From Dichromatic to Trichromatic Images; Implications for Image Recovery and Visualization

Vlad C. Cardei
Simon Fraser University
School of Computing Science
Burnaby, BC, Canada

Abstract

This paper explores the possibility of recovering a lost color channel from an RGB image, based on the information present in the remaining two color channels, as well as on *a priori* knowledge about the statistics of sensor responses in a given environment. Different regression and neural network recovery methods are compared and the results show that even simple linear techniques suffice to obtain a good approximation of the original color channel.

Remote sensing and other visualization applications can benefit from these methods. Desaturated colors and skin tones are faithfully restored, which is important for the compression of still images and video signals.

Introduction

Missing camera sensors lead to loss of information in acquired images. If the sensor failure is localized to a set of pixels, the image can be restored with a high degree of accuracy, based on local information from the neighboring pixels [1] [2]. However, it can happen that a whole color channel may fail, in which case the image cannot be recovered accurately. From a visualization perspective, the issue in this extreme case is to minimize the perceptual distortions caused by the missing sensor. Losing all information present in a color channel of an RGB image strongly degrades the image quality and gives the image a strong color cast (e.g. if **G**, the green channel is lost, the image will look purple because of the remaining **R** and **B** components). In the RGB color space, the image gamut is located in a plane defined by the remaining two color channels.

The loss of a whole color channel could be caused by sensor failure, an image transmission problem or because the camera is dichromatic in the first place. As shown in [3], the Mars Orbiter Camera (MOC) has only two narrow-band sensors, in red and blue wavelengths. To visualize images acquired with this camera, the green channel is synthesized by averaging the red and blue channels. In this case, the recovered colors are used only for visualization purposes; they do not correspond to the 'true' colors that can be seen on Mars, not only because of the uncertainty introduced by synthesizing the green channel, but also because the camera was not calibrated for human observers. The issue of image reproduction [4][5] introduced by the relationship between camera sensor responses and visual perception (as partially determined by the eye sensitivity curves) will not be addressed in this paper.

In the general case, where statistics about the likelihood of surfaces and illuminants is not known, synthesizing the lost channel by simple averaging is the obvious choice. However, if one takes into account *a priori* knowledge about the scenes that are being taken with the camera, there are better methods than simple averaging. Some of these methods, based on knowledge about the statistics of 'the world' (defined as the environment in which the camera is used) will be explored in the following sections.

The uncertainty in the recovery of the lost channel is also reduced due to the fact that the set of all possible camera responses is limited. This happens because the RGB values produced by the camera are correlated, due to overlapping and broadband sensors (as encountered in most CCD cameras) and to the low dimensionality of the received color signal.

The RGB camera responses are obtained by integrating the color signal with the camera sensor sensitivity functions.

The color signal C is defined as the product between the spectrum of the illumination I(λ) and the spectral reflectance S(λ):

$$C(\lambda) = I(\lambda) \cdot S(\lambda) \qquad (1)$$

Previous studies have shown that common light sources, such as daylights [6], and surface reflectances [7] can be modeled with three dimensional basis functions. Thus, RGB triplets like (255,0,0) are highly unlikely.

Recovery Methods

To simplify the presentation, the following methods will assume that the green channel is to be recovered, based on the remaining red and blue channels. In case of differences in the recovery method, based on which color channel is missing, all three recovery cases will be discussed.

Averaging

This method works in the absence of any knowledge about the world. If the green channel is lost, it is computed by averaging the two remaining channels; if either the red or blue channels are to be recovered, they are linearly extrapolated from the other two, as shown in Figure 1.

