

Gamut Expansion for Video and Image Sets

Hyrum Anderson, Eric K. Garcia and Maya R. Gupta
University of Washington
Dept. of Electrical Engineering
Seattle, WA 98195

Abstract

A semi-automated gamut expansion method is proposed for transforming the colors of video and images to take advantage of extended-gamut displays. In particular, a custom color transformation is learned from an expert's enhancement of a single image on an extended gamut display. This methodology allows for the gamut-expansion to be defined in a contextually appropriate way. From the user-enhanced image, we compare defining the gamut expansion by one linear transformation, or by a multi-dimensional LUT which we learn via local linear regression. We show that using the estimated multi-dimensional LUT with trilinear interpolation (a standard workflow for ICC profiles and color management modules) leads to significantly more pleasant reproduction of skin tones and bright saturated colors.

1 Introduction

Several new advances in display technology have led to increased display color gamuts. For example, a recently adopted multimedia specification called the High-Definition Multimedia Interface (HDMI) version 1.3 includes support for the xvYCC color standard, which can support 1.8 times as many colors as the sRGB color space [5]. Multiprimary displays are commercially available that have color gamuts with more saturated colors than the sRGB color space. Media encoded in sRGB and other standard colorspace will require gamut expansion in order to take advantage of these larger display gamuts.

A simple approach to gamut extension is to map the sRGB extreme primaries to the extended gamut extreme primaries, and linearly transform the color space given these sample color pairs. However, linearly stretching the sRGB colors to better fit an extended gamut leads to oversaturation of skin tones, pastels, and neutrals [9]. Instead, we show in this paper that a more pleasing gamut expansion can be implemented using the flexible multidimensional look-up ta-

ble (LUT) ICC profile architecture, which enables different parts of the colorspace to be enhanced differently. Multi-dimensional LUT ICC profiles are standard for characterizing printers, and have been shown to be a flexible architecture for implementing a wide range of color enhancements [1–3, 10].

We propose a gamut-expansion method in which an expert provides native-to-extended gamut sample color pairs from which a scene-specific color transformation is learned using regularized local linear regression (LLR). The learned color transformation is then applied to each image frame comprising the scene via a multidimensional LUT. More generally, one LUT for a multi-scene video could be learned based on a few different user-enhanced scenes, or one LUT for a set of images could be learned based on a subset of user-enhanced images. This user-enhanced method is a compromise between frame-by-frame user-enhancement and fully-automated methods. It is expected that this semi-automated method will provide a balanced trade-off between expert interaction and color enhancement quality, and provide further insights into automating gamut expansion.

2 Related Work

The objective of gamut-mapping can be summarized by two main motivations: accurate reproduction and pleasant reproduction [9]. Partly to preserve the perception of skin colors, MacDonald et al. have explored methods involving a *core gamut* that remains unchanged by gamut mapping, however explicitly defining such a core gamut is a challenge [8].

Although there has been considerable research in gamut compression [9], little work has been published about color gamut expansion. Kang et al. developed a computer-controlled interactive tool with which observers adjusted color characteristics to make images more pleasant [6]. A global linear fit to the average of the user-defined transformations formed the basis for a gamut expansion algorithm. That approach aims at a universal gamut expansion method, and its broad focus and globally linear model are restrictive.

3 Proposed User-trained Gamut Expansion

We propose to learn a gamut expanding transform based on a single user-enhanced frame. The user begins with an $N \times M$ sRGB image rendered as an sRGB image on the extended gamut display, then enhances the colors to use a greater portion of the extended gamut colorspace (EGC). The result is a user-defined mapping from the original sRGB image colors to EGC image colors that is appropriate to apply to similar images, e.g., in the movie scene.

Manually enhancing a video frame allows the user to enhance the sRGB colors in the context of the image, as opposed to choosing input and output colors without spatial or semantic context. Enhancing the image can be done with standard tools such as Adobe Photoshop. We hypothesize that a user enhancing an image is a faster method than asking a user to gamut-expand a large set of colors. For example, enhancing one 720×480 image results in 345,600 sample color pairs that can be used to train the gamut-enhancement. Morovic discusses the importance of image context to gamut mapping in [9].

The user-enhanced frame and its pre-enhanced counterpart provide NM sample color pairs that define a mapping from sRGB to EGC. We convert all colors to CIELab and all color processing is done in the CIELab space. We form a baseline gamut expansion by fitting the hyperplane that minimizes the squared-error to the NM training sample color pairs.

