Linköping Studies in Science and Technology
Licentiate Thesis No. 1147

Using Multicoloured Halftone Screens for Offset Print Quality Monitoring

Lars Bergman

Linköpings universitet
INSTITUTE OF TECHNOLOGY

Department of Science and Technology
Linköping University
SE-601 74 Norrköping
Sweden

HÖGSKOLAN HALMSTAD

School of Inf. Science, Computer and Electrical Eng.
Halmstad University
Box 823, SE-301 18 Halmstad
Sweden

LiU-TEK-LIC-2005:02

Halmstad January 2005

Using Multicoloured Halftone Screens for Offset Print Quality Monitoring

PPP-group
Intelligent Systems Laboratory
School of Information Science, Computer and Electrical Engineering
Halmstad University, Box 823, SE-301 18 Halmstad, Sweden

ISBN 91-85297-25-9 ISSN 0280-7971
Printed in Sweden by UniTryck, Linköping, 2005

Malcolm *(From Bra Böckers lexicon, 3rd edition 1987, my translation)*

Malcolm III Canmore was the king of Scotland from 1059 to his death 1093. He conquered Macbeth who had earlier murdered his father Duncan I. Malcolm supported the Anglosaxsons against the Norman kings in England. He was however, forced to temporary acknowledge the superiority of Wilhelm the conquerer.

—

In the early nineties we needed a name for the instrument we were developing at Halmstad University for measuring print properties. Our cooperation partner at that time, Swedish Newsprint Research Center (TFL), owned a densitometer from Macbeth. I therefore named the instrument Malcolm to show that we were using new approaches different from the traditional ones, when developing the instrument and methods.

Abstract

In the newspaper printing industry, offset is the dominating printing method and the use of multicolour printing has increased rapidly in newspapers during the last decade. The offset printing process relies on the assumption that an uniform film of ink of right thickness is transferred onto the printing areas.

The quality of reproduction of colour images in offset printing is dependent on a number of parameters in a chain of steps and in the end it is the amount and the distribution of ink deposited on the substrate that create the sensation and thus the perceived colours. We identify three control points in the offset printing process and present methods for assessing the printing process quality in two of these points:

- Methods for determining if the printing plates carry the correct image
- Methods for determining the amount of ink deposited onto the newsprint

A new concept of colour impression is introduced as a measure of the amount of ink deposited on the newsprint. Two factors contribute to values of the colour impression, the halftone dot-size and ink density. Colour impression values are determined on gray-bars using a CCD-camera based system. Colour impression values can also be determined in an area containing an arbitrary combination of cyan magenta and yellow inks. The correct amount of ink is known either from a reference print or from prepress information. Thus, the deviation of the amount of ink can be determined that can be used as control value by a press operator or as input to a control system. How a closed loop controller can be designed based on the colour impression values is also shown.

It is demonstrated that the methods developed can be used for off-line print quality monitoring and ink feed control, or preferably in an online system in a newspaper printing press.

Contents

Chapter 1

Introduction

Printed media is one of the most common ways of communication. It covers a variety of products such as books, brochures, magazines, newspapers and others. Despite the exploding amount of information available on the internet, the diversity of printed products is not declining. Newsprint takes a special place in this collection of printed media. In the newspaper printing industry, offset is the dominating printing method. Offset printing is an indirect lithographic printing technology spreading since the 1970's and is the far most used printing method today.

Traditionally the printing industry has been regarded as a craftsman like industry. This is still true to some extent, but modern newspaper printing is more of an advanced process industry where care is taken to make the process consistent and predictable. The demand for increased print quality requires better ways of judging the print quality then what is possible using only the naked eye. We need objective means for determining values characterizing those properties of the printed result that can be translated into actions to take for decreasing the deviation from the desired result. The objectivity means that we have to measure the properties in some way.

The result obtained from the offset printing process is dependent on a variety of parameters. We can divide the parameters into nine different groups depending on the origin: press, ink, inking system, blanket, printing plates, dampening solution, dampening system, substrate and print job. Most of the parameters usually change slowly over a long period of time such as days or months. These changes can be kept under control by overview and maintenance of the equipment. Some of the parameters change from one print job to another, others during the print run. To compensate for variations of the parameters, the operator has to constantly monitor the print and take appropriate actions during the print run.

The use of colour printing in newspapers has increased rapidly during the last decade.

It is not wrong to state that newspaper producers today have the capability to produce colour print on the majority, if not on all pages. In colour printing, four layers of different colours are printed separately, one on top of the other, to give the impression of large colour gamut through colour mixing. In four colour printing the inks are cyan, magenta, yellow and black. To print a colour image it has to be separated into the four basic colours, or *separations*. Printing in colour increases the demand for measuring even more and raises the need for methods capable to measure not only each single ink, but also the overprints of the inks.

In the offset printing process, where there is a need of measuring multiple properties, a imaging based monitoring system is advantageous since such a system gives a possibility to implement methods for measuring a variety of properties with only one sensor. Therefore, the methods developed in this work utilize an imaging based system.

The offset printing process relies on the assumption that for each separation an uniform film of ink of right thickness is transferred onto the printing areas. Print quality is a complex matter dependent on a variety of factors. The amount and the distribution of ink deposited on the substrate are two key properties affecting the print quality. In this thesis, we present methods for assessing the amount of ink transferred onto the the substrate. From only one measurement on a multi coloured halftone area, the average amount of cyan, magenta, yellow, and black inks in the area is provided. In contrast to the colour shift measurements, such as ΔE, the obtained values are easy to interpret for the press operator. It is demonstrated how the obtained values can be used directly for print quality monitoring and manual as well as automatic ink feed adjustments.

The thesis is organized as follows: In Chapter 2, the offset printing process is described with the focus on newspaper printing. The related work is discussed in Chapter 3. Chapter 4 presents the proposed way of monitoring the amount of ink in well known screen areas, as well as the way of determining the amount of ink deposited on the substrate. In Chapter 5, the method for monitoring the amount of ink in an arbitrary inked area is presented. The estimated deviation of the amount of ink has to be interpreted and translated into control actions taken to compensate for the deviation. In Chapter 6, we show how a closed loop controller can designed based on the measures described in Chapter 4 and 5. Finally discussion and conclusions are presented in Chapter 7 along with the suggestions of further improvements.

Chapter 2

Offset printing

Offset printing is an indirect lithographic printing technology. It is the far most used printing technology today. The offset printing technology is used for a wide varity of product types ranging from art print to mass produced advertisements.

The lithographic printing method was invented in the end of the 18th century. It uses stone material as a printing surface. Lithography originates from the Greek word ***lithos*** that stands for stone. In lithographic print, the areas we want to transfer ink – *the printing areas* – and non printing areas are on the same level. This is possible by exploiting the feature that oil and water are mutually repellant. The non printing areas are water accepting and the surface in the printing areas is altered to be water repellant and ink accepting. By dampening the surface we can apply the printing ink with a roller and be certain that it will only be transferred to the printing areas. When the substrate is pressed onto the surface the ink will be transferred onto it. Additionally will some of the water be transferred onto the substrate, which can cause problems in some cases, as we will see later.

In offset printing, contrary to the lithographic printing, the substrate is not put in a direct contact with the printing surface. The ink is instead transferred from the printing surface to the substrate via a flexible intermediate carrier called blanket, therefore the name offset. In modern offset printing, printing plates have replaced the stone material as the printing surface. The surface of the plates can be altered to be either ink accepting or water attracting, often photographically. For further reading about printing technologies see [1, 2].

2.1 Ink

Ink contains three basic ingredients: pigment, vehicle and additives. The pigment is made of solid particles that give the colour to the ink. The vehicle is the fluid that carries the pigments and makes them adhere to the substrate. The additives are included to give the right characteristics to the ink.

After the ink has been transferred onto the substrate it has to be anchored to it. This process is called *drying* and is a two stage process. In the first stage, the ink must *set* or stick to the substrate. In the second stage, a stable linkage between the ink and the substrate is created while the ink solidifies. This is achieved by either a chemical or physical reaction. In heatset printing external energy for example in the form of hot air, IR, UV or microwaves is applied to dry the ink. In coldset printing the physical interaction between ink and paper or evaporation of a fluid is used to dry the ink. To print newspapers coldset printing is usually used.

2.2 Substrate

The substrate can be fed through the press in the form of sheets or as a web. Web-fed presses can run at a higher speed than sheet-fed presses, since the substrate guides it self through the press. But web feeding is associated with the following problems not found in sheet-fed presses:

i To keep the web stable in its path through the press a certain amount of tension has to be applied. If the applied tension is to low the web will not run stable in the press and the result is decreased printing quality, unwanted wrinkles, and folds, or the web will break. On the opposite, if the applied tension is to high the web will break.

ii The applied tension will cause dimension changes of the substrate.

iii The dampening solution will change the dimension of the substrate such as paper and will thereby influence the web tension. Changes in the printing speed will indirectly change the web tension. Printing starts and stops will be especially critical.

iv Defects in the substrate can cause web breaks. Web breaks cause rather lengthy and thereby costly stops in the printing process. The term *runability* is used to sigify how the substrate behaves in the press.

Newsprint is inexpensive paper made of mechanical pulp (grained wood). The wood contains lignin and therefore the paper yellows rather quickly. Today newsprint often contains a large portion of recycled fibers even up to 100 %.

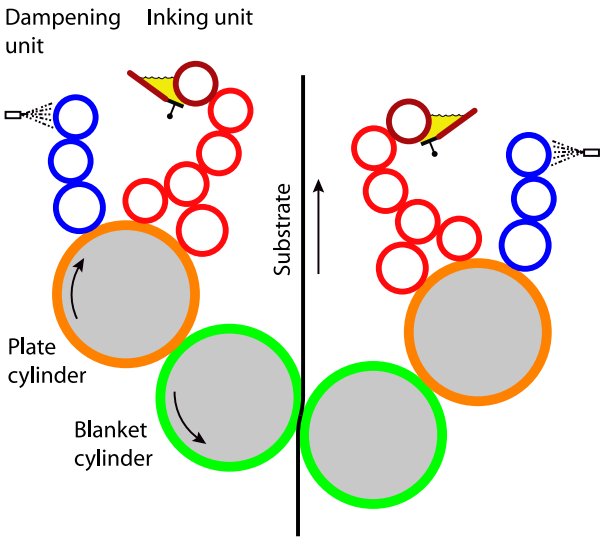

Figure 2.1: Example of a double sided offset printing unit.

2.3 The printing unit

The offset printing unit, the scheme of which is shown in Figure 2.1, consists of the following parts:

- The inking unit.
- The dampening unit.
- The plate cylinder with the printing plates.
- The blanket cylinder with the blanket.
- The impression cylinder. In a double sided printing unit, the opposite blanket cylinder replaces the impression cylinder.

The inking and dampening units are shown in detail in Figure 2.2.

2.3.1 The inking unit

The inking unit supplies ink to the printing areas on the printing plate to maintain a constant film thickness on the plate. The inking unit must replace the "used up" ink.

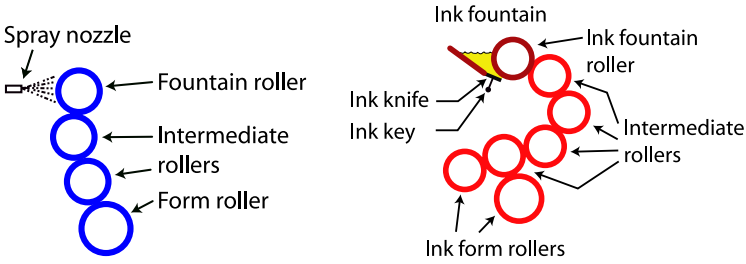

Figure 2.2: Left: Example of a dampening unit. Right: Example of an inking unit.

In a typical newspaper press inking unit, the ink is fed from a tray or a fountain by a fountain roller trough an adjustable aperture formed by a metal blade, the *ink knife* or *ink doctor*, via a series of intermediate rollers to the ink form rollers onto the plate. The ink knife is divided into sections called *ink zones* across the press. Figure 2.2 shows a typical newspaper press inking unit.

