SCIENTIFIC PAPERS
OF THE UNIVERSITY OF PARDUBICE
Series A
Faculty of Chemical Technology
19 (2013)

PREDICTION OF RGB CAMERA VALUES BY MEANS OF ARTIFICIAL NEURAL NETWORKS

Miha LAZAR[a], Aleš HLADNIK[1a], Dejana JAVORŠEK[a] and Tim JERMAN[b]
[a]Faculty of Natural Sciences and Engineering,
[b]Faculty of Electrical Engineering,
University of Ljubljana, SL–1000 Ljubljana,

Received September 30, 2013

*Artificial neural networks (ANN) enable modelling of complex nonlinear systems that cannot be easily described using formal equations and have been implemented in many fields of science and technology for pattern recognition, clustering or data fitting. The goal of our study was to create a system that transforms XYZ and L*a*b* values into arbitrary camera RGB values in stable — but without strict knowledge of — photographing conditions, by means of the ANN data fitting ability. We adopted a two layer feed-forward neural network with sigmoid hidden and linear output neurons, that can fit multi-dimensional mapping problems quite well, when using enough neurons in the hidden layer and being fed by congruent learning set of data. The network was trained with Levenberg–Marquardt backpropagation algorithm. Learning data sets consisted of input XYZ or L*a*b* values and output RGB values. Input data were calculated from the reflectance values of Gretag Macbeth Digital ColorChecker SG test chart obtained by spectrophotometric measurements, by taking into account three*

[1] To whom correspondence should be addressed.

different standard illuminants (A, D50 and D65) and two standard colorimetric observers (2 ° and 10 °). Output data were RGB values of test chart ColorChecker SG acquired by Nikon D50 digital camera. Our goal was to find answers to several questions, such as what is an optimal number of hidden layer neurons, what degree of accuracy can we obtain by training ANN with a limited number of color samples, how does number of neurons affect ANN learning time and also which type of input data is more suitable for the prediction of RGB values. Since each ANN learning epoch starts with a random weight distribution and random training, validation and testing data selection, every learning cycle stopped in its local minimum. To assess the representative values of difference between the actual and the predicted values, learning cycle for each number of hidden layer neurons and for each learning data set was repeated many times and average ANN training time and average, median and minimal error rates for training, validation and testing data were recorded.

Introduction

Unambiguous and accurate color description of printed documents requires that their spectral information is known. Majority of common scanners, cameras and also human color vision are trichromatic. In a quest for describing color documents by their spectral data satisfactorily, one cannot overlook the possibility of using ubiquitous inexpensive RGB cameras as input devices for primary data acquisition. Different methods have been proposed for trichromatic to spectral conversion [1,2]. Some of more recent ones are based upon artificial neural networks (ANN) [3,4], and such feed-forward neural networks can be used as universal approximators for continuous functions under appropriate conditions [5-7]. In a process of development of ANN-based system for higher dimensional spectrum/reflectance data approximation from lower dimensional RGB camera input data, many steps have to be overcome and questions answered. Some of them are "how to get congruent and big enough learning data set?", "what is the optimal number of hidden layer neurons?" and "What degree of accuracy can be achieved by a limited learning data set?".

The main task, when developing ANN-based reflectance recovery system from consumer camera readouts, is to efficiently learn the system. Such ANN can later satisfactorily reconstruct pixel wise image reflectances from RGB values. Most of capturing conditions can be unambiguously measured, verified and controlled, but some, like RGB camera optical characteristics and spectral sensitivity, are hidden or inaccessible. ANN learning set consists of input-output pairs, namely RGB camera responses and adequate reflectances captured from the set of standard patch samples. Reflectances are measured *via* spectrophotometer, while RGB values are obtained from camera responses in illumination-controlled environment. For this purpose it is important to obtain large enough learning set containing pairs of camera RGB and spectrophotometrically measured reflectance values for each color patch.

In our study we focused on an inverse problem which enabled us to map reflectances, XYZ, and L^*a^*b values into RGB camera responses via ANN as discussed in details below.