We also form a multi-dimensional LUT gamut expansion by estimating the gamut-enhanced colors for regularly spaced grid points in CIELab space. Given NM sample color pairs $\{x_i, y_i\}_{i=1, \dots, NM}$ that map the compressed gamut color x_i to the extended gamut color y_i , estimate the extended gamut color \hat{y} for each grid point g of a 3D LUT by fitting a hyperplane to the neighboring sample pairs of g . For each gridpoint in the LUT, *local ridge regression* [4] fits a least-squared error hyperplane to a set of neighbors while penalizing the slope of this hyperplane. This regularization reduces the variance of the estimate and has the effect of smoothing the estimated function. Regularizing using ridge regression has been shown to work well in estimating color transformations [3], as it avoids steep extrapolations that can lead to clipping at the gamut boundaries.

For a given LUT grid point g , local ridge regression is used to estimate each of the three CIELab components of the extended gamut color separately. For example, let \hat{y}_L denote the L^* component of our estimate \hat{y} , and let y_{i_L} denote the L^* component of the i th training extended gamut color y_i . Then the local ridge regression estimate based on a neighborhood \mathcal{N}_g of g is given by $\hat{y}_L = \hat{\beta}^T g + \hat{\beta}_0$, where

$$(\hat{\beta}, \hat{\beta}_0) = \arg \min_{(\beta, \beta_0)} \sum_{i \in \mathcal{N}_g} (y_{i_L} - \beta^T x_i - \beta_0)^2 + \lambda \beta^T \beta.$$

The enhanced grid points form a LUT that can be stored in an ICC profile and used by any standard color management module to enhance the gamut of a new image. Each pixel of an image is enhanced by finding the closest LUT values and interpolating them to estimate an enhanced color for that pixel. Color management modules may use different methods to interpolate the LUT values; in this work we use trilinear interpolation, which is a standard choice in color management [7].

Two factors control the accuracy of the estimated LUT: the size of the neighborhood and the resolution of the LUT. If the neighborhood size used to estimate the enhancement for each grid point is small compared to the density of training sample colors, then some training colors would not be included in the neighborhood of any grid point, and hence would not be taken into account in the estimation of the gamut-expansion LUT. On the other hand, too large a neighborhood relative to the density of training sample colors will fit hyperplanes to large sections of the colorspace, potentially smoothing desired nonlinearities in the gamut expansion. The second factor is the resolution of the LUT. If the desired gamut expansion is highly non-linear, then using a LUT with large spacing between grid points will not make it possible to capture the desired nonlinearities.

4 Experiments

To validate our approach, we designed two experiments to evaluate the possible pleasantness and accuracy of the proposed gamut expansion. In both experiments, we compared the LLR LUT proposed above against the baseline globally-linear transformation. Based on preliminary experiments, we used a neighborhood size of $k = 100$ for the local linear regression and a grid resolution of $21 \times 21 \times 21$. (We found the results to be fairly robust to neighborhood size and grid resolution, but neighborhoods of size $k = 10$ or smaller led to a few objectionably-enhanced pixels in some video sequences due to ill-posed local hyperplane fits.)

In the first experiment, an artist used Adobe Photoshop to enhance the first frame of a video, increasing the saturation of various colors, while preserving the natural color of the skin tones. The enhanced image was used to learn the color transformation and a LUT was generated and applied to all the images in the video. This experiment was applied to the 1080p test sequence *Walking Couple*. We compared the subjective quality of the gamut-expanded images produced by the LUT and linear transform methods; a representative frame is shown in Fig. 1.

In the second experiment, to alleviate the need for artistically enhancing every frame to create ground truth (for which consistency is infeasible), an original video was taken as truth data for an expanded gamut. Compressed

gamut images were generated artificially from the original video sequence by applying a color selective desaturation filter that nonlinearly desaturated all colors except skin tones. Then, the gamut expansion transforms (linear and LUT) were estimated using sample color pairs that map the first gamut-compressed frame to the first original frame. The two estimated gamut-expansion transforms were then applied to all frames in the sequence. The enhanced frames were qualitatively compared to the original sequence by the authors looking for objectionably-enhanced colors, and quantitatively compared to the original sequence by computing the ΔE CIELab metric, $\Delta E = \sqrt{L^*{}^2 + a^*{}^2 + b^*{}^2}$, where roughly 2.5 ΔE is considered “just noticeably different.” This experiment was applied to the 1080p test sequence *Pedestrian Area* and the 720p test sequence *Night*.

