Improving the Yule-Nielsen modified spectral Neugebauer model using Genetic Algorithms

Conference Paper · June 2012

3 authors:

Yiming Qian
York University

Nawar Mahfooth
Ryerson University

Matthew Kyan
York University

Improving the Yule-Nielsen modified spectral Neugebauer model using Genetic Algorithms

Yiming Qian[1], Nawar Mahfooth[2], Matthew Kyan[1]
[1] *Ryerson University, Canada*
[2] *Smartia Printing Technologies Inc., Canada*
yqian@ryerson.ca, nawar@smartia.ca, mkyan@ee.ryerson.ca

Keywords: Colour Printing, Dot Gain, Spectral Prediction Model, Yule-Nielsen modified Spectral Neugebauer model, Genetic Algorithm

Abstract

Yule-Nielsen Modified Spectral Neugebauer model is widely used in colour prediction but different physical factors such as light scattering in substrates, ink chemical properties and ink drop velocity all influence the accuracy of the prediction. In this paper, we are presenting a method that uses the standard IT8.7/4 CMYK test chart and YNSN dot gain enhanced model to construct paper-ink-printing machine dependent prediction models. IT8.7/4 was printed on an offset press with 100lb coated paper. PressSign software was used to control the ink keys along with an X-Rite eXact to ensure the densities were within GRACoL specifications; 10 measurements were obtained after 24 hours using the iSis (a spectrometer). The measured data was smoothed and averaged to obtain clean spectral data. The 16 Neugebauer Primary colours (NPs) and 4 dot-gain curves were extracted from spectral data. Genetic algorithms were used to fine-tune the dot gain curves and at the same time to find and optimize the spectral n value (different n for different wavelength range). The final model used YNSN model as a base, instead of using only one n value, 36 n values based on the spectral measurements were used. The dot gain model was used to correct the input value before feeding to the YNSN model. Our contribution in this paper was to improve the Yule-Nielsen modified Neugebauer model by integrating a genetic algorithm for computing the dot gain of the CMYK's and finding a set 36 spectral n-values.

Introduction

Yule-Nielsen Modified Spectral Neugebauer (YNSN) model is widely used in colour prediction but different physical factors such as light scattering in the substrate (paper), ink absorbing into the paper, ink spreading out into the paper, all influence the accuracy of the prediction. Different paper types have different ink absorption rates: for example, uncoated papers can absorb more ink than coated ones. As paper is passed through a printing press the pressure of the plates can squeeze the ink out of its dot shape causing gain and result in different shapes on the paper. Different inks under the same pressure will also obtain different dot gain due to its own characteristics [1]. Some of those physical properties appear in the printing result in the form of dot gain. The dot gain describes the phenomenon associated with ink spreading around halftone dots imparted to the substrate, and leads to a darker colour appearance than intended [2]. In the YNSN model, those properties are partially corrected by applying a single n factor (equation 1) across all wavelengths, when associating the measured with theoretical reflectance of a particular ink set. However, ink reflectance at different light wavelengths may exhibit very different behaviour with a given substrate. As such, a spectral n-value model is used to further improve the YNSN model [3]. The spectral n-value model assigns different n factors on each wavelength but those n factors are often hard to determine. The most popular way is using training data and linear regression to estimate the parameters based on the least square error [4]. Linear estimation methods cannot accurately estimate the inks' nonlinear features due to those nonlinear parameters' complex relationships.

The method proposed in this paper offers a novel approach to fine-tune the dot gain and estimate the spectral n factor parameters at the same time by using an optimization procedure based on the Genetic Algorithm (GA). The advantage of this algorithm is that it is capable of finding a set of optimal parameters simultaneously, instead of finding them individually [4,5]. In the proposed method, the genetic algorithm is applied to process a set with 123 parameters at once. The flow chart of the process is shown in Figure 1.

