LEARNING COLOR APPEARANCE MODELS

E. Boldrin, P. Campadelli§, R. Schettini

ITIM, National Research Council, Milano, Italy

§DSI, Universita' degli Studi di Milano, Milano, Italy

Abstract

We present a method for faithfully approximating the Hunt94, LLAB and RLAB color appearance models by means of feed-forward neural networks trained with the error back-propagation algorithm. In particular we present experimental evidence that in eight "standard" viewing conditions the same network architecture is capable of learning quite satisfactorily the transformations performed by the three models.

Introduction

The reproduction of color across different media and viewing conditions requires color appearance modeling. It demands device-independent color description (which can be achieved by colorimetric characterization of the hardware equipment) and viewing condition independence, which requires a color appearance model. Unfortunately, while most color management systems already incorporate device profiles that make colorimetric color reproduction possible, they often neglect color appearance modeling. And there is not even general agreement to date as to the choice and use of a color appearance model [1,9].

We describe here a method for faithfully approximating the Hunt94 [6,7,8], LLAB [10] and RLAB [3] color appearance models, by means of feed-forward neural networks trained with the back-propagation algorithm on training sets derived from the ANSI IT8 7.2 color target, to obtain an effective and efficient mapping of color appearance [2]. The complexity of the neural networks devised for the different models and viewing conditions is quite contained. Since the networks permit color mapping in real time, it is possible to build a library of neural network mapping functions, that plugged into color management systems, effectively transform the color stimuli perceived on one device to the corresponding stimuli required to produce the same appearance on a second device.

Experimental design

The computational complexity of a color appearance model is often a drawback to its use in color image reproduction. Starting from this consideration, we recently defined a method for learning by examples the mapping obtained, once the viewing conditions were set, by the combination of the forward and reverse Hunt94 models [2]. The learning was done by feed-forward neural networks trained with the back-propagation algorithm [5, 12]. We now apply the same method for learning LLAB and RLAB models, and compare our results with those already obtained with the Hunt94 model.

The color mappings to be learned, specified by sets of pairs of corresponding color stimuli, have been defined according to the guidelines of the CIE Technical Committee 1-27: "Specification of Colour appearance for reflective media and self-luminous display comparisons" [1]. We considered three different illuminants D50, D65 and A for hardcopy. For the softcopy device we have assumed a monitor with white chromaticities matching those of the standard D65 and D93 illuminants. The maximum luminance of the light was set at 80 cd/m^2 for both hard and softcopies. In total eight sets of viewing conditions were considered: Booth A → CRT D65, Booth D50 → CRT D65, Booth D65 → CRT D93, Booth D50 → CRT D93, CRT D65 → Booth A, CRT D65 → Booth D50, CRT D93 → Booth D65, CRT D93 → Booth D50. Both the adapting field and the background were set at a 20% reflecting neutral gray, while a white proximal field of the same chromaticities as the illuminant was used.

After setting the viewing conditions that define the mapping to be learned, we chose the colors for building the appropriate training, validation and test sets. In particular we selected the Agfa™ ColorReference color target, designed according to the ANSI IT8.7 standard, to construct the networks' training sets, while we used the Macbeth ColorChecker to construct the networks' validation sets [13]. A set of 1250 samples, taken at regular intervals from the whole Munsell Atlas served as the test set [11]. For each color, in the set viewing conditions, we computed the corresponding color by applying the combination of the forward and reverse LLAB and RLAB models. The Cartesian CIELAB coordinates of the pairs of corresponding color stimuli constituted the sets of input/output examples used in the design of the neural networks.

The colors were represented in Cartesian CIELAB coordinates in order to permit a straightforward application of the back-propagation learning algorithm, which uses the Euclidean distance between the desired and obtained outputs to measure the network error and, consequently, update the network weights. The coordinates of the colors were normalized, those of input colors in the range of [-1, 1], and those of output colors in the range of [0.1, 0.9], for processing by the networks. The entire lightness range [0,100] was considered in normalizing the CIELAB values while, to avoid *"out of range"* problems during both testing and the common use of the procedure, we increased the maximum value of chroma of the Agfa data set by 15%.

To learn the Hunt94 model satisfactorily we performed a systematic search for a good network architecture [2], comparing the behavior of 25 different networks. All of these networks had 3 input and 3 output neurons, while the number of hidden layers and of units in each layer varied, so that there were about 60 to 160 weights to be learned. A higher number of parameters would not have been compatible with the size of the training set: in fact, an empirical rule states that the size of the set must be at least 1.5 times the number of free parameters in the network. In all the networks, weights and thresholds were initialized randomly with values in the interval of [-1, 1], and the neuron transition function was the logistic mapping on [0,1]. Back propagation learning was performed by pattern. The architecture which produced the best performance was a network with 3 hidden layers of 7 neurons each, which implies a network of 27 neurons and 164 links in all. Considering that the LLAB and RLAB models are less complex then the Hunt94 model, we used the same architecture for learning both of them.

