Model-reference color reproduction for video system based on back-propagation networks

Po-Rong Chang, Chih-Chiang Tai and *Bao-Fuh Yeh*

Dept. of Communication Engineering, National Chiao-Tung University

Hsin-Chu, Taiwan, Fax: 886-35-710116

ABSTRACT

This paper presents a feedforward multi-layered neural network approach, which is capable of characterizing the colorimetric model of a color cathode ray tube (CRT) monitor driven by a digital red, green, and blue signal source. The model consists of forward and inverse transforms. The forward transform predicts the CIE display color for a given triad of digital RGB video signals. The inverse transform determines the digital RGB drive signals necessary to produce a color of specific CIE coordinates. Nevertheless, the identification of both forward and inverse transforms becomes intractable when the nonlinear distortions are involved in the colormetric model. These distortions result from the violations of the assumptions of phosphor constancy, gun independency, spatical uniformity and monitor characteristic function with an expression of power function. The feedforward neural networks have the capability to learn arbitrary nonlinearity and show great potential for determining the forward transform model of a color monitor. To avoid inverting the complicated nonlinear forward transform directly, an alternative approach to the inverse transform identification is proposed on the basis of Widrow and Winter's neural-based inverse system control scheme. The performance of both transforms was evaluated by the prediction and measurement of 29 test colors chosen to adequately sample the color gamut reproducible by the CRT phosphors.

1 INTRODUCTION

For computer graphics and video service applications, it is desired to specify and reproduce colors on CRT displays accurately in order to produce high-quality images in electronic imaging systems. The colors produced by a CRT display can be expressed in terms of CIE (Commission Internationale de l'Eclairage) colorimetry which is a standardized and widely accepted color-specification system. The system is based on the 1931 CIE color matching functions, from which a set of tristimulus values can be derived. From these tristimulus values, all the CIE color standards can be calculated. Since the CIE color standards are based on additive color mixture, the video monitor produces its colors by the linear combination of red, green, and blue phosphors. The coefficients of the linear combination are termed as the phosphor matrix. The distinct levels of excitation in the three phosphars of a color video monitor are generated by the digital video drive values which are applied to the three guns of the monitor. The drive values are controlled by digital-to-analog converters whose inputs come from application software. The relationship between the video signal level and monitor light output in each of its three channels is called the monitor characteristics function. Typically, it is expressed as the form of a power function. The forward transform model of a color monitor is defined as a composite system of both the monitor characteristic function and phosphor matrix [1]. The forward transform model acts as an emulator which can predicts the CIE display color for a given digital RGB video drive signals. The characteristics of the forward transform model can be identified by measuring all the colors a monitor can produce and then storing them in a large table [2]. Hartmann and Madden [1] used the first-order approximation to fit those colors and represent the forward model. They showed that the approximation model for their particular monitor is valid when the drive levels are above 30 IRE (Institute of Radio Engineers) units. However, the forward tranform became nonlinear for input drive values below 30 IRE. In addition, Cowan [2] indicated that the commonly used parametric expression of a power function

0-8194-1465-4/94/$6.00

did not suffice to describe the monitor characteristics function and sometimes may make large identification errors. To tackle this difficulty, he used a complicated two-parameter expression to improve the performance.

A color video monitor is calibrated completely when the forward transform model is known and then gives the CIE tristimulus values of the colors produced in terms of the drive signals applied to the three guns of the monitor. The inverse transform model, which determines the digital RGB video drive signals needed to produce a color pspecific CIE tristrimulus values, is usually identified by taking the inverse operator on the estimated forward model [1][2][3]. This means that the color difference between the color produced by the inverse transform model and the expected color is dependent on the estimation accuracy of forward transform models. Alternative approach to the inverse transform identification is on the basis of Widrow and Winter's inverse system control scheme using feedforward neural networks [4]. Feedforward multi-layered neural networks have been widely used in the realization of complex nonlinear systems for which the system architecture is not necessarily known. Recent research works [5][6][7] showed that neural networks can uniformly approximate any continuous function with arbitrarily desired accuracy. This leads to a promosing method for modelling both the forward and inverse transforms of a color video monitor.

The results of applying the feedforward neural network to the identification of both forward and inverse transform models were discussed. The model was evaluated by the prediction and measurement of 29 test colors chosen to adequately sample the color gamut reproducible by the monitor phosphors. A comparison of both neural-based method and Hartmann's first approximation model is illustrated in the last section. It is demonstrated that the performance of neural-based method is superior to the first approximation method.