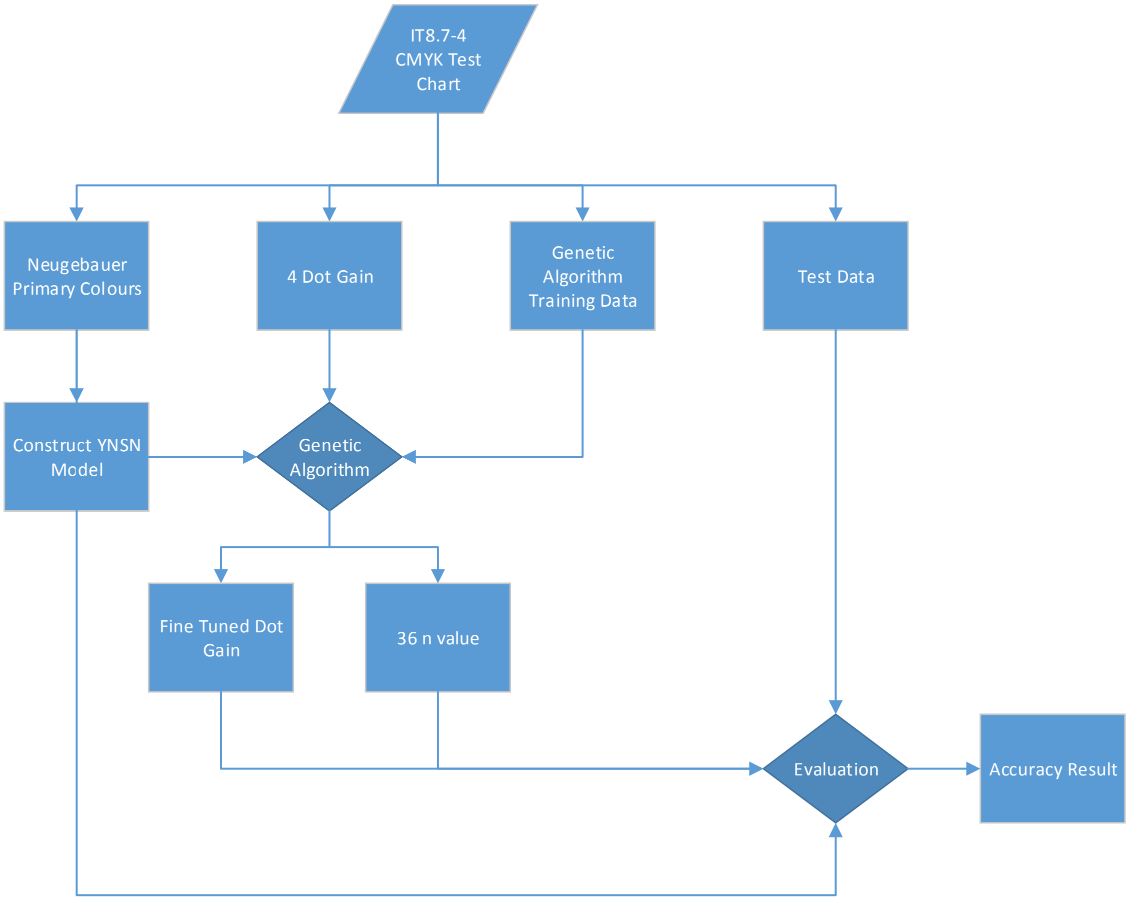

Figure 1: Dot Gain Fine-Tuning and n Factor Estimation Process Flow Chart

Yule-Nielsen Modified Spectral Neugebauer Model

The Yule-Nielsen Modified Spectral Neugebauer (YNSN) Model considered the effect that light laterally scatters within the paper which leads to an increase in the probability of the ink dots absorbing the light [6]:

$$R_{\lambda} = [\sum_{i=1}^{N} F_i R_{\lambda,i}^{1/n}]^n \tag{1}$$

Where

F_i is the fractional area coverage of the Neugebauer Primary colour
$R_{\lambda,i}$ is the measured reflectance
n is the Yule-Nielsen factor
R_{λ} is the calculated reflectance of the ink mix

Typically, four situations are considered for how light is reflected from paper. First, the light hits a printed dot and reflected from it. Second, the light hits the paper and reflects through a printed dot. Third, the light hits the paper and reflects without interacting with any printed dots. The last situation is that the light is absorbed by the paper, and thus never reflected out [7]. The model in equation (1) attempts to predict the reflectance of a particular ink mix given the measured reflectances of individual ink primaries used, while accounting for these four reflectance situations (embodied in the parameters F_i and n respectively).

The primary inks produce solid ink elements as well as overprints. Those colours are referred to as Neugebauer Primary Colours (NPs). The traditional CMYK system has 2^4 Neugebauer Primary Colours (16 NPs). Each colour has a fractional area coverage, which represents the probability that each Neugebauer Primary Colour covers the paper. The fractional coverage can be calculated by following equation:

$$F_i = \prod_{j=1 \to N} \text{(If ink } j \text{ is in Neugebauer Primary } I\text{, then } a_j \text{ Else , (1- } a_j\text{))} \quad (2)$$

Where

a_j is the effective area coverage of ink j.

