Color Space Selection for JPEG Image Compression

Nathan M. Moroney † and Mark D. Fairchild

Munsell Color Science Laboratory, Center for Imaging Science
Rochester Institute of Technology, Rochester, New York
†currently at the RIT Research Corporation

The Joint Photographic Experts Group's image compression algorithm has been shown to be a very efficient and powerful method of compressing images.[1,2] However, there is little substantive information about which color space should be utilized when implementing the JPEG algorithm. Currently, the JPEG algorithm is set up for use with any three component color space.[3] The objective of this research was to determine whether or not the color space selected will significantly improve image compression capabilities. The RGB, XYZ, YIQ, CIELAB, CIELUV, and CIELAB LCh color spaces were examined and compared. Both numerical measures and psychophysical techniques were used to assess the results. The final results indicate that the device space, RGB, is the worst color space to compress images. In comparison, the nonlinear transforms of the device space, CIELAB and CIELUV, are the best color spaces to compress images. The XYZ, YIQ, and CIELAB LCh color spaces resulted in intermediate levels of compression.

Introduction

Digital color images are becoming increasingly more common as color scanners and printers become more affordable and reliable. However, color images require at least three times as much memory and take three times as long to transmit as monochrome images. As a result, considerable research has gone into color image compression.

One of the most widely used techniques for color image compression is the JPEG algorithm. This technique is composed of three basic forward steps and three analogous inverse steps. The first step is transforming from a spatial representation to DCT frequency representation. This transfromation results in an image in a highly compacted form. Next, the frequency data is selectively quantized based on properties of the human visual system. This step is where information is actually discarded and the image is further compressed. Finally, the quantized frequency coefficients can be Huffman encoded.

The JPEG algorithm can compress images anywhere from 2:1 to 20:1 or even higher. However, there is generally a trade-off between image compression and image quality. Higher compression results in lower image quality and vice versa. Visually lossless compression or compression with no perceptible deterioration in image quality usually results in compressions of 4:1 to 8:1.[4] The primary intention of this research was to determine if any one color space could be used with the JPEG algorithm in order to consistently improve the algorithm's performance.

Experimental

The Baseline JPEG algorithm was used to perform all of the image compressions. The example luminance and chrominance quantization and Huffman encoding tables listed in the 1990 ISO/CCITT Draft were used as the default tables for the compression. Four experimental images were chosen and can be briefly described as follows: a still-life with fruit, a profile view of two birds, a group portrait of three musicians, and a pasture with mountains in the distance. This selection of images was used so that a wide range of natural and man-made colors and textures would be included in the experiment. A Sony GDM-1950 monitor was used as the display device. The images were viewed at a distance of three feet away and an effective resolution of 24 bits/min^2. All of the computations were performed in floating point in order to minimize round-off error. Two psychophysical experiments were conducted and the images generated for the second experiment were analyzed using various image processing and colorimetric error metrics.

The luminance and chrominance quantization tables were applied according the quantization scheme outlined in Table I. Basically, the luminance quantities, Y and L*, were quantized using the luminance quantization table and all of the other channels were quantized using the chrominance quantization table. The only exception was for the RGB color space in which all of the channels were quantized using the luminance quantization table.

Table I. Quantization schemes used for different color spaces.

Color Space	Channel 1 Q-Table	Channel 2 Q-Table	Channel 3 Q-Table
LAB	Luminance	Chrominance	Chrominance
LCH	Luminance	Chrominance	Chrominance
LUV	Luminance	Chrominance	Chrominance
RGB	Luminance	Luminance	Luminance
XYZ	Chrominance	Luminance	Chrominance
YIQ	Luminance	Chrominance	Chrominance

Experiment One: Deriving Visually Lossless Thresholds of Compression

The first psychophysical experiment was a paired-comparison forced-choice experiment in which 20 observers participated. The original four images were compressed to seven different bits/pixel levels in each of the six color spaces for a total of 168 images. An original image and a compressed and decompressed version of that image were presented to the subject in a random successive manner. The observer was then instructed to select which image in the pair was the compressed image. The original image was always present in each of the pairs. Of course, the higher the level of compression, the worse the resulting image quality and the easier it was for the subject to perceive the compression.

The overall results for all of the observers and all of the images were then analyzed using probit analysis. This analysis basically fits a cumulative normal curve to data points representing frequency response at various levels of some experimental parameter.[5] Therefore for a given image and color space combination, a cumulative normal curve was fit to the frequency that all of the observers could correctly identify the compressed image at various compression levels. The probit analysis will then yield a threshold for visually lossless compression and also an estimate of the uncertainty of this threshold. Specifically, this threshold value is the level of compression at which the compressed image can correctly be identified by 50% of the time.

