CS510 Computational Photography – Final Project

*Scott Rubey*

Abstract

This paper presents a method for combining *N* > 1 high dynamic range (HDR) photographic images into a composite low dynamic range (LDR) image, which can subsequently be adjusted for brightness, color saturation and contrast for the purpose of producing an aesthetically pleasing result that can be viewed through an LDR medium. Modern digital sensors are currently unable to capture the dynamic range present in many natural scenes in a single frame; attempting to lighten shadows or darken highlights can lead to unnatural looking photos, higher noise and digital artifacts while failing to recover information in clipped portions of the image. Using the functionality available in the OpenCV image processing library to combine and tone map a number of input photographs covering the entirety of the scene’s dynamic spectrum, one can generate a composite image that maintains detail in bright and dark sections. While the information held in the resulting LDR photograph can be output to a common device such as a computer monitor or printer, it may lack the contrast, clarity and color rendition necessary to be deemed pleasing to some users. In such situations, post-processing controls may be necessary to recover a degree of artistic quality, as determined by the user. This paper shall present results for several photographs processed using the above described methods that would not otherwise contain a reproducible dynamic range nor appear aesthetically pleasing without such interventions.

1: Introduction

Photographers have been developing innovative methodologies for controlling exposure for well over a century. Studio photographers, such as those who specialize in portraiture and product photography, have countless approaches for balancing light across a frame. Soft boxes, reflectors, umbrellas and artificial lighting mechanisms such as flashes have all become commonplace in such environments. Outdoor photographers specializing in landscapes, architectural and sports photography, often find themselves at a disadvantage when it comes to controlling exposure. The juxtaposition of a bright, colorful sunrise behind an unlit rocky foreground, for instance, can be deceptive to a camera’s metering system, at best; at worst, one might not be able to capture the entire scene in a single frame due to the extreme contrast.

The film photographers of yesteryear had an even greater uphill battle than modern digital photographers have today. Color-positive slide films such as Fujichrome Velvia 50 had a dynamic range of merely 4-5 stops, while Kodachrome 64 weighed in closer to 8. Yet, serious nature photographers bought up Velvia en masse due to its crisp sharpness and punchy color rendition. Graduated neutral-density filters were the tool of the trade for managing high-contrast scenes, particularly when shooting directly into the sun. Such “grad filters,” as they were known (aka “grad-ND’s”), consisted of a glass plate, either circular or rectangular; half of the filter was optically clear, while the other half was treated to be several stops darker, yet color-neutral. Photographers could contain extreme contrast by placing the bright section of the frame behind the darker portion of the filter, thus properly exposing both the bright background and dark foreground. Further interventions could then be administered in the darkroom through burning and dodging.

Fast forward to the modern era, and things are a bit less cumbersome. Cameras are lighter weight and optically sharper, and photo development is accomplished without the use of toxic chemicals. Even today’s most advanced digital sensors, however, have trouble capturing greater than 14-stops of dynamic range; while this is a dramatic improvement over film technology, to be sure, it can still present problems given scenes with extreme contrast. Such scenes are easy to spot given a quick glance to the camera’s histogram. If the user sees values peak on the righthand side, it’s a sure sign of clipped highlights; peaking at the left side denotes clipped shadows. (The photograph shown in Figure 1, made on the modern Nikon D800, has both.) Furthermore, LDR output devices such as computer screens and printers are often unable to realistically render HDR lighting characteristics in a way that’s pleasing to the human eye. While some nature photographers still utilize graduated ND’s in their workflow, technological advancements during the last two decades have made it possible to combine numerous exposures to yield a single low-dynamic-range image. By mounting the camera on a sturdy tripod and making a number of bracketed photographs (i.e. 2 or more images of varying shutter speeds that collectively capture the full dynamic range of the scene), the modern photographer now has

A picture containing mountain, man, looking, front

Description automatically generated

Fig. 1 A photograph whose histogram shows clipping in both highlights and shadows.

the ability to create realistic images of high-dynamic-range scenes without juggling a host of expensive optical filters.