Table 1 Calculation of the area coverage for each Neugebauer primary given the concentration of inks

Index	Neugebauer Primary (I)	Area Coverage (F_i)
1	W	(1-c)(1-m)(1-y)(1-k)
2	K	(1-c)(1-m)(1-y)k
3	Y	(1-c)(1-m)y(1-k)
4	YK	(1-c)(1-m)yk
5	M	(1-c)m(1-y)(1-k)
6	MK	(1-c)m(1-y)k
7	MY	(1-c)my(1-k)
8	MYK	(1-c)myk
9	C	c(1-m)(1-y)(1-k)
10	CK	c(1-m)(1-y)k
11	CY	c(1-m)y(1-k)
12	CYK	c(1-m)yk
13	CM	cm(1-y)(1-k)
14	CMK	cm(1-y)k
15	CMY	cmy(1-k)
16	CMYK	cmyk

Dot Gain Model

During the printing process, the ink that is dropped on the paper will spread out and increase its size. This ink property is well known as dot gain. There are two types of dot gains "Mechanical Dot Gain" and "Optical Dot Gain". The mechanical dot gain is the ink size gain due to the physical movement of the liquid ink and is the dominant effect in any dot gain process. The optical dot gain is caused by light reflection between the ink dot edge and paper, which is usually seen as a blur effect of the ink drops. In this paper, the dot gain reflectances are measured by spectrometer on the printed result of an IT8.7/4 test chart. The traditional way to calculate dot gain is the Murray-Davies equation:

$$DotGain = \frac{R_0 - R_N}{R_0 - R_{100}} \times 100 \quad (3)$$

Where

R_0 is the reflectance of paper

R_{100} is the 100% coverage reflectance

N is the applied ink coverage

This equation gives out an estimation of the dot gain value but it is not very accurate. In the proposed method, a linear regression algorithms have been combined with YNSN model to accurately calculate the real dot gain value. The reflectance of Neugebauer Primary Colours are measured from the test chart then fed into the YNSN model. Different individual ink effective areas are applied to the YNSN model regressively to calculate its reflectance Root Mean Square (RMS) error against the measured value (equation 4). The effective area with the minimum RMS error is the dot gain value of the respective ink. For the CMYK model, four look-up dot gain tables in a range of 0 to 100% are constructed where each look-up table represents one ink channel.

$$Error = \sqrt{\sum_{i=1}^{36} (R_{ai} - R_{bi})^2} \quad (4)$$

Yule-Nielsen Modified Spectral Neugebauer with spectral n-value Model

The traditional Yule Nielsen n factor is a single value in the YNSN model but in the practical situation different light wavelengths propagate and reflect in slightly different ways due to both the paper and ink. We conjecture that the single Yule Nielsen n factor can only partially represent this phenomenon. The spectral n-value approach is proposed to solve this problem by assigning different n value to different wavelengths. The spectrometer iSis is capable of sampling 36 samples from 380nm to 730 nm wavelength in a single reflectance measurement, so in this paper 36 n- values are used to improve the traditional YNSN single n factor. The YNSN model is then changed to the following expression.

$$R_{\lambda} = [\sum_{i=1}^{N} F_i R_{\lambda,i}^{1/n_i}]^{n_i} \tag{5}$$

Where

n_i is the n Yule Nielsen n factor for the i^{th} wavelength.

Genetic Algorithm Optimization

Genetic Algorithm is an optimization method inspired by the evolution of populations of a given species which is a perfect algorithm for spectral-based printer characterization [8]. The principle is founded on the notion that an individual in the population is encoded in a particular way (by their DNA) so as to be more or less capable of survival in a given environment. The more capable, the fitter or more likely the individual will be to reproduce and thus transfer useful DNA to later generations. Over time, the population will be comprised of individuals with DNA highly suited to surviving in the environment (as individuals with unsuitable DNA will die out). The process models both reproduction, the exchange of genetic 'material' and random mutations, to ensure that many possible encodings for individuals are explored. Some measure of 'fitness' is usually designed to evaluate the individuals against a particular objective, so that when individuals randomly select pairs and reproduce children: the children will be evaluated along with their parents and only the best candidates will survive (survival of the fittest principle).