Theoretical Background

RGB camera imaging controlled environment system, usually dark room with standard illumination setup, as illustrated in Fig. 1, is described by spectral radiance of the illuminant $I(\lambda)$, spectral reflectance of the surface $r(\lambda)$, spectral transmittance of the optical system $o(\lambda)$ and spectral sensitivity of the camera $a(\lambda)$. Camera output d_{RGB} described for each of the three camera channels is formulated as

$$d_i = \int_{\lambda_{\min}}^{\lambda_{\max}} I(\lambda) r(\lambda) o(\lambda) a_i(\lambda) \mathrm{d}\lambda \qquad i \in \langle R, G, B \rangle \tag{1}$$

and $[\lambda_{\min}, \lambda_{\max}]$ represents the interval of visual spectrum wavelengths or, in the case of known spectral sensitivity of the camera, the interval inside which any of the three channels spectral sensitivities is greater than zero. Spectral transmittance of the optical system and spectral sensitivity functions of the consumer camera color channels are usually, as in our case, undefined.

Colorimetric values, on the other hand, are well defined, as in the case of CIE XYZ, through color matching functions (CMFs)

$$t_i = \frac{1}{N} \int_{\lambda_{\min}}^{\lambda_{\max}} I(\lambda) r(\lambda) c_i(\lambda) \mathrm{d}\lambda \qquad i \in \langle X, Y, Z \rangle \quad \text{and} \quad c_i \in \langle \bar{x}, \bar{y}, \bar{z} \rangle$$

$$\text{and} \quad N = \frac{1}{N} \int_{\lambda_{\min}}^{\lambda_{\max}} I(\lambda) \bar{y}(\lambda) \mathrm{d}\lambda \tag{2}$$

$[\lambda_{\min}, \lambda_{\max}]$ represents the interval of wavelengths outside of which all the CMFs are zero.

CIE L^*a^*b values are easily calculated from XYZ, considering chosen illuminant white point.

Typical artificial neural network system for spectral reflectance recovery from arbitrary camera output (Fig. 2) is trained from a set of standard colour patch camera output and spectral reflectance vector pairs. When trained, spectral reflectance can be obtained pixel wise from every camera RGB colour pixel readout, and colorimetric values could be calculated from spectral reflectance vectors.

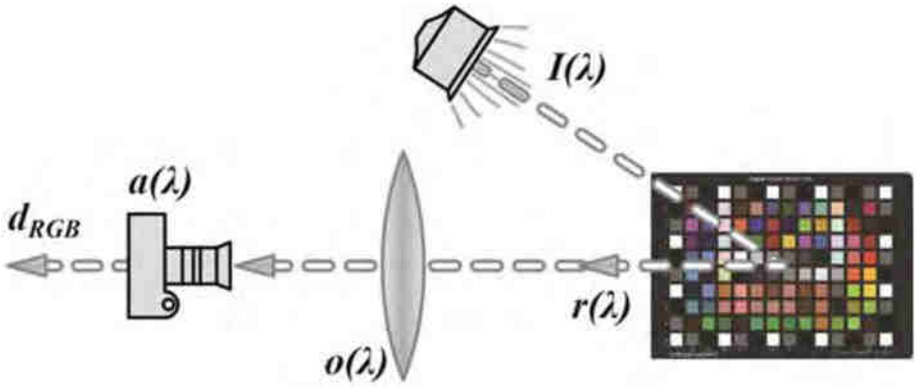

Fig. 1 Outline of the consumer camera RGB image acquisition system

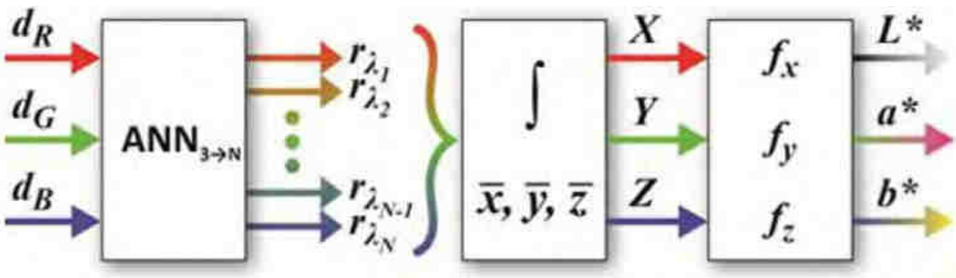

Fig. 2 ANN for spectrum (r) recovery from arbitrary camera output (d)