The amount of ink fed is controlled by adjusting the width of the aperture with a mechanism often referred to as an *ink key*, and by adjusting the speed of the fountain roller. The intermediate rollers are used to make the ink film uniform, and have the effect that ink is transported not only in the printing direction but also in the cross direction. The amount of ink needed in an inking zone is determined by the inked area in that zone. So, the intermediate rollers have to make the ink uniform without disturbing the ink supply in different zones. To maintain an even ink film thickness in the printing direction, the arrangement of the ink form rollers is vital.

2.3.2 The dampening unit

The dampening unit supplies a constant film of dampening solution to the non printing areas of the printing plates. Since some of the dampening solution is printed via the blanket, some is emulsified into the ink and printed with it, and some is evaporated. A constant supply of the solution has to be maintained to the printing plate.

The dampening solution can be supplied in a variety of ways:

- Like in the inking unit, from a fountain with a roller except that the solution is lifted out with the roller so there is no knife.
- With a brush system that sprinkles the dampening solution onto the roller.
- With a spray system where water is sprayed onto the roller.

The spray system is the one, where it is most easy to control the amount and dis-

tribution of the dampening solution. The film of dampening solution is then made uniform by intermediate rollers and fed to the printing plate with one form roller, as shown in Figure 2.2.

The dampening solution is water with additives to reduce the surface tension. The chemical and physical interaction between the dampening solution and the ink is vital to keep the printing and non printing areas separated. If the separation can not be maintained, *toning* occurs, the printing plate carries ink also in the non printing areas. If the supply of dampening solution is to high *scumming* occurs. To much dampening solution will be emulsified into the ink making the print grainy. The water in the dampening solution makes the fibers in paper swell, causing a growth in the dimension of the paper. The effect on newsprint paper is especially large.

2.3.3 Plate cylinder with the printing plates

An offset printing plate is typically an aluminum plate with the surface grained to be hydrophilic and coated with a thin ($1\mu m$) layer of photo alterable material. The plate making process consists of removing the coating from the non printing areas.

2.4 Monochromatic printing

An image such as a photograph is a *continuous-tone (contone)* image built up of shades of gray or different *tones*. The shade of gray is often expressed in terms of *density* characterizing the ability to absorb light.

In offset printing, it is not possible to vary the ink thickness or *density* locally in different print areas. To be able to reproduce a continuous-tone image, the image has to be converted into a *halftone* image. In a halftone image, the different tones are made up of dots of different size but equal *density*. The halftone dots can be modulated in a variety of ways.

The traditional way is to divide the image into a structured grid – a *screen* – and in each screen cell an image element – a *halftone dot* – is varied to create the different tones [3]. This is often referred to as *amplitude modulated (AM) screening*. Another way is to modulate the number, or frequency, of equal sized small dots, referred to as *frequency modulated (FM) screening*. There also exist combinations, *hybrid screens*, of these two.

The term *screen ruling* is used to express the size of a screen cell. The more narrow the screen is the more natural the image appears. However a higher screen ruling put higher demands of the printing process and the substrate. A typical newsprint screen is an AM screen with a screen ruling, or actual screen frequency, of 85 lines/inch

(LPI), meaning there are approximate 3 dots per millimeter.

The process of converting the continuous image into the halftone image is referred to as *screening* or *halftoning*. Traditionally amplitude modulated or monospaced clustered single dot screens are used in newspaper printing. Each screen cell contains one single, often a round dot, that is size modulated. When tone values become high enough, approximately 50 %, the dot touches the cell boundaries and thus has to change the shape. When the tone value increases further, we can talk about modulating holes (that is what are left in the corners) instead.

One disadvantage of single dot screens is that the size of the dot in low tone values or the hole in high tone values becomes very small. Printing small dots or holes requires a well controlled and calibrated printing press and high quality substrate. This is one reason why frequency modulated screens are not of common use in newspaper printing. There exist hybrid screens – amplitude modulated screens for the mid tones and frequency modulated fixed size dots or holes for the lower and higher tone values. The advantage of such a screen is that the otherwise required small dots can be avoided. But such screens are not of common use yet.

Modern halftoning is performed digitally. Digital representation means finite spatial resolution and discontinuity in the tone steps. The area coverage and thus tone levels are fixed to discrete steps. The possible screen angles are also fixed to discrete steps. Higher number of steps needs higher output resolution and thus higher bandwidth.

2.5 Colour printing

Using a single ink we are able to reproduce a monochrome continuous-tone image. The far most used ink is the black ink. It has a hight contrast to the substrate and produces sharp edges well suited for text and thin lines.

To be able to reproduce a continuous tone colour image, we have to use more than one ink. By printing inks of different colours on top of each other we are able to produce colour prints. The far most used combination of colours is cyan (C) – a greenish blue, magenta (M) – a purplish red, yellow (Y), and black (K, where K stands for "Key" colour). These four colours are often referred to as *process colours*, and multi colour print using these CMYK process colours are often named four-colour print. By adding black, lines, letters and such can be printed in a shaper way, and the amount of the three chromatic inks is reduced. Coldset printing, such as newspaper printing, is specially sensitive to high amounts of ink because the ink drying relies on the fact that one or more components of the ink are transported into the bulk of the paper. Figure 2.3 shows a colour halftone screen produced using cyan, magenta, yellow and black inks. Correct colour reproduction requires that the inks are printed in the correct proportions.

Printing four colours means that the substrate has to pass four printing units. Of course it is possible to use only one printing unit and let the substrate pass four times each time with different ink, but this is not practical, especially in newspaper printing. One common way of arranging the printing units in newspaper printing is to stack the four printing units on top of each other. This pile of printing units becomes rather high, therefore the name four-high. In a four-high printing press, there is a relatively long distance between the printing units.

There are some problems in connection with four-colour printing that do not exist in single colour printing:

- The prints have to be aligned to each other, referred to as *registering* the prints.
- Each colour has to be printed with the right amount of ink or the resulting colour will differ from the desired one.
- For each printing unit the substrate passes it becomes less receivable for new ink. This is because in the areas where ink is printed, new ink has to stick to an inked surface and not to a unprinted substrate. Moreover, the bulk of substrate such as paper can only contain a certain amount of liquid. In each printing unit, dampening solution is applied to the substrate. The more moisture the substrate is the less receivable it is for new ink and will eventually not be able to receive more ink at all.
- The dampening solution will cause paper fibers to grow. This effect is called *fan-out*. The *fan-out* effect is dependent on a variety of parameters mainly: the substrate, the amount of dampening solution applied, the printing speed and the distance between the printing units, e.g. the run-time between the printing units. The relatively long distance between printing units in a four-high press causes a correspondingly large fan-out.

To print a colour image, the original image, often represented in the RGB-colour space, has to be converted to the CMYK colour space. This process of converting the original continuous-tone image into the four continuous-tone colour images is called *colour separation*, and the corresponding images are called *separations*.

The total space of reproducible colours with a given set of process colors is called *colour gamut*. The *colour gamut* of a typical set of process colours is rather small in comparison to the volume of the RGB-colour space of colour photos and of most input devices such as scanners and cameras. The colour separation process must therefore be able to handle colours that can not be printed. This technique is called *gamut mapping*. *Gamut mapping* is a complicated process, still a field for research [4, 5, 6]. Figure 2.4 illustrates colour gamuts attainable by different reproducing techniques. The gamuts are given in the $CIELAB$ colour space [7].

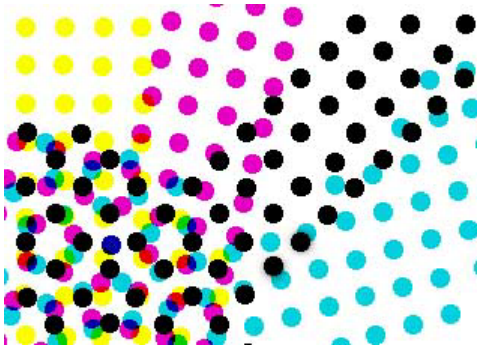

Figure 2.3: Example of a four colour screen with typical screen angles. Note the rosette formation.

Due to different distortions in the printing process, the printed halftone dots are larger than the ones found in the corresponding separations. Two effects can be recognized which affect the size of a halftone dot: mechanical and optical change of the dot-size. The mechanical dot-size is the active printing area on a printing plate or the actual inked area in the print. The optical dot-size is the apparent optical size of the dot. Due to optical effects such as light spreading, mainly in paper but also in ink, optical dot-size is most often larger than the mechanical. The colour separation process therefore has to include compensations for the optical and mechanical effects as well as the optical properties of the substrate and ink, and the interaction effects between those. The erroneous mechanical size of the halftone in the printing plates can be avoided or heavily reduced with the right control and maintenance of the plate making equipment of today.

Due to a high pressure in the printing nip (the contact area between two rollers) the ink is squeezed and the halftone dots will therefore mechanically grow slightly in the print.

In newspaper printing, the dominating effect to compensate for is the optical enlarged appearance of the halftone dots referred to as *Dot Gain*. The cause of this optical enlargement of the halftone dots was first described by Yule and Nielsen [8]. They also proposed a way of modelling this effect. Several other attempts have been made to model this optical dot gain [9, 10]. The optical dot gain in newsprint can be as much as over 30 %.

Interference or Moiré and rosette patterns are two undesirable effects present in multi colour printing. The screens of the individual colour separations are therefore often rotated with respect to each other with angles chosen so that these effects are avoided as much as possible. Figure 2.3 shows an image of a typical halftone screen used in newspaper printing.

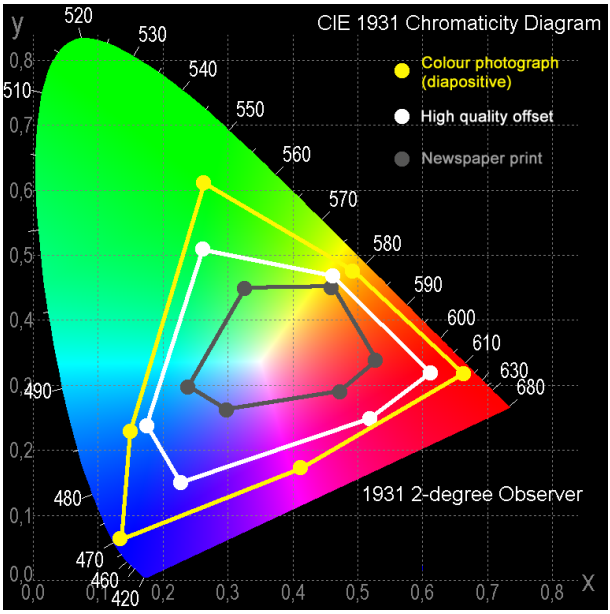

Figure 2.4: Colour gamuts in the $CIELAB$ colour space for different reproduction processes. *Yellow* – Colour photograph (diapositive); *White* – High-quality offset printing; *Gray* – Newspaper printing

Finally, the colour separated halftoned image has to be rendered to the output image, which is an exact image of the printing plate. This output image is our *print original*. All the steps from the original image to the print original are included in the concept *prepress*. In digital prepress, the halftone image rendering is done in a process called *Raster Image Processor* or *RIP*.

The task of the plate making process is to make an exact copy of the print original on the printing plate. Checking the size and shape of the halftone dots on the printing plate is vital for achieving the high print quality. An overview of the colour reproduction in the printing process can be found in [11].

2.6 Plate making

In modern newspaper printing, the dominating method for printing plate making is the so called Computer to Plate or *CTP* technique. CTP is the term used for the technique of computer-controlled direct imaging of printing plates from digital data. In a typical CTP-device a computer controlled laser prints the image on the plate by

altering the coating of the plate in the printing or alternatively the non-printing areas. In the post processing step, the coating in the non-printing areas is removed and the plate is treated as being light resistant.

It is vital that the laser is kept focused, otherwise the light energy applied is spread, which leads to growth of the imaged areas causing dot gain and decreased exposure of the coating. This can lead to that the coating is not developed properly and will be unstable in the printing process.