2 NEURAL-BASED MODEL IDENTIFICATION OF COLOR MONITORS

Neural networks have become a very fashionable area of researh with a range of potential applications that spans AI, engineering and science. All the applications are dependent upon training the network with illustrative examples and this involves adjusting the weights which define the strength of connection between the neurons in the network. This can often be interpreted as a system identification problem of estimating the system that transforms inputs into outputs given a set of examples of examples of input-output pairs.

This section focuses on the feasibility of neural networks and their learning algorithms for training the networks to represent forward and inverse transform models of nonlinear color monitor systems. Training a neural network using input-output data from a nonlinear plant can be considered as a nonlinear functional approximation problem. Approximation theory is a classical field of mathematics. From the well-known Stone-Weierstrass theorem [7], it shows that polynomials can approximate arbitraily well a continuous function. Recently, the approximation capability of networks has been investigated [6][7][8] by using the similar concept based on the Stone-Weierstrass theorem. A number of results have been published showing that a feedforward network of the multilayer perceptron type can approximate arbitrarily well a continuous function [5][6]. To be specific, these research works prove that multilayer feedforward networks with as few as single layer and an appropriately smooth hidden layer activation function are capable of arbitrarily accurate approximation to an arbitrary continuous function.

Before applying the feedforward neural networks to the model identification of color monitor system, it is improtant to establish their approximation capabilities to some arbitrary nonlinear real-vector-valued continuous mapping $\mathbf{y} = \mathbf{f}(\mathbf{x}) : D \subseteq R^n \rightarrow R^m$ from input/output data pairs $\{x, y\}$, where D is a compact set on R. For example, the colorimetry of a monitor can be represented by $\mathbf{f}(\mathbf{x})$ when the domain D and its image space $\mathbf{f}(D)$ are defined in the RGB and $CIEXYZ$ color spaces respectively. Consider a feedforward network $NN(\mathbf{x}, \mathbf{w})$ with $\mathbf{x}$ as a vector representing inputs and $\mathbf{w}$ as a parameter weighting vector that is updated by some learning rules. It is desired to train $NN(\mathbf{x}, \mathbf{w})$ to approximate the mapping $\mathbf{f}(\mathbf{x})$ as close as possible. The Stone-Weierstrass theorem [7] shows that for any continuous function $\mathbf{f} \in C^1(D)$ with respect to $\mathbf{x}$, a compact metric space, an $NN(\mathbf{x}, \mathbf{w})$ with appropriate weight vector $\mathbf{w}$ can be

found such that $\| NN(\mathbf{x}, \mathbf{w}) - \mathbf{f}(\mathbf{x}) \|_{\mathbf{x}} < \epsilon$ for an arbitrary $\epsilon > 0$, where $\| \mathbf{e} \|_{\mathbf{x}}$ is the norm defined by

$$\| \mathbf{e} \|_{\mathbf{x}} = \sup_{\mathbf{x}} \{ \| \mathbf{e}(\mathbf{x}) \| : \mathbf{x} \in D, \ \| \cdot \| \text{ is the vector norm } \} \tag{1}$$

For network approximators, key equations are how many layers of hidden units should be used, and how many units are required in each layer. Cybenco[5] and Wang et al have shown that the feedforward network with a single hidden layer can uniformly approximate any continuous function to an arbitrary degree of exactness–providing that the hidden layer contains a sufficient number of units. However, it is not cost-effective for the practical implementation. Nevertheless, Chester [8] gave a theoretical support to the empirical observation that networks with two hidden layers appear to provide high accuracy and better generalization than a single hidden layer network, and at a lower cost (i. e. fewer total processing units). Since, in general, there is no prior knowledge about the number of hidden units needed, a common practice is to start with a large number of hidden units and then prune the network whenever possible. Additionally, Huang and Huang [9] gave the lower bounds on the number of hidden units which can be used to estimate its order.

2.1 Forward and Inverse System Model Identification by Feedforward network

In general, system identification is usually recognized as a process to train networks to represent nonlinear dynamical systems and their inverses. This would be distinctly helpful in achieving the desired output signal of the system. The issue of identification is perhaps of even greater importance in the field of adaptive control and signal processing systems [4]. Since the plant in an adaptive control varies in operation with time , the adaptive control must be adjusted to account for the plant variations.