In our study we focused on a reversed dataflow. First, from a set of reflectances under particular standard illuminant and observer we calculated colorimetric tristimulus – CIE XYZ – and CIE $L^*a^*b^*$ values. Second, we used pairs of these XYZ (or Lab) and camera RGB values to train neural network $\mathrm{ANN}_{3\div3}$ to generate camera specific RGB output estimates (Fig. 3).

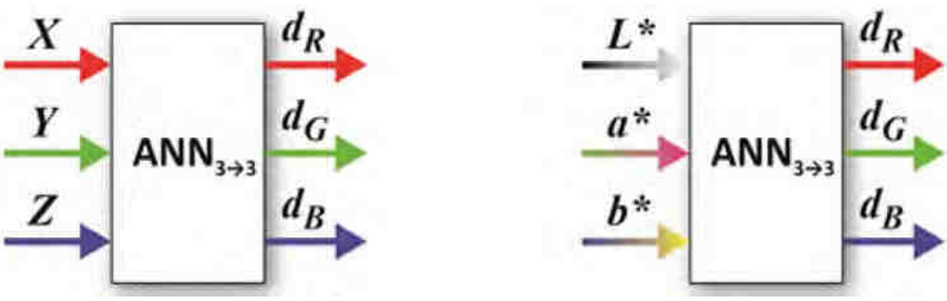

Fig. 3 ANNs for arbitrary camera RGB values recovery from colorimetric values

If such $ANN_{3 \div 3}$ modelling would turn out to give accurate outputs, we could later use it to generate a larger set of camera output-reflectance pairs, and afterwards train $ANN_{3 \div N}$ more accurately.

Experimental

For our experiment we adopted a two layer feed-forward artificial neural network with sigmoid hidden and linear output neurons. This ANN configuration can fit multi-dimensional input-output mapping quite well, when using enough neurons in the hidden layer and being fed by congruent learning set of data. The network was trained with Levenberg–Marquardt backpropagation algorithm.

Learning data sets consisted of input XYZ and $L^*a^*b^*$ values and output RGB values. Input data were calculated from the reflectance values of Gretag Macbeth Digital ColorChecker SG test chart by taking into account A, D50 and D65 standard illuminants and 2° and 10° standard observers. Colour patch reflectances were obtained with "Eye-One iO" spectrophotometer. Output data were averaged central quarter (35-by-35 pixel) patch area RGB values of test chart ColorChecker SG acquired by Nikon D50 digital camera. Test chart consists of 59 greyscale and 81 color patches as shown in Fig. 4.

