OBTAINING LINEAR EQUATIONS BY REGRESSION METHOD FOR COLOR MANAGEMENT

[1]VOLKAN YAVUZ, [2]SEMIH ÖZDEN, [3]YUSUF SÖNMEZ

Gazi University, Polatlı Technical Sciences Vocational College Department of Printing and Publishing Technologies Ankara Turkey , Gazi University Technical Sciences Vocational College Electronics and Automation Department Ankara Turkey, Gazi University Technical Sciences Vocational College Electrical and Energy Department Ankara Turkey
E-mail: [1]vyavuz@gazi.edu.tr, [2]sozden@gazi.edu.tr, [3]ysonmez@gazi.edu.tr

Abstract - In this study, cyan, magenta, yellow and black colors prints were produced in 4 different digital printing machines and the results were determined with color measuring device and transferred to computer. The main aim is to approximate these printed colors to their original reference values and thus reducing the color variations caused by the differences between of the printing machines. In this study, the L*a*b* values were used as the color space of the prints obtained from the printing machines. This color space is geometric arrangement in visible areas (L*a*b*). A defined color space model provides an objective classification criterion so that colors can be expressed numerically. Linear regression method was used to approximate the print color values to the original value. Here, the coefficients of the second order equations were determined. The Gravitational Search Algorithm (GSA) method was selected to determine the coefficients of the equation. This method calculates the coefficients of equations so that the difference between the L * a * b * values of the colors obtained from the printing machines and the L * a * b * values of the original colors is minimized. All process was performed in MATLAB software. The obtained results show that the different printing presses, which were different from their original values, are approximated to their original values thanks to the defined equations. These equations can be used in different printing machines for producing closer printing results to original colors.

Index Terms - Color management, linear regression method, Gravitational Search Algorithm.

I. INTRODUCTION

Color management is one of the important issues of graphic arts. After the design phase, suitable printing methods for the products to be directed to the press are to be found and suitable materials should be used when applying this printing method. Perhaps the most important material to achieve the desired result is ink. Production is carried out by printing machines produced by many companies using different printing techniques. The latest technology of these printing systems is digital printing machines. In color digital printing machines, Cyan, Magenta, Yellow and Black (CMYK) are used as basic colors as in other printing systems. All colors are obtained with a mixture of these four colors.

Achieving the desired result in print depends on the use of correct colors in the graphic design phase, selection of the appropriate printing technique and machine, and printing with the right ink. This process is also paralleling the concept of color management. Color management is a concept used to get the right colors in all the devices used. At this stage, the last part of the printed material is covered by the screen of the computer on which the graphic design is made [1]. When different printing machines are used, the results differ a big problem from the original color value. The aim of color management is to eliminate this difference and ensure that every machine produces results that are close to the original colors.

In this study, it is aimed to approximate the color output obtained by using different printing machines to the original colors. For this reason, the Lab values of the CMYK colors of the output from four different presses were measured by means of the color-measuring device and a mathematical model was created to approximate these values to the Lab values of the original CMYK colors. The generated model is a second-order linear equation that computes the L, a, and b values of each fundamental. The coefficients of the equation are calculated by Gravity Search Algorithm (GSA). The calculation was performed in the MATLAB program. Experimental results show that the difference between the L*, a* and b* values of the CMYK colors of the output from different printing machines and the L*, a* and b* values of the original CMYK colors is reduced by the created equations.

II. MATERIAL AND METHOD

Color management also assumes a communication function between devices as it aims to give different devices the same results. The most important device that provides this communication is the color measuring device. The standards set by the International Color Consortium, which is called ICC Profiles, apply for color measurement [2]. This consortium takes place in partnership with a large number of reliable and field related organizations. All color measuring devices are also measuring with these standards.

Color measuring devices can display values such as a RGB, CMYK and CIE L*, a* and b* (Lab) as a result. The use of Lab colors was deemed suitable for this study. The CIE Lab color model is based on the human perception of color. The numeric values in the Lab

model define all the colors seen by a person with normal vision. The Lab model is considered to be a color model independent of the device, since the Lab model identifies how a particular colorant will appear to be colorless enough to create colors for a device (such as a monitor, desktop printer, or digital camera). Color management systems use the Lab model to transform results from one color space to another so that results can be predicted [3]. The L * a * b * color model can use simple values to describe the adjectives red / green and yellow / blue. When CIE L * a * b * shows a color, L * lightness shows a * red / green value and b * yellow / blue value [4-6]. In this study, a visual with CMYK colors was prepared and their printings were taken from four different digital presses. The CIELab values were measured via the spectrometer of these prints. Then, Lab values of the original CMYK colors were accepted as references and a mathematical model was created to approximate the values of the print colors to the original color values. The generated mathematical model is a linear regression method and there is a second order linear equation which will approximate the L*, a* and b* values of each fundamental color to the original values. In the study, the calculation of the coefficients of these equations is considered as an optimization problem. The problem is to find the coefficients that will minimize the difference between the original values and the output values. For this purpose, Gravitational Search Algorithm (GSA), which is a powerful heuristic optimization method, is used.

