A Content-Driven Color Adjustment System

Robert Amantea, Jane Asmuth, David P. Bortfeld, Martin H. Czigler
Gerald E. Nostrand and Richard M. Peterson
National Information Display Laboratory at the David Sarnoff Research Center
Princeton, New Jersey

Eugenio Martinez-Uriegas
SRI International, Menlo Park, California

David Zokaities
USGS/EROS Data Center, Sioux Falls, South Dakota

Abstract

The USGS EROS Data Center has the task of taking multi-spectral satellite data and converting it into useful images that are easily interpreted and unambiguous. Information from several spectral bands are combined into red, green, and blue channels to create a pseudo-colored image which often takes on the realistic look of a very high quality photograph. We describe the design of and preliminary results with a software system that will assist in the coloring and upon completion, produce color hard copy that preserves fine detail, hues, and color texture so that the information content of the original image is conserved in the hard copy. The system uses conventional color look up table methodology to match the CRT and film recorder color gamuts. It differs from existing approaches in that it will use a variety of gamut stretching methods for maximizing the use of the film recorder and CRT gamuts. The selection of the gamut stretching method will be controlled by minimizing the amount of change in an image-content measure that is based on the human visual system.

Introduction

USGS EROS Data Center (EDC) provides a variety of types of remotely sensed image data to users with a variety of applications. These applications range from vegetation characterization for land use assessment to environmental impact analysis EDC produces color images from multi-spectral satellite data by assigning individual or combined spectral bands to red, green and blue (R, G, B). The colors are chosen to produce a pleasing image while maintaining the visibility of color texture, i.e., differences in both contrast and hue that allow the user to visualize localized variations in the spectral information. The mapping of the multi-spectral data to R, G, B varies with the spectral bands chosen and the kind of presentation desired. Foliage, for example, may be represented as red or green depending on the client. A typical image of Southern California processed from satellite data by USGS EDC is shown in our poster.

In Fig. 1 we diagram the current process used to produce color transparencies. Multi-spectral data is displayed on a workstation with the desired R, G, B to-multi-spectral-band assignments using the LAS (*Land Analysis System*) software package. Histograms of the gray scale values in each of the three colors are produced by LAS. Using this output and his or her experience, the analyst devises a non-linear mapping between the raw data and the R, G, B representation (stretch functions) that should result in a near-optimal use of the full luminance-color gamut of the transparency film recorder (*MacDonald Dettwiler Color Fire 240*) to produce informative hardcopy that is also a pleasing picture.

A test transparency is made and developed photographically, a process that can take up to two days. If the output is not what is desired, the stretch functions are modified. It may take several iterations through this sequence to obtain a suitable output. In some cases the workstation is used to inquire about local histograms in the image, but, in general, this iterative fine tuning proceeds without further reliance on the computer.