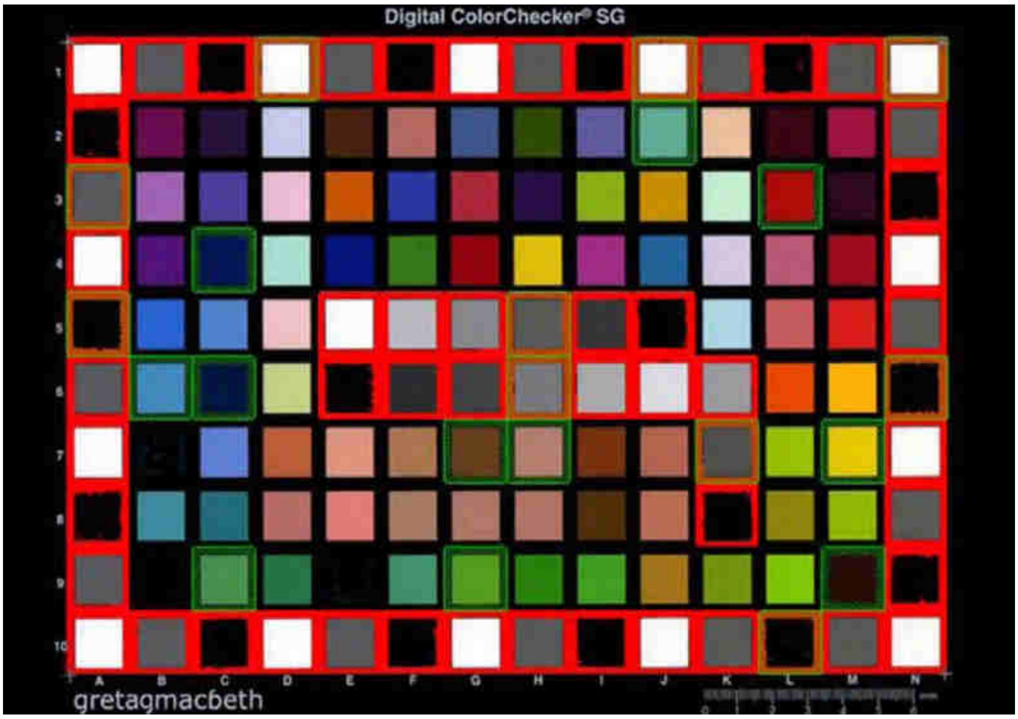

Fig. 4 GretagMacbeth Digital ColorChecker SG with red marked 59 greyscale patches and diagonal grey hatching marked 21 randomly chosen testing set patches (see Fig. 9) (coloured image in electronic version)

Learning of ANNs with 1 to 20 hidden layer neurons were repeated 21 times for each number of neurons. Each repetition generated a different set of training/validation/testing patches. ANN training, validation and testing performances were evaluated *via* normalized Mean Squared Error (MSE).

Results

Every ANN learning set was divided into training, validation and testing set in 70/15/15 percent ratio. ANN is first trained with the training set, and its performance is then verified by MSE gradient of validation set. If gradient is negative, learning algorithm assumes that ANN's performance improves. Just before validation set's MSE gradient starts to ascend, learning is stopped, despite descending training set's declining gradient in order to prevent overfitting.

As we had a limited and quite small learning set, it was reasonable to assume that for each size and complexity of input-output learning ANN data there exists an optimum number of neurons. When we trained ANN with different numbers of hidden layer neurons changing from 1 to 20 with repeating learning for each number 21 times, there was always some quantity range of optimal

