Epiphenomenon of Color in Visual Perception

A. P. Petrov
Russian Research Center, Kurchatov Institute, Moscow, Russia

Abstract

It is concluded from presented experimental and theoretical consideration that the notion of "light source" is inadequate for description of the color perception and an alternative way to define and to describe the illumination is suggested. Color constancy as phenomena of perceiving the same color of curved and homogeneously painted surfaces is considered and it is shown that surface colors can be adequately represented by 3x3 matrices. There discussed a hypothesis that the dimension of the color space (the colorimetric space) is determined by the dimension of the ordinary geometric space of the visual scene and that the number of types of cones should be considered only as an implementation parameter. A straightforward solution of the image irradiance equation based on the concept of surface color and the above description of illumination is suggested and discussed.

Why is color an epiphenomenon? This question is answered in connection with the main question of what kind of vital tasks require explicitly perceived color.

Introduction

Surface is a key-note not only for color perception. It is universally significant across species and perceiver-independent because the surfaces are the places where vision and other senses interact with each other and with motor control systems to avoid obstacles during motion. The best definition of *surface* in our science would be as *a set of terminal points of free movement* and detection of the surfaces and computation of their shapes are vitally important task to be solved by the visual system equipped with mechanisms for color perception. It means that for better understanding we have to consider the task of color computation as a part of the problem of surface and space perception (segmentation, shape-from-shading, illuminant determination, etc.).

Illumination

As a problem of inverse optics the task of recovering illumination conditions from image data is extremely ill-posed and if to describe the illumination in terms of light sources, etc. then, in general, it does not have a unique solution. What kind of variables are used by the human visual system for description of illumination is the topic of a future research. Here we formulate a new approach to the task which seems very promising as a base for development of a new concept of perceived illumination.

Let us place a small white Lambertian sphere at some point P of the scene and measure the intensity, $I(\mathrm{n};P,\lambda)$, of the light scattered from different parts of this probe sphere. Here λ and n, as usual, denote the wavelength of light and the normal vector on the sphere.

It is important that for a scene illuminated by a combination of a diffuse source and a set of arbitrarily located and spectrally different point sources, $I(\mathrm{n};P,\lambda)$ is a spherical function of a special kind

$$I(\mathrm{n};P,\lambda) = \mathrm{s}(P,\lambda)\cdot \mathrm{n} + I^d(P,\lambda) \qquad (1)$$

for each region G of the probe sphere illuminated by the same set of the sources visible from all points of the region, the last term is a diffused light component. Note that illuminations (1) form a three-dimensional family of spectral functions independently of the number and spectral quality of the light sources. This fact allows to expect a relation between the dimensions of the regular physical and color spaces.

Instead of using spectral energy distributions, we can write the illumination formulae as a vector-function mapping our probe sphere into the color space

$$\mathrm{h}(\mathrm{n};P) = \mathrm{H}_0 \cdot \mathrm{n} + \mathrm{h}_0^d \qquad (2)$$

where h, h_0^d are vectors in the color linear space H and H_0 is a 3x3 matrix. The matrix H_0 depends on P and n but is constant as a function of n in each region G.

Color Constancy

The basic meaning of *color constancy is keeping unchanged the perceived color when bending a surface observed or perceiving the same color of curved and homogeneously colored surfaces*. If we imagine that this color constancy is achieved in a visual system then it will obviously demonstrate Helmholtz-Hering constancy also in some conditions, for instance, the perceived color of some moving object will be the same while the object will be exposed to illuminations of different spectra and intensity during the motion.

This statement of the color constancy problem settles the argument about the range of allowed changes of illumination the choice of which is often questionable[1]. A natural set of illuminant variations consists of the ones which occur when a surface rotates or undergoes deformations in a scene with fixed illumination conditions (sources, etc.). It is clear that such a range is determined by the scene and in different scenes chromaticities, brightnesses and other characteristics of possible incident light will be different.

