

# Xfuse HDR Technology

## High Dynamic Range Video Rendering Pipeline for Digital Cameras and Machine Vision Applications

By Igor Vanyushin, Ph.D. & John Omvik

### 1. Capturing high-contrast scenes

One of the major areas of interest in Digital Image/Video and Machine Vision applications in recent years has been the capturing and rendering of high-contrast and low-light scenes.

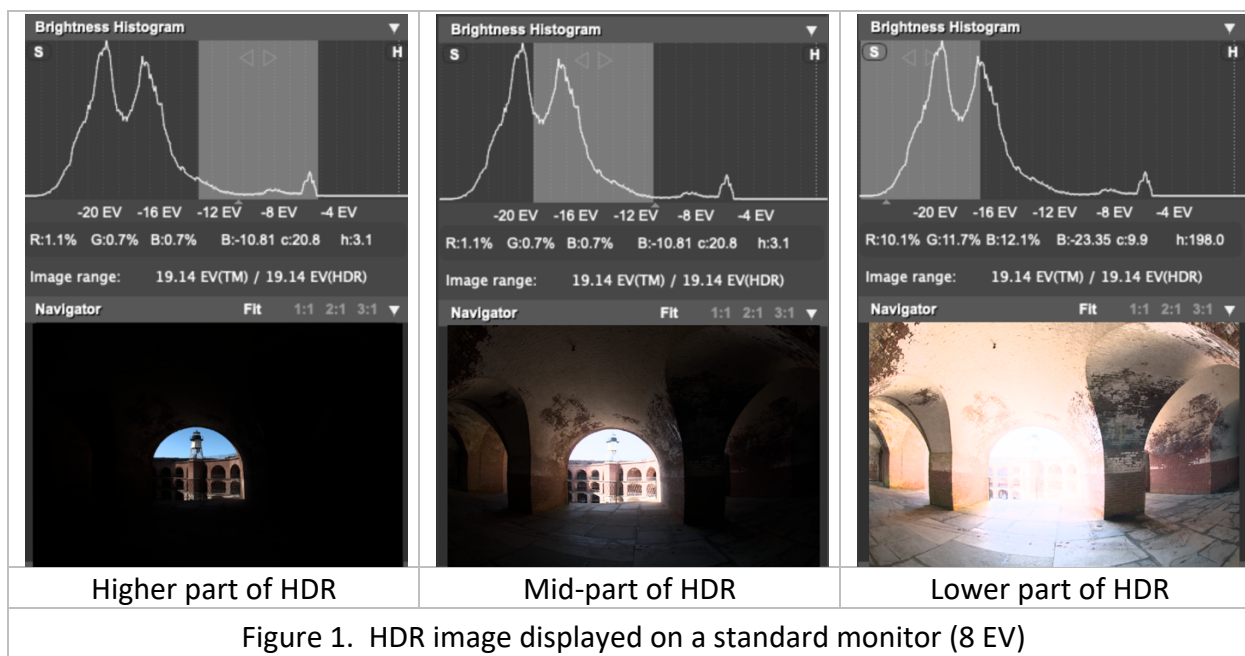
Many new breakthrough image sensors have been introduced that are capable of capturing up to 23 EV/140dB, thus, coming very close to the human vision instantaneous contrast range known to be within 22 EV/132 dB. It is also known that despite the announced dynamic range for such sensors, their characteristics are still far from the true human eye dynamic range. “Real” 24-bit discretization sensors are also rather expensive, especially when they must work at reasonable framerates.

HDR image sensors commercially used in mobile and machine vision applications perform a multi-capture HDR operation. This approach lowers the average sensor’s cost but comes with some obvious drawbacks. The multi-capture operation can also lead to such problems as nonuniform quantization of tones, uneven signal-to-noise ratios (SNR), known as SNR dips, and motion artifacts.

Specific postprocessing is required for these sensors. Non-uniform quantization is usually not noticeable after a tone mapping/compression due to a gamma-like compression curve. But tone mapping does increase the visibility of noise and motion artifacts.