To fix the plate in a right position in the press, holes and slits are punched in the plate. Positioning of these is vital otherwise the corresponding ink will be miss-aligned with respect to the others.

2.7 Printing

A typical newspaper printing press is a web fed coldset offset press without ink monitoring or a closed loop controller system. During the printing process the operator has to monitor and adjust the amount of ink deposited on the substrate to maintain a correct colour reproduction. The amount of ink deposited on the substrate is dependent on several parameters.

There are five key issues affecting the amount of ink deposited on the substrate an operator has to deal with during a print run:

- The size of the inked area. The size of the inked area is job dependent and will not change during the print run.
- The feed of ink.
- The feed of dampening solution.
- The substrate.
- Environmental parameters, such as temperature and humidity. The temperature and humidity will stabilize after a while and normally the variations can be neglected.

The behaviour of the substrate, newsprint, will vary during the whole print run. Newsprint is fed to the press from reels and a print run will consume several reels of newsprint. The intra reel substrate variation is usually less than the inter reel one.

The amount of ink to be fed is dependent on the size of the inked area, while the amount of dampening solution is dependent on the amount of ink. Increased amount of ink requires increased amount of dampening solution. The relationship is not well

understood yet. Unfortunately there is no way of determining a good balance of the ink and dampening solution . Usually the operator has to relay on his experience.

2.8 Summary

The offset printing process relies on the assumption that an even film of ink of the correct thickness is transferred to the substrate in the printing areas. Thus it is vital for achieving high print quality that ink and dampening solution can be supplied evenly and steady to the printing plates. Controlling the amount of ink deposited on the paper is very important.

The quality of reproduction of colour images in offset printing is dependent on a number of parameters in a chain of steps and in the end it is the distribution and the amount of ink deposited on the substrate that creates the sensation and thus the perceived colours. In this chain we can identify three check-points:

I The print original.

II The printing plates.

III The print.

If we can identify relevant measurable properties and requirements for the properties in each check-point, we can use them to secure the process in that point. If all of the properties fulfill the requirements, we can assure a correct reproduction of colour images.

The output from the prepress is the print original, and it carries the image of the printing plate. The task of the printing process is to make an ink copy with the correct ink film thickness of the printing areas of the print original on the substrate by using the printing plates.

The methods presented in this thesis are not intended to solve any of the problems in the prepress part of the newspaper offset process. Consequently the print original is assumed to be correct.

To fulfill the requirements of check-point II we need methods to investigate the printing plates to assure that they carry a correct copy of the print original. To fulfill the requirements of check-point III we need methods to determining the amount of ink deposited on the substrate.

In Chapter 4 and 5 of the thesis, we present methods for assessing the process quality in check-point II – monitoring the printing plates – and check-point III – monitoring the print regarding the amount of ink deposited on the substrate.

In the first part of Chapter 5, we present a method of determining the dot-size on printing plates. The proposed method does not require any calibration and is not sensitive to the colour of the surface. The method can also be used for determining the approximate mechanical dot-size of the printed halftone dot as well. In the second part of Chapter 4 and in Chapter 5, we present a method of determining the amount of ink deposited on the substrate.

Chapter 3

Related work

In the graphic arts industry, traditionally, a densitometer has been the main measurement tool to obtain information for quality evaluations and process control. In the newspaper printing industry, instruments have mainly been used to calibrate the process in order to obtain information for adjusting the prepress part of the process. The use of instruments to monitor for example the amount of ink during normal production is not very common. One reason can be that traditional methods need some form of designated measuring areas, and such are regarded superfluous. In the newspaper printing industry, the operator has traditionally relied on his eyes for monitoring the quality of the print.

Bearing in mind the high printing speed in modern presses and the increased use of four-colour prints we realize that a printing operator has difficulties in monitoring and controlling the process off-line and can often only manage to tackle the huge deviations. Monitoring the print quality parameters on line, by putting a monitoring system in the press will support the operator in the decision process. By using a closed loop system to control the printing process, the workload of the operator will be decreased significantly and the operator will get more time to manage the overall print quality.

Surprisingly, few approaches and tools for monitoring print properties or controlling the print quality online can be found in the literature. In the existing solutions, the most common print property used is ink density, therefore, designated measuring areas are needed. However, commercial systems for monitoring and controlling registration are more common.

3.1 Traditional methods of print quality monitoring

There are two main traditional approaches to quality monitoring in the newspaper printing industry, namely monitoring by measuring density and the approach based on the concept of *gray-bars*.

3.1.1 Density

Traditionally light absorbtion ability of a material is described as *density* [1]. High density means high absorbance. The instrument used to measure density is called a *densitometer*. Density is given by:

$$D_{ink} = \log_{10}(1/R) \tag{3.1}$$

$$R = I_m/I_i \tag{3.2}$$

where R is the reflectance of the sample, I_m is the reflected light intensity, and I_i is the intensity of the incident light. The use of logarithm makes the density values additive or subtractive in respect to light intensity.

Usually an operator is interested in density in the inked area in respect with the unprinted paper. This is often called *relative density*. Relative density can be obtained by subtracting the density measured on the unprinted paper from the density in the inked area. It can also be estimated by replacing I_i in (3.2) with the intensity measured on the reference area, in our case the unprinted paper.

Density in chromatic printing inks is measured in a narrow spectral band where the ink has a high absorbance and thus a high contrast to the substrate. Figure 3.1 shows reflectance spectra for the cyan, magenta, yellow, and black inks, the corresponding densitometer filter transmittance functions as defined by DIN 16536.

Several print property measures are based on density. The halftone *dot-size* A_p is one and is defined according to Murray and Davis [12] as:

$$A_p = D_{halftone}/D_{solid} \tag{3.3}$$

where D_{solid} is the density measured on the solid ink area and $D_{halftone}$ is the density measured on the halftone area. A_p is often referred to as optical dot-size. From A_p we can determine the dot-gain A_{gain}:

$$A_{gain} = A_p - A_{nom} \tag{3.4}$$

where A_{nom} is the nominal dot-size.

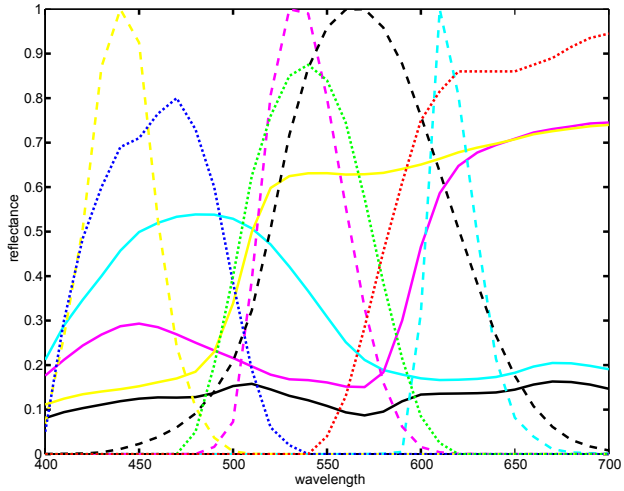

Figure 3.1: Reflectance spectra for cyan, magenta, yellow, and black inks shown as solid lines, and the corresponding densitometer transmittance filter functions, as defined by DIN 16536 shown as dashed lines. Example of transmittance spectra for the red, green, and blue filters for a CCD-camera are shown as dotted lines.

3.1.2 Gray Bars

Gray Bars are test patches made up of two fields. One gray balance field and one black field. *Gray-bars* are assembled to be an easy aid for an operator to, without the use of any instrument, determine the amount of the inks deposited on the substrate in respect to each other. Figure 3.2 shows a *gray-bar* commonly used in newspaper printing industry. The *gray-bars* are not suited for determining the overall level of ink using only the naked eye.

The idea behind the gray balance field is that the right amount of the three chromatic inks, cyan, magenta, and yellow, will produce a neutral colour. If cyan, magenta, and yellow inks are printed with the same halftone screen the result is a brownish tint. The solution in that case is to increase the amount of cyan with approximately 5-10%.

The halftone value in the black field has to be selected so that the black field appears as dark as the gray balance field when the two fields are printed with the right amount of inks. One common combination of halftone values for a *gray-bar* used in newsprint is 30% cyan, 22% magenta, 22% yellow, and 33% black. The use of *gray-bars* for monitoring the amount of ink has become very popular in the newsprint industry lately [13] even though using the the naked eye only an unbalance between

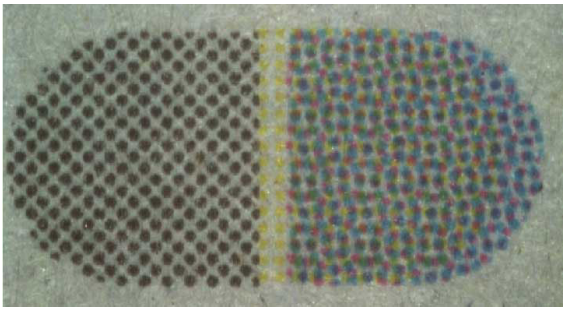

Figure 3.2: A *gray-bar* commonly used in newspaper printing. The left patch is made of 33% black, the right of 30% cyan, 22% magenta, 22% yellow ink.

the amount of inks is easily detected. The approach of using the *gray-bars* for monitoring the amount of ink is to:

1. Detect an erroneous amount of one or more of the chromatic inks, seen as a colour cast in the gray balance field.
2. Detect an erroneous amount of all the chromatic inks if compared to the black ink seen as a difference in darkness between the gray balance field and the black field.

The use of gray balance fields does only reveal an erroneous relative amount of inks in respect to each other. The overall level of the inks have to be determined in another way.

A result of measuring the amount of ink on *gray-bars* is sensitive to erroneous dot-sizes. It is therefore vital that dot-sizes on the printing plates are correct. Otherwise the operator can be misled to compensate for an erroneous halftone dot-size by changing the *density*. This will cause incorrect reproduced colours in areas with other halftone dot-sizes.

Experience shows that despite the easiness for anyone to detect an unbalance between the inks it is not always easy, even for a skilled operator, to decide what action to take to reduce the deviation from the correct amount of ink. This leads to a conclusion that operator needs support in form of an instrument suggesting what to do.

3.2 CCD-camera based monitoring systems

The print quality in newspaper printing is constantly changing due to variations in the process. To reduce these quality changes the operator has to constantly monitor and

control the process. Taking into account the high printing speed, the large amount of pages and the high amount of colour pictures printed in modern newspapers one realizes that it is very difficult for an operator to achieve high quality level by monitoring and controlling the process manually. The solution is to put a system in the press for monitoring the print quality parameters online and to control the process by a closed loop controller.

A densitometer is not well suited to be put online for two reasons. First, there is a problem of synchronizing the measurement with the print, a misalignment is not possible to detect. Secondly, it is not possible to detect defects in the measuring area. Examples of defects are: non printed areas, dirt, and density variations in the printing area. A CCD-camera based imaging system is more suited for online implementation. Such systems have the possibility to detect the misalignment and adjust the measuring area accordingly. It is also possible using such a system to detect defects in the measuring area.

Recently measuring instruments based on a colour CCD-camera imaging system have been increasingly used as alternatives to the traditional instruments such as densitometers for colour printing quality control. Despite the fact that colour CCD-cameras are currently not primarily designed to measure colours, there are several advantages of using colour CCD-camera based systems for colour print quality monitoring:

- Ability to measure multiple printing quality parameters simultaneously.
- Possibility to record an image, which can be further analyzed for obtaining various parameters characterizing the printing process.
- Greater ability to detect measurement errors.

However, despite the aforementioned advantages, the following two drawbacks impede the use of colour CCD-cameras for colour print quality monitoring:

1. The dynamic range of most of the CCD-cameras limits the range of densities that can be measured accurately. CCD-cameras with high dynamic range exist but are expensive.
2. Measuring *density* with colour CCD-cameras is not possible without an additional technique, because spectral responses of filters implemented in colour CCD-cameras differ from those used in traditional densitometers. The filters in a CCD-camera is designed to give as good colour reproduction as possible using three channels, and are rather broad. The filters in a densitometer are designed to give as good contrast as possible between the ink and the substrate, are more narrow, and their transmittance filter functions covers different parts of the spectra, see figure 3.1.