In the study, the GSA algorithm is used as a standard regression method, which is used to write the mathematical correlation between two physical quantities that change in relation to each other as real as possible. For use in the method, the simple second-order linear regression equation is written as follows.

$$Y_{iL} = w_{0L} + w_{1L}X_{Li} + w_{2L}X_{Li}^2 \quad i=1,2,3\ldots,n \ (1)$$
$$Y_{ia} = w_{0a} + w_{1a}X_{ai} + w_{2a}X_{ai}^2 \quad i=1,2,3\ldots,n(2)$$
$$Y_{ib} = w_{0b} + w_{1b}X_{bi} + w_{2b}X_{bi}^2 \quad i=1,2,3\ldots,n(3)$$

Here, Y_{iL} is the dependent variable (original L value), X_{Li} independent variable (L taken from the machine output), w_{0L},w_{1L}, and w_{2L} are the unknown parameters and the total number of samples in the n value data set. In the study, separate functions for the L, a and b values of each color value are created and these functions are given in equations 1, 2 and 3.

For each sample in the data set, the sum of the errors is expressed as the sum of the differences between the true and the estimated values as follows.

$$\sum_{i=1}^{n} e_i = \sum_{i=1}^{n} \left| \left(Y_{i,actual} - Y_{i,estimated} \right) \right| (4)$$

In the study, the GSA method finds w_0, w_1 and w_2 as follows, making the sum of absolute errors smallest.

$$F_i = \min \left(\sum_{i=1}^{n} e_i = \sum_{i=1}^{n} \left| \left(Y_{i,actual} - Y_{i,estimated} \right) \right| \right) (5)$$

GSA is a powerful optimization algorithm, which simulates Newtonian law of gravity and law of motion. It was produced in 2009 by Rashedi et al [5]. In the GSA, there are agents and masses of them. While agents corresponds to objects, their masses measure performance of them. According to the GSA each object interacts with a gravitational force and move toward objects with heavier masses. Object with heavier masses corresponds better solution. The flow diagram of the GSA is shown in Figure 1.

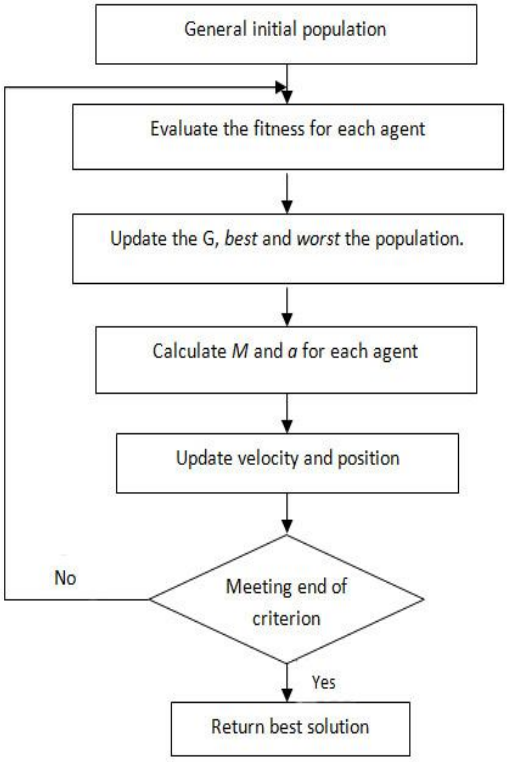


Figure 1. General flowchart of the GSA [5].

In the study, it is aimed to arrange the basic colors before printing according to the form obtained by using the coefficients found by the proposed method and to print the correct colors in this number.

III. EXPERIMENTAL RESULTS

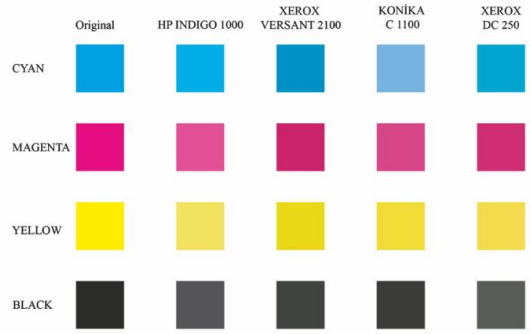
In this study, the reference L, a and b values (Y_{iL}, Y_{ia}, Y_{ib}) of four fundamental color (CMYK) were obtained to find the w coefficients given in equations 1,2 and 3 using GSA. The data to be used for the calculation of the coefficients consists of the L, a and b values obtained from the color measuring device of the outputs from 4 different presses. Examples of independent variable data to be used in the calculation of the reference values and coefficients used are given in Table 1.

