# Color Classification Using Color Vision Models

*Tuija Jetsu*[*], *Yasser Essiarab*[*], *Ville Heikkinen*[**], *Timo Jaaskelainen*[**] *and Jussi Parkkinen*[*]
[*] *Department of Computer Science and Statistics,* [**] *Department of Physics and Mathematics, University of Joensuu, Finland*

## Abstract

*According to recent physiological research results, there are lots of individual differences already at the detection level of our color vision system. It is not completely clear yet, how the other levels of color vision system compensate the detection differences. It is known that people with different kinds of color vision system properties still experience similar color sensations. This could be explained by the fact that instead of detecting and analysing colors in exactly the same way, we all just have learned to classify colors in a certain way. In this article, we experiment with four models developed for replicating certain properties of human color vision. We examine the color classification abilities of these models and show the differences and similarities in their behavior.*

## Introduction

After the light reflecting from an object has travelled through the external and intermediate layers of the human eye, it enters the first part of our color detection system: the retina. It has been shown by measurements of a living human eye [1, 2] that people with a very divergent distribution of short-, middle- and long-wavelength sensitive cones on the retina can still have perfectly normal color vision. Also the spectral sensitivities of cones on the retina differ between individuals [3, 4].

Existing color vision models (for example Ingling and Tsou [5], Bumbaca and Smith [6], De Valois and De Valois [7] and Guth [8]) assume that all people process the same color signal exactly in the same way. However, according to the previously mentioned physiological experiments, there are lots of individual differences already at the detection level of our color vision system. How the following neural levels of color vision compensate the detection differences, is not completely clear yet. One approach to explaining the similar color vision of people with different kinds of color vision system properties could be that instead of detecting and analysing each color in exactly the same way, we all just have learned to classify colors in a certain way, which eventually gives similar results to everyone else's color sensation.

In this article, we examine the color classification abilities of four existing color vision models [5, 6, 7, 8]. All these models have the cone responses for color signals as their starting point and red-green and blue-yellow opponent values as output. Processing steps of the models differ from each other, and we wanted to test, whether systems with different properties (like people with differently structured visual system) are able to classify colors in the same way. Our goal in this paper is to find out, whether it is easy to divide the output of the models into color classes or not. It is possible that with more complicated classification algorithms and/or with very careful pre-processing of data the classification results could still be improved. However, our main interest was not to find an optimal classification algorithm for color, but to see if it is possible to perform reasonable classification for model outputs with a reasonably simple method.

## Models

The color classification is performed using output of the following models: Ingling and Tsou [5], Bumbaca and Smith [6], De Valois and De Valois [7] and Guth [8].

Ingling and Tsou model is a simple linear one-opponent stage model. They present two different sets of formula, one for dark- and another for light-adapted conditions (threshold and suprathreshold forms). We have used the suprathreshold form, where impact from S cones is present also in red-green channel. This approach is analog to the fact of indiscriminate receptive field surround mentioned in De Valois and De Valois model description. The processing of color signal in Ingling and Tsou's model starts by multiplying the incoming signal by Smith and Pokorny cone fundamentals in order to get cone responses $L$, $M$ and $S$. After that, opponent stage responses are calculated by summing up cone responses $r-g$, $b-y$, and $V_\lambda$ (Eq. 1). We use $r-g$ and $b-y$ responses in our experiments.

$$\begin{bmatrix} r-g \\ b-y \\ V_\lambda \end{bmatrix} = \begin{bmatrix} 1.2 & -1.6 & 0.4 \\ 0.24 & 0.105 & -0.7 \\ 0.6 & 0.4 & 0 \end{bmatrix} \begin{bmatrix} L \\ M \\ S \end{bmatrix} \quad (1)$$

Bumbaca and Smith's color vision model for computer vision applications also has only one opponent stage in their model, but unlike Ingling and Tsou, they are using logarithmic cone sensitivity functions which causes their model to have nonlinear features. Their goal was to build a computer vision system, which would take the advantage of color vision discrimination capabilities of the human color vision. In Bumbaca and Smith's model, the incoming signal is first multiplied by Smith and Pokorny cone fundamentals in order to get cone responses $L$, $M$, and $S$. The nonlinear response of cones, $L^*$, $M^*$, and $S^*$ , (Eq. 2-4) is simulated by taking logarithm of $L$, $M$ and $S$ signals. The nonlinear cone responses are summed in order to form achromatic and chromatic channels $A$, $C_1$, and $C_2$ (Eq. 5). We use $C_1$ and $C_2$ responses in our experiments.