Today off-line CCD-camera based systems are commercially available. Example of instruments based on CCD-camera imaging systems are the ones designed to measure print properties on the specially designed Minitarget [14] test patch.

A review on colour measuring systems used for newspaper printing can be found in [15]. As to the research related literature, we have found only three papers devoted to online print quality monitoring and/or closed loop control systems. The CONES system [16] is an off-line expert system designed to capture the knowledge of an experienced press operator and to "learn" the actions he takes to compensate for the colour deviations in the print. The system uses the deviation in ink *density* between a reference print and the corresponding actual print as input to a neural network, which is trained to determine which ink needs adjustment and the amount of change needed. A rule based expert system uses the rules to determine which ink keys need adjustment and how much to adjust these ink keys. The output is ink key adjustment suggestions to the operator.

The CLOOP [17] is a closed loop system that measures solid ink *density* in a colour bar. The authors have solved the problem of synchronizing and detection of defects in the measuring area by using a CCD-camera. The *density* is then measured with a spectrophotometer. The output from the system is control signals to the inking system.

The ARGUS system [18] is a more complete guidance tool for the operator. It offers four colour measurement values to the user: ink *density*, registry information, grey balance, and colour deviation information. The first three values are measured in areas automatically determined from the Post Script page description. The colour deviation is measured in areas defined by the user. Additionally the ARGUS system can calculate ink preset values from the post script information. The ARGUS system uses a matrix 3-chip colour CCD-camera and stroboscopic light mounted on a traversing unit making it possible to acquire images anywhere from the print. To enable positioning at a certain point in the print, the system relies on printed reference points. The ink *density* show the highest accuracy amongst the parameters measured. Colour deviation and grey balance values are given in the CIE $L^*a^*b^*$ colour space. The accuracy achieved on tested areas of halftone newsprint is moderate . The calculation of the registry values is based on the cross-correlation of the CMYK-separations. The results are not accurate enough for registry control – mainly due to difficulties in estimating the CMYK-separations from the RGB-signals.

Some commercial online ink amount monitoring and closed loop control systems exist, but they are not mainly aimed towards the newspaper industry. Two examples of online systems are "Color Quick" and "Intelligent Density System". These systems show a high resemblance with two of the systems presented above. More information on these two systems can be found at www.gmicolor.com and *www.qipc.com.*

3.3 Summary

The oldest way of quantifying the amount of ink in a print is to measure density and the measurement is normally obtained in solid print areas. But obtaining print quality properties in solid ink areas has one drawback, the properties do not reveal how the process affects the print in halftone areas.

It is easy to obtain reference values for monitoring the amount of ink by measuring ink density in solid ink areas. The values are determined once and for all when the prepress parameters are determined. Obtaining reference values for monitoring the amount of ink in halftone areas is more complicated [19].

All the methods and systems mentioned are dependent on the availability of areas with special properties. This limitation is fatal if such areas can not be found or must not be printed.

Monitoring the amount of ink in solid print areas does not exploit information on how the process influences the dot-size, only the *density*. Monitoring the amount of ink in halftone areas reveals the influence of printing process on both the *density* and halftone dot-size, but requires a way of determining the correct amount of ink for example by using a proof print or from prepress information, such as the print original.

An off-line monitoring system is a helpful tool when controlling the process manually. However a closed loop control system requires an online monitoring system. An advanced online monitoring system should:

1. Be able to determine the amount of ink in any inked area. Such a system is not restricted to specially dedicated areas.
2. Be able to determine the correct or expected amount of ink from prepress information, such as the print original. In this way it is possible to determine the deviation of the amount of ink in respect to the correct amount anywhere in the print.

To be able to make measurements in more than one position in the cross-direction of the press, an online monitoring system have to either offer the possibility to move the sensor to the desired positions, a *traversing* system, or have multiple sensors, one for each desired position. CCD-camera based monitoring systems are especially well suited to be put in an online traversing system, because they offer the possibility to measure multiple printing quality parameters simultaneously, to compensate for misalignments, and to detect defects in the measuring area.

In the following two sections, it is shown how we design a colour CCD-camera based off-line ink monitoring system. First monitoring on *gray-bars* is considered. Then

we extend the approach by making it applicable on an arbitrary inked area of CMY inks.

Chapter 4

Monitoring the amount of ink in well known areas

We make two assumptions in the definition of a well known area. First, we assume the printed halftone area to have the same properties, halftone values, from print to print. Meaning that we have the same target values each time, and that we can obtain reference values in advance by measuring a reference print in the form of a "perfect" or "golden" print. The determined values can be easily interpreted and translated to control actions since the operator can expect the same behavior each time. We also assume the area to contain the same halftone value, one for each separation, all over the area. This reduces the need for being very accurate when positioning the measuring device.

Because the offset process relies on the assumption that each separation is printed with the correct ink film thickness our objective is to determine the deviation of the amount of ink separately for each separation. A colour shift measure, such as ΔE, does not reveal the deviation for each separation. Thus, for the press operator it is difficult to exploit ΔE for adjusting the ink flow.

Monitoring the amount of ink in well known areas consists of two steps. First, we determine the amount of ink for each separation on the print. The determined amount of ink is then compared with the target values in order to assess the deviations.

The aforementioned gray-bars fulfill our assumptions about the suitable measuring area. Of course, before using an area such as a gray-bar to monitor the amount of ink we must assure that the halftone dots on the printing plate have the expected size. We do not measure the mechanical dot-size in the print because the mechanical size of the halftone dot on the print is heavily dependent on the amount of ink fed to the printing plate [18]. An increased amount of ink gives an increased halftone mechanical dot-size on the print and the vice versa.

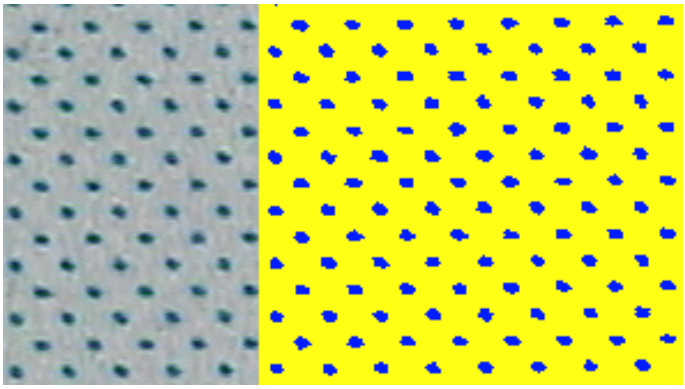

Figure 4.1: Image of halftone dots on a printing plate. On the right, the detected halftone dots are shown.

4.1 Determining the inked area

Check-point II in the graphic process is the printing plates. To fulfill the objective of assessing the quality of the printing process we need to determine the dot-size on the printing plate. By comparing it with the dot-size in the print original we can get the discrepancy value.

To solve the task we have developed a calibration free segmentation technique which can be used to determine the dot-size on any kind of media, regardless of the colour of the halftone dots or the background **Paper A**. Since the method is calibration free, the result obtained does not contain errors due to calibration media or calibration procedure. The segmentation is accomplished in two stages through classification of image pixels. In the first stage, rough image segmentation is performed. The results of the first segmentation stage are then utilized to collect a balanced training data set for learning refined parameters of the decision rules. Figure 4.1 shows the analysis result obtained using the method proposed for an image of a printing plate. From the segmented image, determining the average dot-size is easily done.

The reflected light on printing plates origins from surface reflections only, thus there is no interior light spreading and therefore no optical dot-gain. Consequently, measuring the dot-size with our method on printing plates, the revealed dot-size can be assumed to show the approximate mechanical dot-size of the actual print area. However, measuring dot-size on a substrate such as newsprint, due to the optical effects such as light spreading in the interior of the paper, the measured dot-size values are highly dependent on the method used. Dot-size measured on newsprint with an optical instrument, such as a densitometer, will reveal a significantly larger dot-size value than the mechanical one. Because we determine a local inked area in an image, we

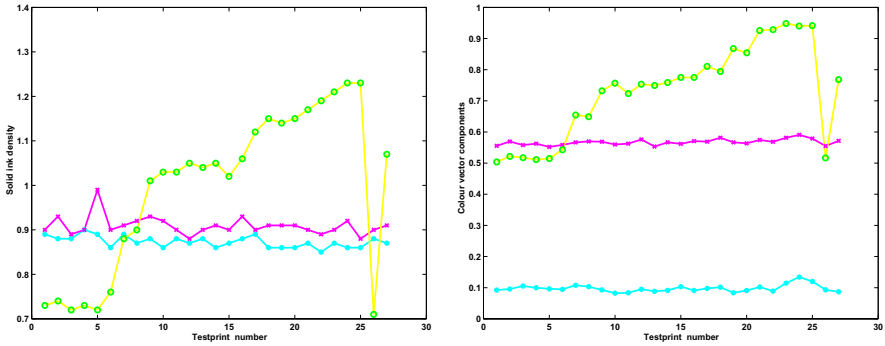

Figure 4.2: Left: Densities measured on a series of sheets printed with constant densities of cyan and magenta and varying density of yellow. Right: The corresponding *colour impression* values obtained in an area of approximately 8% cyan, 50% magenta, and 65% yellow.

can expect being able to measure a dot-size more closely to the true mechanical one, than if it is done with a densitometer. To determine the magnitude of the difference between the dot-sizes obtained with our method and the "true" mechanical dot-sizes, we compared the results obtined from our method and the non-optical instrument *PIXE*, **Paper A** [20]. The results have shown that our method is accurate enough to measure mechanical dot-size on newsprint as well as on printing plates.

4.2 The concept "Colour Impression"

The *colour impression* is a term used to characterize the perceived amount of ink deposited on a substrate, and should not to be confused with the term impression used in the printing industry. Two factors contribute to values of the *colour impression*, the halftone dot-size and ink density, **Paper B**. The values represent the observed average amount of ink in an area separately for cyan, magenta, yellow, and black. *Colour impression* is presented as a CMY or $CMYK$ colour vector, where each component represents the amount of the corresponding ink. The space chosen is easy for the operator to interpret. A value of 0 corresponds to the unprinted paper and a value of 100 to a solid ink area with the correct density.

Values of the colour vector components are obtained by registering the RGB image from the measuring area, transforming the RGB image to its device independent $L^*a^*b^*$ counterpart. Then mapping the set of the average $L^*a^*b^*$ values for the area to the triplet or quadruple of CMY or $CMYK$ values, respectively. Artificial neural networks are used to learn and perform the mapping $Lab \rightarrow CMY(K)$, **Paper B**

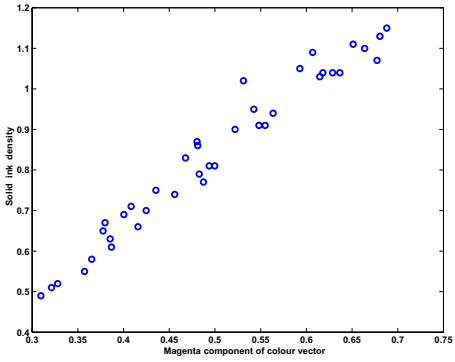

Figure 4.3: Relation between magenta component of the *colour impression* and magenta solid ink density.

[21, 22]. To train the networks a test-print containing areas of varying dot-size is used. The test-print is printed with the correct ink density, and the true amount of ink is therefore assumed to be the one determined on relevant single ink step wedges.