The procedure of training a neural network to represent the forward dynamics will be referred to as forward model identification. The basic configuration for achieving this is shown schematically in Fig. 1. A feedforward neural network with a single hidden layer is placed in parallel with the system and receives the same input $\mathbf{x}$ as the system. The system output provides the desired response $\mathbf{d}$ during training. The purpose of the identification is to find the appropriate weights $w'_{ij}s$ of the network with response $\mathbf{o}$ that matchs the response $\mathbf{y}$ of the system for a given set of inputs $\mathbf{x}$. During the identification, the norm of the error vector, $\| \mathbf{e} \| = \| \mathbf{d} - \mathbf{o} \|$, is minimized through a number of weight adjustments by the delta-bar-delta learning rule. In our case, those weights are updated by minimizing the system error, $E = \sum_{k=1}^{N} \| \mathbf{e}_k \|^2$, by the same algorithm, where $\mathbf{e}_k = (\mathbf{d}_k - \mathbf{o}_k)$.

Fig. 1 shows use of a feedforward neural network for direct modeling of an unknown system to obtain a close approximation to its responses. By changing the configuration, it is possible to use the feedforward network for inverse modeling to obtain the reciprocal of the unknown system's transfer function when the system is invertible. In constrast to forward system characteristics identification, the system output $\mathbf{o}$ is used as neural network input, as shown in Fig. 2. The unknown system's input, $\mathbf{x}$, is the desired response of the feedforward network. Thus, the error vector of network training is computed as a $\mathbf{x} - \mathbf{o}$. The system error to be minimized through learning is therefore $\tilde{E} = \sum_{k=1}^{N} \| \mathbf{x}_k - \mathbf{o}_k \|^2$. The neural network trained by the delta-bar-delta algorithm will implement the mapping of the system inverse. Once the network has been successfully trained to mimic the system inverse, it can be used directly for feedforward inverse control. In other words, the inverse model is cascaded with the controlled unknown system in order that the composed system results in an identity mapping between desired response (i.e. the network inputs) and the contolled system output.

As mentioned, it is assumed that the system is invertible. Then there exists an injective mapping which represents its inverse. If it is not true, a major problem with system inverse identification arises when the system invesre is not uniquely defined.

A second approach to inverse modeling which aims to overcome these problems is known as specialized inverse

learning[12]. As pointed out in Psaltis et. al. [12], the specialized method allows the training of the inverse network in a region in the expected operational range of the system. On the other hand, the generalized training procedure produces an inverse over the operating space which may be uniquely defined. Fortunately, the mapping of a color monitor system may have a unique inverse. Thus, we could apply the direct invesre method as illustrated in Fig. 2 to find the approximation of the inverse.

3 ILLUSTRATED EXAMPLES

A Minolta chroma meter CS-100 is used to measure the CIE tristimulus values X, Y, and Z. Overall 512 measured training data points are generated from the current commonly used sampling algorithms. The algorithms sample the monitor coordinates linearly with respect to the monitor coordinates. For example, if $512 = 2^9$ samples are to be taken, $8 = 2^3$ values on each of the red, green and blue guns would be taken, and the 512 samples would consist of all combinations of the gun values. The gun values would sample the range of possible gun values linearly. Thus, if there were 256 possible values for each gun, running from 0 to 255, the 8 values chosen would be 0, 36, 72, 109, 145, 182, 218, and 255. If the full range is not to be used because of non-zero turn-on values for guns or blooming, the values would be scaled between the end points of the range. The linear sampling in monitor color coordinates roughly approximates the sampling in a perceptually uniform color space when monitor is designed to have its device color coordinates relatively equally spaced visually.

Two feedforward neural networks with one hidden layer and two hidden layers are designed to learn the forward and inverse system models respectively. According to Huang and Huang's suggestions [9], it is possible to estimate the lower bounds on the numbers of hidden units in both neural networks. The numbers of hidden units for the networks associated with the forward and inverse models are chosen as 50 and 30 for each hidden layer respectively. To perform delta-bar-delta learning algorithm, the parameters related to weight update rule and learning rate update rule are chosen as $\eta(t)|_{t=0} = 0.03, \alpha = 0.8, \kappa = 0.00005, \phi = 0.4$ and $\theta = 0.3$. Since the degree of reproduction is judged by the color difference metric defined in perceptual color space $CIELUV$, instead of $CIEXYZ$ from Fig. 1, the system error E of the three-layer network is given by

$$E = \Delta \bar{E}^{*^2}_{uv} = \frac{1}{N} \sum_{k=1}^{N} (\Delta E^*_{uv})^2_k \tag{2}$$

where $\Delta \bar{E}^*_{uv}$ = the root mean square of the color difference between the network output (o) and the desired output ($y = d$), $\Delta E^*_{uv} = \sqrt{\Delta L^{*^2} + (\Delta u^*)^2 + (\Delta v^*)^2}$, N = number of sample color patches, and L^*, u^*, v^* are the coordinates defined in $CIELUV$ space. Notice that the color coordinates in $CIELUV$ space can be obtained from the three tristimulus values X, Y and Z.