The overall results for all four of the experimental images averaged together are shown in Fig. 1. In this figure, the x-axis is a nominal scale representing the six color spaces tested and the y-axis is the perceptual threshold in bits/pixel. A low value for the threshold corres-ponds to greater degree of compression and a higher threshold value coincides with a lesser level of compression. Essentially, this axis represents how far the images could be compressed in each of the color spaces before the observers could perceive the compression. The lower the threshold values indicated preferable color spaces for JPEG compression. The overall threshold levels are shown as squares and plus or minus two standard errors are plotted around each of the thresholds. Based on these results, the RGB color space was the worst color compression space. The XYZ and YIQ color spaces were next and the CIELAB and CIELUV color spaces were the best color compression space. The CIELAB LCh space was very inconsistent as is shown by the large error bars around the threshold.

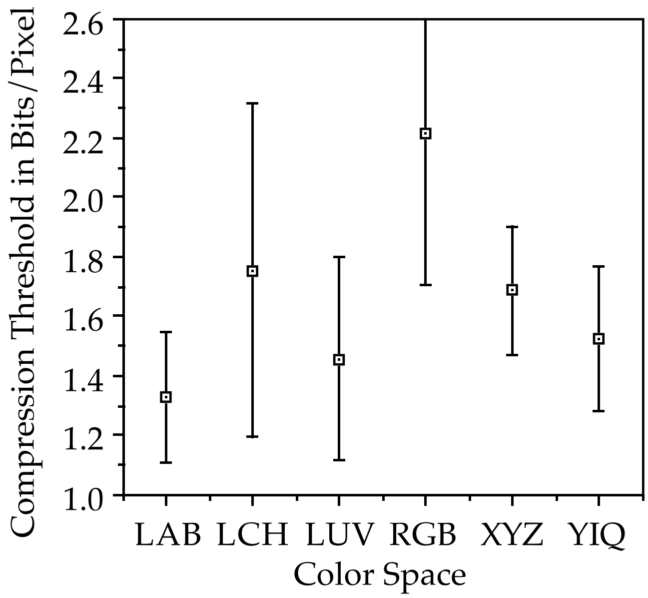

Figure 1. Averaged perceptual compression thresholds for four images

Experiment Two: Supra-Threshold Color Compression Space Comparisons

The second psychophysical experiment was also a paired-comparison experiment. However, this time the four experimental images were all iteratively compressed to one aim compression level for all of the color spaces, approximately 1.25 bits/pixel. This aim compression level was selected by evaluating the results from the first experiment. The purpose of this experiment was to examine compression at higher, or supra-threshold levels, as well as to compare the quality of image compressions using the different color spaces. A total of 25 subjects viewed the 24 experimental images. The observer was once again sequentially presented with pairs of images. This time the subject was asked to determine which image in the pair had the better image quality. The pair always consisted of the same image compressed in two different color spaces.

The Law of Comparative Judgements was then applied to the combined observer responses. This analysis allows an interval scale to be derived out of a collection of paired-comparison ordinal responses. The computations begin by deriving a frequency matrix and then converting to a proportionality matrix. Lastly, the data from this matrix was transformed through the logistic function. The columns of this matrix were summed in order to calculate the interval scale for image quality. Typically, a Z-score transform is used as the last step but, because there were so many instances where one image was worse to 100% of the observers, the logistic function was more approriate.[5]

The resulting interval scale of perceived image quality is shown in Fig. 2. The horizontal axis is a nominal scale delineating the color spaces and the vertical axis is the interval scale representing image quality. A large scale

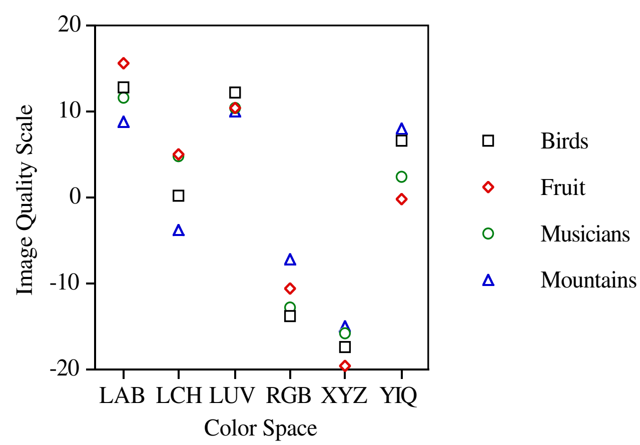

Figure 2. Image quality scale derived for each of the color spaces

Table II. Image Processing Error Metrics Averaged for Each of the images.