In the current work, the genetic algorithm is used to evolve populations over several generations (where an individual's "DNA" encode different parameter sets for F_i and n_i). In this way a population emerges in which a more global optimum (best choice that fits the model to the observed data) results, as compared to that discovered via an exhaustive, individual search (which is more subject to discovering local optima).

The Genetic Algorithm consists of 5 steps (outlined in the following):

1. Initial population
2. Selection
3. Cross Over
4. Mutation
5. Evaluation

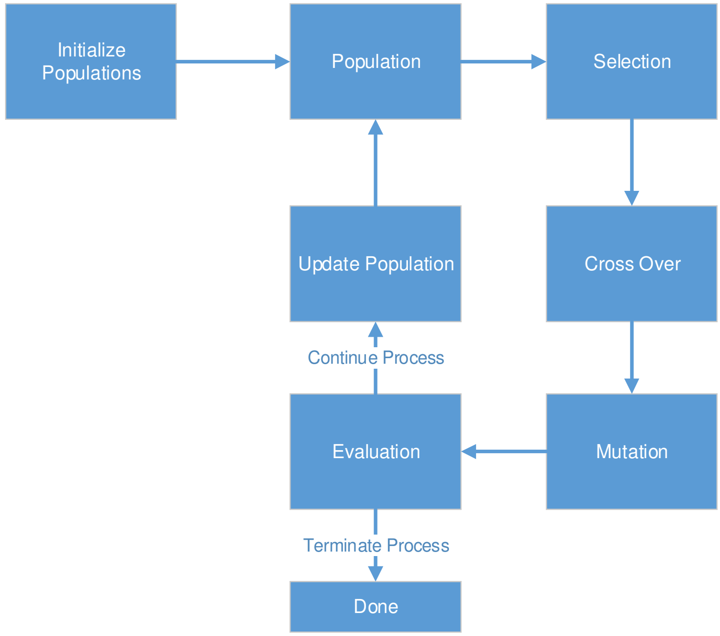

Figure 2: Genetic Algorithm Flow Chart

1. Initialize Population

The dot gain data are arranged into an array with 123 attributes (columns) in an order of Cyan, Magenta, Yellow, and Black with 22 dot gain samples for Cyan, Magenta, Yellow and 21 samples for Black colour in an order from 0 to 100% of the nominal coverage. The 36 spectral n factors are attached to the end of array with its initial value set to 1 (specifically, each row in the data set encodes a particular combination of F_i's and n_i's to fit the model to observed data). The reflectance data from test chart were processed by the dot gain extraction algorithm and constructed into array format as mentioned above. After this process, a final array set was obtained and then duplicated itself 50 times to form a population of 50 parameter sets as an initial population for the genetic algorithm.

2. Selection

The 50 population sets are randomly paired into 25 pairs as the parents of the new generation. Each pair of parents will create two children in the next step.

3. Cross Over

The crossover process is inspired from the natural DNA crossover process, in which two pieces of DNA randomly exchange information and become two new DNA sets. In the experiment, two parents uniformly evaluate each data column (parameter) for exchange with a probability of 50%. The crossover method tries to allow for new combinations of solutions (parent parameter sets) that are already known to be effective from the last generation. The 25 pairs of parents create 50 children and those children, similar to the natural produced children, obtained features from both parents.

4. Mutation

In addition to parental features, children also have their own unique features that are different from both parents, for example child's nose is different from both parents. This process is simulated as a Mutation operation. The algorithm randomly selects 50% of data and changes its value randomly in a range of -1 to +1%. We introduce a few constraints to guide this mutation process for the current work:

1. The dot gain has to equal or larger than original coverage
2. 0% and 100% coverage does not have dot gain
3. n is in a range from 0 to 100

5. Evaluation

After the mutation, 50 children with unique features are produced. The next step is to evaluate those children together with parent generations and rank them in terms of the minimum RMS error between the calculated results against the measured data. The top 50 individuals will be selected as population for the next process round and rest of them will be discarded (die out). This process is simulating the rule of natural selection that only the individuals with the best features will have chance to survive. The evaluation process is similar to the dot gain effective area calculation, which is applying the dot gain data and *n* value into the YNSN model to calculate the RMS error against the measured data. Each individual candidate contains the CMYK 4 channel dot gain look up table and 36 single *n* values. The inputs of the algorithm are the theoretical ink combination and its measured reflectance. The theoretical ink combinations first are corrected from the 4-channel dot gain look up table. The corrected ink combination and estimated *n* value are then applied to the Yule-Nielsen Modified Spectral Neugebauer with spectral *n*-value model to calculate its reflectance.