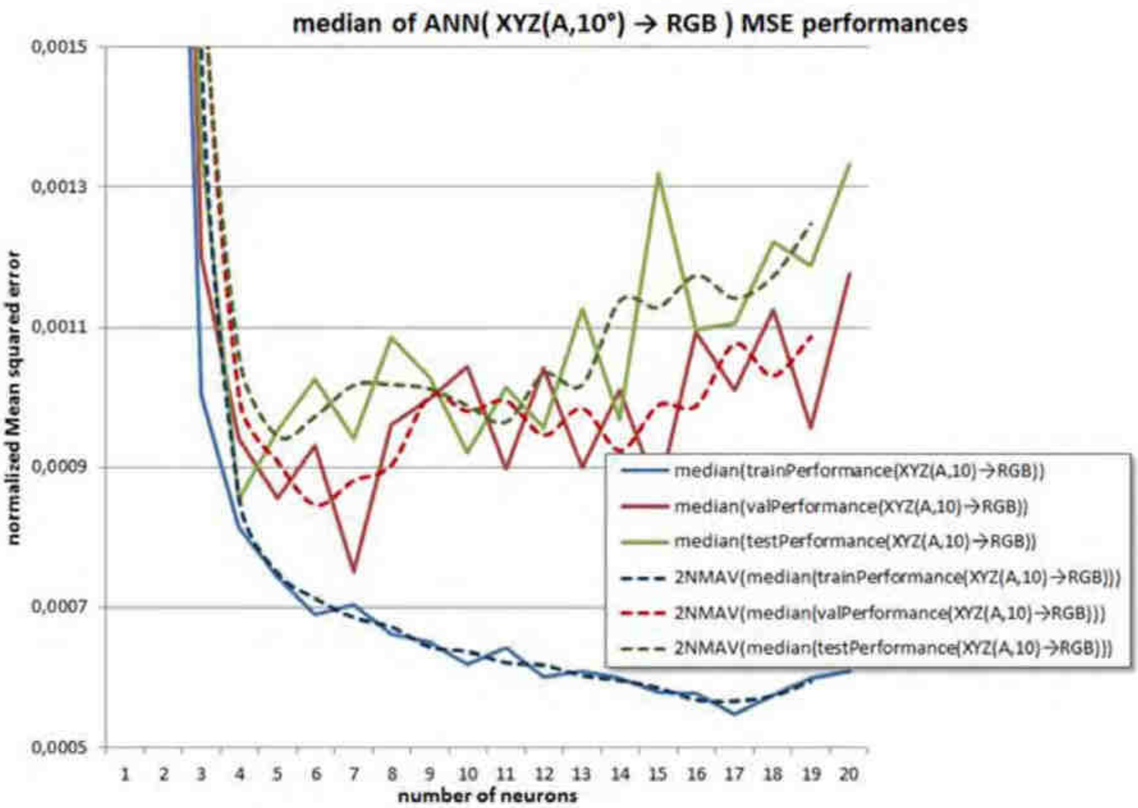

Fig. 5 GretagMacbeth Digital ColorChecker SG with red marked 59 greyscale patches and diagonal grey hatching marked 21 randomly chosen testing set patches (see Fig. 9)