$$L^* = log\, L \quad (2)$$

$$M^* = log\, M \quad (3)$$

$$S^* = log\, S \quad (4)$$

$$\begin{bmatrix} A \\ C_1 \\ C_2 \end{bmatrix} = \begin{bmatrix} a & 0 & 0 \\ 0 & u_1 & 0 \\ 0 & 0 & u_2 \end{bmatrix} \begin{bmatrix} \alpha & \beta & 0 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} L^* \\ M^* \\ S^* \end{bmatrix} \quad (5)$$

Parameters $a$, $u_1$ and $u_2$ are adjusted so that just-noticeable difference in perception in $AC_1C_2$ space is a sphere of radius 1. $\alpha$ and $\beta$ are scaling parameters for the achromatic channel. In our calculations, we used the following values defined in the original paper of Ingling and Tsou: $a = 22.6, u_1 = 41.6, u_2 = 10.5, \alpha = 0.7186$ and $\beta = 0.2814$.

De Valois and De Valois have built their three-stage linear model based on the biological facts of human color vision. Their paper presents two different possibilities for receptive field behavior: discrete and indiscriminate model. Models differ from

each other so that in discrete version cells with L or M cone center are assumed to have no effect from S cones in the surroundings. Indiscriminate version sums together all kind of cells in the receptive field surroundings. We have considered in this paper only the indiscriminate version. The processing in De Valois and De Valois's model starts also by multiplying the incoming signal by Smith and Pokorny cone fundamentals in order to get cone responses $L$, $M$ and $S$. From cone responses, cone opponency signals $L_O$, $M_O$, and $S_O$ are calculated using the receptive field theory (Eq. 6). Cone opponency signals are still summed together to get perceptual opponency signals $RG$, $BY$, and $A$ i.e. red-green, blue-yellow, and achromatic channels (Eq. 7). We use $RG$ and $BY$ responses in our experiments.

$$\begin{bmatrix} L_O \\ M_O \\ S_O \end{bmatrix} = \begin{bmatrix} 6 & -5 & -1 \\ -10 & 11 & -1 \\ -10 & -5 & 15 \end{bmatrix} \begin{bmatrix} L \\ M \\ S \end{bmatrix} \quad (6)$$

$$\begin{bmatrix} RG \\ BY \\ A \end{bmatrix} = \begin{bmatrix} +1 & -1 & +1 \\ -1 & +1 & +1 \\ +1 & +1 & +1 \end{bmatrix} \begin{bmatrix} 10 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} L_O \\ M_O \\ S_O \end{bmatrix}$$

$$= \begin{bmatrix} 90 & -115 & +25 \\ -130 & 95 & 35 \\ -10 & -5 & 15 \end{bmatrix} \begin{bmatrix} L \\ M \\ S \end{bmatrix} \quad (7)$$

Guth's nonlinear three-stage model, ATD95, collects together the results of many previous versions. After ATD95, it has still been deveoped further the latest release being ATD04 [9], but the changes made for the later revisions of Guth's model have no big effect on the color discrimination properties of the model. In this article, we use an implementation based on Guth's ATD95 model. For practical experiments in this article in order to make model comparable to others, we are assuming self-adaptation (test stimulus was also the adapting stimulus) i.e. there is no effect from the surrounding signals. Calculations for Guth's model are presented in Equations 8 - 14), where Equations 8 - 11 describe the calculation of nonlinear cone responses, adding noise and gain control. In Equations 12 and 13 initial responses for the first and second ATD stages are calculated. Initial responses are still further compressed using Equation 14. We use responses $T_2$ and $D_2$ for describing the color in our experiments.

$$L = [0.66(0.2435X + 0.8524Y - 0.0516Z)]^{0.70} + 0.024 \quad (8)$$

$$M = [1.0(-0.3954X + 1.1642Y + 0.0837Z)]^{0.70} + 0.036 \quad (9)$$

$$S = [0.43(0.04Y + 0.06225Z)]^{0.70} + 0.31 \quad (10)$$

$$R_g = R\frac{\sigma}{\sigma + R}, \quad (11)$$

where $R \in \{L, M, S\}$ and $\sigma = 300$.