By inputting the $N(= 512)$ training color patches into the NEC 5-D color monitor and performing the delta-bar-delta algorithm on the neural-based forward models, the training errors can be found as 0.85 for $CIELUV$ color space.

In contrast to the identification of forward modelling, from Fig. 2, the system error of the four-layer network for the inverse modelling is defined in RGB space and given by

$$E = \bar{d}^2_{RGB} = \frac{1}{N} \sum_{k=1}^{N} (\Delta R^2 + \Delta G^2 + \Delta B^2)_k \tag{3}$$

where $\| x_k - o_k \|^2 = \frac{1}{N} (\Delta R^2 + \Delta G^2 + \Delta B^2)_k$.

Meanwhile, the color difference defined in RGB space cannot provide the metric in uniform color spaces of which the corresponding measure is in agreement with the human visual system (HVS). The HVS allows us to process color reproducton in a controlled way in accordance with the visual perception of the reproduced color images. The $CIELUV$ space is related to the $CIEXYZ$ color space by known formula. However, the RGB to XYZ transformation is dependent on the display device. If the training system error defined in RGB space is 0.62, its corresponding root mean square of the color difference defined in $CIELUV$ space is 0.54. By the transformation formula, the root mean square of the color difference defined in $CIEXYZ$ space can be found as 0.43.

Next, we would like to validate the estimated forward and inverse models by test colors which consists of (1) the 24 patches of the MACBETH COLOR CHECKER chart, (2) one baryta white patch, (3) the colors produced by the red, green, blue, and white full monitor drives. The chromametric measurements of MACBETH COLOR CHECKER chart and baryta patch were made under incandescent tungsten illumination to their CIE chromaticities and illuminance factors.

The test colors CIE chromaticities and luminances served as input to the inverse-transform model, which can compute the digital red, green, and blue video drive values required to produce the specified CIE stimuli on the color monitor. [1] introduced three indexes to evaluate the model's color reproduction capability, that is, chromaticity difference, luminance factor difference, and color difference metric defined in $CIELUV$ space. The first two indexes are defined as follows

$$CHROM.DIFF. \triangleq [(x_M - x)^2 + (y_M - y)^2] \tag{4}$$

and

$$\begin{aligned} \%Y &\triangleq \text{Luminance factor difference} \\ &\triangleq 100 * \mid (Y_M - Y)/Y_M \mid \end{aligned} \tag{5}$$

where $x_M y_M Y_M$ are the resulting CIE chromaticities and luminance factor measured on the CRT, and xyY are the specified CIE chromaticities and luminance factor input to the inverse model (x_d, y_d, and Y_d) or the calculated chromaticities and luminance (x_f, y_f, and Y_f) from the forward model. The mean values of the chromaticity differences (or %Y) over the 29 test colors for both the neural-based forward model and Hartwann's forward model are equal to 0.0057 (1.734) and 0.0124 (2.949) respectively. For the inverse transform, the mean values of the chromaticity differences (or %Y) for neural model and Hartwann's model are 0.0067 (1.0987) and 0.02719 (3.3286) respectively. Figures 3 and 4 show a comparison of the chromaticity differences for the Hartwann's and neural-based inverse models in the CIE x,y coordinate system. It is seen that the chromaticity differences of neural-based inverse model are less than Hartwann's inverse model. As a result, this implies that the degree of color reproduction of neural inverse model is better than Hartwann's inverse model. Similarly, a comparison of both Hartwann's and neural-based forward models is illustrated in Figures 5 and 6.