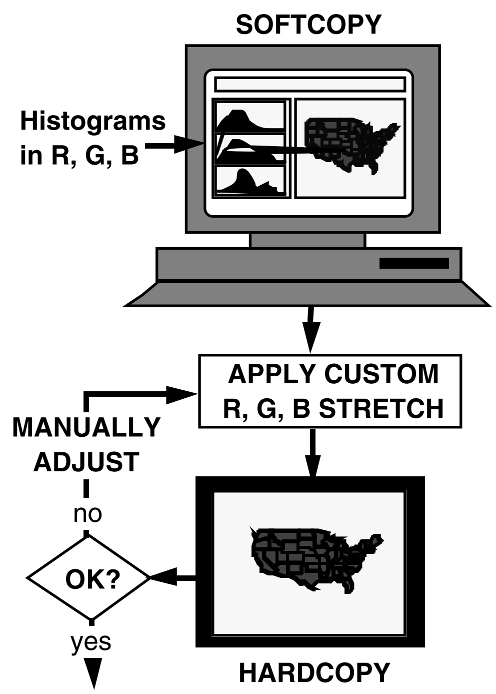

Figure 1. R, G, B histogram adjustment currently relies on manual feedback

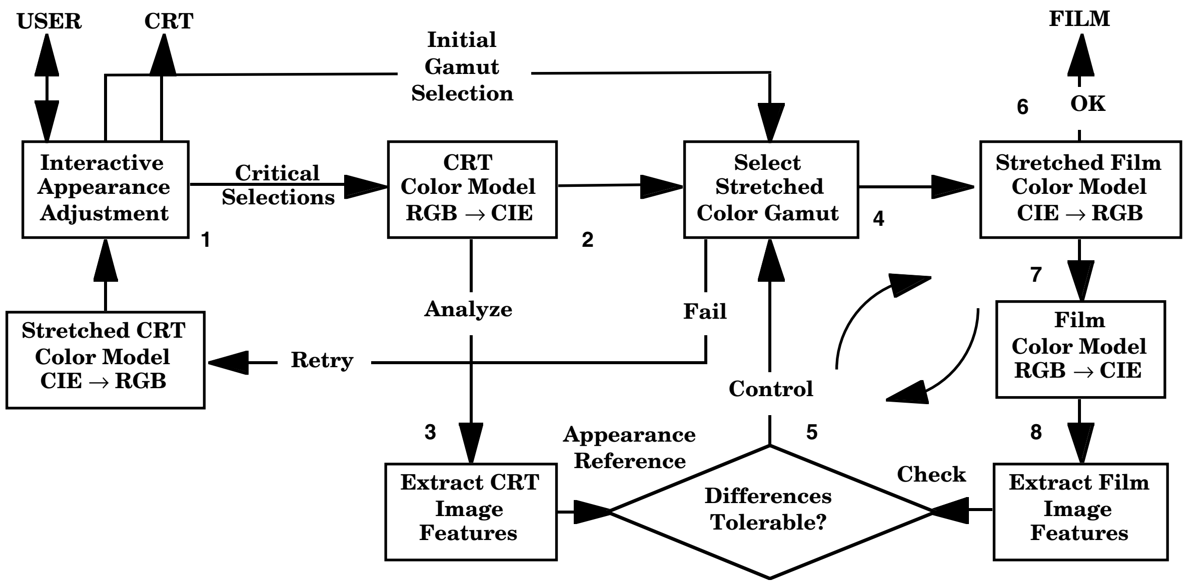

Figure 2. Content-driven color adjustment system

Our work will develop transforms that will provide the analyst with an image on the monitor that shows hue and textural contrast (in both color and luminance) similar to that expected on the output hardcopy device. The goal is to allow the analyst to interactively optimize an image without having to resort to multiple hardcopy iterations. The intent is to produce color presentations on the monitor and the hardcopy, which, while somewhat different, are sufficiently similar to permit the analyst to make the mental translation from the displayed image to the final output with confidence.

Approach

Our approach to optimizing the use of device color gamut while preserving image detail, texture, and colorfulness is summarized by the flow chart shown in Fig. 2. The steps are:

- The user optimizes the coloring of the image on the CRT by adjusting the mapping of the pixel values to colors in each of the red, green, and blue planes (1).
- A CRT device model then simulates the rendering of the mapped RGB values into a perceptual CIE color space, e.g., CIELuv (2).
- Critical image features (to be described later) are extracted (3).
- One of several possible gamut transformations is applied to the data in perceptual color space to map colors in the CRT gamut but not in the film gamut to the film gamut (4).
- The mapped image is then converted from CIE color space into the film RGB color space by applying a film color model which converts either from RGB to CIE or CIE to RGB (6).
- The film color model is then applied to bring the image back into CIE color space (7) and the critical image features are extracted (8).

At this point the image has undergone four transformations: a transformation from CRT RGB to CIE, a gamut mapping, a transformation from CIE to film RGB and a transformation from film RGB to CIE, therefore, we expect some loss to the fidelity of the image.[1] We measure this loss by comparing the features extracted at Step (3) to those extracted at Step (8). We apply a criterion that is empirically derived through the application of psychophysical measurements on the users of the system. If the results are acceptable then the film is presented, if not then a new gamut transformation is selected.