$$\begin{bmatrix} A_{1i} \\ T_{1i} \\ D_{1i} \end{bmatrix} = \begin{bmatrix} 3.57 & 2.64 & 0 \\ 7.18 & -6.21 & 0 \\ -0.70 & 0.085 & 1.00 \end{bmatrix} \begin{bmatrix} L_g \\ M_g \\ S_g \end{bmatrix} \quad (12)$$

$$\begin{bmatrix} A_{2i} \\ T_{2i} \\ D_{2i} \end{bmatrix} = \begin{bmatrix} 0.09 & 0 & 0 \\ 0 & 0.43 & 0.76 \\ 0 & 0 & 1.00 \end{bmatrix} \begin{bmatrix} A_{1i} \\ T_{1i} \\ D_{1i} \end{bmatrix} \quad (13)$$

$$R_j = \frac{R_{ji}}{200 + |R_{ji}|}, \quad (14)$$

where $R \in \{A, T, D\}$ and $j \in \{1, 2\}$.

## Experiments

We tested the classification performance of the four color vision models mentioned above. Classification of colors was performed using the responses of the last stage of each model using only the color opponent channels (red-green and blue-yellow) forming a 2-dimensional color space. For the classification, we used a simple subspace method [10] and represented each color class by its first eigenvector. Also the location of the class related to the origin of the color space was taken into account by using the mean of each class. Algorithm used for the subspace classifier:

```
c = sample to be classified
for each class j
  e = eigenvector representing class j
  m = mean of class j
  e_c = abs(dotproduct(e,c))
  m_c = sign(dotproduct(m,c))
  res[j] = e_c*m_c
end
classify c to class j, where res[j] is largest
```

Classifier was trained separately for each color vision model. If necessary, the data was always centered so that the origin of training color set was located at point (0,0). The classifier was trained with half of the color samples of the Munsell Book of Color - Glossy Collection. Munsell colors were divided into 10 classes according to the Munsell hue categorization (B, BG, G, GY, Y, YR, R, RP, P, PB) and each class was randomly divided into two. The whole book contains in total 1600 samples, which means that model was trained with 800 samples. The rest of samples were used as test set. Implementations of color vision models and classification algorithms were done with Matlab. Spectral data with 1 nm accuracy was used as initial input to the models. Munsell data at CIE A*B* space is shown in Figure 1.

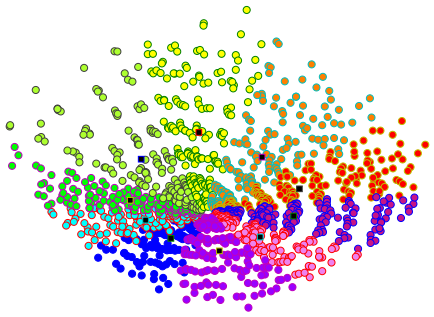

***Figure 1.*** *Munsell data at CIE A*B* space. Dark squares are the representatives of each color class.*

Figures 2 - 5 show the color distribution of Munsell colors using the 2-dimensional output of the four color vision models. Numerical data representing the accuracy of classification is shown in the form of confusion matrices separately for each model. Values are percentages of samples classified to a certain hue class. Each row shows how samples belonging to one color class are classified as different colors. Percentages don't necessarily sum up to exactly 100 procent because of rounding of the results.


©2008 Society for Imaging Science and Technology

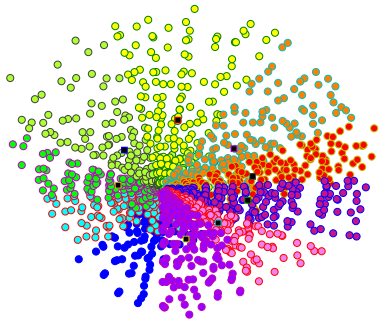

***Figure 2.*** *Munsell data at the 2nd stage of Ingling and Tsou model. Dark squares are the representatives of each color class.*

**Confusion matrix for classification of test samples with Ingling and Tsou model.**

| | B | BG | G | GY | Y | YR | R | RP | P | PB |
|---|---|---|---|---|---|---|---|---|---|---|
| B | 49.3 | | | | | | | | | 50.7 |
| BG | 50 | 35.9 | | | | | | | | 14.1 |
| G | 19.4 | 47.2 | 19.4 | | | | | | | 13.9 |
| GY | | 1.4 | 20.8 | 51.4 | 11.1 | | | 2.8 | 11.1 | 1.4 |
| Y | | | | | 55.6 | 35.6 | | 8.9 | | |
| YR | | | | | | 39.5 | 50 | 10.5 | | |
| R | | | | | | | 46.4 | 45.4 | 8.2 | |
| RP | | | | | | | | 65.5 | 34.5 | |
| P | | | | | | | | | 100 | |
| PB | | | | | | | | | 12.8 | 87.2 |