Dual conversion gain, split diode, staggered HDR and similar technologies have decreased motion artifacts significantly, contrary to multiple-frame HDR capturing. When motion artifacts are well suppressed, the processing of HDR images can focus more on noise reduction algorithms and tone mapping (tone compression) techniques. It is worth mentioning, that for such sensors residual motion artifacts and noise are required to be suppressed as well, but this postprocessing can be much simpler and less hardware intensive, compared to multi-frame HDR capturing.

High dynamic range capturing produces images, whose contrast range extends outside a reproduction range of a standard monitor (see the illustrations on the Figure 1).



The images shown on the Figure 1 show the HDR scene displayed on a standard 8-bit monitor (8EV). To display the whole brightness range of the HDR scene in a single image, a Tonal Compression (known as "Tone Mapping") is required.

## 2. Tonal Compression of HDR images (Tone Mapping)

Tonal compression is intended to decrease the bit-depth of an output image (for example, to decrease data transfer bitrates or render the images on low dynamic range devices like displays or printers). Since any tonal compression always leads to image data losses, it is important to preserve all the necessary information from the image.

Main goals of Tone Mapping (TM), or Tonal Compression, can be described as follows:

- The total input contrast range of HDR image should be fully mapped into a shorter range of the image tones
- Image details should be visually preserved; this means a visually unchanged perception of the details:
  - during their compression; there should be no noticeable halos around the details after the compression.
  - when visually compared to the scene captured.
- Colors of the image details should be visually preserved; this means a visually unchanged perception of the colors:
  - during their compression,
  - when visually compared to the scene captured.
- Local and global contrast of the image should be balanced with the details' micro-contrast.
- Noise (and especially — transition noise) should be suppressed or at least not amplified.

- f. Optical distortions, such as Veiling Glare, Haze, aberrations, and others should be corrected.

Unique features of the human vision system allow us to create image capturing, processing and compression techniques, which enable us to reduce the amount of image data to be transferred or stored. Summarizing the above, it can be said that for visual applications the tonal compression should be oriented on the human visual system as well — as the compression increases, the most (visually) meaningful details should be maintained and less distorted, at the expense of the remaining data.

The Tone Mapping algorithm intended for a camera ISP is usually a hardware intensive process for all image details of any size. To simplify the computations, one of the main human vision properties is used — human eye sees the edges of details in the Luminance first, and then the details that are relatively much smaller compared to the image size. So, the process is separated into two parts: Local Processing (details) and Global Processing (remaining total contrast). There will always be a trade-off between the maximal detail size (better quality) and the computation resources available (higher cost).

### Global HDR Processing

The Global HDR Processing (a.k.a. Global HDR Operator) is intended for the whole dynamic range compression of HDR image. In its original purpose, the processing applies a function

$$I_{gm} = F(I_{in}) \quad (1)$$

where an intensity  $I_{in}$  of each pixel at (x,y) position is being mapped into a new value  $I_{gm}$  at the same position; the mapping is being performed regarding the original intensity  $I_{in}$  of a pixel, but disregards intensities of its neighboring pixels. The function always maps intensities from a bigger (HDR) range to a smaller (LDR) range.

The shape of the compression curve  $F(I_{in})$  varies depending on the target application.

For machine vision, the mapping function can vary depending on the intended post-processing (recognition, depth mapping and so on). For example, in a robotic night-vision objects recognition, it can totally suppress headlights of cars, but increase a contrast from low or middle range of intensities. The output data can be more suitable for an object recognition, but unlikely suitable for displaying a natural scene. For human vision applications, the mapping should provide a perceptually natural compression of pixels' intensities.

