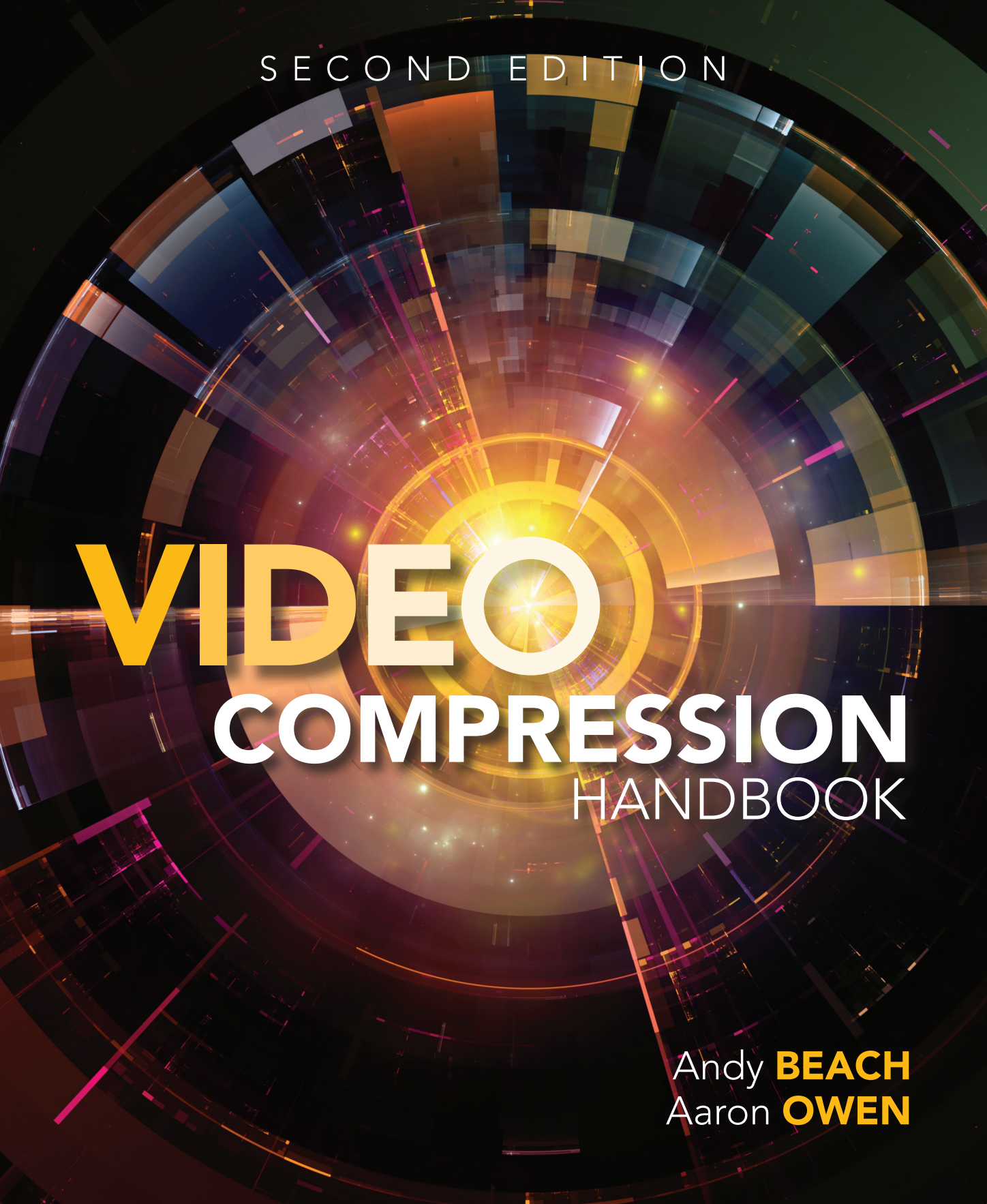
SECOND EDITION
VIDEO
COMPRESSION
HANDBOOK
Andy BEACH
Aaron OWEN

SECOND EDITION

VIDEO COMPRESSION HANDBOOK

Andy **BEACH**
Aaron **OWEN**

I'm really surprised at the influence of YouTube and Facebook and the interest of many to deliver live streaming on their social media networks. I'm also surprised that we live in a day where everyone knows what streaming is and that it's so prevalent around the world. In 2002, I remember having to do a search for streaming video, and you could barely find it on most websites. At that time, NASA and the White House were two of the early adopters of delivering streaming media. Today, it seems that almost every website has streaming media available.

How has video compression changed in the time you've been working with it?

Video compression has changed and become much more advanced. There's also a greater need for compressionists in many companies and organizations. From my view, there are more individuals involved in the video compression and streaming space, and money dedicated to encoding and streaming appears to be more extremely plentiful.

What's the one thing you wished you had known about video compression when you were starting out?

I wish I would have known that video compression and streaming media would become more mainstream. I also wish I would have known that everyday people (nontechnical folks) would be viewing streaming using computers, tablets, mobile devices, and OTT devices, and that large social media companies would be driving the growth of the streaming media marketplace.

What's the next big thing we should be watching in the world of video compression?

