AIC LISBOA 2018
colour & human comfort

Colorization of Dichromatic Images

Brian Funt* and Ligeng Zhu

Simon Fraser University, Vancouver, Canada
**E-mail funt@sfu.ca*

ABSTRACT

This paper explores the color information dichromatic vision provides in terms of its potential for colorization. Given a greyscale image as input, colorization generates an RGB image as output. Since colorization works well for luminance images, how well they might work for dichromatic images? Dichromatic images are colorized using a modification of the colorization method of Iizuka et al. (Proc. SIGGRAPH 2016, 35(4):110:1-110:11). In particular, an sRGB image is converted to cone LMS and M is discarded to yield an LS image. During training, the colorization neural network is provided LS images and their corresponding LMS images, and it adjusts its weights so that M is predicted from the L and S. One does not easily recognize that a colorized dichromatic image is, in fact, based on only L and S, and is not a regular full-color image. This is stark contrast to the dichromatic simulations of Brettel et al. (Brettel, Viénot, Mollon, *JOSA A* 14, 2647-2655, 1997).

Keywords: *Color blindness, dichromat, computational color vision, colorization.*

INTRODUCTION

What it is like to be color blind? What colors do dichromats see? How much 'color' is really missing from a dichromats experience. Since it has been reported that many people do not realize they are color blind until they are many years old, perhaps the difference is not so significant. Of course, we really can never know what another person experiences, but we can explore what color information dichromatic vision provides. In recent years, many colorization methods have been described in the computer vision literature. Given only a greyscale (i.e., luminance) image, these computer-based colorization methods generate a color image with very believable colors. Colorization methods are generally based on 'deep learning' the connection between luminance, the context, and probable color. It appears they encode knowledge about the world such as clear sky is blue, clouds are grey, beaches are a sandy color, forests and grass are green, and so forth. Since colorization works for luminance images, we explore how well they might work for dichromatic images.

The results of colorizing dichromatic images can be expected to give us more insight into what color information is present, as well as missing, for the dichromat at a more experiential level than the standard statements that deuteranopes—observers lacking M cones—cannot distinguish some reds and greens from one another. Brettel et al. [1]developed a simulation of dichromatic vision based on projecting the given LMS stimulus onto a reduced stimulus surface defined by the neutral axis along with the LMS locations of the monochromatic stimuli that are perceived as the same hue by normal trichromats and dichromats. Psychophysical experiments with dichromatic observers validate their method. Although the full colour image and its dichromatic simulation look quite different to the trichromatic observer, they appear indistinguishable to the dichromatic observer. See Figure 1 (left and middle). This example of the dichromatic simulation was produced using Vischeck [7], a web-based implementation of the Brettel et al. [1]algorithm. Note that the Vischeck result only roughly approximates the true Brettel et al. result since it requires the image to be displayed on a calibrated CRT monitor, not an uncalibrated LCD display.

Figure 1: (Left) input image; (Middle) dichromatic simulation that appears the same to the dichromat as the input image when viewed on a calibrated display; (Right) colorized LS image. The dichromatic simulation lacks the green of the apple and the reds in the background; however, the colorization result shows that much of the colour information can still be correctly inferred (but note the grey suit and tie) from the LS input.

BACKGROUND

Well prior to the current interest within the computer vision community concerning the colorization of greyscale images, Cardei [2] extrapolated RB images into RGB images using three different strategies: a very small (by today's standards) 3-layer, 9-neuron neural network, linear regression and polynomial regression. He reported quite good results, with the neural network and polynomial regression methods leading to similar errors in the prediction of G from RB input. By comparison, modern colorization networks use 30 layers with 1,500,000 parameters. These colorization strategies are based on deep convolutional neural networks [6], which have been shown to be able to learn complex mappings from large amounts of training data. In particular, Iizuka et al. [4] showed that deep Convolutional Neural Networks colorize greyscale images very well. However, their proposed model is very large and takes weeks to train and hundreds of megabytes to store. Johnson et al. [5] proposed an alternative network structure that was originally developed for transferring the style of one image to another. For dichromatic colorization, we use the basic method of Iizuka et al. but modify it to make use of the structure that Johnson et al. proposed.