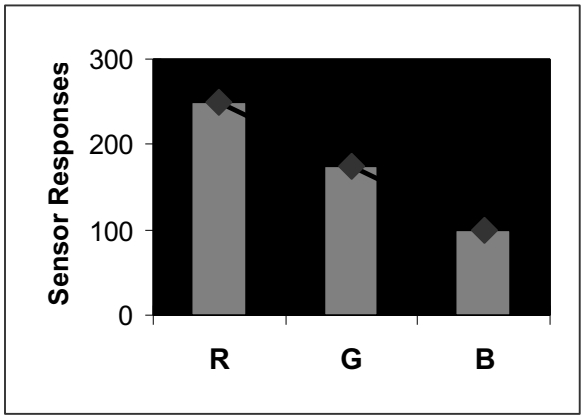

Figure 1. Color channel recovery by averaging

Linear Regression

This least-square regression maps the R and B color channels into G, such that it minimizes the residual squared error for the calibration data. The goal is to find a transformation matrix M, such that:

$$G \cong \begin{bmatrix} R & B \end{bmatrix} \cdot M \qquad (2)$$

where R,G and B are column vectors containing the camera responses for the respective color channels. In this case, M is given by:

$$M = \begin{bmatrix} R & B \end{bmatrix}^{*} \cdot G \qquad (3)$$

where the * operator denotes the pseudo-inverse. In fact, M is a 2x1 matrix that assigns global weights to the R and B vectors. M is computed on a calibration data set and is then tested on a different data set.

Polynomial Regression

This method adds polynomial terms to the calibration data. For the experiments shown in the next sections, the transformation matrix M is computed as described above, such that it minimizes the least-square error of the following equation:

$$G \cong \begin{bmatrix} R & G & R^2 & G^2 & R \cdot G \end{bmatrix} \cdot M \qquad (4)$$

A major disadvantage of this method is that, due to its non-linear form, it is sensitive to the magnitude of the pixel brightness. In the case of linear regression, the lost color is computed as a weighted average between the remaining two colors. Thus the chromaticity of a recovered RGB pixel does not depend on its brightness. Polynomial regression, on the other hand, introduces non-linear terms and thus makes the recovery dependent on the absolute pixel brightness.

Neural Networks

The neural network method was implemented in order to capture possible non-linear complex relationships between R, B and the lost G channel. The neural network used in the experiments is a Perceptron [8] with two hidden layers, as shown in Figure 2.

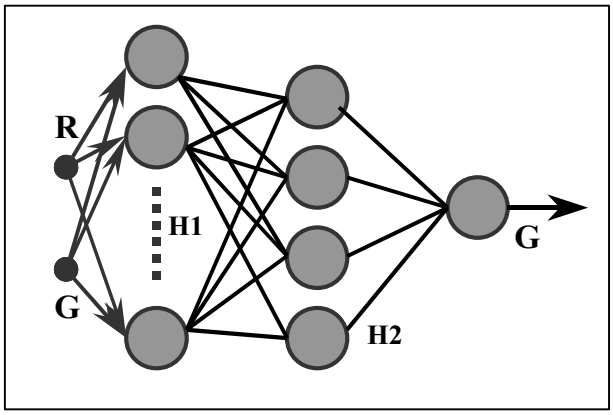

Figure 2. Multi-layer Perceptron

The two inputs to the network correspond to the red and blue channel and the output corresponds to the green channel. The first hidden layer contains nine neurons and the second one contains four. In a first phase, the neural network is trained by Backpropagation [8] to estimate the green channel from the red and blue inputs. The training set contains a large set of RGB values corresponding to normal camera outputs. After training, the network is tested on a different data set. It is presented with the R and B values and it produces an estimate of G. For a particular RGB triplet, the estimation error is defined as the absolute value of the difference between the network estimate and the correct answer. The average estimation error on a given test set is taken as a measure of the neural network performance.

Experiments

Experimental Setup

The recovery methods described above were tested on synthesized data as well as on real images. Each experiment consists of a training (or calibration) phase, during which the regression matrices and neural network parameters are computed, and a test phase, during which each method is tested on a different data set. All methods are compared against each other and against a 'do-nothing' method. This is in fact a pseudo-method: it assumes that the estimate of the recovered channel is actually zero. Thus, the average error for this method is equal to the average brightness of

that particular lost channel. This provides a comparison of the other methods with a worst case scenario, where nothing is done to improve the image quality.