Surface Color

Above to describe the illumination we used a standard probe body, the white Lambertian sphere. It is obvious that if instead we take another uniformly colored spherical surface, we get the same formula with different matrix H_0 and vector h_0^d

$$h(n;P) = H \cdot n + h^d \tag{3}$$

The function h(n;*P*) on the left hand side of (3) can be thought of as a color image after substitution of coordinates in the retina for the variable *P* and hence, Eq. (3) is the image irradiance equation for the case of color image and complex illumination

$$h(x,y) = H \cdot n(x,y) + h^d \tag{4}$$

Here *x,y* are coordinates in the image h; the normal vector n of the surface is the function of *x* and *y*.

By definition, let us call *surface color* the matrix C in the following equation

$$C \cdot H_0 = H \tag{5}$$

C is the color matrix defined in [2] as a mapping $H \to H$ which describes changes in color images when the white surface is substituted with the colored one. The black surface will obviously have the null color matrix and the white one will get the unit color matrix independently of the basis in the color space.

Equation (4) is analogous to the image irradiance equation of [3] but written for complex color illumination, color objects and in psychophysical variables. Let us make a note about the left hand side of Eq.(4), i.e. the image h(*x,y*). In biological visual systems the cell responses are nonlinear for photooptical stimuli and moderate contrasts. Hence it makes no sense to consider the retina response as an element of a linear space, and we prefer to define the input image as the results of a colorimetric procedure carried out upon the scene.

From (5) it follows that we define the surface color as a linear operator acting in the color space. To each surface patch there corresponds its color, i.e. a numeric color matrix C. The set of surface colors is a region in the corresponding space of matrices. It can be easily shown that:

- the elements $c_{i,j}$ of the color matrix are bounded;
- the color set is convex;
- the color set is symmetrical with the center **E**/2;
- the dimension of the color set is 9;
- there is a partial order relation in the color set;
- there is no natural metrics in the color set.

In [2] it is shown that the color set represents all the perceived colors adequately. This means that any two surfaces with equal color matrices look alike for a viewer with normal color vision and vice versa two samples looking the same color have equal color matrices.

On Dimension of the Color Space

Trichromacy in man and some other species is sometime explained by the fact of limited number of dominant light sources creating terrestrial illumination. Dimension of the color space and its linearity should be considered together because dimension is a characteristic of the space structure. From this point of view we have to find an explanation for both facts, linearity of the color space and trichromacy of color vision. As for the linearity, the visual system puzzles us very much because there is no level or element in the visual brain that demonstrates linear behavior but in spite of this, colorimetry shows perfect linearity of the whole system. The same can be expected about dimension of the color space, namely, the eye obtains one number of cone types but the dimension equals a different number.

The suggested concepts of surface color and color constancy create a new explanatory framework for these questions. To start with, in the above explanation we can substitute terrestrial illumination assumed to be approximately a 3D-space, with the illumination defined in Section 2 whose dimension is exactly equals three and thus, eliminate the argument of poor approximation of terrestrial illumination by limited three degrees of freedom.

This substitution changes the explanation seriously. The formula of illumination description given by Eq. 2 is based on the structure of the ordinary physical space where the observer actively solves his orientational tasks and where shapes are defined as objects. So, we have replaced ecological-level constraints by the most fundamental physical-level ones which are observer-independent and valid for trichromats, tetrachromats, pentachromats and others. The same we can tell about the linearity of the colorimetrical space. Its origin is connected with the linearity of the image irradiance equation, Eq. 4, with respect to the color matrix.

Concluding this section let us formulate a hypothesis of color space dimension: *All the visual diurnal species have 3D linear color spaces defined by any possible procedure of colorimetry. The dimensions of the color spaces are not directly determined by corresponding numbers of cone types or other features of receptors.*

Color Computation

The assumptions used here are as follows:

(1) the illumination contains no diffuse component;
(2) interreflections do not contribute significantly into illumination;
(3) the light sources are located at distances much greater than the characteristic size of the surface;
(4) the surface is Lambertian with piece-wise constant;
(5) optics corresponds to distant observation or orthogonal projection.

For a segment of the input image which contains data measured at a uniformly colored patch of the surface and all points of which are illuminated with the same set of light sources, i.e. for the one to be finally determined, the IIE can be written in the form:

$$\mathrm{A} \cdot \mathrm{n}(x, y) = \mathrm{h}(x, y) \tag{6}$$

where the matrix $\mathrm{A} = \mathrm{C} \bullet \mathrm{H}_0$.