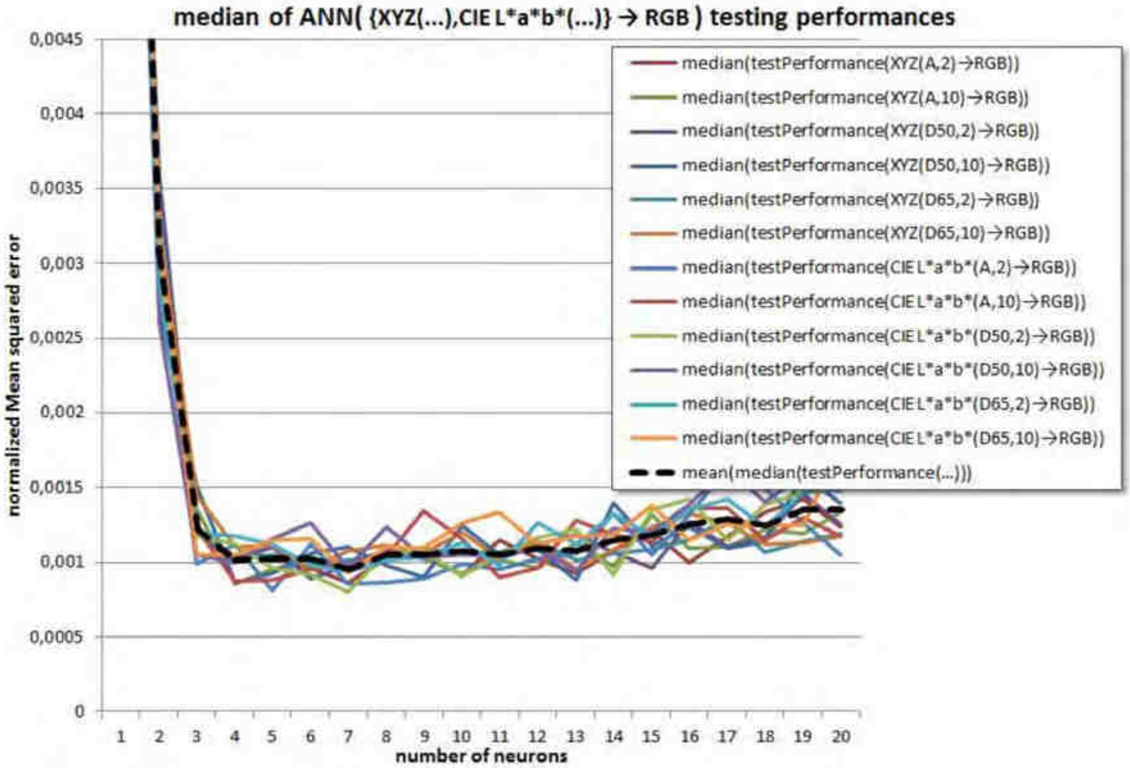

Fig. 6 Normalized MSE testing set performances median values of 21 ANN learning cycles for each number of neurons and each of selected colorimetric values (XYZ or Lab), standard illuminant and observer combination

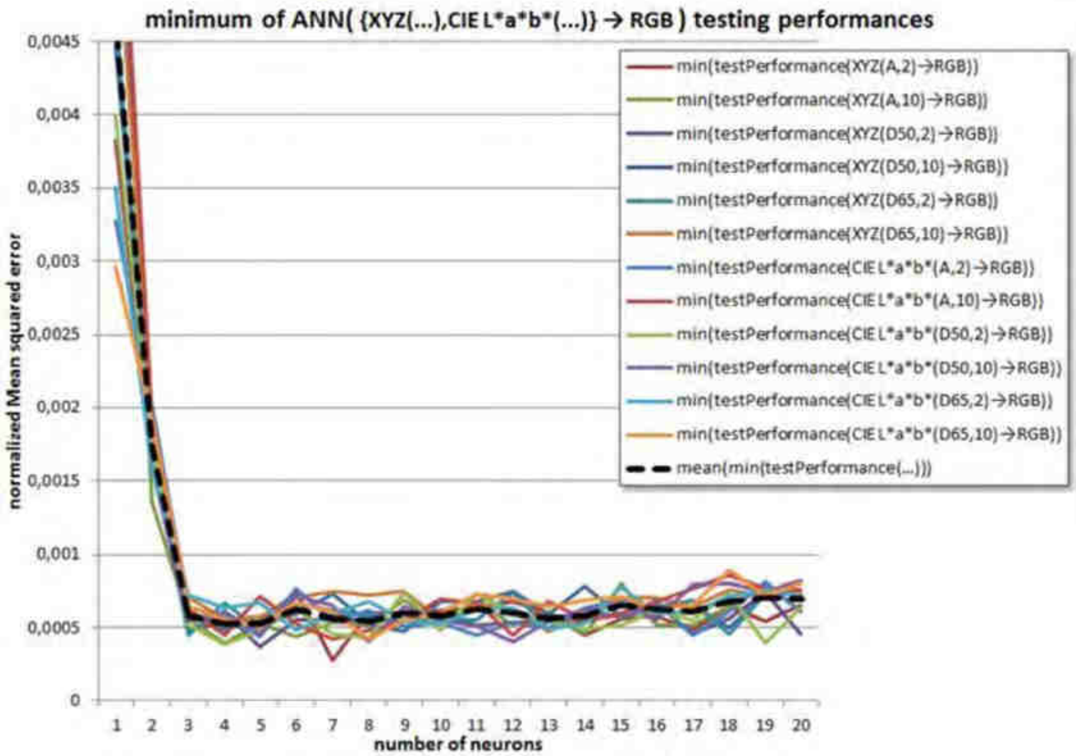

Fig. 7 Normalized MSE testing set performances minimum values of 21 ANN learning cycles for each number of neurons and each of selected colorimetric values (XYZ or Lab), standard illuminant and observer combination