Among others [1], (including sigmoid-like functions, homography and so on), one of the best classical approaches in the human eye behavior simulation is a usage a gamma-curve

$$F(I_{in}) = I_1 \left( \frac{I_{in}}{I_0} \right)^\gamma \quad (2)$$

Here, the gamma function mimics a local adaptation of a human eye for darker areas of a scene, while an adaptation to the whole scene brightness is maintained by the eye's pupil. Since the dynamic range of the HDR capture system and mapping are hardware limited, the simplest HDR process "capture-mapping", which would simulate the human eye, can be performed by two basic steps:

- a) Capturing: adjusting a global exposure for HDR image capture to provide the brightest pixels of the image to be at the highest possible output intensity (or capture all possible intensities in the range), see Figure 2;
- b) Mapping: compress pixels' intensities of the HDR image to LDR by making all pixels brighter in accordance with the gamma-function (2), except for the brightest pixels in the image, see Figure 3.

Some image sensors can perform both such steps on-chip [2-4].

Drawbacks of a simple mapping (2) are obvious:

- a) The gamma function "knows nothing" about details in the image. According to the gamma function, the details will have less contrast in highlights and higher contrast in shadows.
- b) If there is no chroma preservation, colors will look washed out.

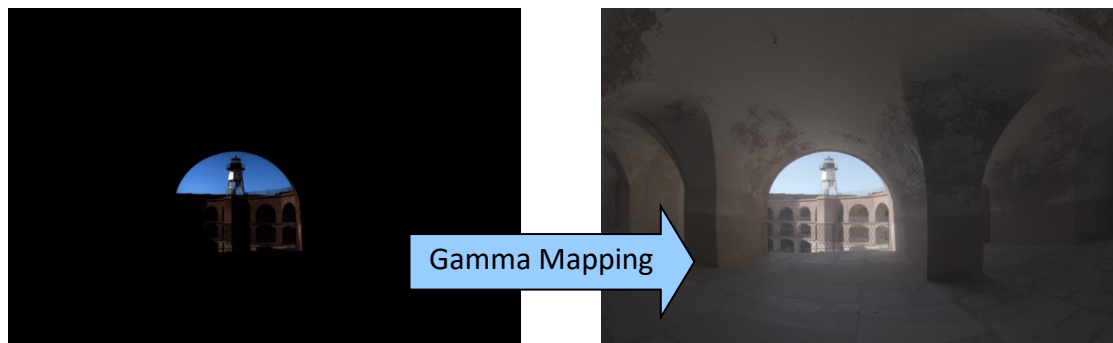


Figure 2. Initial HDR image

Figure 3. Tone Mapping with Gamma

To improve the situation with the insufficient contrast, an additional local modification of the gamma curve can be performed. The "local" term means a partial modification of the curve (2), which allows increasing a mid- or high-level contrast.

The curve modification is obtained from the image statistics, mainly from image histograms.

One of the well-known ways of tone mapping curve modification is a histogram equalization — a method, which tries to increase a global contrast using local or global peaks of the histogram.

Here is an example of a histogram equalization produced by GIMP software (see Figure 4):



Figure 4. HDR image after a Histogram Equalization

Since total histogram equalization can obviously lead to an excessive contrast and losses of details, some tone mapping techniques use just a *local histogram equalization* (LHE) to provide better contrast rendering at the meaningful ranges of intensity (with higher local densities of histogram statistics), as in [5].

Such “meaningful” ranges are represented as distinct peaks of the histogram on the plots shown on the Figure 5.