EXPERIMENTS

Training and testing a neural network requires a large number of images. For this we use the Microsoft COCO [3] image dataset, which contains images in sRGB format. The first step is to convert the sRGB data to LMS (i.e., cone response space). The non-linear sRGB input is converted first to linear XYZ (CIE XYZ space) according to the sRGB standard, and then from XYZ to LMS cone space using the Hunter-Pointer-Estevez transformation matrix. To simulate a deuteranope, the M channel is discarded, yielding a 2-channel, LS image.

The network structure is outlined in Figure 2, with the details listed in Table 1. An LS image is input and then processed by the deep colorization model, which is composed of an encoder, shave blocks and a decoder. The corresponding M channel is predicted from the patterns found in the training data.

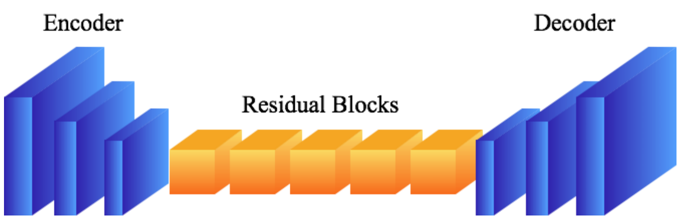

Figure 2: Blue indicates a composition of a convolutional (conv) layer, a batch normalization (BN) layer, and a rectified linear unit (ReLU). Orange indicates a block that includes a shave layer in parallel to a composite Conv-BN-ReLU-Conv-BN convolutional layer. The results from the parallel paths are then summed.

Encode

type	kernel	stride	output
conv	9x9	1x1	32
conv	3x3	3x3	64
conv	3x3	3x3	128

Residual Block

type	kernel	stride	output
conv	3x3	1x1	128
conv	3x3	1x1	128
residual	-	-	128
sum	-	-	128

Decode

type	kernel	stride	output
deconv	3x3	2x2	64
deconv	3x3	2x2	32
conv	9x9	1x1	1

Table 1: Specifications of the different components used in the dichromatic-colorization network. Kernel indicates the size of the convolution (conv) kernel. Stride controls the subsampling of the input data. Output refers to the number of convolution filters of the given size and stride used in the respective convolutional layer.

The network training involves inputting 83K LS images from the COCO dataset along with their corresponding M-channel pixel values. During training, network's weights are adjusted so that M is predicted from the L and S. The training converges to a low average 'loss' measured by the L2 norm in LMS space. The network's LMS output is then converted back to sRGB for display. The remaining COCO images are used for testing. Figure 3 shows some typical examples.

RESULTS AND DISCUSSION

As is apparent from the right-hand column of Figure 3, the colorization results are surprisingly believable. In other words, one does not immediately recognize that a colorized dichromatic image is, in fact, 'dichromatic' (i.e., in the sense that it is derived from only two channels of color information), and not a regular full-color image. While it is impossible to draw a formal conclusion from such examples about the nature of dichromatic colour perception, it does indicate that the dichromatic images do contain more colour information that might be expected from trichromatic viewing of the Brettel et al. simulations. As might be anticipated from the fact that the L and M cone sensitivity functions significantly overlap, much of the information is there, it simply has to be extracted computationally. Whether or not dichromats do extract such information is a question for further study.

ACKNOWLEDGEMENT
Financial support provided by the Natural Sciences and Engineering Research Council of Canada.

REFERENCES

[1] Brettel, H., Viénot, F. and J. Mollon. 1997. *Computerized Simulation of Color Appearance for Dichromats*, J. Optical Society of America. A 14, 2647-2655, 1997.

[2] Cardei, V. 1999. *From Dichromatic to Trichromatic Images; Implications for Image Recovery and Visualization.* PICS 99 Proc. of the Conf. on Image Processing, Image Quality and Image Capture Systems.

[3] Tsung-Yi Lin, T., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., and L. Zitnick. 2014. *Microsoft COCO: Common Objects in Context*, Proc. ECCV Part V, LNCS 8693, 740-755.

[4] Iizuka, S., Simo-Serra E. and H. Ishikawa. 2016. *Let there be color!: Joint end-to-end learning of global and local image priors for automatic image colorization with simultaneous classification,* Proc. SIGGRAPH 2016, 35(4):110:1-110:11.

[5] Johnson, J. Alexandre, A. and F. Li. 2016. *Perceptual losses for Real-time Style Transfer and Super-resolution*. Proc. European Conference on Computer Vision.