The implementation described herein utilizes the technology described above for the purpose of creating realistic high-dynamic-range images, allowing the user to define several key parameters. It additionally allows the user to vary certain aesthetic properties of the photograph upon successful generation of the multi-image composite. This paper briefly describes the process and theory behind HDR imaging and demonstrates the capabilities of this implementation using high-contrast image sets captured by the author.

2: The Process

The source material for this process almost invariably requires a set of two or more photographs made from a tripod-mounted camera and covering a dynamic range of, give or take, three stops at a minimum. (This coverage is highly dependent upon the lighting conditions of the scene, so this should not be taken as a hard-and-fast rule.) The camera must be mounted on a sturdy tripod for two reasons: 1) a handheld camera is unlikely to be steady enough to produce adequately aligned frames, and 2) a handheld camera is likely to yield motion blur in the brightest exposures (i.e. those with the longest shutter speeds).

For the given implementation, the user must provide the program with their set of source images, along with a .txt file containing each image’s name and shutter speed. The images are then aligned using alignment functionality made available through OpenCV. This can be viewed as a pre-processing step and is necessary to account for minor variances in camera position. (While a tripod can greatly reduce human-caused variances in camera position, there may still be environmental phenomena, such as wind, that must be considered, or micro-movements due to the photographer’s adjustment of camera settings.) From here, one can view the output process as composed of two overarching steps: 1) merging, and 2) post-processing. These steps are covered in detail in the following subsections.

2.1: Merging

In order to create a composite image that covers the entirety of the dynamic range present in the scene, we must first capture the camera’s response curve. The “response curve” refers to a mapping of intensity values as a conversion from scene radiance to 1’s and 0’s. Scene radiance and human perception thereof are non-linear, which is why this is approached as a mapping problem rather than a simple linear function. We turn to OpenCV’s “getRespCurve()” functionality to yield the curve information. This function takes as input a set of images, along with their exposure values (in the form of our .txt file, as mentioned above), and produces a composited curve covering the entirety of the dynamic range present in the bracketed source images. A calibration measure is applied prior to output; for the purposes of this project, Debevec’s algorithm was found to produce the most aesthetically pleasing result (cv2.createCalibrateDebevec()).

The HDR image is computed through analysis of the response curve as it applies to a pixel (x, y) across the range of source photographs. By taking the pixel with a radiance value that falls within an acceptable upper/lower bound and placing that pixel in the resulting output image, we yield a composite image consisting of pixels that can then be appropriately mapped to a low-dynamic-range device. This merging of source images takes place via the use of OpenCV’s createMergeMertens() function, which uses Merten’s HDR processing algorithm, and which

A body of water surrounded by trees

Description automatically generatedA large body of water surrounded by trees

Description automatically generated

Fig. 2 (left) shows the HDR output that results from compositing 5 source images using OpenCV’s createMergeMertens() function. Fig. 3 (right) shows the tone-mapped image.

takes our response curve and exposure values as parameters. The resulting composite for one of this project’s test images is shown in Fig. 2.

The final step in the merging phase is to tone-map the HDR image. As the astute reader will notice in Fig. 2, the highlights and shadows in our resulting HDR image still appear clipped. This is because the camera captures source images at a bit-depth beyond what a traditional output device can display; it is at this source bit-depth that our output pixel values are calculated, thus the resulting image contains values that are outside the constraints of most monitors/printers. The purpose of tone-mapping is to take the HDR pixel values and map them into an 8-bit range.

Numerous studies have been conducted on the subject of tone-mapping, and many of these studies resulted in new (and sometimes improved) algorithms. Several of these algorithms are made available to developers in OpenCV; due to resulting aesthetics of the source content, this project made use of Reinhard’s algorithm (cv2.createTonemapReinhard()), which takes four parameters: gamma, intensity, light-adapt and color-adapt. Meaningful differences in output were only witnessed when adjusting the gamma and light-adapt parameters.