Unfortunately, both chromaticity and luminance factor differences cannot provide a measure in accordance with HVS to the reproduced color images. Another index of merit to color reproduction is the color difference metric defined in $CIELUV$ space, $\Delta \bar{E}^*_{uv}$. Table 1 shows a comparison of the $CIELUV$ color differences resulted from neural-based and Harwann's inverse models over the 29 test color patches. Similarly, the comparison of both neural-based and Hartwann's forward models is shown in Table 2. It is shown that that both the neural-based inverse and forward models yield the color difference r.m.s. errors 2.9073 and 2.6375 respectively which are less than those (15.2664 and 9.8553) obtained from Hartwann's methods by about an order of magnitude. An useful indication is the size of the minimum perceptible color difference and this is of the order of 0.0004 in the CIE 1976 u'-v' uniform chromaticity coordinates. In television assessments, a unit ten times this value, i.e. 0.004 has often been used as a just noticeable difference (JND) under critical conditions of viewing. This corresponds to 2 to 3 units in $CIELUV$ space. Undoubtedly, the values of $\Delta \bar{E}^*_{uv}$ for both neural-based models fall in the range of JND.

4 CONCLUSION

The use of feedforward multi-layered neural networks promises to bring a new level of accuracy to the prediction of video display colorimetry and monitor calibration. The process of monitor calibration is performed by training a feedforward neural network to represent the forward transform of the target color CRT monitor. This would allow the monitor's output for any set of input values to be predicted from the input values. The inverse transform has been obtained by performing Widrow and Winter's inverse system identification. The trained inverse transform is simply cascaded with the color CRT monitor in order that the composed system results in an identity mapping between desired CIE coordinates (i.e. network inputs) and the output display colors. The results have shown that the $CIELUV$ color difference r.m.s. errors of both neural-based calibration and prediction models are less than the Hartwann's models by an order of magnitude and also fall the range of JND.

References

[1] W. T. Hartmann and T. E. Madden, "Prediction of display colorimetry from digital video signals," *Journal of Imaging Technology,* vol. 13, no. 4, August 1987.

[2] W. B. Cowan, "An inexpensive scheme for calibration of a color monitor in terms of CIE standard coordinates," *Computer Graphics,* vol. 17, pp. 315–321, 1983.

[3] W. B. Cowan and N. L. Rowell, "On the gun independence and phosphor constancy of color video monitors" *Color Research and Application,* vol. 11, pp. S34–s38, 1986.

[4] B. Widrow, and R. Winter, "Neural nets for adaptive filtering and adaptive pattern recognition," *IEEE Computer Mag.,* vol. 21, pp. 25–39, March, 1988.

[5] G. Cybenko, "Approximations by superposition of a sigmoidal function," *Math. Control Systems Signals,* vol. 2, no. 4, 1989.

[6] K. Hornik, M. Stinchcombe, and H. white, "Universal approximation of an unknown mapping and its derivatives using multilayer feedforward networks, "*Neural Networks,* vol. 3,pp. 551–560, 1990.

[7] Cotter, M. E., "The Stone-Weierstrass theorem and its applications to neural nets." *IEEE Trans. Neural Networks,* vol. 1, pp. 290–295 1990.

[8] D. Chester, "Why two hidden layers are better than one," in *Proc. Int. Joint Conf. Neural Networks,* Washington, DC, IEEE, June 1989, pp. I-613–I-618.

[9] S. C. Huang and Y. F. Huang, "Bounds on the number of hidden neurons in multilayer perceptrons," *IEEE Trans. Neural Networks,* vol. 2, no. 1, pp. 47–55, Jan. 1991.

[10] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning internal representations by error propagation," in *Nature,* vol. 323, pp. 533–536, 1986.

[11] R. A. Jacobs, " Increased rates of convergence through learning rate adaption," *Neural Networks,* vol. 1, pp. 295–307, 1988.

[12] D. Psaltis, A. Sideris, and A. A. Yamamura, "A multilayer neural network controller," *IEEE Control Systems Magazine,* vol. 8, pp. 17–21, 1988.