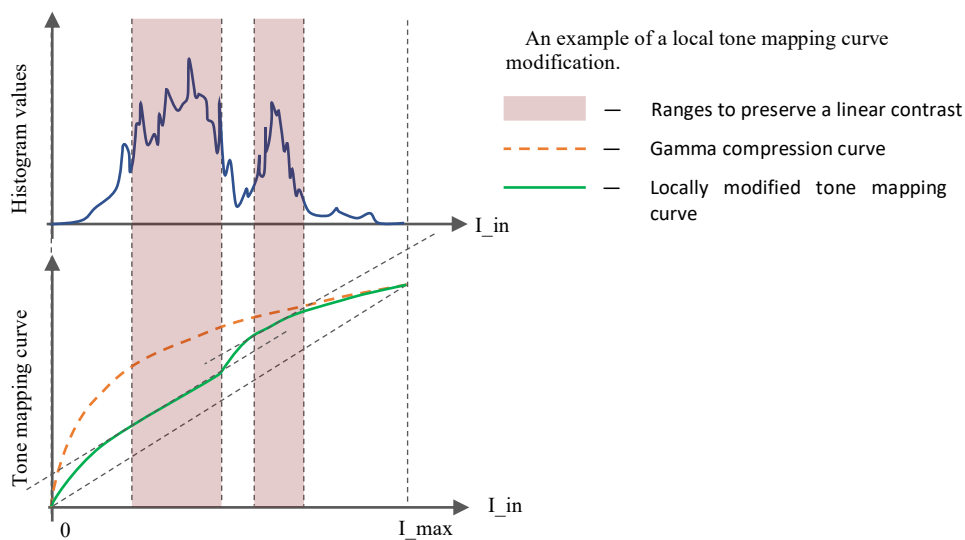


Figure 5. Local modification of a Tone Mapping Curve

The result of LHE Tone Mapping is shown on the Figure 7 vs a simple Gamma, Figure 6.



Figure 6

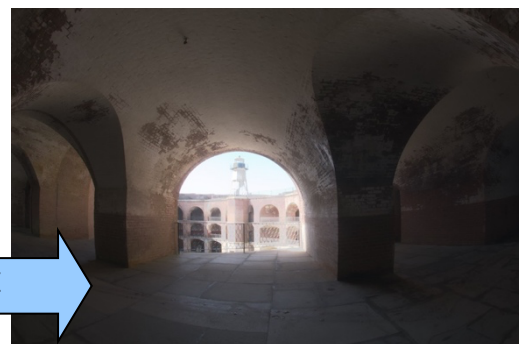


Figure 7

Simple local histogram equalization (LHE) (without a usage of spatial regions of interest), Figure 7, vs. simple gamma tone mapping, Figure 6

As shown here, the result is improvements are minimal, since it is still a single TM curve applied to each pixel independently, regardless its neighborhood. To further improve the quality of LHE, in some techniques the statistics is measured over regions of interest or image blocks [6-8].

The methods of the tone mapping curve modification have limitations implied by the target mapping range — the total sum of local statistical ranges plus compressed ranges should not exaggerate the target LDR.

### Local Processing (a.k.a. Local Operator)

Like any global statistics, any single-curve mapping method has drawbacks, which cannot be avoided, especially for high compression rates. As it was pointed out, tone curves (i.e. histograms) “know nothing” about details of the image, so the local contrast modification is applied regardless details positions in the image. Another drawback is — an inversion, which can appear for high compression rates, when darker areas of the image become brighter than neighboring brighter areas. On the illustrations, a subsequent increase of a contrast for LHE TM is shown: no LHE (Figure 8), LHE mid-contrast (Figure 9), LHE high-contrast (Figure 10).



Figure 8. Gamma TM



Figure 9. LHE: mid-contrast



Figure 10. LHE: high contrast

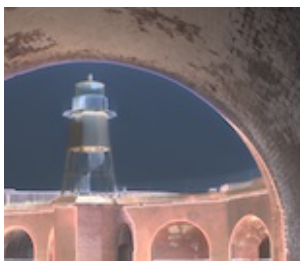


Figure 10a. LHE TM: inversion of contrast

Unlike Global Operator, the Local (spatial) Processing comprises a rendering of image details pixel-wise, depending on each pixel's position. This, for sure, requires more computation resources and their optimization. Nevertheless, when the resources are limited, a combination of both Global and Local operators is used.





Figure 11. Machine-like rendering of HDR image (simulated with GIMP software)

For machine vision applications, it is usually sufficient to preserve image details in any suitable form (Figure 11), while for photo and video applications the visual quality of the preserved details (and their colors) is much more important. This task is not very difficult for 10–12-bit images but required to be more sophisticated for higher bit depths — the higher compression, the better HDR visual rendering of the image details and colors should be. In addition, the

wider dynamic range of a human vision system also comes from the eye's very fast motion, which should also be considered, when naturally looking tone-mapped images are being produced.