I think AV1 is the next big thing in the world of video compression. I'm excited to see what will occur when this new codec is available in encoding tools, in streaming servers, and on devices.

This page intentionally left blank

4 Preprocessing Video and Audio

You have shot and edited your video, and now it's time to compress, encode, and deliver it. But one more step is left before final encoding can begin. This step is called *preprocessing*, and it consists of a variety of optimizations you need to perform on the video and audio before you can hand them off to the encoder. These optimizations can include deinterlacing, inverse telecining, cropping, scaling, aspect ratio adjustments, noise reduction, brightness and color corrections, and audio adjustments.

Preprocessing of some sort is almost always necessary to get your video to look its best. The goals of preprocessing are to clean up any noise in the video and to optimize its basic elements for playback on the devices you are targeting (TVs, mobile phones, computers, and so on). Preprocessing is sometimes referred to as the "magic" part of compression because it takes a lot of practice to achieve the desired results. This is the most artistic and frequently misused part of the compression process. It's easy to go overboard with the changes until you get a feel for how much you really need to do to your video and audio to optimize it for encoding.

Fortunately, preprocessing is a craft that can be learned, and practice will only make you better at it. Understanding why you preprocess video and audio and how the various types of optimization that occur at this stage affect your final product will help you make your preprocessing choices.

When and What to Preprocess

To some degree, every piece of video will need some preprocessing, but the amount is wholly dependent on both the source video and the format you are creating. There are basic things such as scaling and changing the frame rate that will often be required, but color and luma changes should be used judiciously to avoid taking away from the intent of the content. Often, new compressionists will want to pull out all the bells and whistles when preprocessing, and the results may be negligible at best. Add to that the fact that any additional processing you are asking the machine to do to the content will lengthen the overall encoding process.

Remember, the goal of preprocessing is to improve the quality of the video and audio and to maintain the original creative intent. So, it's best to use a light hand most of the time.

Spatial/Geometric Preprocessing

Whether it's simply changing the frame size and aspect ratio to match the destination or reframing the video entirely, spatial preprocessing is one of the most common types of preprocessing you'll encounter as a compressionist.

Cropping: Title and Action Safe Zones

Cropping is a way of identifying a specific region of the video to use in the compression, excluding the other areas of the source frame. Cropping can be used to change the aspect ratio of a wide-screen 16 × 9 video to the older 4 × 3 format, or it can be used to crop out unwanted picture areas at the edge of the frame.

Most TVs do not display the entire image that is transmitted. Instead, they are *overscanned*, meaning slightly larger than the viewable area of a consumer-grade television. This is done for several reasons, most of which culminate in needing to hide irregularities that exist in the edges of the video frames. Production people are aware of this and have created three

regions of a video image that affect how they frame shots and incorporate graphic overlays: overscan, action safe, and title safe. **Figure 4.1** will be familiar to anyone who has looked through a video camera viewfinder. It depicts the action- and title-safe areas that a camera operator needs to be conscious of when framing a shot.

Figure 4.1 *Compressionists need to be aware of the action- and title-safe regions of the frame and use them as a general guideline for cropping nonbroadcast content.*

The outermost region of the video is known as the *overscan* region. This area may not appear on standard consumer TV screens, and it often contains the edge of the set or cables and other equipment. Professional-grade monitors have a mode that allows this overscan area to be viewed, known as *underscan* mode. These monitors may also include action-safe and title-safe indicators.

The *action-safe* area is the larger rectangle within the image area. This area displays approximately 90 percent of the video image and is where camera operators will make sure to keep the primary action framed to keep it viewable on TVs.

The smaller rectangle in the image is the *title-safe* area (comprising about 80 percent of the visible image). It is far enough in from the four edges of a standard TV that text or graphics should show neatly without being cut off or distorted. The title-safe area started out as a guide for keeping text from being distorted by the rounded corners on old cathode ray tube (CRT) TVs. Most modern TVs display a lot more of the area outside of the title-safe zone

than their CRT counterparts. However, as a rule of thumb, text and titles should still be contained within the title-safe area.

Cropping can be a useful tool to ensure that the video displays the same regardless of delivery and playback format. While TVs routinely overscan, mobile devices and online video players do not. If you want the video to appear to display the same on all devices, it may be necessary to crop the overscan area on the formats that do not overscan.

Scaling

Scaling is another key part of the preprocessing process, and it simply means either enlarging or shrinking the frame size. A *proportionate scale* means that the same scaling factor is applied to both the horizontal and vertical axes. Many times you'll crop an image and then scale it up so that the frame size stays the same as it did when you started. Other times, it may mean shrinking a video from the original size to a size more appropriate for delivery (**Figure 4.2**).

Figure 4.2 *These images demonstrate how a 720 × 480 source clip (left) can be scaled down to 320 × 240 (right) for web delivery with no loss in quality.*

Scaling video up is called an *upconvert*. Going from a small frame size to a higher frame size (as shown in **Figure 4.3**) is not recommended but can be necessary in certain situations. In these situations, the different scaling algorithms in different tools can produce dramatically different results. Some tools will simply scale pixels, while others will *interpolate*, or create new pixels. There are lots of options for tools in this area ranging from lower-cost