Traditionally, the solid ink density is regarded to be the best print property for determining the amount of ink. Therefore, we investigated the correlation between the *colour impression* and ink density values. Three series of test-prints were produced with constant density for two of the inks and the density for the third ink was varied in controlled steps. We call the series cyan, magenta, and yellow test series. To assess the correlation solid ink densities were measured on colour bars and the *colour impression* in the corresponding column in a halftone area. Figure 4.2 shows the solid ink densities and the corresponding *colour impression* values obtained on the yellow test series. The left hand part of Figure 4.2 shows the solid ink density for the cyan, magenta, and yellow inks. The corresponding *colour impression* values obtained in a halftone area with approximately 8% cyan, 50% magenta, and 65% yellow are shown on the right hand side of Figure 4.2. Observe that the densities do not carry information on the halftone values but the *colour impression* values do. Figure 4.3 illustrates the high correlation between the solid ink density and the *colour impression* values for the magenta test series.

Gray-bars are well suited for measuring the *colour impression* – the amount of ink. The *gray-bars* have approximately the same properties each time they are printed, so it is possible to obtain reference values from a "golden" print. The *gray-bars* are made of halftone areas, and therefore information on how the printing process influences the halftone dots is taken into account when measuring on the *gray-bars*. The bars contain all inks. The black is separated from the chromatic print, which increases the measuring accuracy.

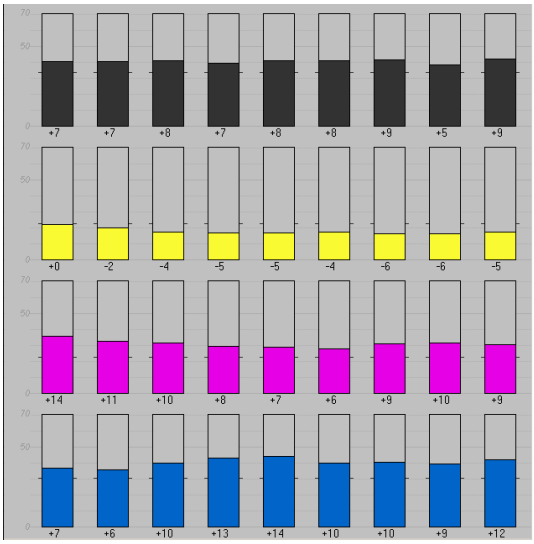

Figure 4.4: Example of *colour impression* values as presented to the operator in the instrument Malcolm. A *colour impression* value of 0 corresponds to the unprinted paper and a value of 100 to a solid ink area with the correct density. The values shown under the bars represent the deviation of the *colour impression* from the desired.

The advantage of adopting the *colour impression* is that we can use a colour CCD-based system to monitoring ink in halftone areas. Provided we have reference values, we can measure the deviation of the amount of ink from the desired in an area containing any combination of the chromatic inks or black ink.

The measuring approach is implemented as a tool in the commercially available instrument Malcolm [23]. Figure 4.4 shows how *colour impression* is presented to the operator in the instrument Malcolm. The bars from the top to the bottom show the *colour impression* values obtained for black, yellow, magenta and cyan inks respectively. From the left to the right the bars represent values for nine different inking zones. The numbers shown below the bars represent the deviation of the *colour impression* values from the desired ones.

4.3 Obtaining target values

When using the well known measuring areas we can easily obtain the target values, or correct amount of ink, by measuring the *colour impression* on a "perfect" or "good enough" print. The target values have to be obtained using the substrate, printing inks

and the press to be used in the printing process. Fortunately the target values do not change over time and can therefore be obtained in advance in a print run when special care can be taken of determining and adjusting the amount of inks.

4.4 Summary

The CCD-camera based calibration free technique allows us to monitor the accuracy of halftone dot-sizes on the printing plates. Therefore we have the possibility to fulfill the desired control task in the check-point II.

The CCD-camera based calibration free technique also allows us to estimate the approximate mechanical halftone dot-size on newsprint.

The *colour impression* values measured using a colour CCD-camera based system provides an estimate of the amount of ink in a newspaper halftone print area for each separation. The *colour impression* values integrate both the halftone dot-size and ink density. Values of the *colour impression* space are easily interpreted by the press operator.

Provided gray bars and target values we can obtain *colour impression* deviation values, which are easily translated into control actions. Thus we can fulfill the control task in the check-point III.

The target values can be obtained from a reference print, or from the prepress information.

Chapter 5

Monitoring the amount of ink in arbitrary inked areas

In this chapter, we extend the methods presented in the previous chapter to be usable for measuring the amount of ink in an arbitrary inked area, **Paper C**. When measuring in arbitrary areas, we can not expect having any "golden" print to obtain target values, as in the former case. Instead, we want to use information from the print original to calculate the target values.

When using *gray-bars* as measuring areas, the problem of positioning the measuring device is a simple task. The *gray-bars* do only contain one single halftone value for each separation and the area is well limited by the edges.

Extending the methods for measuring the amount of ink in an arbitrary inked area, therefore, poses two problems to solve: How to obtain the target values from the print original, and how to assure comparing the same areas in the actual print and the print original.

5.1 Estimation of reference values

In the print original, we have information in the form of halftone dot areas or nominal values for each separation, cyan, magenta and yellow. To obtain the target values these nominal CMY values have to be mapped to the expected amount of ink or *colour impression* values $\widehat{CMY}_t$. To obtain the mapping, we apply the spectral Neugebauer equations based modelling in the way similar to the one suggested in [24]. By using the dot-gain curves obtained from the modelling we get the designated target $\widehat{CMY}_t$ values:

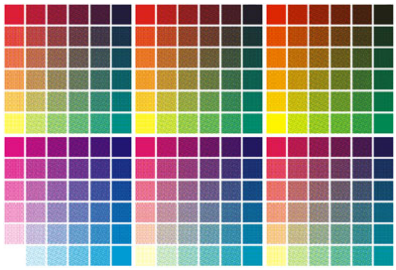

Figure 5.1: Print original CMY image used to evaluate the method described. The image contains 216 test-patches with all combinations of cyan, magenta, and yellow inks in 20% steps.

$$\hat{C}_t = G_c(C) \tag{5.1}$$

where $\hat{C}_t$ is the target value for the cyan ink, C is the halftone cyan dot-size in the print original, and G_c is the dot-gain function for the cyan ink. The target values for the magenta and yellow inks are obtained likewise.

The actual *colour impression* values in the print are obtained in the same way as it was explained in Chapter 4, except that $\widehat{CMY}_t$ values are now used as target values when training the artificial neural network.

5.2 Comparing Printed Result and the Target

To assess the actual *colour impression* on an arbitrary area of a multicolour print, first an RGB image of the area is recorded. The RGB values are transformed to their $L^*a^*b^*$ counterparts and the mapping $L^*a^*b^* \rightarrow CMY$ is then performed by the neural network. To be able to compare the actual *colour impression* values with the targeted ones, we have to match the image pair – the recorded image of the print and the corresponding area of the print original image.

The images can be shifted, rotated, scaled, and skewed in respect to each other. During our test we assumed the skewing is negligible. To perform the matching of the image pair, we exploit properties of the Fourier transform [25] allowing to obtain the shift, rotation, and scaling parameters needed. The method has shown good performance even on images with quite large discrepancy, such as the images used in our tests – the digital target bitmap image the printed result image obtained by scanning the printed result.

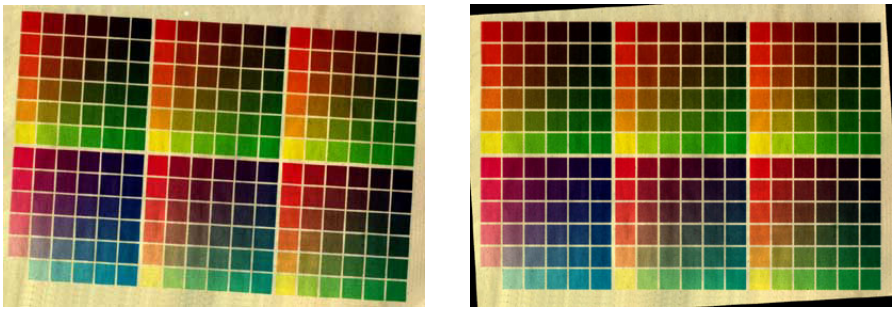

Figure 5.2: Left: Scanner image before matching obtained from the test print containing the 216 test-patches. Right: Scanner image matched to the print original shown in Figure 5.1.

5.3 Results

To evaluate the proposed methods we used a test image containing 216 test-patches with all combinations of cyan, magenta, and yellow inks in 20% steps. The CMY print original image is shown in Figure 5.1. The print was scanned and matched against the print original. Values where then obtained for each test patch from the print original and the scanned image. The scanner image obtained from the test print is shown in Figure 5.2: to the left – before matching with the print original image, and to the right – after the matching.

To evaluate the accuracy of the estimated $\widehat{CMY_p}$ values, the maximum error values were obtained for the yellow ink. The size of the maximum errors or deviation from the estimated target values were $+3.6\%$ and -3.2%. This can be regarded as dot-size errors. For cyan and magenta inks, the maximum errors obtained were approximately 30% lower. Thus, the estimate is accurate enough for the practical use in newspaper printing. Figure 5.3 shows the error map for the yellow ink – the deviation of the actual $\widehat{CMY_p}$ values from the values of the correct amount of ink $\widehat{CMY_t}$, obtained from a test-print with the correct ink density.

5.4 Summary

By employing the methods developed we can estimate the deviation of the amount of ink deposited on a newsprint from the correct amount obtained from the print original in an area containing any combination of the chromatic inks. We can assure comparing the same areas in the actual print and the print original by matching the areas. Thus, we can fulfill the control task up to the check-point III for any combina-

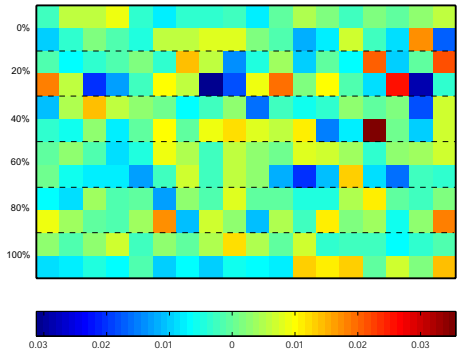

Figure 5.3: Yellow separation error map showing the deviation of the actual $\widehat{CMY}_p$ values from the target $\widehat{CMY}_t$ values. The yellow separation nominal halftone values are shown on the left in percentage.

tion of the three chromatic inks. Since the methods developed utilize a CCD-camera based system, they are well suited for online implementation in a printing press.

Chapter 6

Controlling the amount of ink in the printing press

In this chapter, we discuss how the proposed methods for monitoring the amount of ink can be used for controlling the ink feed.

The basic parameter determining the amount of ink to be fed, the setting of the ink keys and the ink fountain roller speed in an inking zone, is the *ink demand*. In its simplest form *ink demand* can be expressed as the relative printing area on the print original and thus on the printing plate. High *ink demand* means high amount of ink to be fed through the inking system and vice versa. The setting of the ink keys and the ink fountain roller speed are mainly determined by the ink demand in that zone.

The operator controls the amount of ink deposited on the substrate by controlling the flow of ink to the printing plates, the *ink feed*. The ink feed is primarily controlled by adjusting the ink keys but also by adjusting the speed of the ink fountain roller.

Ink does not flow only in the machine direction in the inking unit. Depending on how the ink knives, intermediate rollers, and ink form rollers are organized in the inking unit, some amount of ink will flow in the cross direction as well. The ink will flow towards the ends of the rollers in the inking unit and ink will flow from one ink zone toward the adjacent zone where the ink feed is lower. The operator has to take this issue into consideration when adjusting the ink feed in the press. Adjusting the ink feed in one inking zone will effect the amount of ink deposited on the substrate in the neighboring ink zones as well. Therefore we have to include information from the adjacent ink zones when modelling ink feed in the actual ink zone, aiming to use the developed models for ink feed control.

6.1 Controlling the printing process manually

Visual inspection and manual adjustment of the ink feed is the usual way of controlling the printing process. An operator learns by experience how the printing process behaves and how the printing press reacts upon his adjustments. The operator creates unawarely an intuitive way of modelling the printing process. However the operator will not be consistent in his judgements. His behaviour will be affected by a number of parameters such as lightning conditions, tiredness, eyesight and so on. Different operators make different judgements about the print and different adjustments under the same conditions. It is also known that operator's knowledge of an offset printing press is specific to that particular press.