5 Results

Results from the artist-enhanced experiment are shown in Fig. 1. The LLR method closely reproduces the greens and yellows of the artistically enhanced image but less accurately reproduces the magenta. The linear method also cannot capture this bright magenta, though it also yields insufficiently vibrant greens and yellows and gives the entire image a slight yellow color cast.

Results from enhancing a gamut-compressed video are shown in Fig. 2 and 3. Fig. 2 demonstrates that LLR preserves skin tones while saturating other colors. The linear method tints the faces towards red and does not reproduce the saturation of the blue shirt in the foreground. Fig. 3 shows a bike that was not present in the frame used to train the LUT. The LLR method reproduces the bright saturated red of the bike where the linear method fails.

Quantitative results from the gamut-compressed images are shown in Fig. 4. The median and 95% worst ΔE CIE errors as a function of image number are shown for both LLR and linear methods in Fig. 2 and 3. Notably, on both sequences, median error of the colors produced by the LLR method are less than one ΔE unit apart from truth data. In *Night*, 95% of the colors are less than five ΔE units for all frames, while in *Pedestrian Area* these are within three ΔE . The jump in the 95% error for the linear method near frame 175 in *Pedestrian Area* corresponds to the appearance of a man in bright red.

6 Discussion

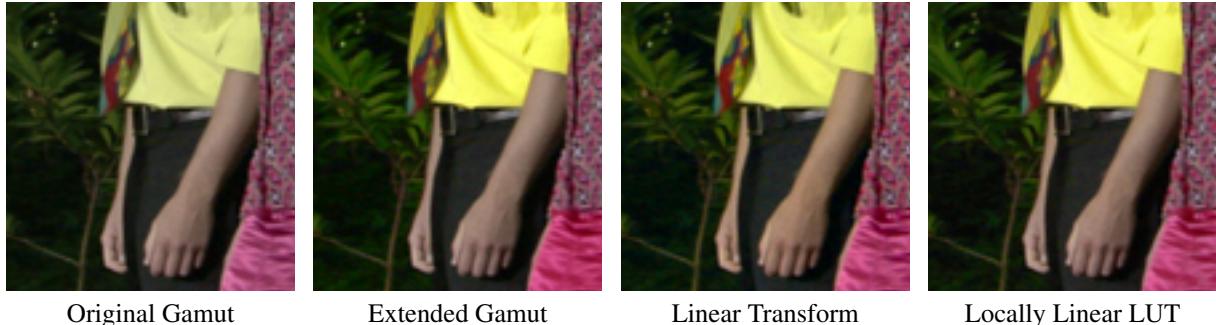
We have demonstrated that a LUT transformation can produce significantly better gamut-expansion than a globally-linear transformation, while still being compatible with ICC standardized color workflows. The proposed semi-automated method may be appropriate for off-line re-

mastering of digital video intended for extended-gamut displays. Video-specific gamut enhancements would allow different movies to best retain their specific color feel. We explored learning one LUT for each scene, and further work is required to determine how broadly a particular LUT can be applied to different scenes within a movie. We hypothesize that one LUT learned from a few different scenes would be suitable to gamut-expand all episodes of a television sitcom, but that preserving the artistic feel of a movie might require multiple LUTs. The proposed semi-automated method could also be used for gamut-compression. For example, a film may be created in an EGC, but an sRGB gamut version may be desired for those viewers with only sRGB displays. We hypothesize that one gamut-expansion ICC profile could be estimated for general use from a diverse set of user-enhanced images.

While we have presented one method (k Nearest Neighbor LLR) for estimating the multidimensional LUT, we would expect other regression methods to also perform well for this application, though some empirical-risk minimization methods might have difficulty processing the large number of training samples.

References

- [1] C. Gatta, S. Vacchi, D. Marini, and A. Rizzi. Proposal for a new method to speed up local color correction algorithms. *Proc. of the SPIE*, 5293:203–212, 2004.
- [2] M. R. Gupta. Learning icc profiles for custom color enhancements with adaptive neighborhoods. *IEEE International Conference on Image Processing*, pages 968–971, 2005.
- [3] M. R. Gupta, E. K. Garcia, and A. Stroilov. Enclosing neighborhoods for local learning with application to color management and color image enhancement. *In review by the IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2006.
- [4] T. Hastie, R. Tibshirani, and J. Friedman. *The Elements of Statistical Learning*. Springer-Verlag, New York, 2001.
- [5] K. Kakinuma. Technology of wide color gamut backlight with light-emitting diode for liquid crystal display television. *Japanese Journal of Applied Physics*, 45, 2006.
- [6] B. Kang, J. Morovic, M. Luo, and M. Cho. Gamut compression and extension algorithms based on observer experimental data. *ETRI Journal*, 25(3):156–170, June 2003.
- [7] H. Kang. *Color Technology for Electronic Imaging Devices*. SPIE Press, United States of America, 1997.
- [8] L. MacDonald, J. Morovic, and K. Xiao. A topographic gamut compression algorithm. *Journal of Imaging Science and Technology*, 46(3), 2002.
- [9] J. Morovic. *Digital Color Imaging Handbook*, chapter 10, pages 639–686. CRC Press, Boca Raton, FL, 2003.
- [10] A. Rizzi, D. Marini, and L. D. Carli. Multilevel Brownian retinex colour correction. *Machine Graphics and Vision*, 2002.