All RGB values are scaled in the range of 0 to 1. The errors computed in the RGB color space, are in the same range. To provide a correlation with the magnitude of errors perceived by a human observer, the errors are also reported in the perceptually uniform CIE Lab space. The conversion from the RGB space to the Lab space is done in two steps. The first step is done within the sRGB framework [9]: it is assumed that all images are in sRGB format and that they are displayed on a sRGB compliant monitor. sRGB values are converted into CIE XYZ coordinates. The second step involves the transformation from the XYZ space to Lab.

Thus the errors in Lab space are relative to what a human observer sees on a standard sRGB monitor. It must be noticed that the algorithms were calibrated and thus optimized for the RGB color space and not for the Lab space. Because the transformation from the RGB to the Lab color space is non-linear, equal errors in the RGB space will yield errors with different magnitudes in the Lab space.

Experiments Done on Synthesized Data

Working with synthetic data has the advantage of a perfectly controlled environment. Moreover, an arbitrarily large number of images can be generated, such that, when testing the algorithms, the results are stable.

This experimental setup illustrates the performance of the recovery methods under ideal conditions.

For any RGB triplet, its values for the three color channels are computed from a randomly selected surface reflectance S_j, the spectral distribution of an illuminant E_k and by the spectral sensitivities of camera sensors for that color channel ρ^G, according to the following equation:

$$G = \sum_i E_i^k \cdot S_i^j \cdot \rho_i^G \qquad (5)$$

The index i is over the wavelength domain, in the range of 380nm to 780nm. Two identical experiments were performed with sensor sensitivities functions from two digital cameras: SONY DXC-930 and Kodak DCS-460. For each experiment, the data was generated from 34 illuminant power spectra, 260 surface reflectances and a set of sensors.

The whole data set, composed of 8,840 RGB triplets was divided into two equal parts, one used for training and one for testing. The recovery methods were calibrated and tested for each set of sensors, for all of the R, G and B channels. The following two tables present the recovery errors obtained on the test sets, for both sets of sensors. The average errors in the RGB and Lab space are μ_{RGB} and μ_{Lab}.

Table 1. Recovery Errors for the Kodak DCS Sensors

Lost Channel:	**R**		**G**		**B**	
Method	μ_{RGB}	μ_{Lab}	μ_{RGB}	μ_{Lab}	μ_{RGB}	μ_{Lab}
Do-nothing	0.359	50.83	0.334	128.3	0.269	70.09
Average	0.107	13.93	0.076	21.16	0.108	17.79
Linear Regression	0.104	11.28	0.073	20.60	0.060	11.28
Polynomial Regression	0.097	13.49	0.065	18.48	0.060	10.93
Neural Network	0.097	13.49	0.069	20.29	0.058	10.94

Table 2. Recovery Errors for the SONY DXC Sensors

Lost Channel:	**R**		**G**		**B**	
Method	μ_{RGB}	μ_{Lab}	μ_{RGB}	μ_{Lab}	μ_{RGB}	μ_{Lab}
Do-nothing	0.241	76.74	0.184	104.4	0.241	76.74
Average	0.159	32.41	0.073	30.92	0.159	36.55
Linear Regression	0.145	32.41	0.053	24.29	0.145	43.05
Polynomial Regression	0.144	32.99	0.052	23.96	0.138	34.54
Neural Network	0.138	31.30	0.052	23.80	0.137	31.48

The results, in RGB as well as in Lab space, are better for the Kodak than for the SONY sensors. This is due to the fact that the Kodak sensors are broad and overlapping while the SONY sensors are quite sharp. Because of this, the SONY camera can produce a larger gamut of RGB values, which in turn increases the uncertainty associated with the recovery of a particular color channel.

Including implicit knowledge about possible RGB values improved the recovery methods that were calibrated, relative to the simple averaging method. However, since all surfaces and illuminants were selected at random and thus they are equally probable, the improvement the other methods over averaging was not large, with the notable exception of the B recovery for the Kodak sensors (see Table 1).

Experiments done on Real Images

Experimenting on real data has the disadvantages that the number of images is not very large and that image artifacts, such as noise or flare, corrupt the sensor responses. On the other hand, working on real data permits the extraction of statistical information about the distribution of RGB triplets in the world.