There are different techniques, which are targeted to produce as best result as possible for Local Contrast processing. They include a wide range of approaches from a simple Gaussian filtering and Gaussian and Laplacian pyramids to AI and object recognition techniques, as in [9-14].

Main targets of the local contrast processing to keep a visual naturality of details are:

- a) Maintaining sufficient contrast of details along with lowering noises
- b) Avoiding halos and contrast losses by an appropriate compression of the details,
- c) Maintaining visual appearance of colors.

### **Xfuse HDR rendering**

Xfuse's HDR rendering solution was created using a proprietary approach, where the Tone Compression processing balances between lower compression for details (keeping visually meaningful elements of the image) and higher compression of the remaining visual data. Local-depth visual perception model (LDVP) mimics a shading technique used in realistic paintings. It is known, that in many cases paintings are actually tone-mapped images, and the details can be naturally perceived from any reasonable distance to the painting.

In the Xfuse locally adaptive HDR rendering, the local histogram equalization has been replaced with a local contrast preservation (LCP) in the exposure domain (logarithmic representation of pixels' intensities), which works in both Local and Global contrast. As in the classical approach, the LCP separates Global and Local components (layers) of the image. Details are rendered locally, and for the remaining data (global contrast layer) a histogram statistic is being obtained. As it is known, in the logarithmic domain the gamma curve (2) is linear. On the figure below, a linear additive function of the curve is shown. For Global Contrast statistics, local histogram peaks are detected, and the linear mapping curve is modified by decreasing the curve's slope at the peaks of "statistical local contrast". Then, the function is applied additively to each pixel's logarithmic intensity  $\log(I_{in})$ , depending on  $\log(I_{in})$ . As shown on the Figure 12, the peaks will keep their width, thus preserving the

contrast in exposure manner. Illustratively, this is equivalent to an additional illumination of darker areas of the scene with the same intensity of an additional light.

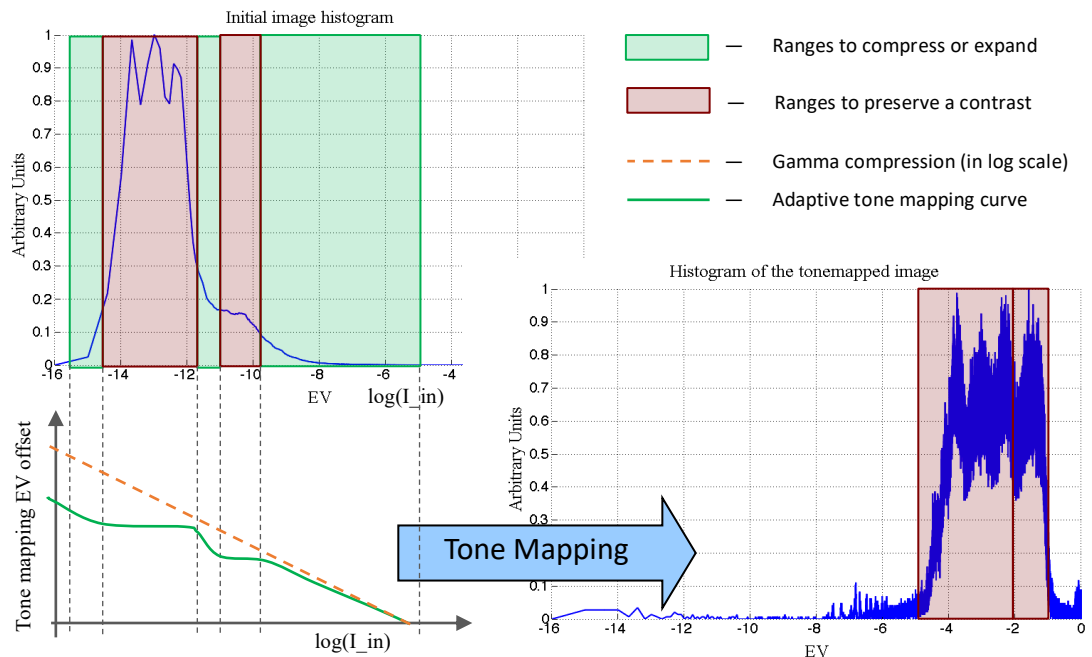


Figure 12. Tonal Curve in EV ( $\text{Log}(I_{in})$ ) scale

In the Local Contrast component, the implemented locally adaptive rendering of image details based on LDVP model has an ability to render details of any size in the image within a single pass (depending on hardware or software resources), while not producing noticeable halo artifacts.