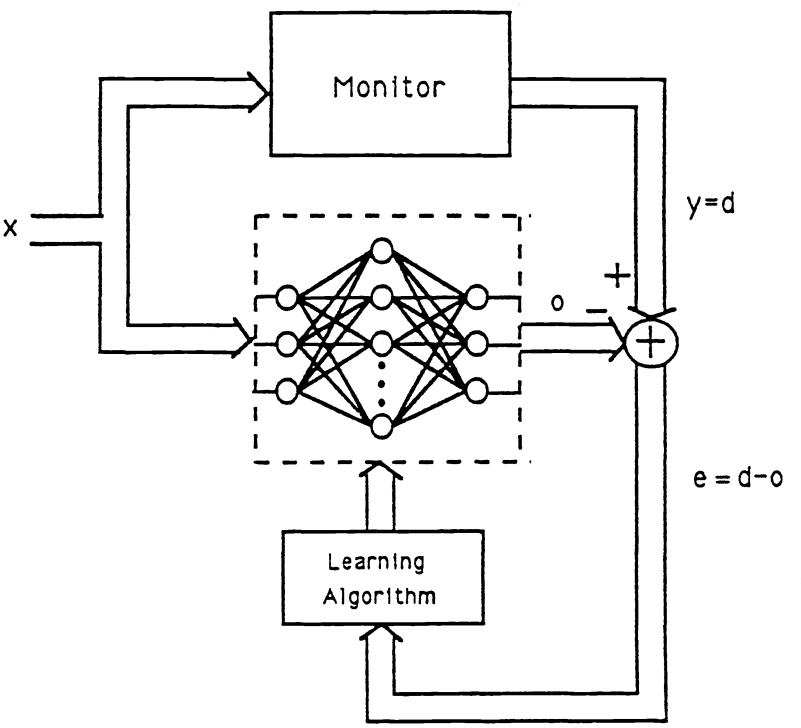

Fig. 1 Forward model identification of color monitor.

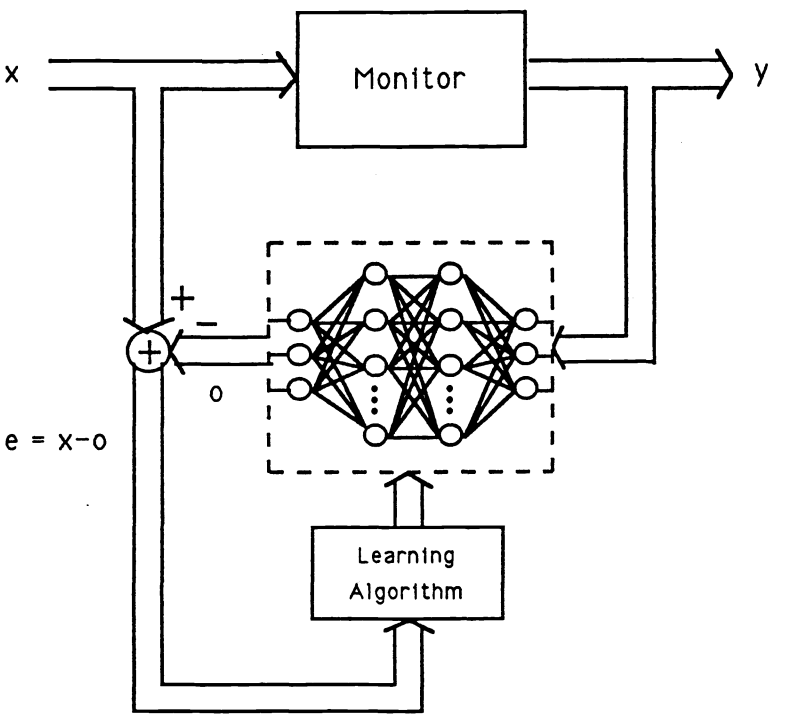

Fig. 2 Plant inverse identification of color monitor.