In this experiment, the recovery methods were calibrated on RGBs taken from 46 images that were acquired with a Kodak DCS-460 camera. The test was done on a different set of RGBs taken from other 42 images, taken with the same camera. All images were linearized, to compensate for the built-in gamma correction of the camera and were resampled, to reduce the noise.

A total of 118,223 RGBs were used for calibration and 96,688 RGBs were used for testing. The recovery methods were also compared against the neural network trained on synthetic data, generated with the Kodak sensors. The results are presented in Table 3.

Table 3. Recovery Errors for Real Data

Lost Channel:	R		G		B	
Method	μ_{RGB}	μ_{Lab}	μ_{RGB}	μ_{Lab}	μ_{RGB}	μ_{Lab}
Do-nothing	0.212	28.73	0.209	68.04	0.205	54.58
Average	0.082	9.54	0.042	9.91	0.074	12.72
Linear Regression	0.071	8.81	0.042	9.72	0.057	10.05
Polynomial Regression	0.069	8.37	0.042	9.95	0.055	9.71
Synthetic Network	0.084	10.54	0.055	12.52	0.069	12.32
Real Network	0.073	8.76	0.051	11.53	0.056	9.86

All errors are smaller than the ones reported for tests done on synthetic data. One reason for this is that image recovery is more accurate on desaturated RGBs and, statistically, this type of RGB triplets is more frequent than saturated RGBs in real images. Figure 4 shows the relationship between recovery error and pixel saturation. The average recovery error is plotted against the pixel saturation, defined as the distance from white in the rg-chromaticity space. The rg-chromaticity space is a normalized RGB space, in which the coordinates are: r=R/(R+G+B) and g=G/(R+G+B).

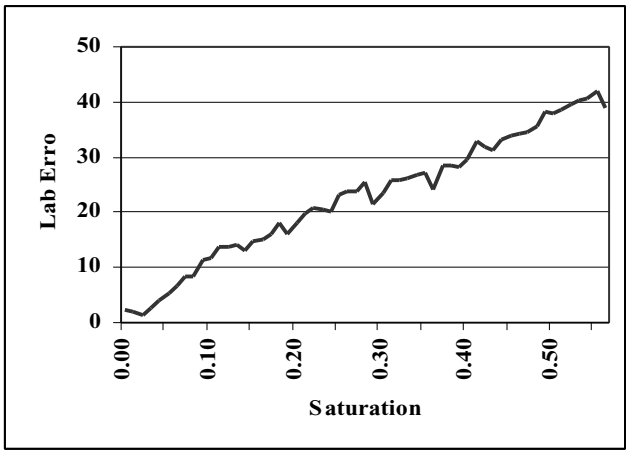

Figure 3. Recovery error as a function of saturation

Moreover, the recovery methods that take into account the statistics of the world perform better than the simple averaging method. The neural network that was trained on synthetic data performed worse than the other recovery methods because it was trained on a different, equally probable distribution of RGB triplets.

Experiments of Color Recovery of Face Images

As noted before, desaturated colors (such as skin, for example) are recovered with more accuracy than saturated ones. Based on this observation, a separate test was done on images containing only human faces. Eight images, taken with a DCS camera under different illuminants, were tested using the methods calibrated for the previous test (i.e. on real data). Table 4 shows the results of the recovery.

Table 4. Recovery Errors for Human Faces

Lost Channel:	R		G		B	
Method	μ_{RGB}	μ_{Lab}	μ_{RGB}	μ_{Lab}	μ_{RGB}	μ_{Lab}
Do-nothing	0.751	51.53	0.559	92.67	0.505	75.11
Linear Regression	0.154	8.16	0.05	6.14	0.053	4.85

Implications for Image Compression

As shown in the previous experiment, the appearance of human faces and skin color in particular is faithfully recovered. This aspect could have implications for the transmission of still or video images (such as video-conferencing) in which the fidelity of the color tone of human faces is considered to be more important than the rest of the image, and in which spatial resolution is also important. Temporary bandwidth limitations can be partially compensated by transmitting dichromatic images instead of trichromatic ones.

