Malcolm – a New Partner in Graphic Art Quality Control

Kerstin Malmqvist*, Antanas Verikas*, Lars Bergman*, Lennart Malmqvist**

Keywords: Newspaper, Printing, Quality, Colorimetry

Abstract: Driven by a need for a new versatile system for quality control in graphic arts the Malcolm concept has been developed. The basis for the Malcolm concept is a new method for analysis and control of color. The printed result on paper is stripped color by color into different pictures in e.g. a 4-color print. The color classification is made by artificial neural net algorithms. These pictures, now in a digital form, can then be used for calculations of different properties of the printed result for each one of the separate printing cylinders. In an application specially designed for work in the printing plant a new measure is introduced which expresses a measure directly related to each one of the very inks used in the press. The measurement can be made in an almost arbitrary part of a picture. One advantage with this concept is that the corrections for the values for each color which are needed for the optimal result can be directly related to adjustments necessary on the press.

Introduction

Our challenge has been to develop a new concept for fast and robust color measurements during the printing process directly in printed multicolor pictures. The measurements can be used for direct control of ink amount in order to achieve an optimal reproduction that corresponds to the demands from, for example, an advertising agent.

The Physics of Paper and Print Group, **PPP-Sweden**

* Centre for Imaging Science and Technologies, Halmstad University, Sweden
URL: //http:www.hh.se/dep/cbd/cist.html

** Dep. of Physics, Lund Institute of Technology, Sweden
URL: //http:www.fysik.lth.se

The impression of a multicolor print is the result of complicated light scattering in printed and unprinted parts of the area studied. In the printed parts we can find sub-areas with both single and multiple layers of ink. Consequently, there are many factors which influence the color impression:

- the amount of ink inside the dots
- the size of the dots
- the printing order
- the color and structure of the unprinted paper
- the properties of the ink used
- etc

To get a high quality print it is important to find the optimal combination of these factors with respect to the properties of the printing press. However, in the very printing process when the press is running, the factor controlling the printed result is the amount of ink. The press operator controls the final result by adjusting the ink feed. It leads to the conclusion that the press operator needs a fast and robust tool, which can give him direct information about the adjustments necessary to get an optimal result.

The concept of color impression

The traditional way to control the print result in newsprint is to use solid print test areas for the four inks cyan, magenta, yellow, and black. The four ink densities are measured using a densitometer and the ink feed is adjusted towards a set of standard densities. But measuring the solid print density does not take into consideration the effect of dot gain or any of the light scattering phenomena which appear in a multicolor screen area. The color space used by a densitometer has a limited correspondence with the actual colors of the inks used and lacks sensitivity with regard to color changes, e.g. when ink gets dirty.

It is important to get a direct correspondence between the result from a measurement and the effect obtained by adjustments of the different colors in the press. One consequence is that the characterisation of the color impression should be expressed in a 4-dimensional color space where the unit vectors coincide with the very colors of the inks actually used and the type of paper. Even the properties of the actual press will influence the color space to some extent. It is easy to understand that such a color space rarely can be orthogonal or even linear.

We have introduced the concept of color impression as a color vector in the color space described above. The components of the color vector integrate information from both actual tonal values and ink densities.

To be able to create such a true color space we had to find algorithms that are possible to calibrate in an easy way to the specific combination of ink – paper – press used. The color space should work for the whole range of tonal values and all combinations of multicolor screens. Then we also had to take into consideration some light scattering phenomena like the fact that the observed color inside the dot changes with the size of the dot. Another phenomenon that had to be taken into consideration is that the observed color inside a dot varies due to interference with other colored areas even if no direct overlapping is present.

For creation of the color space and for interpretation of the results from the measurements artificial neural net algorithms have been used. The advantage of using such algorithms is the ability to train them without having full information about ink, paper, etc.

The Malcolm concept

The methods for color measurements described above have been implemented in an instrument, called Malcolm. For the image acquisition a specially designed measuring device is developed containing a CCD camera and six photo diods as the light source. The recorded image consists of 512x512 pixels covering an area of 5x8 mm^2.

For the aim of controlling the ink feed during the printing process a special software tool is designed. The measurement can be done in an almost arbitrary part of a multicolor picture. However, mostly the press operator wants a fast answer giving direct adjustment signals for each one of the colored inks. To obtain this, the measured values are compared with beforehand given reference values and the differences between them are presented as recommendations to adjust the ink feed. Such a measurement takes only 1 second.