Table 1. L * a * b * Color Equivalents of Measured Print Values

	Original			HP INDIGO 10000			XEROX VERSANT 2100			KONIKA C1100			XEROX DC 250		
	L	a	b	L	a	b	L	a	b	L	a	b	L	a	b
C	57	-40	-52	64	-31	-43	53	-33	-42	70	-11	-31	61	-32	-35
M	50	80	-2	56	64	-6	45	69	4	52	63	-3	47	68	2
Y	94	-5	102	90	-6	65	86	-7	86	88	-5	80	88	-4	72
K	11	0	1	34	1	-2	25	-2	1	21	-1	2	37	-2	3

When comparing the Lab values of the original document with the values obtained in the print result, it is seen that there are big differences. The color equivalents of the values given in Table 1 are shown in Table 2. The large differences in color reveal the importance of practice. Table 3 shows the w coefficients found using the GSA algorithm.

Table 2. Comparative Color Table



These coefficients and the Lab values presented in Table 1 are substituted in equations 1,2 and 3 and the estimated Lab values for each color value are calculated. The error values between these calculated values and the original values are shown in Table 4.

Table 3. Obtained parameters of the second order linear equations

		Coefficients		
		w0	w1	w2
Cyan	L	3,4959	1,6477	-0,0126
	a	6,7567	5,7678	0,1350
	b	-1,8130	2,6377	0,0341
Magenta	L	1,9517	1,8704	-0,0182
	a	6,1826	2,1476	-0,0156
	b	-2,0627	-0,0009	-0,0015
Yellow	L	3,9644	2,1405	-0,0127
	a	1,9646	2,8889	0,2936
	b	1,4223	2,4684	-0,0151
Black	L	-1,9767	0,9585	-0,0170
	a	-0,2548	0,1903	0,1958
	b	0,7884	0,1796	0,0183

Table 4. Experimental results obtained from proposed model

		Real	HP INDIGO 10000	XEROX VERSANT 2100	KONIKA C1100	XEROX DC 250	MAE	MAPE	RMSE
C	L	57	57	55	57	57	0,50	0,88	1,00
	a	-40	-42	-37	-40	-40	1,25	3,13	1,80
	b	-52	-52	-52	-51	-52	0,25	0,48	0,5
M	L	50	50	50	50	50	0,00	0,00	0,00
	a	80	80	80	80	80	0,00	0,00	0,00
	b	-2	-2	-2	-2	-2	0,00	0,00	0,00
Y	L	94	94	94	94	94	0,00	0,00	0,00
	a	-5	-5	-4	-5	-5	0,25	5	0,5
	b	102	98	102	102	101	1,25	1,23	2,06
K	L	11	11	11	11	10	0,25	2,27	0,50
	a	0	0	0	0	0	0,00	0,00	0,00
	b	1	1	1	1	1	0,00	0,00	0,00
SUM OF ERROR			6	12	7	3			

According to Table 4, the results obtained from four different printers and the actual color values are evaluated with MAE, MAPE and RMSE. Each color L, a, b values are calculated independently. From a color perspective, the lowest error is found in the magenta. The highest error occurred in Cyan color. The highest error on the Xerox VERSANT 2100 from the point of view of the printer, the lowest error was obtained on the XEROX DC 250 device. In general, it is seen that the error value is close to zero for the color values of all the machines. This makes it possible to create a standard print quality by approximating the color output to the original color regardless of the machine through the created mathematical model.

RESULTS

In this work, it is aimed to eliminate the color differences in the output from different printing machines.Therefore, a mathematical model was created to minimize the difference between the machine output and the original color values. The coefficients of the second order linear equation are found by Gravitational Search Algorithm. According to the results obtained from the experimental studies, the color values estimated through the equation are approximated to the original color values. Thus, when a visual image is printed without considering the printer difference, it is aimed to update the color values according to this model and to approximate the printout to the original, thus creating standardized print quality independently of the machine.

REFERENCES

[1] M. Mahy and D. DeBaer, "HIFI color printing within a color management system", Proc. 5th IS&T and SID Color Imaging Conf., pp. 277-283, 1997.
[2] ICC Profile Format Specification, Version ICC.1:1998-09, International Color Consortium, September 1998.
[3] https://helpx.adobe.com/tr/photoshop/using/color-modes.html
[4] S. Brues, “Postscriptium on Color Management” GretagMacbeth, Switzerland, 2000.
[5] H. Speirs, “Introduction to Prepress”, BPIF, Pira International, UK, 1998.
[6] “The Basic Principles of Color and Lab for Computer Publishing”, Linotype-Hell,2003.
[7] E. Rashedi and H Nezamabadi, "GSA: A Gravitational Search Algorithm", Information Sciences, Vol. 179, No. 1, pp. 2232-2248, April 2009.

★★★