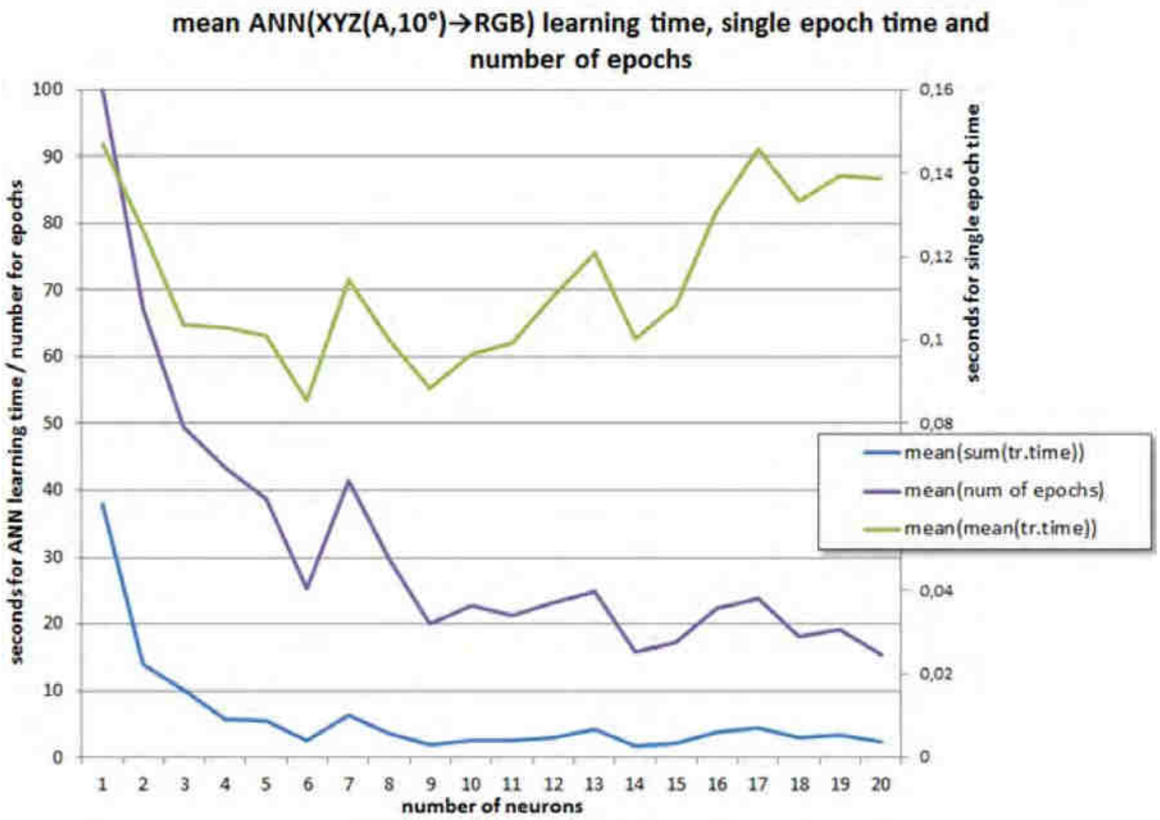

Fig. 8 Average learning time, average number of epochs and average single epoch time of 21 ANN (10°, A illuminant, XYZ to RGB mapping) learning cycles for each number of neurons