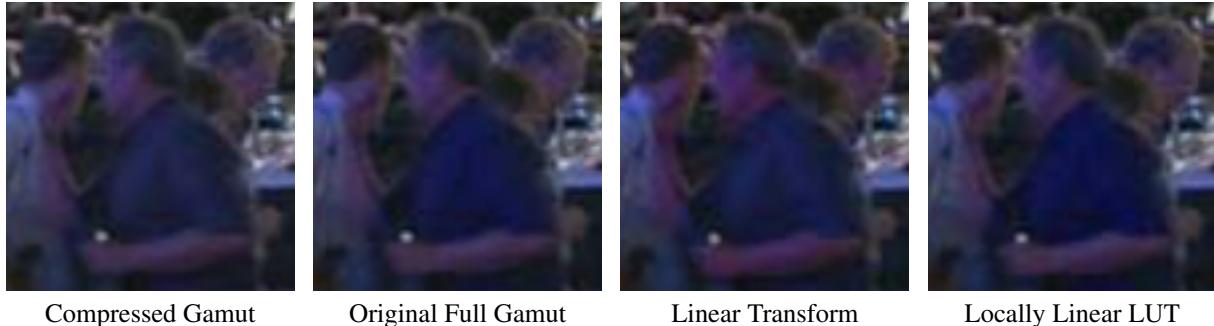
Original Gamut

Extended Gamut

Linear Transform

Locally Linear LUT

Figure 1. Frame 0 from *Walking Couple* for the original frame, the artist's gamut-enhanced frame, and the estimated gamut-enhanced frame using both the linear transformation and the locally linear LUT.



Compressed Gamut

Original Full Gamut

Linear Transform

Locally Linear LUT

Figure 2. Frame 222 from *Night* for a compressed gamut version, the original full gamut frame, and estimates of the full gamut frame using the linear transform and the locally linear LUT. The estimates were trained on Frame 0 of *Night*. A key difference between the estimates is the artificial reddening of the skin tones with the linear transform.



Compressed Gamut

True Full Gamut

Linear Transform

Locally Linear LUT

Figure 3. Frame 96 from *Pedestrian Area* for a compressed gamut version, the original full gamut frame, and estimates of the original full gamut frame using the linear transform and the locally linear LUT. The estimates were trained on Frame 0 of *Pedestrian Area*, the bike pictured here was not present in Frame 0.

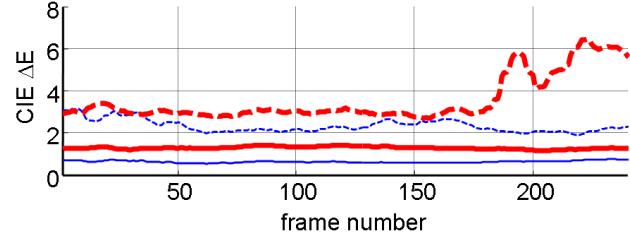
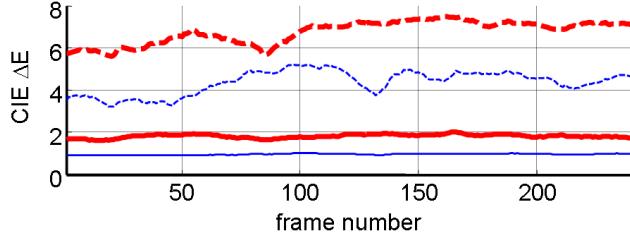


Figure 4. Error plots generated for *Night* (left) and *Pedestrian Area* (right). The light blue lines represent the median (solid) and 95% (dashed) errors for the locally linear LUT and the heavy red lines represent median (solid) and 95% (dashed) errors for the linear transformation.