There are four critical parts to this technology: device color models, gamut mapping, feature extraction, and feature comparison.

Device Models

The generation of a device color model proceeds according to the following steps:

1. The RGB device-dependent color space is sampled uniformly and color patches corresponding to each sample are rendered on the device. The CIE xyY chromaticity coordinates of each patch are then measured with a colorimetrer.
2. The chromaticity coordinates are transformed to one of the CIE perceptual color spaces, CIELuv for this example.[2]
3. A three-dimensional cubic polynomial is used to fit the CIELuv coordinates to the RGB coordinates by the least squares method.
4. This polynomial is then used to generate a CIELuv coordinate for each RGB coordinate, thereby forming a color look-up table (CLUT).
5. An inverse CLUT, which maps CIELuv coordinates back into RGB coordinates, is created by sampling and Interpolating from the original CLUT.

The errors between measured data and the device models are:

ERROR	CRT Model	Film Model
max $\|\Delta u^*\|$	1.4397	0.6166
max $\|\Delta v^*\|$	1.8675	0.6969
max $\|\Delta L^*\|$	0.5744	0.3258
max ΔE	1.9257	0.7203
rms ΔE	0.7606	0.3355

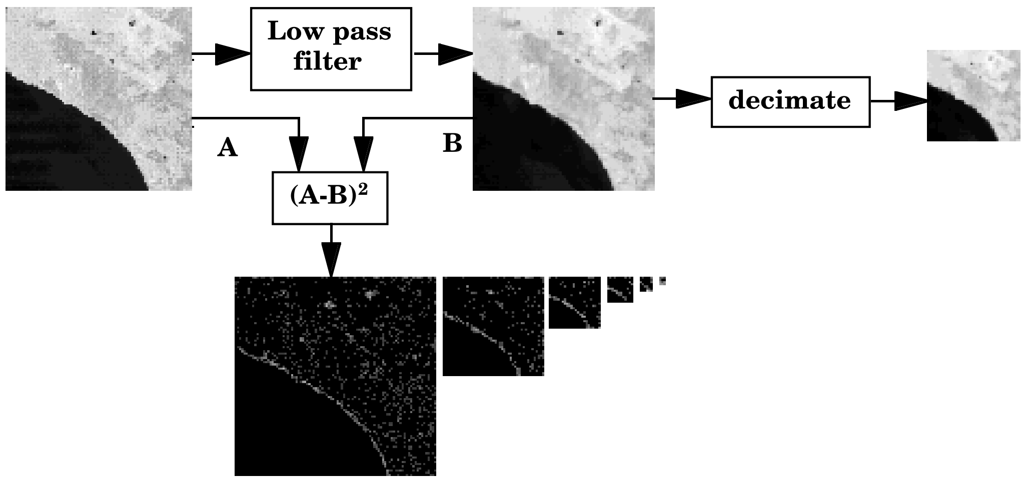

Figure 3. Flow chart of image pyramid creation

Feature Extraction

In common parlance, the term *image features* refers to some characteristic object or region in an image, for example, farm land, city streets, mountains, etc. However, these are really interpretations of the image, we need something more fundamental. We believe there is a relationship between the above conventional features and the spatial frequency content of the perceived image. Rather than develop a complete human vision model, we apply a very simple approximation. We separate the CIELuv planes of the image into octave-related spatial frequency bands. Pyramid technology is used for this separation.[3]

A flow chart for constructing one level of the image pyramid is shown in Fig. 3.

- The image is first low pass filtered.
- The result is then subtracted from the original thereby producing a band-pass filtered version of the original.
- This result is then squared to form the first, highest spatial frequency level of the pyramid.
- The low pass filtered image is then decimated, where every other pixel is removed in both directions, so image size is reduced by a factor of two in each direction.
- The decimated low-pass filtered image becomes the input to the same process to form the next level of the pyramid.