[6] Krizhevsky, A., Sutskever, I. and G. Hinton. 2012. *ImageNet Classification with Deep Convolutional Neural Networks.* Proc. Advances in Neural Information Processing Systems 25, 1097–1105

[7] Vischeck, http://www.vischeck.com/ Accessed June 1, 2018.

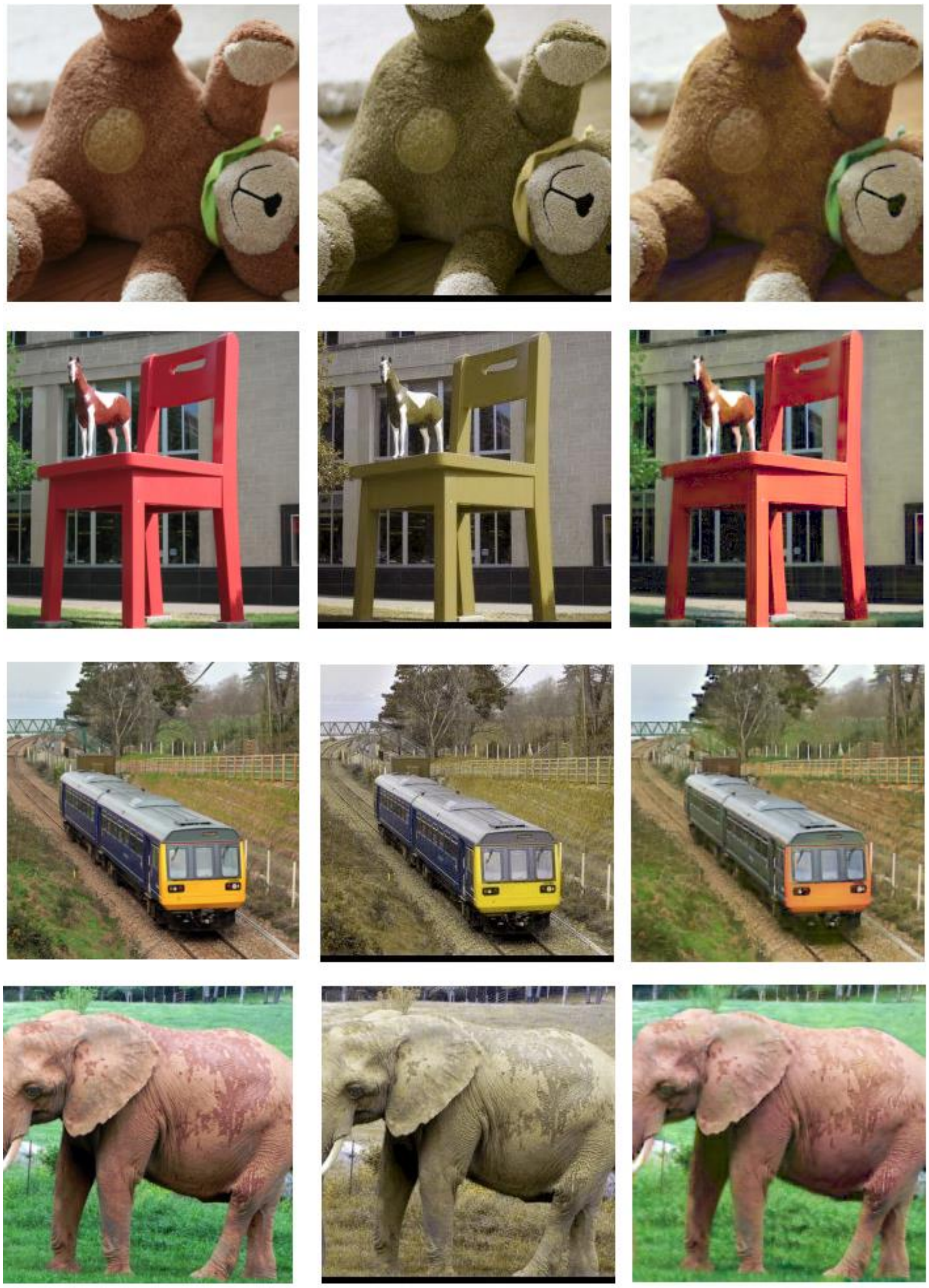

Figure 3: Left column contains the sRGB input images; Middle column are the Vischeck corresponding dichromatic simulation results; Right column are the LS to LMS colorization results.