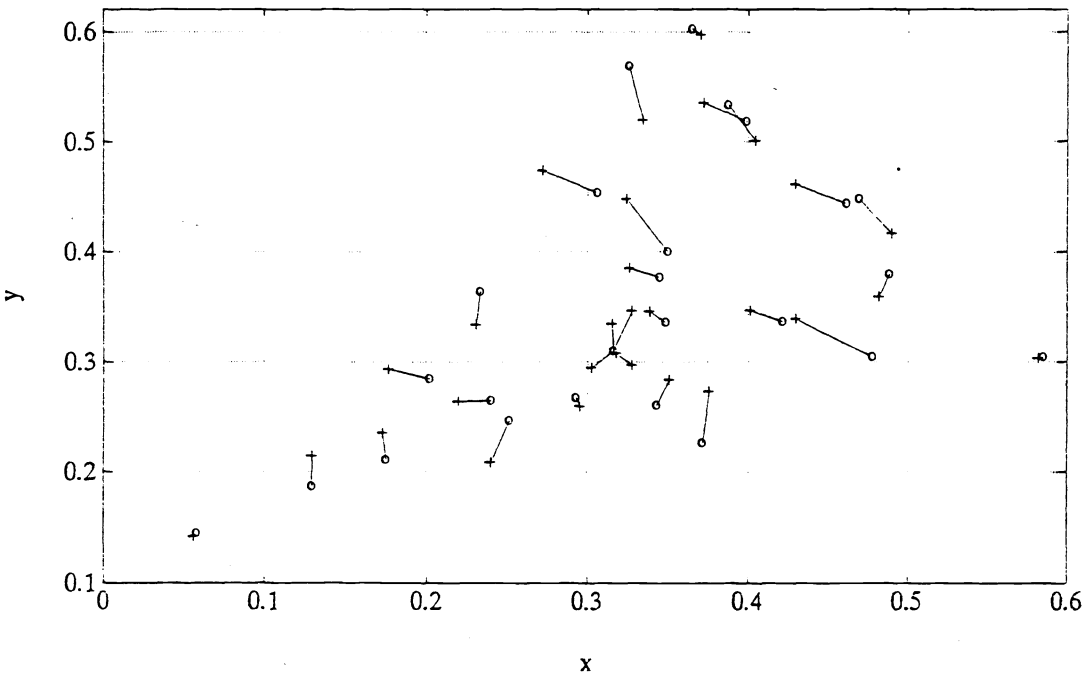

Fig. 3 Chromaticity diagram of measurement data ('o') and Hartwann's prediction model outputs ('+').

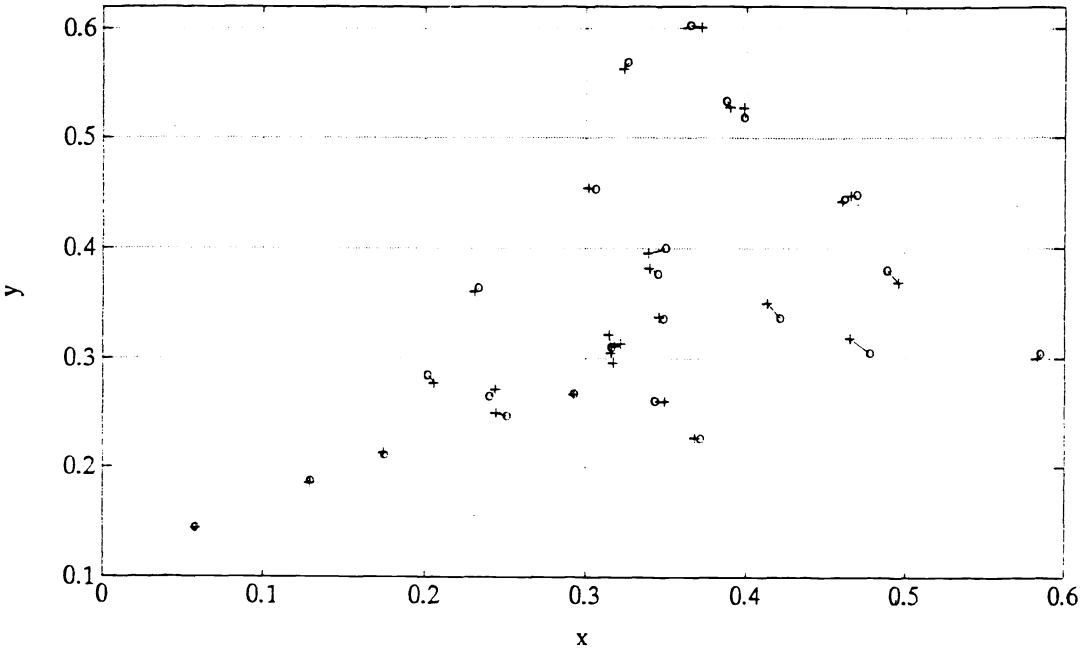

Fig. 4 Chromaticity diagram of measurement data ('o') and neural-based prediction model outputs ('+').