Each level of the pyramid contains the spatial spectral energy in octave related spatial frequency bands. The average energy of each pyramid level can be plotted as a function of frequency as shown in Fig. 4. These points make up a low-dimensional characteristic vector of the selected region (in this case a 6-dimensional vector). This particular analysis is based solely on the achromatic part

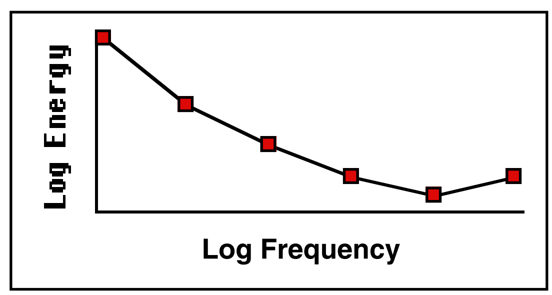

Figure 4. Average spectral energy as a function of frequency

of the image (L* data). If we also apply this to the chromatic planes we generate twelve more numbers to be added to the feature vector.

We believe that if we preserve this vector during gam-ut processing to an acceptable tolerance level, to be determined by psychophysical experiments, (a just tolerable difference – JTD) then we can preserve the image quality. Since these vectors are determined by the user's choice of important regions, we get content-based JTD color gamut matching.

Gamut Transformation

We plan to include several different of gamut mapping paradigms in this system:

- maintain lightness and hue; map saturation along a ray passing through the lightness axis,
- maintain hue; map lightness and saturation along a ray passing through the L*=50 point,
- minimize ΔE (e.g., map along the perpendicular to the color gamut volume boundary).

Each of these methods can be parameterized in terms of the sharpness or strength of the saturation map. As an example consider the constant lightness-constant hue map. There are two simple saturation maps that could be made. Assume that Device 1 has greater available saturation than Device 2.

- We could map linearly from Device 1 to Device 2, e.g., $s_2 = ks_1$
- We could clip from Device 1 to Device 2, e.g.,

$$s_2 = \begin{cases} s_1 & s_1 \le s_{2\max} \\ s_{2\max} & s_1 > s_{2\max} \end{cases}$$

Depending on the input image, either of these may prove to be better than the other.

This behavior can be generalized by developing a parameterized map whose strength varies with a control parameter. Note that in all cases we would like the saturation transformation to be smooth, map gray to gray, and to approach the extremes of the gamut asymptotically. Equation 1 shows this behavior.