“Gamma,” simply put, refers to the overall output exposure of the tone-mapped image. When the parameter holds a value of 0, the image is rendered entirely black. The higher the Gamma value, the brighter the image. The “Light-adapt” naming convention seems a bit of a misnomer, particularly when compared to its neighboring argument “color-adapt.” Higher light-adapt values result in greater color saturation in the output image; lower values, conversely, result in flatter color values. Because differences in these parameter values yielded such varying results in the final output images, it was deemed necessary to allow the user some modicum of control over these values. In this implementation, the user is allowed control over Gamma and Light-adapt values via track sliders; Light-adapt is renamed Vibrance for ease of understanding. Since these values are applied as function parameters for the purpose of creating a low-dynamic-range output image, the user has no way of seeing the results of their adjustments in real time; optimal results (to the user’s aesthetic taste) must be arrived at through experimentation. Testing of numerous source composites produced pleasing results with Gamma set in the 1-3 range and with Vibrance set in the 3-4 range (as scaled to the domain of the slider).

A tone-mapped output image can be seen in Fig. 3; Gamma was set to 2, and Vibrance was set to 4.

2.2: Post-processing

While the resulting tone-mapped image’s pixel values fall well within the abilities of an output device to display/print, it remains somewhat flat in the contrast and color arena. This is not unexpected; the sole purpose of tone-mapping is to produce a lower-contrast image than its source. The human eye, however, often sees what is processed by the brain as pure white and pure black at any given moment in time, be it straight in front or in the periphery of their vision. For this reason, as well as the notion that photography can be viewed as both an art and a science, this implementation allows the user to adjust a number of aesthetic parameters in the low-dynamic-range image prior to final file output, including the ability to add back white and black pixels. This step has become known as “post-processing” in the digital photography community, and the parameters included have long been standard in such applications as Adobe Lightroom and Photoshop. Brightness, Saturation, Whites and Blacks parameters can be adjusted in real time via track slider to the user’s aesthetic taste. These parameters were developed as follows.

“Brightness,” as one might expect, refers to the overall brightness of the final image. Adjusting brightness is a simple task given OpenCV’s convertScaleAbs() function. By simply capturing the slider value and applying it to the “beta” argument, one can brighten or darken a scene to their taste. The algorithm (if one may call it an “algorithm”) adjusts all pixels equally,

A picture containing text, sign, street, city

Description automatically generatedA picture containing text, sign, city, street

Description automatically generated

Fig. 4 shows the effects of the Brightness slider when used to darken the overall image. Fig. 5 shows what happens when Saturation is used to excess. While the overall color saturation of a given image is a matter of aesthetic taste, garish artifacts can be introduced if taken too far. Brightness and Saturation are global parameters that should be used sparingly, as a general rule of thumb.

however, thus must be viewed as more of a sledgehammer than a scalpel. Testing found that Brightness should be used sparingly, if at all. Fig. 4 shows the result of darkening an image using the Brightness slider.

Saturation also benefits from a simple implementation. While there are a number of ways to alter the color saturation of an entire image, this project made use of OpenCV’s functionality for converting to/from the HSV color space. By capturing the slider’s value at any given time and applying it to the S (saturation) channel prior to converting back to the working color space, one can easily add some life into an otherwise dull image. Once again, this parameter impacts the pixel values of the entire image, thus may produce garish results if used to excess (dependent, of course, upon the source content). Fig. 5 shows the effects of Saturation taken to the upper extreme: one can start to see problematic color artifacts in the sky when this parameter is abused.

The Whites and Blacks sliders are where the user can really add some punch to the final image. These sliders can be used to add contrast back into a flat scene. Whereas a global “Contrast” slider may have been of some limited benefit here, the availability of independent sliders for adjusting relative pixel brightness and darkness was deemed far more expressive from an artistic perspective. Since these parameters were intended to alter each pixel’s value individually (versus globally, as with Brightness and Saturation), their implementation was slightly (but only slightly) more complicated. An explanation of the Whites parameter follows.

To begin, we convert the image to L\*a\*b color space, separating the channels so that we can target each pixel’s luminance (L) value. Then, we capture the track slider’s position and add it to the source image’s maximum L value (Lmax). (This may be a positive or negative addition, depending upon whether the slider is moved to the right or to the left.) Third, we normalize the image’s pixel values from Lmin (the source image’s lowest L value) to our newly calculated Lmax. Lastly, we recombine the channels and convert the image back to our working color space. Normalizing the values between our original lower bound and our newly calculated upper bound has the effect of targeting brighter areas of the image while still having some impact on the luminance of darker pixels. Pixel values are altered proportionally to their original luminance. One may view the Blacks slider as the inverse of the Whites slider, thus it requires no further explanation. Figures 6 and 7 are provided to show the effects of the Whites and Blacks sliders. For best results, these are often used in tandem with each other in order to achieve greater contrast and tonal separation.