The primary step to increase the print quality when controlling the printing process manually is to use an instrument for assessing the printed result, independently of the operator.

Malcolm measuring *colour impression* values is an example of such an instrument. The instrument is designed to aid the operator in determining the amount of ink. Provided *gray bars* and reference values the instrument gives the deviation from the desired amount of ink for all four process colours with one single measurement. The instrument is also designed to present the results obtained from several measuring points and for all inks in a clear and easy to grasp way. Figure 4.4 shows an example of how the *colour impression* values for nine ink zones and all inks are presented to the operator in the Malcolm instrument.

Experience has shown that operators are quit cautious in their adjustments of the printing process. Figure 6.1 shows the actions taken by the operator based on *colour impression* values and the result after his adjustments. The deviation from the correct amount of ink is large even after the adjustments.

Off-line measurements are rather time consuming. Bearing in mind the high printing speed and large number of pages in the modern newspaper printing industry, it is easy to realize that off-line monitoring can only be used for tackling huge deviation related problems.

6.2 Controlling the printing process automatically

Three basic choices have to be made when attempting to automate the ink feed control. More specifically, one must choose the way to quantify the degree of deviation of the amount of ink deposited on the substrate from the desired one, the way to model the process, and the way to control the process.

Building models of dynamical systems by observing input and output data (system

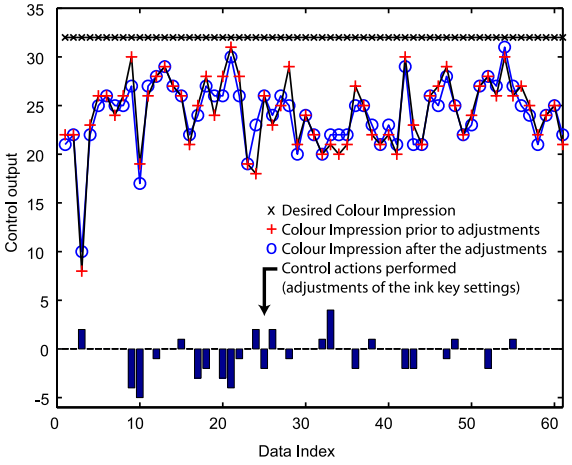

Figure 6.1: The desired *colour impression* values, the *colour impression* values before and after the operator adjustments and the actions taken by the operator based on the *colour impression* values. Note the moderate adjustments and thereby the rather large deviations from the correct *colour impression* values even after the adjustments.

identification), is a necessary step in investigating the system behaviour and constructing system controller, **paper D**.

Numerous parameters need to be measured or estimated for modelling the printing process. We categorize the parameters of the printing process into three groups, namely observable (output), adjustable, and additional parameters. Our observable parameters are the *colour impression* values. The adjustable parameters are the parameters used to control the ink feed. The primary adjustable parameter is the ink key level. Additionally it may be necessary to control other press parameters such as ink fountain roller speed, the feed of dampening solution and others. In addition to the observable and adjustable parameters, we use a set of additional parameters characterizing the printing process. Some of them are determined in advance and do not change during the print run, others constantly change during the print run. The following set of parameters have been used to model the printing process:

1. The copy number, used as a measure of printing time and therefore indirect indicator of the temperature of the printing rollers.

2. The printing speed in copies per hour.

3. The ink fountain roller speed – in most presses the ink fountain roller speed is changed when print speed is changed in order to compensate for the change in ink flow due to the speed change.

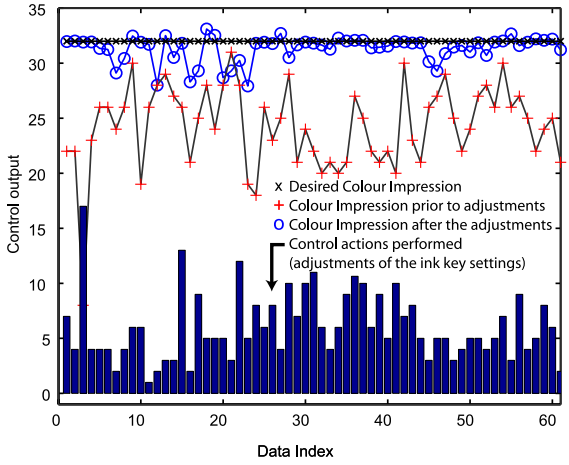

Figure 6.2: The actions of the inverse model based controller and the resulting *colour impression* values after the adjustments, using the same input data as in Fig. 6.1.

4. The level of ink in the ink tray. The level influences the pressure at the outlet and therefore it does influence the ink flow in the press.
5. The estimated ink demand – the desired amount of ink for a given area (ink zone).
6. The estimated ink demand for the left-adjacent ink zone.
7. The estimated ink demand for the right-adjacent ink zone.
8. The ink key setting.
9. The ink key setting for the left-adjacent ink zone.
10. The ink key setting for the right-adjacent ink zone.
11. The *colour impression* – our resulting parameter.

We can use different approaches when building a press controller. The controller can be based on a forward model which predicts the observable parameters, or an inverse model which predicts the adjustable parameters.

We built and tested four different controllers based of the aforementioned set of parameters for modelling the printing process. The controllers investigated were: The model predictive controller – controller which has been recognized as being an effective tool for tackling some of the difficult control problems in industry, the fuzzy logic based controller, the inverse press model based controller, and the press

operator model based controller. The *press operator model based controller* models the behaviour of the press operator, while the inverse model based controller models the inverse dynamics of the press. There is an analogy between these two controllers. An ideal operator would behave as a good *inverse model based controller*.

The inverse printing press model based controller did show the best performance among the controllers tested. The obtained controlling accuracy of the ink flow was higher than that obtained from the experienced printing press operator. Figure 6.2 shows the controller actions and the resulting *colour impression* values after the adjustments, using the same input data as in Fig. 6.1.

6.3 Summary

Colour impression have shown to be well suited for both manual and automatical ink feed control in the printing process. The inverse printing press model based controller has shown the best performance among the controllers tested. The obtained controlling accuracy of the ink feed was higher than that obtained from the experienced press operator.

Chapter 7

Discussion, conclusions and future work

The key issue for good reproduction of images in offset printing is establishing the correct amount of ink deposited on the paper. The reproduction quality of colour images in offset printing is dependent on a number of parameters in a chain of steps. However, we have identified three control points in the offset printing process to assure the correct amount of ink is deposited on the substrate.

An option for colour image segmentation applied for assessing the size of halftone areas of printing plates or paper has been developed. The segmentation is based on classification of image pixels and does not require any calibration and is not sensitive to the colour of the surface. The method developed outperformed the other segmentation techniques tested.

We have also presented approaches to solving the task of establishing the correct amount of ink deposited on the substrate. The introduced *colour impression* concept is well suited for measuring on *gray bars* and provided that a reference print is available the deviation of the amount of ink deposited on the substrate from the correct one can be determined based on the *colour impression* values. From only one measurement a trained neural network is capable of estimating the amount of each cyan, magenta, yellow, and black inks deposited on paper in the measuring area. To our knowledge, this is the only method capable of integrating the ink density and tone values when measuring the amount of ink on multicolour halftone pictures.

The *colour impression* based approach can be used for an area containing any combination of the three chromatic inks. Reference *colour impression* values for an arbitrary area can be obtained from the print original. We have demonstrated that it is possible to match the measuring and the reference areas. Thus, we can determine the deviation of the amount of ink deposited on the substrate from the correct amount

in an arbitrary inked area that does not contain any black ink. This gives the opportunity to measure the amount of ink deposited on the substrate without any specially designated measuring areas.

Because the offset process relies on the assumption that each separation is printed with the correct ink film thickness, it is desired to determine the deviation of the amount of ink separately for each separation. A colour shift measure, such as ΔE, does not reveal the deviation for each separation. By contrast, each component of the proposed *colour impression* vector characterizes one separation and, therefore, is easily interpreted by the operator. To our knowledge, this characteristic property of the method developed is unique if compared to other techniques used in the printing industry. In contrast to ΔE and density, the obtained values are easily translated into ink flow adjustments, performed manually as well as automatically.

We have demonstrated the possibility of monitoring the amount of ink deposited on the paper in the newspaper printing process. It has been shown how the determined amount of inks can be used in combination with other process information to generate control signals for the inking system of the press.

In the offset printing process, where there is a need of measuring multiple properties, an imaging based monitoring system is appreciated, since such a system gives a possibility to implement methods for measuring a variety of properties with only one sensor. Therefore, the methods developed in this work utilize a CCD-camera based imaging system.

Future work

Measurement of all process colours

To be able to measure the amount of cyan, magenta, yellow, and black inks, each in an arbitrary inked area, methods for determining the *colour impression* values must be modified. Preliminary tests with methods used to measure the amount of ink deposited on the substrate in a picture containing all four process colours have shown promising results.

Registration

Another key issue for a correct reproduction of images in offset print is the registration of the separations. We believe that the information from the print original on the image recorded from the printed result can be used for measuring the miss-alignment of the separations.

An online system

One way to reduce the rather large quality variations during the print-run and achieve a high quality print through the whole edition, is to control the amount of ink using an online close loop system. The methods presented are originally developed for off-line use. However, they are well suited for use in an online ink monitoring system as well. Monitoring the observable parameters online is essential for a closed loop ink controlling system.

Colour reproduction

The methods presented in this thesis are not intended to solve any of the problems in the prepress part of the newspaper offset process. Consequently the print original is assumed to be correct. The implementation and use of the methods developed give the possibility of establishing the quality of the printing process along the steps from the print original to the printed result. Information on the deviation of the printed result from the intended one can then be used in the process of adjusting the prepress part of the offset printing process.

Publications

[A] L. Bergman, A. Verikas M. Bacauskiene, Unsupervised Colour Image Segmentation Applied to Printing Quality Assessment, Image and Vision Computing, Vol. 23, 2005 (available at *www.sciencedirect.com*).

[B] A. Verikas, K. Malmqvist, L. Malmqvist, L. Bergman, A New Method for Colour Measurements in Graphic Arts, Color Research and Application, Vol. 24, No. 3 1999, pp. 185-196.

[C] L. Bergman, A. Verikas, Intelligent Monitoring of the Offset Printing Process, Neural Networks and Computational Intelligence proceedings 2004, pp. 173-178.

[D] L. Bergman, A. Verikas, C. Englund, J. Kindberg, J. Olsson, B. Sjögren, Modelling and Control of the Web-Fed Offset Newspaper Printing Press, TAGA Proceedings 2003, pp. 179-194.

References

[1] M. Adams, D. Faux, L. Rieber, Printing Thechnology, 4th edition, Delmar Publishers 1996.

[2] H. Kipphan, Handbook of Print Media, Springer 2001.

[3] H. Kang, Digital Colour Halftoning, SPIE 1999,

[4] S. Gustavsson, Color gamut of halftone reproduction, Journal of Imaging Science and Technology Vol. 41 1997, pp. 283-290.

[5] C. Christoffer, S. Yang, H. Kwok, Efficient gamut clipping for color image processing usin LHS and YIQ, Optical Engeneering, Vol. 42, No. 3 2003, pp. 701-711.

[6] E. Montag, M. G. Fairchild, Psychophysical evaluation of gammut mapping techniques using simple rendered images in artificial gamut boundarys, IEEE Transaktions on Imgae Processing, Vol. 6, No. 7 1997.

[7] G. Wyszecki Color Science, Johan Wiley & Sons 1982.

[8] J.A.C Yule and W.J. Nielsen, The penetration of light into paper and its effect on halftone reproduction, TAGA Procedings 1951, pp. 65-67.

[9] B. Kruse, M. Wedin, A new approach to dot gain modelling, TAGA Procedings 1995, pp. 329-338.

[10] L. Yang, S. Gooran, B. Kruse, Simulation of optical dot gain in multichromatic tone production, Journal of Imaging Science and Technology Vol. 45, No. 2 2001, pp. 198-204.