We have experimentally found that the best learning results are obtained when the momentum term is set at 0.9, and the learning rate, beginning from an initial value of 0.5, is changed with a step of 0.1 every 1900 epochs until a final value of 0.05 is reached; learning is then continued without further changes.

Experimental results

For each function to be learned (that is the mapping performed by an appearance model for a given viewing condition), network training was stopped when the error, while still decreasing on the training set, began to increase on the validation set. Different learning processes required different numbers of epochs (presentation of the complete training set) to meet this condition for termination. The learning procedures, depending on the number of epochs have required several hours of processing on a IBM Risc 6000 320H. However, once trained, the networks perform the desired color mapping very efficiently. The quality of the performance of each trained network was measured by the mean error, that is the mean Euclidean distance in the CIELAB color space, on the training, test and validation sets. Experimental results for the Hunt94, the LLAB and the RLAB models are summarized in Tables I, II and III. The mean CIELAB errors on the training (Agfa^{TM} ColorReference color target), validation (The Macbeth ColorChecker) and test (1250 samples, taken at regular intervals from the whole Munsell Atlas) sets are shown for each color appearance model and for each mapping. The minimum and maximum error values are reported in round brackets below the means. The last column records the number of learning epochs required to obtain these results.

These results demonstrate that the approximated transformations are very good: the greatest color errors are very small, especially when compared with inter-observer variability in color appearance judgment, which often exceeds 20 CIELAB units, and intra-observer variability, which may be the on order of 10 CIELAB units [4, 8].

Conclusions

In order to obtain efficient mapping of color appearance, we have defined a method to approximate by learning the combination of the forward and reverse color appearance models for each desired experimental set-up. Experimental results confirm the feasibility of this method, we believe that this learning method will function equally well for different viewing conditions.

It could also be used to approximate further extensions of the models considered and new ones. Also attractive is the possibility of using neural networks to learn experimentally defined corresponding stimuli directly. This would mean that the networks could constitute an efficient way of realizing the desired transformation even in the absence of a satisfactory model.

References

[1] P.J. Alessi, CIE guidelines for coordinate research on evaluation of colour appearance models for reflection print and self-luminous display image comparison, Color Research and Application **19**, 48-58 (1994).

[2] E. Boldrin, P. Campadelli, R. Schettini, Effective and efficient mapping of color appearance, Color Research and Application, 1997 (in print).

[3] M.D. Fairchild "Refinement of the RLAB color space" Color Research and Application, **21**, 338-346 (1996).

[4] M.D. Fairchild, R.L. Alfvin, Precision of color matches and accuracy of color-matching functions in cross-media color

reproduction, Proc. Third Color Imaging Conf., Scottsdale (USA), 18-21, (1996).

[5] J. Hertz, A. Krogh, R.G. Palmer, Introduction To the Theory of Neural Computing, Addison-Wesley, New York, **1** 115-162 (1991).

[6] R.W.G. Hunt, Revised colour appearance model for related and unrelated colours, Color Research and Application, **16,** 146-165 (1991).

[7] R.W.G. Hunt, An improved predictor of colourfulness in a model of colour vision, Color Research and Application, **19,** 26-26 (1994).

[8] R.W.G. Hunt, MR Luo, Evaluation of a model of colour vision by magnitude scaling: discussion of the collected results, Color Research and Application **19**, 27-33 (1994)

[9] M.C. Lo, M.R. Luo, P.A. Rhodes, Evaluating colour models' performance between monitor and print images, Color Research and Application **21**, 277-291 (1996).

[10] M.R. Luo, M-C Lo, W-G Kuo, The LLAB (l:c) Colour model, Color Research and Application **21**, 412-429 (1996).

[11] J.Parkkinen, J.Haanpalo, Color Spectra Database, Lappeenranta University of Technology, http://www.lut.fi/ltkk/tite/research/color/lutcs_database.html

[12] D.E. Rumelhart, G.E. Hinton, R.J. Williams, Learning internal representations by error propagation. In D.E. Rumelhart, J.L. McClelland Eds, Parallel distributed processing, MIT Press, Cambridge (MA), **1**, 145-168 (1986).

[13] S. McCamy, H. Marcus, J.G. Davidson, A color-rendition chart, J. of Appl. Photo. Eng., **2,** 95-99 (1976).