Color Space	Maximum Error (0 to 255)	Minimum Error (0 to -255)	RMS Error (0 to 255)	PSNR decibels	Entropy Difference Bits/Pixels
CIELAB	50.75	-75.92	6.75	31.90	-0.12
CIE LCh	70.25	-86.00	8.01	30.24	-0.12
CIELUV	55.17	-76.67	6.75	31.86	-0.09
RGB	85.83	-92.00	8.66	29.37	-0.03
XYZ	103.42	-105.42	12.75	26.26	0.04
YIQ	59.33	-58.33	7.07	31.39	0.05

Table III. Colorimetric error metrics averaged for each of the images.

Color Space	Average RMS DL^*_{ab}	Average RMS Dc^*_{ab}	Average RMS DH^*_{ab}	Average DE^*_{ab}
CIELAB	1.86	6.29	6.16	2.80
CIE LCh	2.20	7.41	7.27	2.93
CIELUV	1.92	6.50	6.34	3.01
RGB	2.78	7.96	7.61	4.48
XYZ	3.18	10.61	10.39	7.50
YIQ	2.20	7.04	6.91	4.30

value for a given image indicates a greater perceived image quality for that color space. Remember that all of these images have been compressed to the same level and that if the color spaces were equivalent then all of the images in all of the colors spaces should have roughly the same image quality. The results for all four images are shown at once and it is evident that there is a drastic and systematic variation in the image quality for the six color spaces and only a small scene dependency. The CIELAB and CIE-LUV color spaces were the best color compression spaces. The XYZ space is the worst color compression space and RGB is next. The YIQ and CIELAB LCh spaces produce intermediate levels of image quality. Except for the last place finish of the XYZ space, these results are similar to the results of the first experiment.

Finally, image processing and colorimetric error metrics were computed for the images used in experiment two. The image processing metrics computed were maximum and minimum error, RMS error, PSNR and entropy difference. These quantities provide some sort of objective measure of the image quality of the images. The average values for each of the images are listed in Table II shown below. In a majority of the cases, the nonlinear color spaces performed best and the XYZ and RGB color compressions were the worst. The YIQ color space was close third and the CIELAB LCh space was a distant forth.

The colorimetrc error metrics computed were average RMS DL^*_{ab}, average RMS Dc^*_{ab}, average RMS DH^*_{ab} and average DE^*_{ab}. These values provided some measure of the errors in the lightness, chroma, hue and total color of the image. These values were averaged for each of the images and the results are listed in Table III. Once again, the CIELAB and CIELUV color spaces had the smallest errors and the XYZ and RGB largest errors. The YIQ and CIELAB LCh spaces were third and fourth respectively.

Conclusions

Both the psychophyiscal experiments and the numerical error metrics suggest that CIELAB and CIELUV are the best color spaces for JPEG image compression. The YIQ color space is in third and CIELAB LCh is an inconsistent fourth. The RGB and the XYZ color spaces were the worst for JPEG compression.

It is likely that most significant reason that the non-linear transforms of device space performed so well is the L* transformation. The L* transform preserves the darker lightness or shadow regions of the image in a much more perceptually uniform manner. In addition, the non-linear color spaces also reduced the redundancy in the channels and resulted in a much more compacted representation of the image.

Finally, the quantization schemes listed in Table I were probably not optimal. One definite example is the quantization of h_{ab} in CIELCh space using the chrominance quantization table. The polar nature of h_{ab} makes this quantization very inefficient and also induces larger errors. The XYZ quantization scheme could also have been improved either by transforming from XYZ to xyY space or by using the luminance quantization table for all three channels.

References

1. J. L. Mitchell, W. B. Pennebaker, "Evolving JPEG Color Data Compression Standards", *Standards for Electronic Imaging Systems*, *SPIE Optical Engineering Press*, Bellingham, Washington, 1991, pp 68-97.
2. A. Leger, T. Omachi, G. K. Wallace, "JPEG Still Picture Compression Algorithm", *Optical Engineering*, **30** 947-954 (1991).
3. G. K. Wallace "Overview of the JPEG (ISO/CCITT) Still Image Compression Standard" *SPIE* **1244**, *Image Processing Algoritms and Techniques*, 220-233 (1990).
4. D. Goodenow, "JPEG Compression", *PRINT RIT The Journal of the RIT School of Printing Management and Science*, **1**, 6-9 (1993).
5. C. J. Bartleson, F. Grum, *Optical Radiation Measurements*, Vol. **5**. Academic Press, New York, 368-439 (1984).