[11] S. Rydefalk, M. Wedin, Literature review on the colour gamut in the printing process, PFT-repport No. 32 1997.

[12] A. Murray, Monocrome reproduction in photoengraving, J. Franclin Insst 1936, p. 221.

[13] A. Andersson, Gråbalans i praktik och teori, Examensarbete, Linköping University 2003.

[14] H. Kunzli, MiniTarget for quality control in newspaper printing, TAGA Proceedings 2000, pp. 509-520.

[15] E. Marszalec, I. Heikkila, H. Juhola, T. Lehtonen, Online devices and meassuring systems for the automatic control of news paper printing. SPIE, Vol. 3826 1999, pp. 304-312.

[16] S. Almutawa, Y. B. Moon, The development of a connectionist expert system for compensation of color deviation in offset lithographic printing, Artificial Intelligence in Engineering No. 13 1999, pp. 427-434.

[17] B. Pope, J. Sweeney, Performance of an online closed-loop color control system, TAGA Proceedings 2000, pp. 417-431.

[18] C. Södergård, R. Launonen, Inspection of colour printing quality, International Journal of Pattern Recognition and Artificial Intelligence, Vol. 10, No. 2 1996, pp. 115-128.

[19] K. Malmqvist, L. Bergman, H. Busk, L. Malmqvist, Dot size stability in newspaper printing, Procedings of the 8th Scandinavian Conference on Image Analysis 1993, pp. 643-640.

[20] S. A. E. Johansson,J. L. Campbell, PIXE - a novel technique for element analysis, John Wiley & Sons 1988.

[21] S. Haykin, Neural Networks, Prentice-Hall 1999.

[22] C. M. Bishop, Neural networks for pattern recognition, Oxford University Press 1995.

[23] L. Bergman, A. Verikas, K. Malmqvist, MALCOLM - A new partner in printing industry, SSAB Procedings 2002, pp. 147-150.

[24] R. Balasubramanian, Optimization of the spectral Neugebauer model for printer characterization, Journal of Electronic Imaging Vol. 8, No 2 1999, pp. 156-166.

[25] B. S. Reddy, B. N. Chatterji, An FFT-based technique for translation, rotation, and scale-invariant image registration, IEEE Trans Image Processing Vol. 5, No. 8 1996, pp. 1266Ű1271.

Acknowledgements

First, I would like to thank my supervisor Björn Kruse.

Secondly, I would like to thank my assistant supervisor Antanas Verikas for the assistance and patient guidance in the creation of this thesis. $Labai\ ačiū$

Thirdly, I would like to thank Kerstin Malmqvist for the arrangement of the funding of my research education with the T2F research program.

I would also like to thank my coauthors and colleagues of the IS-lab and the PPP-group at Halmstad University.

This work is funded by the Swedish research program T2F. Thank you for giving me this opportunity.

The PPP-group at Halmstad University have for many years cooperated with the four paper mills Stora Enso Publication Paper – Kvarnsveden Mill, Stora Enso Publication Paper – Hylte Mill, Holmen Paper – Hallsta Paper Mill, and Holmen Paper – Braviken Paper Mill. Thank you for your financial support. Thank you all who have participated in our projects during the years of cooperation.

Thank you also the KK-fundation for your financial support.

Our work would not has been the same without the opportunity to test our ideas in a real printing environment. Therefore ,I would like to thank the management of Hallandsposten AB. I would also like to thank the staff of the former printing shop at Hallandsposten, today VTAB, and especially Fredrik Sköld. Thank you for your interest in our work and for your point of views and comments, they have been highly valuable.

Thank you Springer printing shop at Arhrensburg Germany for your points of view in the initial phase of our work, and the aid by supplying us with test prints.

Thank you EAE gmbh at Arhrensburg Germany for giving us the opportunity to use your communication protocol.

Last, but most important, I would like to thank my parents Inger and Ingvar for their

support and the encourage and finally, if possible even more important, I would like to thank my daughter Lina for the support and love without it, this would not have been possible. I love you!

Paper A – Unsupervised Colour Image Segmentation Applied to Printing Quality Assessment

L. Bergman, A. Verikas M. Bacauskiene, Unsupervised Colour Image Segmentation Applied to Printing Quality Assessment, Image and Vision Computing, Vol. 23, 2005 (available at *www.sciencedirect.com*).

Lars Bergman

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Image and Vision Computing xx (2004) 1–9
image AND vision COMPUTING
www.elsevier.com/locate/imavis

Unsupervised colour image segmentation applied to printing quality assessment

L. Bergman[a], A. Verikas[a,b,*], M. Bacauskiene[b]

[a]Intelligent Systems Laboratory, Halmstad University, Box 823, S 301 18 Halmstad, Sweden
[b]Department of Applied Electronics, Kaunas University of Technology, Studentu 50, LT-3031 Kaunas, Lithuania

Received 26 October 2002; received in revised form 13 August 2003; accepted 8 November 2004

Abstract

We present an option for colour image segmentation applied to printing quality assessment in offset lithographic printing by measuring an average ink dot size in halftone pictures. The segmentation is accomplished in two stages through classification of image pixels. In the first stage, rough image segmentation is performed. The results of the first segmentation stage are then utilized to collect a balanced training data set for learning refined parameters of the decision rules. The developed software is successfully used in a printing shop to assess the ink dot size on paper and printing plates.

Keywords: Colour image segmentation; Fuzzy clustering; Quality inspection; Colour printing

1. Introduction

The motivation for this work comes from the printing industry. Offset lithographic printing is the most widely used commercial printing process. It is used to produce high quality pictures in the production of magazines, catalogs, newspapers, etc. The pictures are created by printing cyan (*C*), magenta (*M*), yellow (*Y*), and black (*K*) dots of varying sizes upon each other through screens having different raster angles [1]. Fig. 1 (left) illustrates an example of an enlarged view of a small area of a newspaper picture that contains dots of all the four inks.

Colour observed in a local area of such pictures depends on proportions of the four inks deployed on that local area of paper [2]. To obtain high quality prints a relatively high precision of determining the proportions of inks is required. Therefore, there is a great need to accurately measure the percentage of area covered by inks of different colours. Since four separate printing plates—one for each ink—are used to print such pictures, the expected and the actual percentages of an area covered by the different inks can be measured separately for each ink. Fig. 1 (right) presents an example of a halftone cyan area, on which the percentage of cyan needs to be accurately determined. Thus, to measure the percentage, we need to solve an image segmentation task. Though only a bi-level segmentation is required, colour information needs to be taken into consideration, since yellow ink is used and, moreover, the four inks can be printed on coloured backgrounds.

Various colour image segmentation techniques have been proposed. The most commonly used approaches include: histogram thresholding [3–5], feature/colour space clustering [6,7], edge detection approaches [8,9], neural network based approaches [10–14], region-based approaches [15,16], Markov random field [17] and mixture-of-Gaussians modelling [18], physics-based approaches [19], and combinations of above [20–23]. A recent survey of colour image segmentation methods can be found in [24].

All the existing colour image segmentation approaches are strongly application dependent and suffer from different characteristic drawbacks. For example, histogram thresholding does not consider spatial details and does not work well for images without obvious peaks and valleys. Feature space clustering based methods do not utilize spatial

* Corresponding author. Tel.: +46 35 167 140; fax: +46 35 216 724.
E-mail addresses: lars.bergman@ide.hh.se (L. Bergman), antanas.verikas@ide.hh.se (A. Verikas), marija.bacauskiene@ktu.lt (M. Bacauskiene).

0262-8856/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.imavis.2004.11.003

Fig. 1. Left: An enlarged view of a part of a newspaper picture that was created by printing dots of cyan, magenta, yellow and black inks. Right: An example of an image containing very small areas of one colour and large areas of other colour.

information too. How to select features for obtaining satisfactory segmentation results remains unclear. Region-based approaches are quite expensive in computation time and sensitive to the examination order of regions and pixels. Edge detection approaches are quite sensitive to noise and do not work well for images containing ill-defined edges. Neural network based approaches usually require long training time and initialization may affect the results. Markov random field modelling is quite expensive in computation time.

In this work, our objectives are a relatively high segmentation speed and accuracy and low sensitivity to noise. Observe that the relatively high segmentation accuracy needs to be obtained even for the images containing very small areas of one colour and large areas of other colour. Fig. 1 (*Right*) provides an example of such an image.

The approach we adopted in this work consists of three phases. In the first phase, we cluster a relatively small number of randomly selected pixels represented by a 3-dimensional colour vector. Parameters of the clusters found are then used for rough image segmentation. In the second phase, the segmented image is preprocessed and based on the segmentation results a new set of pixels is selected. An equal number of pixels from the different colour clusters is included in this new data set used for a refined estimation of parameters of the clusters. In addition to the 3-dimensional colour vector, information from a local neighbourhood is also used to represent pixels in this new data set. In the third phase, the `Minimum Fuzzy Cluster Volume` clustering algorithm is applied to the data set. Then, by assigning image pixels to the clusters found the final image segmentation is obtained.

The remainder of the paper is organized as follows. In Section 2 we briefly describe the colour space used. Section 3 presents the approach proposed. Section 4 describes the results of the experimental investigations. Finally, Section 5 presents conclusions of the work.

2. Colour space used

Colour image acquisition equipment such as a *CCD* colour camera obtains the *RGB* values, which can be directly used for representing colours in the *RGB* colour space

$$\{R, G, B\} = \int E(\lambda)O(\lambda)\mathbf{F}(\lambda)\,\mathrm{d}\lambda \tag{1}$$

where $E(\lambda)$ expresses spectral properties of the illumination source, $O(\lambda)$ is spectral reflectance function of an object, and $\mathbf{F}(\lambda)$ stands for three spectral sensitivity functions of the colour camera. However, different acquisition equipment give us different *RGB* values for the same incident light. One more drawback of the *RGB* colour space is that the metrics does not represent colour differences in a uniform scale, making it difficult to evaluate the similarity of two colours from their distance in the space.

To meet the requirement of uniformity of distribution of colours the Commission Internationale de l'Eclairage (*CIE*) has recommended using one of two alternative colour spaces: $L^*u^*v^*$ or $L^*a^*b^*$ colour space [25,26]. It is a common practice to use the $L^*a^*b^*$ colour space for describing absorbing materials such as pigments and dyes [1]. Therefore, we used the $L^*a^*b^*$ colour space in this work.

To map the RGB values into the $L^*a^*b^*$ colour space, the RGB values are first transformed to the *XYZ* tristimulus values as follows:

$$X = a_{11}R + a_{12}G + a_{13}B \tag{2}$$

$$Y = a_{21}R + a_{22}G + a_{23}B \tag{3}$$

$$Z = a_{31}R + a_{32}G + a_{33}B \tag{4}$$

with the coefficients a_{ij} being determined by a colourimetric characterization of the hardware used. *XYZ* tristimulus values can describe any colour. It is often convenient to discuss 'pure' colour in the absence of luminance. For that purpose, the *CIE* defines x and y chromaticity co-ordinates: $x = X/(X+Y+Z)$ and $y = Y/(X+Y+Z)$. A colour plots as a point in an (x,y) *chromaticity diagram*. However, the distribution of colours observed in the chromaticity diagram is also non-uniform. A dominant wavelength correlates very non-uniformly with the perception of hue and excitation purity with the perception of saturation.

Having *XYZ* tristimulus values the $L^*a^*b^*$ colour space is defined as follows [26]:

$$L^* = 116(Y/Y_n)^{1/3} - 16, \quad \text{if } Y/Y_n > 0.008856 \tag{5}$$

$$L^* = 903.3(Y/Y_n), \quad \text{if } Y/Y_n \le 0.008856 \tag{6}$$

$$a^* = 500[(X/X_n)^{1/3} - (Y/Y_n)^{1/3}] \tag{7}$$

$$b^* = 200[(Y/Y_n)^{1/3} - (Z/Z_n)^{1/3}] \tag{8}$$

where X_n, Y_n, Z_n are the tristimulus values of X, Y, and Z for the appropriately chosen reference white. If any of the ratios X/X_n, Y/Y_n, and Z/Z_n is equal to or less than 0.008856, it is

replaced in the above formulae by:

$$7.787f + 16/116 \tag{9}$$

where f is X/X_n, Y/Y_n, or Z/Z_n, as the case may be [26]. New measures are provided in the colour space which correlate with hue and saturation more uniformly. For example, *CIE hue-angle*: $H_{ab} = \arctan(b^*/a^*)$ [26].