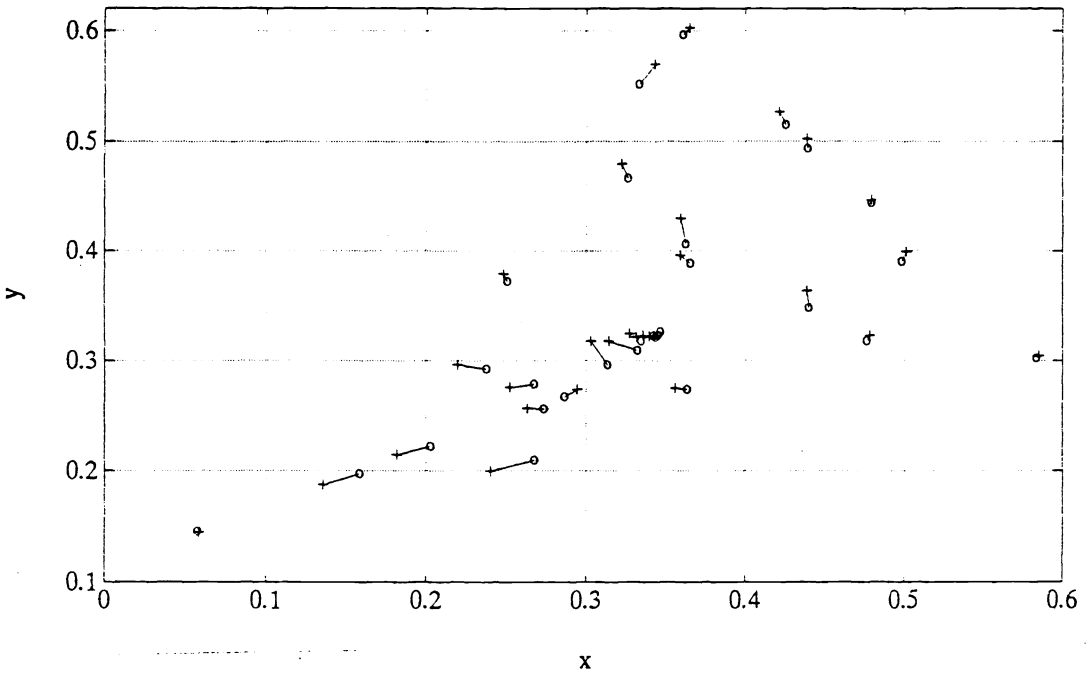

Fig. 5 Chromaticity diagram of measurement data ('o') and Hartwann's forward model outputs ('+').

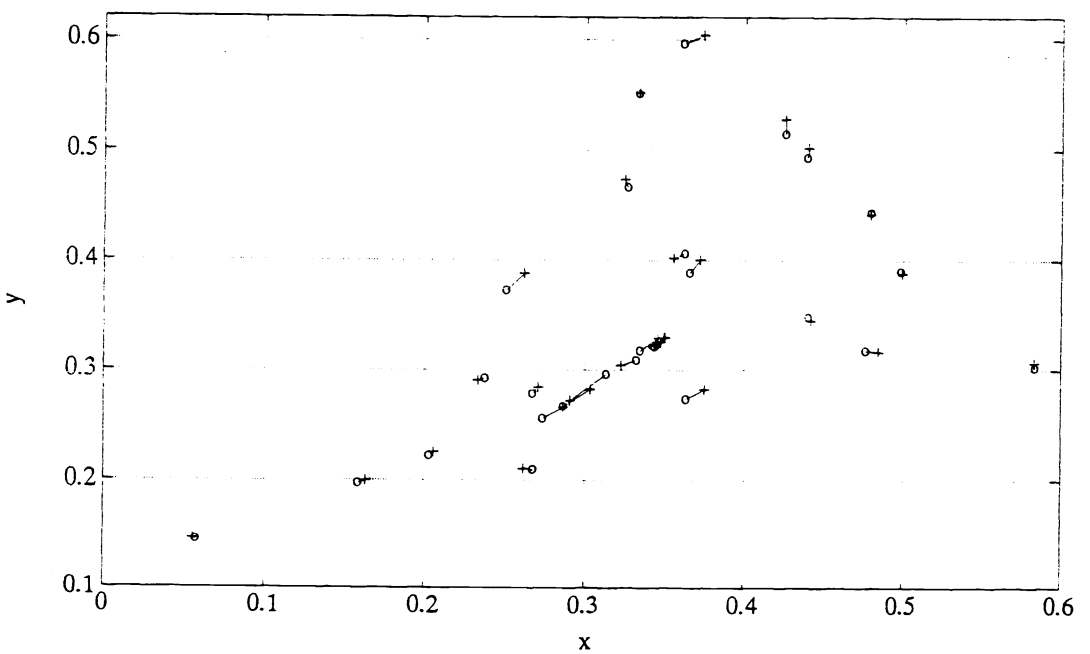

Fig. 6 Chromaticity diagram of measurement data ('o') and neural-based forward model outputs ('+').