When comparing the measured values with the reference values it could be some positioning problem in the case of measuring in an arbitrary part of a picture. Therefore the best way of using Malcolm for controlling purpose is to measure on double greybars where one of the halfs is printed grey by using cyan, magenta, and yellow dots and the opposite half is printed grey as a black halftone screen as shown in Figure 1.

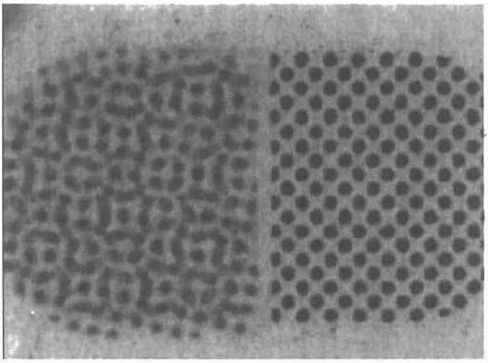

Figure 1. An example of a double greybar. The left part is printed with the three inks cyan, amgenta, and yellow. The right part is printed with black ink only.

The results from the measurements are presented in columns with the reference values marked as shown in Figure 2. The results are saved in a file as well, for later presentation, e.g. as trend curves, and analysis of the actual ink feed.

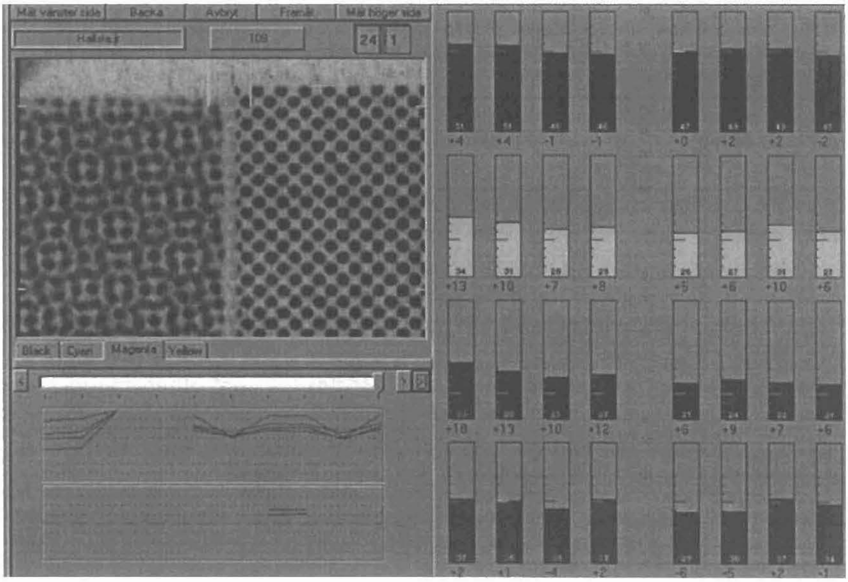

Figure 2. Example of result from Malcolm measurements

Color classification

To be able to scrutinise the printed result from each printed color separately the printed result on paper had to be stripped color by color into 4 different pictures in e.g. a 4 color print. Each one of these 4 pictures, now in a digital form, can be used for calculations of different properties of the printed result for each one of the separate printing cylinders. This stripping of the final result is made by means of artificial neural net algorithms.

A 4 color CMYK print consists of 4 different overlapping dot screens. The black ink is opaque and conceals what is printed beneath, while the three inks cyan, magenta, and yellow are transparent to some extent. Consequently, the printed result is built up by sub-areas belonging to one of the nine following color classes: C, M, Y, K, CM, CY, MY, CMY, and W, where W represents the unprinted paper. A few of these classes are difficult to separate. This is especially true for the two black classes K and CMY. Never the less, with the improved versions of neural net algorithms we have succeeded in separating between also these two classes (Verikas *et al.*, 1997, 1998 and 1999).

Figure 3 shows two classification results. In the left part of the figure the eight classes after a classification of a 3 color print is represented by eight different colors. In the right part the classification result of a print with yellow dots on a magenta background is presented. For the human eye it is almost impossible to identify any dots in the image. The color classification algorithm, however, maps them quite adequately.