	TRAINING SET ΔE average (ΔEmin , ΔEmax)	**VALIDATION SET** ΔE average (ΔEmin , ΔEmax)	**TEST SET** ΔE average (ΔEmin , ΔEmax)	**Number of learning epoches**
D50BOOTH → D93CRT	0.147 (0.018, 0.501)	0.559 (0.029, 2.640)	0.298 (0.016, 2.825)	106,000
D93CRT → D50BOOTH	0.130 (0.011, 0.562)	0.559 (0.065, 2.718)	0.298 (0.012, 2.884)	133,000
D65BOOTH → D93CRT	0.134 (0.013, 0.479)	0.486 (0.029, 1,845))	0.256 (0.010, 2.229)	107,000
D93CRT → D65BOOTH	0.111 (0.009, 0.435)	0.492 (0.055, 2.027)	0.305 (0.006, 3.899)	118,000
ABOOTH → D65CRT	0.126 (0.007, 0.514)	0.594 (0.061, 3.229)	0.282 (0.017, 2.819)	295,000
D65CRT → ABOOTH	0.118 (0.009, 0.444)	0.559 (0.370, 1.910)	0.196 (0.005, 1.737)	202,000
D50BOOTH → D65CRT	0.145 (0.008, 0.510)	1.508 (0.116, 3.036)	0.377 (0.016, 5.023)	96,000
D65CRT→ D50BOOTH	0.110 (0.011, 0.397)	0.467 (0.037, 2.377)	0.266 (0.008, 3.652)	149,000

Table I. Hunt 94 approximation: summary of the expertimental results.

	TRAINING SET ΔE average (ΔEmin , ΔEmax)	**VALIDATION SET** ΔE average (ΔEmin , ΔEmax **)**	**TEST SET** ΔE average (ΔEmin , ΔEmax)	**Number of learning epoches**
D50BOOTH → D93CRT	0.326 (0.015, 2.072)	1.082 (0.099, 8.667)	0.817 (0.049, 13.401)	541,500
D93CRT → D50BOOTH	0.178 (0.011, 0.639)	0.650 (0.092, 3.037)	0.363 (0.006, 4.629)	120,000
D65BOOTH → D93CRT	0.398 (0.035, 5.144)	1.096 (0.084, 6.547)	0.598 (0.024, 5.197)	147,500
D93CRT → D65BOOTH	0.235 (0.0025, 0.803)	0.848 (0.070, 3.819)	0.471 (0.030, 5.324)	23,000
ABOOTH → D65CRT	0.471 (0.052, 5.772)	1.367 (0.070, 7.222)	0.634 (0.048, 7.831)	89,000
D65CRT → ABOOTH	0.184 (0.019, 0.640)	0.590 (0.057, 3.022)	0.330 (0.013, 3.103)	155,000
D50BOOTH → D65CRT	0.331 (0.070, 2.622)	1.332 (0.186, 9.159)	0.718 (0.0156, 5.792)	422,500
D65CRT→ D50BOOTH	0.185 (0.026, 0.699)	0.550 (0.075, 2.897)	0.342 (0.012, 3.697)	106,000

Table II. LLAB approximation: summary of the expertimental results.

	TRAINING SET ΔE average (ΔEmin , ΔEmax)	**VALIDATION SET** ΔE average (ΔEmin , ΔEmax **)**	**TEST SET** ΔE average (ΔEmin , ΔEmax)	**Number of learning epoches**
D50BOOTH → D93CRT	0.187 (0.0126, 0.972)	0.700 (0.048, 3.309)	0.379 (0.019, 3.319)	30,500
D93CRT → D50BOOTH	0.149 (0.017, 0.490)	0.579 (0.112, 2.498)	0.310 (0.008, 5.470)	61,100
D65BOOTH → D93CRT	0.165 (0.008, 0.775)	0.547 (0.026, 3.132)	0.290 (0.705, 6.280)	21,500
D93CRT → D65BOOTH	0.159 (0.023, 0.687)	0.576 (0.060, 2.695)	0.332 (0.011, 4.379)	28,000
ABOOTH → D65CRT	0.152 (0.052, 5.772)	0.557 (0.064, 2.603)	0.298 (0.013, 3.058)	67,500
D65CRT → ABOOTH	0.145 (0.015, 0.641)	0.498 (0.032, 2.603)	0.243 (0.017, 2.177)	56,500
D50BOOTH → D65CRT	0.136 (0.016, 0.621)	0.531 (0.044, 2.592)	0.304 (0.011, 3.601)	68,500
D65CRT→ D50BOOTH	0.136 (0.012, 0.420)	0.495 (0.039, 2.588)	0.274 (0.005, 4.662)	66,500

Table III. RLAB approximation: summary of the expertimental results.