The following chart shows relative file sizes obtained by JPG compression of images with all three RGB channels present and images where one of the color channels (G in this experiment) was set to zero:

$$Ratio = \frac{Size\ of\ dichromatic\ image}{Size\ of\ original\ image} \qquad (6)$$

The compression ratios are expressed as a function of the quality factor. A quality factor of 100 corresponds to loss-less compression.

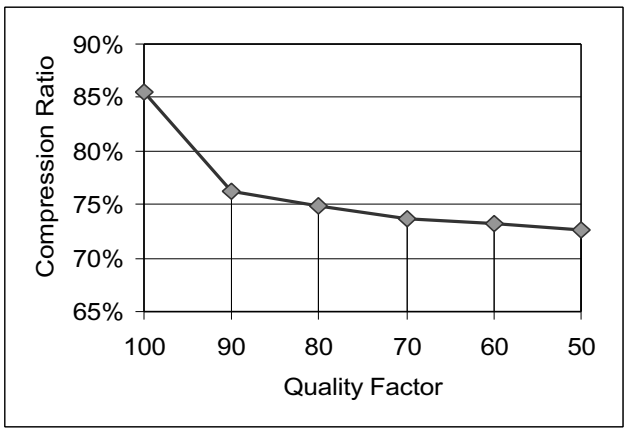

Figure 4. Compression Ratios

Conclusion

This paper explored the possibility of recovering a lost color channel from an RGB image, based on the information present in the remaining two color channels, as well as on *a priori* knowledge about the statistics of sensor responses in a given environment.

A camera's sensors determine the size of the RGB gamut that can be obtained using that camera. Using this information, as well as samples of possible illuminants and surface reflection functions improves the simple averaging approach to color channel recovery.

Using statistics about the environment and calibrating the recovery methods relative to camera responses in a given viewing context further improves the recovery of the missing color channel. From the experiments that were performed it seems that a neural network approach does not improve much the accuracy of recovery over a less complex method, such as linear regression.

Desaturated colors are recovered more faithfully than saturated ones. This has implications in the tone fidelity of faces –and skin in general– and can lead to additional image compression.

Remote guided vehicles as well as remote sensing applications can benefit from the redundancy introduced by this visualization technique.

Acknowledgements

The author would like to gratefully acknowledge the support of the Natural Sciences and Engineering Research Council of Canada and Hewlett Packard Corp.

References

1. A. J. Ahumada Jr. and K. Turano, "Calibration of a Visual System with Receptor Drop-Out," Exploratory Vision: The Active Eye, M.S. Landy, L.T. Maloney and M. Pavel eds., Springer Verlag, New York, pp. 157-167 (1996).

2. F. F. Sabins Jr., Remote Sensing: Principles and Interpretation, W.H. Freeman, New York, pp. 238-244 (1987).

3. M.C. Malin, *et al.*, "Early Views of the Martian Surface from the Mars Orbiter Camera of Mars Global Surveyor," *Science*, **279**, pp. 1681-1685 (1998).

4. B. K. P. Horn, "Exact reproduction of colored images," *CVGIP,* **26**, pp. 135-167 (1984).

5. J. A. Worthey and M.H. Brill, "Heuristic Analysis of von Kries Color Constancy," *J. Opt. Soc. Am.*, **3:10**, pp. 1708-1712 (1986).

6. D. B. Judd, D. C. MacAdam and G. Wyszecki, "Spectral Distribution of Typical Daylight as a Function of Correlated Color Temperature," *J. Opt. Soc. Am.*, **54**, pp. 1031-1040 (1964).

7. J. Cohen, "Dependency of the Spectral Reflectance Curves of the Munsell Color Chips," *Psihonomic Sci.*, **1**, pp. 369-370 (1964).

8. D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning Internal Representations by Error Propagation," in Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Volume I: Foundations, D.E. Rumelhart, J.L. McClelland and the PDP Research Group, eds., MIT Press, Cambridge, MA , pp. 318-362 (1986).

9. M. Anderson, R. Motta, S. Chandrasekar and M. Stokes, "Proposal for a Standard Default Color Space for the Internet – sRGB," *Proc. Fourth Color Imaging Conf.*, pp. 238-246 (1996).