A picture containing text, sign, people, street

Description automatically generatedA picture containing text, sign, people, man

Description automatically generated

Fig. 6 shows the effect of using the Whites slider, while Fig. 7 shows the use of Blacks. Contrary to adjusting Brightness, which is a global parameter, adding to Whites has only proportional impact on darker pixels. Furthermore, bright sunlight is still present in Fig. 7 despite a heavy-handed adjustment to Blacks.

3: Testing and Results

Testing was conducted using a number of source photographs made by the author. Image sets consisted of bracketed exposures numbering anywhere from two to six at the extremes, with four to five being average. It was found that, while HDR image-making can be accomplished with two bracketed source images, gradients are captured and controlled much better with four or more. The ultimate decision maker in how many input images to use came down to the source material being captured. One photograph, made in a cave entrance in the Belizean jungle (as seen in Fig. 1 and 9), warranted a higher number of source images to capture the full dynamic range of the scene, while the photograph shown in Fig. 6, 7 and 8 could be adequately handled with four. Notably, a higher input image count didn’t always result in more aesthetically pleasing results; while best results for the image in Fig. 6, 7 and 8 occurred with four bracketed source photographs, five resulted in a rather “squashed” look that could not be recovered in the post-processing phase. Due to the unscientific nature of artistic creation, best results are likely to be found through experimentation, and are highly dependent upon the source content.

As mentioned in the previous section, user-adjustable settings in the tone-mapping phase (i.e. Gamma and Vibrance) were best with Gamma on the low- to middle-end of the spectrum, and with Vibrance set from the middle- to upper-end. Best results during post-processing were achieved with liberal use of the Whites and Blacks sliders, which added a great deal of aesthetic “punchiness” in the form of contrast and tonal separation to otherwise dull looking images. (“Liberal use” is a relative term in this context, as the sliders values were attenuated a bit to scale appropriately with other sliders.) Future versions of this implementation might include a threshold feature, so that the user may continue making adjustments to mid-tones without clipping highlights or shadows.

Saturation, on the other hand, was problematic. Being that it was implemented as a “one size fits all” tool (i.e. all pixel values were adjusted equally), adjustments yielded pleasing results in some image areas and garish results in others. Future implementations of this feature will be geared toward targeting adjustments to user-defined color channels so that, for instance, the user may increase saturation in red image elements while leaving blues unaltered. Further problems resulted from attempts at desaturation. A user might reasonably expect a black-and-white image to result from moving the slider to the far left, however in actuality images adopted a strange greenish cast; further testing and experimentation will be necessary in order to determine the cause, however it is the author’s assumption that slider values may have been scaled inappropriately, perhaps coupled with faulty separation of color-space channels.

The Brightness parameter was left virtually untouched in the most aesthetically pleasing test results, though this is clearly an unscientific measurement.

Final results for several test photographs, along with the slider settings used to achieve them, can be viewed in Figures 8 through 11.

Conclusion

This project focused substantially more on functionality than performance. While some optimizations were discovered that allowed for real-time adjustments in post-processing, file sizes above 2000x1333px became cumbersome or unmanageable from a usability standpoint. Moving forward, the author would like to implement other post-processing functions such as targeted color adjustments, localized exposure adjustments, as well as parameters for hue and white-balance; further efforts would also be devoted to improving performance, which may prove challenging given a greater volume of user-defined parameters. Nonetheless, the given implementation performs image merging, tone mapping and its given post-processing functions largely as expected and allows for a surprising degree of artistic flexibility with a handful of simple controls.

A sign above a valley

Description automatically generated

Fig. 8 Bryce Canyon National Park

A picture containing photo

Description automatically generated

Fig. 9 Blue Hole Cave, Belize

A picture containing outdoor, sunset, beach, man

Description automatically generated

Fig 10 Monument Valley

A screenshot of a lake

Description automatically generated

Fig 11 Mount Rainier National Park