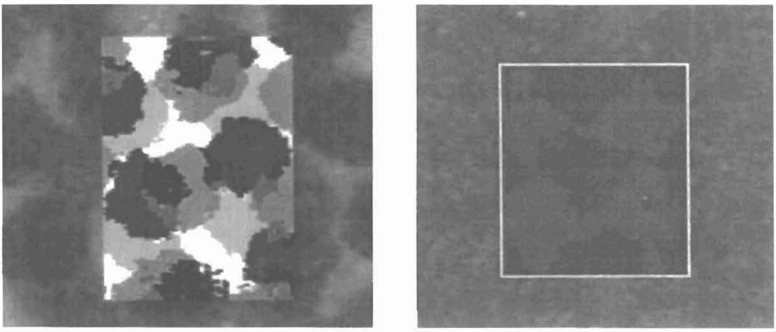

Figure 3. Classification results.
Left: eihgt color classes. Right: yellow dots on magenta background

Figure 4 shows the result from a color classification of a 3 color CMY print stripped into three separate pictures containing the cyan, magenta, and yellow dots respectively. The reconstructed cyan picture, for example, is build up by the color classes C, CM, CY, and CMY..

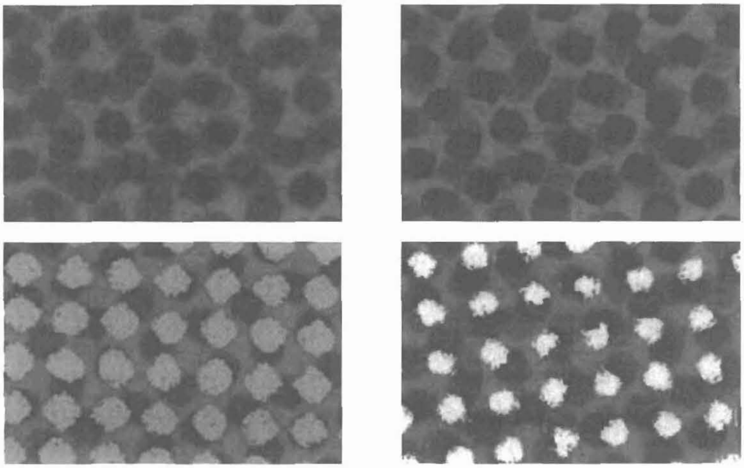

Figure 4. Color stripping of a 3 color CMY print
Top left: original image; Top right: reconstructed magenta dots
Bottom left: reconstructed cyan dots; Bottom right: reconstructed yellow dots

Figure 5 shows the outcome of a separation test between the two black color classes. The image shown in the left part of the figure, is taken from a print that consists of 100% yellow, 100% magenta and cyan dots of 20% tonal value printed on top of the two others. In the right part of Figure 5 the cyan dots are replaced by black dots. It means that in the left part we have CMY-black dots and in the right part we have K-black dots. After classification those pixels which are classified as CMY-black are marked in grey, Figure 5a, and those pixels which are classified as K-black are marked in pure black, Figure 5b. As can be seen from the figure there is no problem in distinguising between the two black color classes. Only 4% of the pixels are not correctly classified, but that small number does not influence the conclusion which black dots are printed by CMY and which by K.

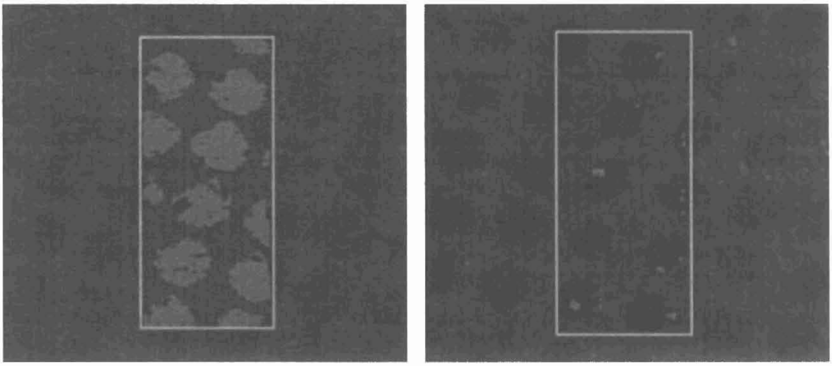

Figure 5. Two black color classes

Missing dots

The color classification technique is used as part of a special tool, developed for studying missing dots in the different printed colors in a multicolor print. In a first step the multicolor print is stripped into separate images for each one of the inks used, as described earlier. These separate images are binary images, where each pixel belongs to a dot or to the background.