$$t(v) = \ln(1 + \exp(v)) - v \qquad (1)$$

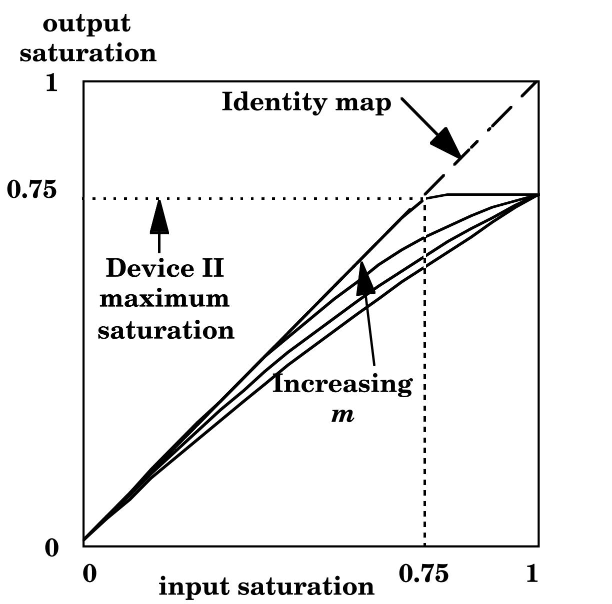

Figure 5. Saturation mapping

For large negative v, t varies linearly with v, however for large positive v, t becomes a constant. Let $v = m(x-k)$, where k is the maximum saturation of Device 2 normalized to that of Device 1, x is the normalized saturation for Device 1 so that x varies from 0 to 1, and m is an arbitrary parameter between 0 and ∞. When m is very small v varies slowly with x so that the transition from linear to constant behavior is gradual. However, when m is large, v varies rapidly with x causing the transition to be abrupt. Incorporating this behavior into a single expression which maps 0 to 0 and full saturation in Device 1 (s_1 = 1) to full saturation in Device 2 ($s_2 = k$) results in

$$g_m(k,x) = k \frac{\ln\left(\dfrac{1+\exp(m(x-k))}{1+\exp(-mk)}\right) - mx}{\ln\left(\dfrac{1+\exp(m(1-k))}{1+\exp(-mk)}\right) - mx} \qquad (2)$$

This expression is shown graphically in Fig. 5 for k = 0.75.

Equation 2 can be used for any of the gamut mapping paradigms. In each case the variable x is interpreted as position along the mapping ray from the gray axis. The inverse of this expression can be used for stretching a gamut outward to attain colors in Device 2 that were not in Device 1.

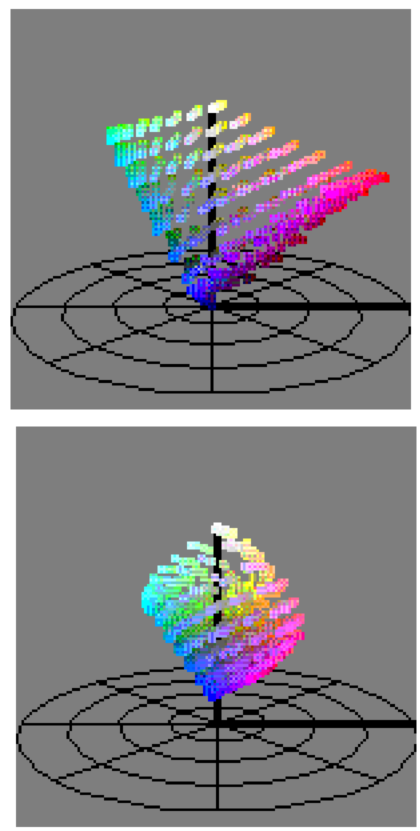

Figure 6. Comparison of the CRT (upper) and film (lower) CIELuv color models

Finally we consider the white and black points of the devices. As can be observed in Fig. 6, the film and CRT gamuts are significantly different. In the above case we have adjusted the CRT white point to 5000K to match that of the film transparency as viewed on a standard light box. The gamma of the CRT was set to 2 which produced a gray scale that closely tracked the gray scale of the film at lightness levels above about L* = 20. Note that there are no dark levels in the film gamut, a particular characteristic of this film.

In testing the constant lightness-constant hue gamut map we found that this mismatch in the device gamuts at low lightness introduced significant mismatch in the rendered images. We are temporarily correcting for this by first mapping the lightness scale from one device to the other.

Conclusion

We have presented a method for a content-driven color adjustment system that preserves color appearance of an image going from a CRT to a film recorder. We have found that a single gamut transformation cannot handle the wide variety of images that occur. By defining image features related to the spatial frequency content of the perceptual image planes we can control which type of gamut transformation is applied and how strong it should be.

References and Footnotes

1. The device color models are lossy in the following sense: If M maps from RGB to CIE then M^{-1} maps from CIE to RGB. Due to discretization effects, $MM^{-1} \neq I$, where I is the identity map.

2. Hunt, R. W. G., *The Reproduction of Color,* Chapt. 8, Fountain Press, Tolworth England, Fourth Edition 1987.
3. Purt, P. J. and Adelson, E. H., "The Laplacian pyramid as a compact image code", *IEEE Trans. Commun.* **COM-31**, (1983), 532-540.