Results and Discussion

In the experiment the IT8.7/4 test chart was printed on an offset press with 100lb coated paper in AM screening. PressSign software was used to control the ink keys along with an X-Rite eXact to ensure the densities were within GRACoL specifications; 5 sheets (each sheet has two IT8.7/4 test chart) were measured after 24 hours using the iSis. Those 10 measurements were then processed by the α-trimmed mean filter to find out the best 3 measurements in terms of the number counts of data in the median range. Those 3 measurements were averaged and form a final data for the later process. In the final data there were 1588 unique samples, 1271 of them were used as genetic algorithm training data and rest 317 samples were used as a test set to evaluate how well the optimized model can predict the observed data. There was no overlap between the training data and test data. The performance benchmarks are based on the Delta-E 1976, Delta-E 1994 and Delta-E 2000. The 95% best average and 5% maximum error are recorded. The testing results are in the following tables.

Table 2 Prediction and Measurement Colour Difference of spectral n vs. Single n

	Spectral n		Single n	
	Best 95% Mean	Worst 5% Mean	Best 95% Mean	Worst 5% Mean
ΔE_{ab}	1.0042	2.7912	1.2361	3.5783
ΔE_{94}	0.6920	1.7924	0.8368	2.2391
ΔE_{00}	0.6877	1.8382	0.8191	2.1911

The spectral *n*-value method considered the fact that different wavelengths of light propagate in the paper and ink slight differently. From the experiment, the prediction result of spectral n method is significantly more accurate than the single *n*. The mean of the 5% maximum error in delta-E 1976 is 2.7912 which smaller the human eyes range to distinguish the colour difference. The spectral *n* factor and the comparison of original dot gain curve and genetic algorithm fine-tuned curve are in the figures below.