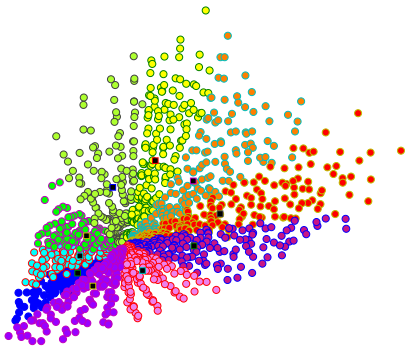

***Figure 3.*** *Munsell data at the 2nd stage of Bumbaca and Smith model. Dark squares are the representatives of each color class.*

**Confusion matrix for classification of test samples with Bumbaca and Smith model.**

| | B | BG | G | GY | Y | YR | R | RP | P | PB |
|---|---|---|---|---|---|---|---|---|---|---|
| B | 77.6 | 10.4 | | | | | | | | 11.9 |
| BG | 6.3 | 92.2 | 1.6 | | | | | | | |
| G | | 30.6 | 69.4 | | | | | | | |
| GY | | | 22.2 | 72.2 | 5.6 | | | | | |
| Y | | | | 1.1 | 71.1 | 27.8 | | | | |
| YR | | | | | 2.3 | 72.1 | 25.6 | | | |
| R | | | | | | | 91.8 | 8.2 | | |
| RP | | | | | | | 10.3 | 83.9 | 5.7 | |
| P | | | | | | | | 1.4 | 91.7 | 6.9 |
| PB | | | | | | | | | | 100 |

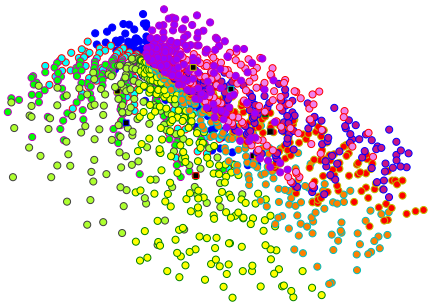

***Figure 4.*** *Munsell data at the 3rd stage of De Valois and De Valois model. Dark squares are the representatives of each color class.*

**Confusion matrix for classification of test samples with De Valois and De Valois model.**

| | B | BG | G | GY | Y | YR | R | RP | P | PB |
|---|---|---|---|---|---|---|---|---|---|---|
| B | 10.4 | 9.0 | | 16.4 | | 6.0 | 7.5 | 29.9 | 19.4 | 1.5 |
| BG | 10.9 | 3.1 | 7.8 | 54.7 | 3.1 | 4.7 | | | | 15.6 |
| G | 5.6 | | 19.4 | 63.9 | 1.4 | 1.4 | | | | 8.3 |
| GY | 5.6 | | 11.1 | 65.3 | 8.3 | | | | | 9.7 |
| Y | 8.9 | | 21.1 | 21.1 | 10.0 | | | | | 38.9 |
| YR | 12.8 | 15.1 | | | | 12.8 | 12.8 | | 45.3 | 1.2 |
| R | | | | | | | 49.5 | 24.7 | 25.8 | |
| RP | | | | | | | 20.7 | 66.7 | 12.6 | |
| P | | | | | | | 15.3 | 73.6 | 11.1 | |
| PB | | | | | | | 12.8 | 59.0 | 15.4 | 12.8 |

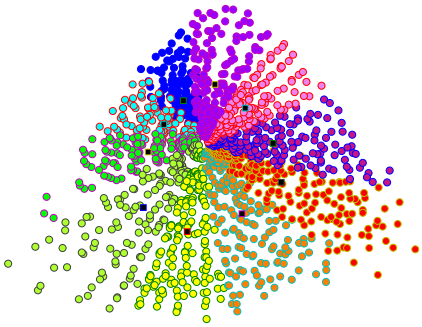

***Figure 5.*** *Munsell data at the 3rd stage of ATD95 model. Dark squares are the representatives of each color class.*