After both components (layers) are rendered, they are combined, thus producing a tone-mapped image (Figure 13).





Figure 13. Xfuse locally tone mapped image

The approach reduces input-to-output delay and creates a scalable design for images of any resolution. An additional benefit of this approach is that noise also reduced (or at least not amplified) after the tone-mapping process.

#### References:

1. Eilertsen, G., Unger, J., & Mantiuk, R. K. (2016). Evaluation of tone mapping operators for HDR video. In *High dynamic range video* (pp. 185-207). Academic Press.
2. Mughal, W., Choubey, B. On Wide Dynamic Range Tone Mapping CMOS Image Sensor. *Sens Imaging* **21**, 33 (2020). <https://doi.org/10.1007/s11220-020-00297-0>
3. Vargas-Sierra, Sonia & Liñan-Cembrano, G. & Rodriguez-Vazquez, Angel. (2015). A 151 dB High Dynamic Range CMOS Image Sensor Chip Architecture With Tone Mapping Compression Embedded In-Pixel. *Sensors Journal*, IEEE. 15. 180-195. 10.1109/JSEN.2014.2340875.
4. <https://www.onsemi.com/pdf/datasheet/ar0331-d.pdf>
5. Boschetti, A., Adami, N., Leonardi, R., & Okuda, M. (2010, July). High dynamic range image tone mapping based on local histogram equalization. In *2010 IEEE International Conference on Multimedia and Expo* (pp. 1130-1135). IEEE.
6. Im, J., & Paik, J. (2014). Spatially adaptive histogram equalization for single image-based ghost-free high dynamic range imaging. *TechArt: Journal of Arts and Imaging Science*, 1(1), 55-59.

7. Im, J., Kim, H., Kim, T., Lee, S., Bae, J., & Paik, J. (2012, May). Ghost-free high-dynamic range imaging using layered exposed images based on local histogram equalization. In *Visual Information Processing XXI* (Vol. 8399, pp. 130-136). SPIE.
8. Rajesh Narasimha, Aziz Umit Batur Texas Instruments Inc, US Patent, [US20140152686A1](#)
9. Jia, W., Song, Z., & Li, Z. (2022). Multi-Scale Exposure Fusion via Content Adaptive Edge-Preserving Smoothing Pyramids. *IEEE Transactions on Consumer Electronics*.
10. Le, C., Yan, J., Fang, Y., & Ma, K. (2021). Perceptually Optimized Deep High-Dynamic-Range Image Tone Mapping. *arXiv preprint arXiv:2109.00180*.
11. Aydin, T. O., Stefanoski, N., Croci, S., Gross, M., & Smolic, A. (2014). Temporally coherent local tone mapping of HDR video. *ACM Transactions on Graphics (TOG)*, 33(6), 1-13.
12. Ahn, H., Keum, B., Kim, D., & Lee, H. S. (2013, January). Adaptive local tone mapping based on retinex for high dynamic range images. In *2013 IEEE International Conference on Consumer Electronics (ICCE)* (pp. 153-156). IEEE.
13. Merianos, I., & Mitianoudis, N. (2019). Multiple-exposure image fusion for HDR image synthesis using learned analysis transformations. *Journal of Imaging*, 5(3), 32.
14. Xu, Y., Wu, X., Wang, J., Dong, H., Wang, Q., Yue, H., & Chen, W. (2021, May). Laplacian Pyramid Based Convolutional Neural Network for Multi-Exposure Fusion. In *2021 33rd Chinese Control and Decision Conference (CCDC)*(pp. 3555-3559). IEEE.