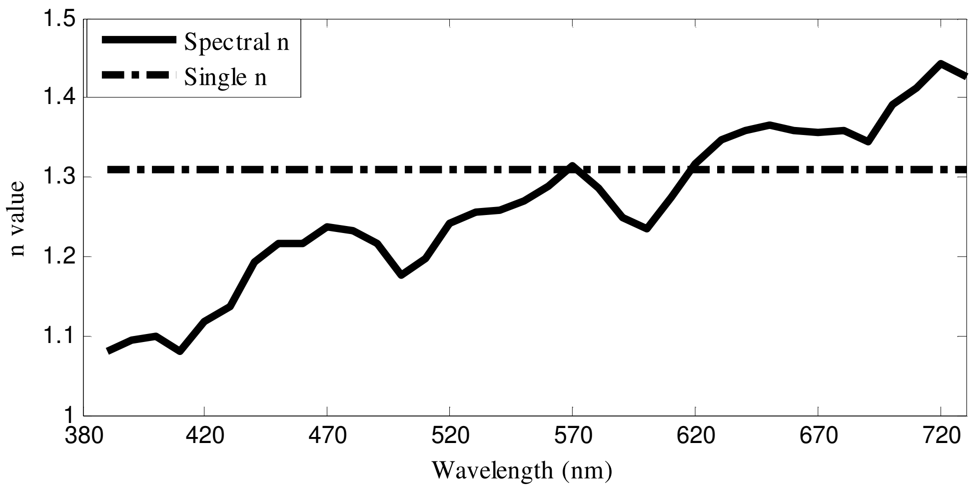

Figure 3 Spectral n values vs. Single n value

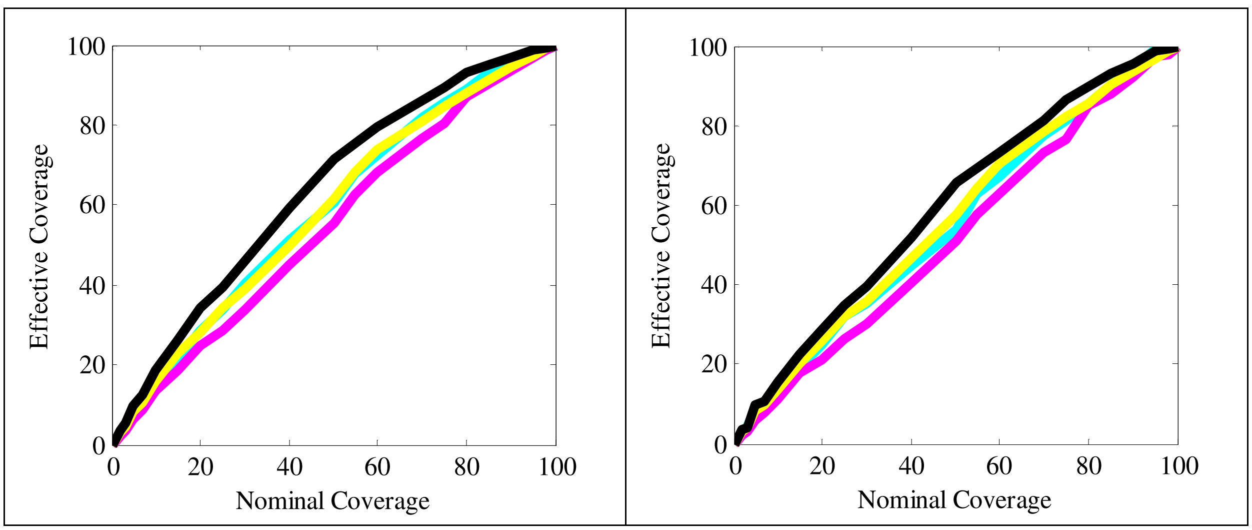

Figure 4 the Original dot gain Curve (left) vs. Genetic Algorithm fine-tuned Curve (right)

Conclusion

This paper presented a method to model dot gain and estimate spectral *n*-value by using genetic algorithm in conjunction with a modified YNSN model. Unlike other algorithms, the genetic algorithm estimates all the parameters in one set to deliver the best representation for the paper-ink-printing model characteristics. The prediction ΔE 5% maximum value in offset press AM screening are smaller than 3, which meets the colour reproduction of printed material requirements. The spectral *n*-value model consistently demonstrates improved ΔE between predicted and observed spectral data, suggesting that different exponents of primary ink reflectances at different spectra might be more suitable in a YNSN-based model to predict the resulting reflectance of particular mixed ink sets. Also, the genetic algorithm proposed has the potential to automatically recover an appropriate model based on measured data from a given ink type and substrate combination, possibly compensating for more subtle effects contributing to dot gain due to ink or substrate properties not explicitly captured by YNSN model.

Acknowledgement

The authors would like to acknowledge everyone who contributed to the success of this work that was supported by FedDev Ontario funding. In particular, thanks to Angus Pady (President, ColourManagment.ca), for his continuous technical support; Mark Sibilia (President, MPP Marketing Group), and Trade Secret Web Printing Inc. for providing access to equipment, onsite press testing and ongoing support.

References

[1] Namedanian M., Gooran S., Nystrom D., *"Investigating the wavelength Dependency of Dot Gain in Color Print"*, Proc, SPIE 7866, Color Imaging XVI, (Jan 2011)

[2] Johansson K., Lundberg P., Ryberg R.,*"A Guide to Graphic Print Production*", Wiley 3 edition (Nov. 2011) ISBN 978-0470907924

[3] Wyble D. R., Berns R. S., *"A Critical Review of Spectral Models Applied to Binary Color Printing"*, Color research and application, p4-19, 25-1. John Wiley & Sons, Inc., New York (Feb. 2000)

[4] Melanie M., *"An Introduction to Genetic Algorithms"*, Cambridge, MA: MIT Press. (1999) ISBN 9780585030944

[5] Urban, P. and Grigat, R.-R. *"Spectral-based color separation using linear regression iteration",* Color Research and Application, Vol 31, p229–238 (June 2006)

[6] Arney J. S., Wu T., Blehm C., *"Modeling the Yule-Nielsen Effect on Color Halftones",* IS&T- The Society for Imaging Science and Technology (1998)

[7] Muehlemann S., *"Spectral Printing of Paintings Using a Seven-Color Digital Press",* Master Thesis, College of Imaging Arts and Sciences of Rochester Institute of Technology, (May 2012)

[8] Zuffi S., Schettini R., Mauri G., *"Using genetic algorithms for spectral-based printer characterization"*, Electronic Imaging 2003 (p 268-275), International Society for Optics and Photonics (2003)

Biography

Yiming Qian received his Bachelor degree in Electrical Engineering from Ryerson University in 2012. Currently he is an M.A.Sc. graduate student at the department of Electrical and Computer Engineering at Ryerson University. His specialty is Machine Learning and Mathematical Analysis/Modeling. He has experience on colour theory and printing. Also he has the knowledge on spectral prediction model and Human Visual System. He previously worked on color printing related projects.