**Confusion matrix for classification of test samples with ATD95 model.**

| | B | BG | G | GY | Y | YR | R | RP | P | PB |
|---|---|---|---|---|---|---|---|---|---|---|
| B | 86.6 | 13.4 | | | | | | | | |
| BG | 4.7 | 92.2 | 3.1 | | | | | | | |
| G | | 26.4 | 73.6 | | | | | | | |
| GY | | | 12.5 | 86.1 | 1.4 | | | | | |
| Y | | | | | 100 | | | | | |
| YR | | | | | 7.0 | 90.7 | 2.3 | | | |
| R | | | | | | 13.4 | 81.4 | 5.2 | | |
| RP | | | | | | | 18.4 | 74.7 | 6.9 | |
| P | | | | | | | | 1.4 | 98.6 | |
| PB | 19.2 | | | | | | | | 1.3 | 79.5 |

## Discussion

There are large differences in the color classification abilities of examined color vision models. If only the distribution of Munsell colors is considered, all models but De Valois and De Valois create quite easily distinguishable color clusters. It is quite clear that De Valois and De Valois model, which has been defined purely based on the biological facts of the human color vision system, is not replicating well the perceptual properties of color vision. Deeper examination of classification results shows that the clusters of Ingling and Tsou model are somewhat overlapping, which causes problems with color classification. Only colors Ingling and Tsou model is able to classify with satisfactory results are purple and purple-blue.

Models whose coefficients have been chosen so that the models are able to replicate the results of certain psychophysical experiments (especially Guth's ATD model) work a lot better in the classification task. Bumbaca and Smith's and Guth's models both give good classification results. Lowest result for Bumbaca and Smith is 69.4 % for green colors and for Guth 73.6 %, also for green colors. Bumbaca and Smith model is able to give 100 % classification for purple-blue colors, and Guth model has full 100% classification for yellow colors. With misclassifications both models behave in a very similar way: always when a color is misclassified, it is classified to one of its neighboring color classes. This kind of behavior is actually typical also for human beings - it is not always easy to tell, where exactly the border line between two different colors is.

It is interesting to see that it is possible to gain similar classification results either by using quite a simple two-stage color model or a more complicated three-stage color model. One common factor for both models is that they introduce some kind of nonlinearity in the performed calculations. Both models also have parameters that have been defined for improving the perceptual properties of the output color space.

## References


[1] Austin Roorda and David R. Williams, The arrangement of the three cone classes in the living human eye, Nature, 397, 520 (1999).

[2] Heidi Hofer, Joseph Carroll, Jay Neitz, Maureen Neitz, and David R. Williams, Organization of the human trichromatic cone mosaic, J. Neurosci., 25, 9669 (2005).

[3] Joseph Carroll, Carrie McMahon, Maureen Neitz, and Jay Neitz, Flicker-photometric electroretinogram estimates of L:M cone photoreceptor ratio in men with photopigment spectra derived from genetics, J. Opt. Soc. Am. A, 17, 499 (2000).

[4] Joseph Carroll, Maureen Neitz, and Jay Neitz, Estimates of L:M cone ratio from ERG flicker photometry and genetics, Journal of Vision, 2, 531 (2002).

[5] Carl R. Ingling, Jr. and Brian Huong-Peng Tsou, Orthogonal combination of the three visual channels, Vision Research, 17, 1075 (1977).

[6] Federico Bumbaca and Kenneth C. Smith, Design and implementation of a colour vision model for computer vision applications, Computer Vision, Graphics, and Image Processing, 39, 226 (1987).

[7] Russell L. De Valois and Karen K. De Valois, A multi-stage color model, Vision Research, 33, 1053 (1993).

[8] S. Lee Guth, Further applications of the ATD model for color vision, Proc. SPIE 2414, pg. 12 (1995).

[9] S. Lee Guth, The constancy myth, the vocabulary of color perception, and the ATD04 model, Proc. SPIE 5292, pg. 1 (2004).

[10] Erkki Oja, Subspace Methods of Pattern Recognition, Research Studies Press Ltd., Letchworth, England, 1983.


## Author Biography

*Tuija Jetsu received her MSc. degree in Computer Science from University of Joensuu, Finland in December 2004. Currently she is a PhD. student and member of Color Research Group at University of Joensuu. Her work is currently related to color vision models.*


©2008 Society for Imaging Science and Technology