Given that the matrix A is found, the separate computation of the color matrix and the illumination matrix requires the decomposition of A. This task cannot be solved with only local operations over the image because together with a solution (C, H_0, n) we have a set of solutions ($\mathrm{C} \bullet \mathrm{M}$, $\mathrm{M}^{-1} \bullet \mathrm{H}_0$, n), where M is any matrix with $det\mathrm{M} \neq 0$. It may be that to achieve a unique solution additional information must be used. This information could be taken from other “channels” such as normalization or calibration procedures for some known and recognizable object in the scene.

The normals to the surface are of unit length by definition and, hence denoting

$$\mathrm{B} = (\mathrm{A} \cdot \mathrm{A}^*)^{-1} \tag{7}$$

we get the equation

$$\mathrm{Bh} \cdot \mathrm{h} = 1 \tag{8}$$

that hold in the entire region *G*. After differentiating this *norm constancy condition* twice with respect to *x* and *y* we get a system of equations allowing a solution in analytic form.

Let us define (a,b,c) as the value of the determinant of the matrix [a, b, c] where a, b, c are vectors of three components. Then denote

$$\mathrm{B}^0 = \begin{pmatrix} (\mathrm{h}, \mathrm{h}_x, \mathrm{h}_y) & 0 & 0 \\ 0 & -(\mathrm{h}_x, \mathrm{h}_y, \mathrm{h}_{xx}) & -(\mathrm{h}_x, \mathrm{h}_y, \mathrm{h}_{xy}) \\ 0 & -(\mathrm{h}_x, \mathrm{h}_y, \mathrm{h}_{xy}) & -(\mathrm{h}_x, \mathrm{h}_y, \mathrm{h}_{yy}) \end{pmatrix} \tag{9}$$

Then the formulae for a direct computation of the matrix B can be written as follows:

$$\mathrm{B} = \frac{1}{\mathrm{h}, \mathrm{h}_x \mathrm{h}_y} \cdot \left[\mathrm{h}, \mathrm{h}_x, \mathrm{h}_y\right]^{-1*} \cdot \mathrm{B}^0 \cdot \left[\mathrm{h}, \mathrm{h}_x, \mathrm{h}_y\right]^{-1} \tag{10}$$

Note that all matrices used in the formulae exist because of the rank[4] of the region equals three (in the regions rank 2 and 1 different formulas would be produced). Meanwhile, for a computer implementation of formulae Eq. 12 it is necessary to use carefully regularized algorithms of matrix inversion and derivative estimation.

Given the matrix **B**, we can produce additional information and compute the shape, i.e. the normals n to the surface. In order to do this we have to find the matrix **A** from Eq. 7. On account of the fact that $\sqrt{\mathrm{B}}$ is not unique we need to compute the matrix **U** of a rotation such that the unit field $\sqrt{\mathrm{B}} \bullet \mathrm{h}$ after the rotation satisfies the integrability condition.

Let us denote $\mathrm{A}_0^{-1} = \sqrt{\mathrm{B}}$ a positive and symmetrical value of square root of **B**. Then the integrability condition can be written this way:

$$(\mathrm{A}_0^{-1} \cdot \mathrm{U}^* \cdot \boldsymbol{\Phi} \cdot \mathrm{U} \cdot \mathrm{A}_0^{-1})\mathrm{h} \cdot \mathrm{h} = 0 \tag{11}$$

where

$$\boldsymbol{\Phi} = \begin{bmatrix} 0 & 0 & -\partial / \partial y \\ 0 & 0 & \partial / \partial x \\ \partial / \partial y & -\partial / \partial x & 0 \end{bmatrix}$$

After we find the matrix U with the help of some optimization procedure the normal field to our surface can be found as the following linear transformation of the input image:

$$\mathrm{n}(x, y) = \mathrm{U} \cdot \mathrm{A}_0^{-1} \cdot \mathrm{h}(x, y) \tag{12}$$

Why is Color an Epiphenomenon?