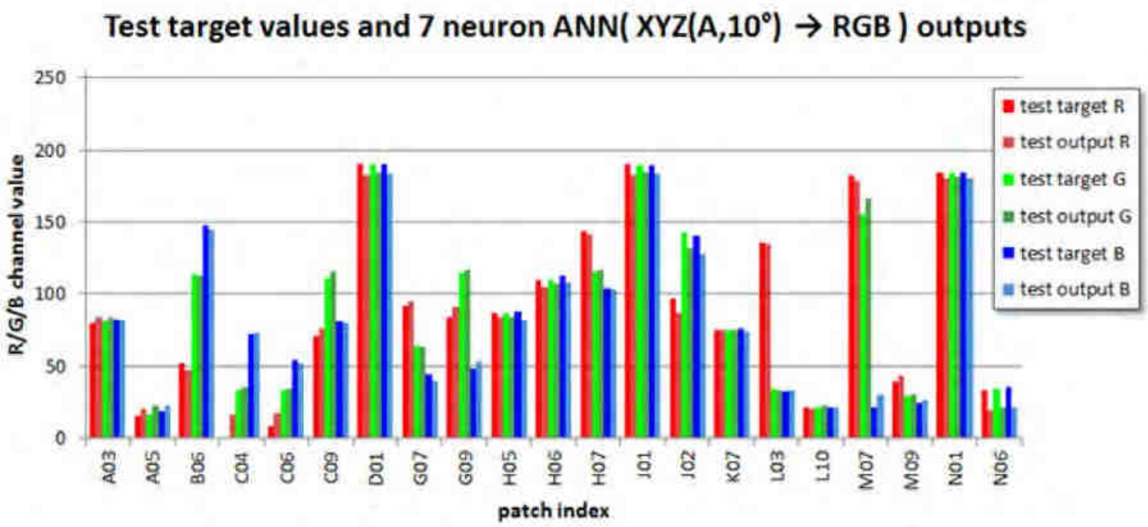

Fig. 9 Side-by-side comparison of RGB test targets and RGB outputs of ANN with 7 neurons in the hidden layer

number of neurons, typically between 4 and 8. Learning cycle's results for 10° standard observer with A illuminant CIE XYZ input and camera RGB output data are shown in Fig. 5.

Similar change of MSE gradient can be found with all learning sets in Fig. 6, where average of median of testing set performance values with lowest area between 4 and 8 is depicted with black dotted line. Average of minimum of testing set performance values (best ANN learning results) shows the same trend (Fig. 7), whereas performance is twice as good.

For a comprehensive experimenting with many repeated learning cycles with different ANN configurations, learning time plays quite an important role. Because of this, we were monitoring learning cycle time parameters. Single epoch time represents one learning step in a complete learning cycle. After many learning epoch repetitions, ANN optimum is reached and learning is stopped. The complete learning cycle time for one learning set is the sum of all single epoch times. As it was expected, we observed that as number of hidden neurons rises, ANN's complexity grows and a single epoch time rises. At the same time, ANN's ability to effectively absorb knowledge rises, and number of learning epochs decreases. Consequently, ANN total training time remains constant at an average of just above 3 seconds. In Fig. 8 average values of these parameters are shown for 21 repetitions of ANN learning cycles. We can also see that learning ANNs with small number of neurons (1, 2, 3) more often needs many epochs to stabilize and training times can be measured in tens of seconds if not stopped by constraint of maximum number of epochs.

Comparison of final results — differences between captured and predicted (ANN outputs) RGB values — is given in Fig. 9, where some values are satisfactorily well matched (for example patches K07, L03) and some differences are still too big (patch N06). Overall results are for our small learning set (Fig. 4 – 140 patches, of which only 81 colour patches) surprisingly good. These results should be improved by further experimenting with larger learning set and additional ANN configurations.

References

[1] Eem J.K., Shin H.D., Park S.O.: Reconstruction of Surface Spectral Reflectances Using Characteristic Vectors of Munsell Colors, Proc. of IS&T and SID's 2nd Color imaging Conference, Nov. 1994, Scottsdale, Arizona, pp. 127-131.

[2] Stigell P.,Miyatab K., Hauta-Kasaric M.: Pattern Recognition and Image Analysis **17**, 233-242 (2007).

[3] Mansouri A., Marzani F.S., Gouton P.: Neural Networks in Two Cascade Algorithms for Spectral Reflectance Reconstruction, IEEE on Image Processing, ICIP, 2005.

[4] Sharma G., Wang S.: Spectrum Recovery from Colorimetric Data for Color Reproductions, Proceedings of SPIE, **4663** (2002).

[5] Cybenko G.: Math. Control, Signal Syst. **2**, 303-314 (1989).
[6] Hornik K., Stinchcombe M., White H.: Neural Netw. **2**, 359-366 (1989).
[7] Ferrari S., Stengel R.F.: IEEE Trans. Neural Netw. **16**, (2005).