Next step is to identify empty areas in each one of the images. Here we use the city block distance transform (Haralick *et al.*,1992). In Figure 6, an original image of a multicolored print and the result after finding the magenta missing dots is presented. For each color the total area and the number of missing dots is determined.

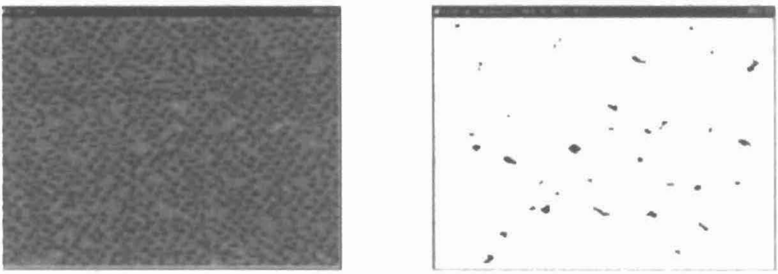

Figure 6. Left: Original image of an multicolored print.
Right: Missing dots of color magenta found in the original image
Artificial neural network methods for color measurements

A neural network is a parallel, distributed information processing structure of processing elements interconnected via signal channels called connections. The strength of the connections are characterised by weight values. The processing elements are sometimes referred to as units, nodes or neurons. Each processing element has a single output connection that branches into as many connections as desired. Most of known neural networks have their processing elements divided into subsets, called layers. The layer related to the input is called an input layer, and that related to the output is called an output layer. The internal layers are referred to as hidden layers. The type of function performed by a network of a given structure depends on values of weights that are determined by minimising some error functional. The estimation process of network weights, which is most often done by using the error backpropagation algorithm (Rumelhart *et al.*, 1986) is called learning or training. See for example (Bishop 1995) for a deeper study of feedforward neural networks.

In our work, we use a feedforward multilayered neural network. The network performs colour classification or colour transformation. The input to the network is given by the *rgb* or $f_1 f_2 f_3$ triplet. The $f_1 f_2 f_3$ triplet is obtained from the r, g, and b by performing a linear transform of the $\{r, g, b\}$ vector. Therefore there are three nodes in the input layer of the network. Since the network has to distinguish between nine colour classes, there are nine output nodes in the classification network. The colour transformation network has three or four output nodes depending on the fact how many printing inks are used.

In order to train the network each input vector must be paired with the desired output vector. The colour classification network performs a pixel-wise image analysis. A considerable amount of pixels from all the nine colour classes is collected for training the classification network. The desired output vector is determined for each pixel used in the training process. Let us assume that the first output node of the classification network stands for the "white paper" class and a "white" pixel from the collected set is considered as a training instance. The $f_1 f_2 f_3$ triplet of the pixel then gives the input training vector. The components of the corresponding desired output vector are set to the following values: $\{1, 0, 0, 0, 0, 0, 0, 0, 0\}$.

The input to the transformation network is given by the mean values of the variables f_1, f_2, and f_3, averaged over a certain area on a printed picture. The variables for training the network are measured on a number of test patches. We recall that output of the transformation network depends on both actual tonal values and ink densities. Since all the test patches were printed keeping the same, nominal, ink densities, we used an estimate of the actual tonal values of the patches as the desired output vector. The estimates of the actual tonal values are easily obtained from the classification network.

Both classification and transformation networks were trained by minimising the average squared error measured as a difference between the actual and desired output values of the network. The Gauss-Newton approximation to Bayesian learning was the training technique applied (Foresee *et al.,* 1997).

The correspondence between color impression and solid print density

As described earlier, we define the concept of color impression as a color vector in the 4-dimensional color space, derived from the described artificial neural net algorithms. The different color components of the vector depend both on the ink densities and the sizes of the dots. In order to be a useful concept in graphic art quality control, the components of the color vector should prove to have a full correspondence with the amount of ink applied on the print. To prove this we made an experiment.

Four series of test prints were produced. The prints contained both solid print bars and multicolored screen areas in the same columns. In each one of the series, containing 40 copies, the density of one colored ink was varied while the densities of the other inks were forced to hold a constant level. In the magenta series, for example, the amount of magenta ink was varied in such a way, that the solid print density, when measured by a densitometer in the solid print bar, varied between 0.5 and 1.15. Figure 7 shows the solid print densities for the cyan, magenta, and yellow inks, measured by densitometer in solid print bars.

Figure 7. Solid print densities in 40 copies of the test print.