What we really enjoy is explicit color being experienced in almost every scene. And it is not mere curiosity to ask why the human visual system gives us colors explicitly together with other important results of its work because there are many dimensions which are actively used by the visual system but are not perceived explicitly. For instance, brightness is such a dimension. Another example of a non-explicit variable is the angular size of an object determining the perceived object size and its distance but never perceived explicitly. Color plays very similar role in perception, namely, it helps to segment images and to recover illumination and shape and yet we do not know any specific goal for the visual system to “output” colors.

The simplest answer to the question is that object recognition can be such a specific function of explicit color. But it is obvious that as a feature for recognition color is very poor, if not deleterious, and we often recognize objects not due to color but in spite of it. If some visual system used color for recognition it would get such a noisy input that the requirements to decision making to be robust would have to ruin all the delicate results of accurate computations of constant surface color. So, we reject this supposed role of color in visual perception. Another possible function of explicit color is to serve as results of segmentation. This means that colors are used as identifiers (codes) of regions in the visual field. However, segmentation is a procedure fulfilled on an image and it is not easy to transfer or to generalize the notion of segmentation on the 3D visual field and, it is most probable that segmentation will be finally claimed inadequate to the task of object detection and separation. Secondly, in our artificial world there are enough uniformly colored objects but they are very seldom to be met in the natural world and an algorithm of segmentation based on constant color would be hardly useful. These two objections make us to drop segmentation as a function of explicit color.

There is a behavioral task which can require using surface color computed as accurately as possible and representing it at the highest level of the visual system. *Following up* the objects is this task. The problem of identification of moving objects could be solved without engaging color or other features if continuous tracking were possible. In reality multiple occlusions and other obstacles often break off the tracking and to resume it the visual system needs reliable data for identification of the object which is followed up. Sameness is detected with using all possible means including color. Surface color in this case works not to join different objects together in a class (the "bananas") but to separate the objects. We believe that this "following up" task brings color up to the representation level of explicit data.

The above brief analysis and the failure of the search for a role that explicit color should play in perception requires a revision of the problem. Our leading idea of the future revision can be schematically outlined as follows.

Experienced color is a gift not purposing any function. And the blue sky over a New Hampshire lake is an unnecessary consequence of the successful tracking of a bird flying over the lake. Hue, saturation and brightness are dimensions of experienced color.

Sensation of color is a socio-cultural phenomenon. To support this claim let us cite from J. J. McCann's[5]. "The experiment to measure the sensation or appearance of the two faces (of a float on the lake) is to ask people to imagine they are visual artists, fine-art painters.... They select a yellow-white paint for the sunlit face and a darker, blue-gray paint for the face in the shade. In this case they have matched the sensation." So, how can people imagine of being fine-art painters? Development of visual artistry has taken 10^4 years of a social life. It is impossible to switch off color constancy inside us and we are only able to overcome it after training or equipped with special instruments. Why should the result of that intellectual work resemble the input data for the color perception level as we sometimes thought of?

Experienced color is an epiphenomenon of human visual perception. The closest analogy of this epiphenomenon is our ability to enjoy music with our hearing designed for quite different sounds to detect, to process and to understand. And if you created a system for understanding of emotional speech you might know that the system liked music. Fine-art, music, ballet, poetry are based on epiphenomena of human perception and explicit color is one of them. Really visible color appears at the aesthetic level and all the contemporary color science still did not touch it.

References

1. V. V. Maximov, *Color Transformations under Changing Illumination*, Nauka, Moscow, 1984. In Russian.
2. A. P. Petrov, Surface Color and Color Constancy, *COLOR research and application*, vol. **18**, No. 4, 1993, pp. 236-240
3. B. K. P. Horn, *Shape from Shading*, M.I.T. Press and McGraw-Hill, 1989.
4. A. P. Petrov and L. L. Kontsevich, Color image field rank and segmentation, *J. Opt. Soc. Am.*, 1994, in press.
5. J. J. McCann, Color Sensations and Color Perceptions, *24th Asilomar Conference on Signals, Systems & Computers*, vol. **1**, November 1990, pp. 408-412.