In the first copy of the test prints, the solid print density of the magenta ink was held 0.5 and after that the density was increased gradually up to 1.15. In copy number 31 we made a sudden decrease of the density down to 0.6 and after that an increase again up to 1.1. The density of the cyan ink was held approximately 0.9 and the density of the yellow ink was held 1.05 through all 40 copies.

The color impression vector was measured by means of Malcolm in a multicolor screen area in corresponding columns of the copies of the test prints where the solid print densities were measured. The screen area we chose was part of a picture showing the face of a girl.

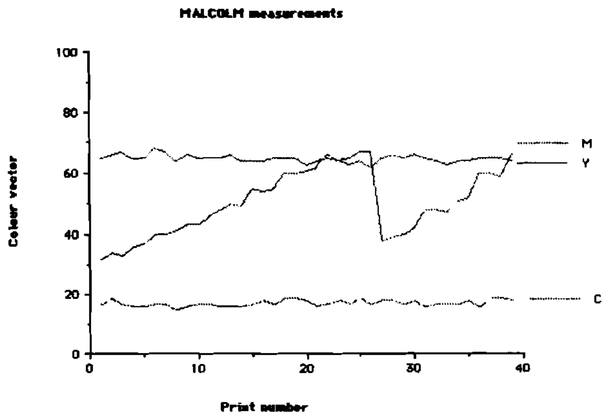

Figure 8. Color impression vector measured in a screen area

As can be seen from Figure 7 and Figure 8 there is a high correspondence between the curves. The correlation coefficient between the two magenta series is 0.96. Furthermore Malcolm gives information about the amount of the colored inks inside the screen area, when the densitometer only reports the solid print density.

Discussion and conclusion

In this paper, we have presented a method for color measurements directly on printed half-tone multicolored pictures. We introduced the concept of the *CMYK* color vector, which lives in the 4-dimensional space of printing inks. Two factors contribute to the values of the vector components, namely, tonal values and ink densities. The color vector expresses integrated information about the tonal values and ink densities. If some reference values of the color vector components are set in advance, for example from a preprint, the color vector then directly shows how much the operator needs to raise or lower the cyan, magenta, yellow and black ink densities in order to correct the colors of the picture being measured. The values of the components are obtained by registering the *RGB* image from the measuring area and then transforming the set of registered *RGB* values to the triplet or quadruple of *CMYK* values respectively. We presented a neural networks based approach for performing such a transformation as well as for inverse color separation.

Experimentally we have shown the validity of the proposed approach of color measurements in half-tone multicolored newspaper pictures. We have found a good correlation between components of the color vector and ink densities.

The Malcolm instrument has during some months been tested at a Swedish newspaper printing plant where the color impression tool has ben used for control of ink feed during the printing. The instrument has proved to significantly facilitate the work with quality control of color reproduction.

The ability of the tool developed to perform the inverse color separation provides a possibility to examine the obtained picture at a microscopic level and to compare the picture with the desired result separately for each ink used. This gives a possibility to study the interaction between different types of paper, ink and printing devices.

The system developed provides a way of fast and objective comparison between two pictures and also enables the user to examine variations in the printing process over time. The comparison can be made even when printing on colored paper. The use of the system can result in reduced consumption of inks and, therefore less severe problems of smearing and printing through.

Literature cited

Bishop C.
1995 "Neural networks for pattern recognition"
Oxford University Press, 1995.

Foresee F.D., Hagan M.D., "Gauss-Newton approximation to Bayesian learning"
1997 Proceedings of the 1997 International Joint Conference on Neural Networks, 1930-1935, 1997.

Haralick R.M., Shapiro L.G.
1992 "Computer and robot vision"
Addison Wesley, 1992

Rumelhart D.E., Hinton G.E., Williams R.J
1986 "Learning internal representations by error propagation"
Parallel Distributed Processing, vol. 1, MIT Press, Cambridge, 1986.

Verikas A., Malmqvist K., Bergman L.
1997 "Colour image segmentation by modular neural network"
Pattern Recognition Letters. **18**, 173-185, 1997

Verikas A., Malmqvist K., Bergman L., and Signahl M.
1998 "Colour classification by neural networks in graphic arts"
Neural Computing & Applications, 1998 (7), 52-64.

Verikas A., Malmqvist K., Malmqvist L., Bergman L.
1997 "A new method for colour measurements in graphic arts"
Color Research & Applications 1999 (3)