The Euclidean distance measure can be used to measure the distance (ΔE) between the two points representing the colours in the colour space:

$$\Delta E = [(\Delta L^*)^2 + (\Delta a^*)^2 + (\Delta b^*)^2]^{1/2} \tag{10}$$

3. The approach

The technique developed consists of three steps. In the first step, rough image segmentation is performed. Then, in the second step, the binary erosion operation is applied to the segmented image. Finally, refined image segmentation based on the data collected from the eroded image is accomplished. We adopted the *Fuzzy Kohonen Clustering* algorithm [27] to learn parameters of clusters utilized for accomplishing the rough image segmentation and the *Minimum Fuzzy Cluster Volume* algorithm [28] to learn parameters of the clusters for the refined image segmentation phase. Next, we describe the proposed image segmentation procedure and the main topics of it.

3.1. Segmentation procedure

The image segmentation procedure is encapsulated in the following eight steps.

(1) Randomly select $N \ll P \times Q$ training pixels, where P and Q are the image sizes in the vertical and horizontal directions, respectively.
(2) Run the `Fuzzy Kohonen` clustering algorithm for the selected pixels using the Euclidian distance measure.
(3) Segment the image by assigning each pixel into one of the clusters found. Use the minimum distance classifier with the Euclidian distance measure for the classification.
(4) Perform the binary erosion operation on all parts of the segmented image.
(5) Using the eroded image, randomly select a new set of N pixels subject to the constraint that each cluster obtained in Step 1 is represented by an equal number N/K ($K=2$ in our case) of pixels. If $N/K > \beta N_{\min}$, where $N_{\min}$ is the number of pixels left in the smallest cluster after the erosion and β is a constant, set the number of pixels selected from each cluster to $\beta N_{\min}$.
(6) For each selected pixel, calculate the local information, as described in Section 3.4, and append this value as an extra component to the colour vector representing the pixel.
(7) Run the `Minimum Fuzzy Cluster Volume` clustering algorithm for the selected pixels.
(8) Segment the image by assigning each pixel into one of the clusters found. Use the minimum distance classifier with the Mahalanobis distance measure [29] for the classification.

The rationale for executing the erosion operation is to prevent the training data collecting process from selecting pixels of `uncertain` colour-pixels located on colour edges.

To speed up the whole image segmentation process, the phase of rough segmentation is based on colour information only. However, to obtain an accurate image segmentation, the local spatial information is also utilized in the final segmentation stage. The rough image segmentation phase is required for collecting a balanced training data set utilized to estimate parameters on which the stage of refined image segmentation is based.

3.2. Fuzzy Kohonen clustering

The clustering algorithm used is summarized in the following four steps.

(1) Initialize the weight vectors $\boldsymbol{w}_i(0)$—centres of the clusters—of all the K (2 in our case) nodes with small random values. Choose the distance $\|\mathbf{x}-\mathbf{w}_i\|_{\mathbf{S}}$ measure between two vectors, where $\mathbf{x}=[L^*,a^*,b^*]^T$, the maximum number of learning iterations $t_{\max}$, a small positive constant $\varepsilon > 0$, and a constant $m > 1$ (the degree of fuzziness).
(2) Compute all the $K \times N$ learning rates $\alpha_{ki,t}$, $k=1,\ldots,K$, $i=1,\ldots,N$, where N is the number of pixels used in the learning process, according to the following formulas:

$$\alpha_{ki,t} = (\mu_{ki,t})^{m_t} \tag{11}$$

$$\mu_{ki,t} = \left(\sum_j \left(\frac{||\mathbf{x}_i - \mathbf{w}_k||_{\mathbf{S}}}{||\mathbf{x}_i - \mathbf{w}_j||_{\mathbf{S}}} \right)^{2/(m_t-1)} \right)^{-1} \tag{12}$$

where

$$||\mathbf{x}_i - \mathbf{w}_k||_{\mathbf{S}} = (\mathbf{x}_i - \mathbf{w}_k)^T \mathbf{S}^{-1} (\mathbf{x}_i - \mathbf{w}_k) \tag{13}$$

where $\mathbf{S}$ is the input data covariance matrix.

$$\mathbf{w}_k = \frac{\sum_i (\mu_{ki})^{m_t} \mathbf{x}_i}{\sum_i (\mu_{ki})^{m_t}}, \quad k = 1, 2, \ldots, K \tag{14}$$

$$m_t = m - t\Delta m \tag{15}$$

$$\Delta m = (m-1)/t_{\max} \tag{16}$$

(3) Update all the K vectors $\{\mathbf{w}_{k,t}\}$ according to the rule:

$$\mathbf{w}_{k,t} = \mathbf{w}_{k,t-1} + \frac{\sum_{i=1}^{N} \alpha_{ki,t}(\mathbf{x}_i - \mathbf{w}_{k,t-1})}{\sum_{i=1}^{N} \alpha_{ki,t}} \tag{17}$$

(4) Compute the difference between the weight vectors in two subsequent cycles:

$$E_t = \|\mathbf{w}_t - \mathbf{w}_{t-1}\|^2 = \sum_{k=1}^{K} \|\mathbf{w}_{k,t} - \mathbf{w}_{k,t-1}\|^2 \tag{18}$$

If $E_t < \varepsilon$ Stop; Else iterate steps 2–4.

3.3. Erosion

For an image I and a structural element W the *erosion* of I by W, denoted $I \ominus W$, is defined as

$$I \ominus W = \{x | (W)_x \subseteq I\} \tag{19}$$

where

$$(W)_x = \{c | c = w + x, \quad \texttt{f} or\ w \in W\} \tag{20}$$

is the *translation* of W by $x=(x_1, x_2)$.

3.4. Local information

The local information $g(x, y)$ at the pixel (x, y) is given by the local average in a small square window w centred at the pixel (x, y):

$$g(x,y) = \frac{1}{(2w+1)^2} \sum_{i=-w}^{w} \sum_{j=-w}^{w} d(x+i, y+j),$$
$$x = w, w+1, \cdots, P-w-1 \tag{21}$$
$$y = w, w+1, \cdots, Q-w-1$$

where $d(x,y)$ is a distance image. The distance image is obtained by calculating the Euclidian distance, given by Eq. (10), between the colour of the pixel being considered and the mean colour of the largest cluster found in the first clustering phase. The mean colour is given by the average L^*, a^*, and b^* values calculated over the pixels assigned to the largest cluster.

3.5. Minimum Fuzzy Cluster Volume Clustering

We initialize the `Minimum Fuzzy Cluster Volume` clustering algorithm with the cluster centres $\mathbf{w}_k$, $k=1,\ldots,K$ and the memberships μ_{kj} obtained from the Fuzzy Kohonen clustering algorithm.

In each iteration of the Fuzzy Volume algorithm, the memberships μ_{ki} and the cluster centres $\mathbf{w}_k$ are updated according to the following rules:

$$\mu_{kj} = \begin{cases} \dfrac{[d_2(\mathbf{x}_j, \mathbf{w}_k)]^{1/(1-m)}}{\sum_{k=1}^{K} [d_2(\mathbf{x}_j, \mathbf{w}_k)]^{1/(1-m)}}, & \text{if } d_M(\mathbf{x}_j, \mathbf{w}_k) - n > 0 \,\forall\, k \\ 0/1, & \text{if } d_M(\mathbf{x}_j, \mathbf{w}_k) - n \le 0 \,\exists\, k \end{cases} \tag{22}$$

where n is the dimensionality of $\mathbf{x}$, m is the degree of fuzziness, $d_M(\mathbf{x}_j, \mathbf{w}_k)$ is the Mahalanobis distance given by

$$d_M(\mathbf{x}_j, \mathbf{w}_k) = (\mathbf{x}_j - \mathbf{w}_k) S_k^{-1} (\mathbf{x}_j - \mathbf{w}_k)^T \tag{23}$$

$$d_2(\mathbf{x}_j, \mathbf{w}_k) = \frac{|S_k|^{1/2} [d_M(\mathbf{x}_j, \mathbf{w}_k) - n]}{\sum_{j=1}^{N} \mu_{kj}^m} \tag{24}$$

where S_k^{-1} is the inverse of the fuzzy covariance matrix S_k

$$S_k = \frac{\sum_{j=1}^{N} \mu_{kj}^m (\mathbf{x}_j - \mathbf{w}_k)(\mathbf{x}_j - \mathbf{w}_k)^T}{\sum_{j=1}^{N} \mu_{kj}^m} \tag{25}$$

and $|S_k|$ is the determinant of the matrix.

In (22), $\mu_{kj}=1$, if

$$k = \arg \min_{i=1,\ldots,K} d_2(\mathbf{x}_j, \mathbf{w}_i) \tag{26}$$

and $\mu_{kj}=0$, otherwise.

$$\mathbf{w}_k = \frac{\sum_i \mu_{ki} \mathbf{x}_i}{\sum_i \mu_{ki}}, \quad k = 1, \ldots, K \tag{27}$$

The same rule (18), as in the Fuzzy Kohonen clustering algorithm, is used to terminate the learning process.

4. Experimental investigations

A one-chip CCD colour camera has been used to capture colour images. The resolution used was such that an image consisting of 768×576 pixels was recorded from an area of approximately 8×6 mm^2. In all the experiments, the data used were normalized to `zero` mean and variance *one*.

To compare segmentation results obtained from the technique developed, we have chosen the Stochastic Expectation Maximization (*SEM*) algorithm [30]. The *SEM* algorithm is a contextual adaptation of the traditional *EM* algorithm. Based on the observation that adjacent pixels are likely to possess the same labels, priors in the *SEM* algorithm are modelled as a second-order *Markov random field*.

As in all image segmentation tasks, there is no objective method to assess segmentation results. An independent evaluation against correct ink coverage values is quite difficult, since these values remain unknown. Therefore, in our tests, we mainly resorted to a careful inspection by a human eye. Additionally, to assess the results obtained from the technique proposed, we employed the *particle induced X-ray emission* (*PIXE*) [31] at a nuclear microprobe

technique [32]. Using the non-optical PIXE based technique a very accurate estimate of the amount of ink pigment on paper can be obtained [32]. However, the PIXE measurements are very expensive and time consuming. Therefore, in [32], only six screen dots have been analysed. To assess our technique, we used the data provided by the authors of [32].

4.1. Parameter settings

There are several parameters to be chosen, namely the degree of fuzziness m, the small constant ε used to stop the learning process, the number of training pixels N, the size of the window w for calculating local information, the size of the structural element W to fulfil the erosion operation, and the parameter β controlling the number of pixels selected for the second learning phase. The values of $m=1.8$, $\varepsilon=0.001$, $N=10{,}000$, $w=1$, $W=3$, and $\beta=0.8$ worked well in all the tests performed. The segmentation results were quite insensitive to the choice of the parameter values, except the size of the local window w. The use of $w>2$ deteriorated the segmentation accuracy of the images containing very small areas of one colour and large areas of other colour. An example of such an image is shown in Fig. 2.

4.2. Segmentation results

Fig. 2 presents a typical example of an image, in which the percentage of the area covered by a printing ink needs to be determined. Figs. 3–5 display intensity histograms of the *Red*, *Green*, and *Blue* bands of the image, respectively. As can be seen from the figures, the histograms contain only one peak. Thus we cannot expect obtaining an accurate estimate of the area by simple histogram thresholding.

Fig. 6 exemplifies an image of an enlarged view of a small printing plate area made aiming to get a halftone with 5% ink coverage—a nominal ink dot size of 5%. However,

Fig. 2. An example of an image containing very small areas of one colour and large areas of other colour.

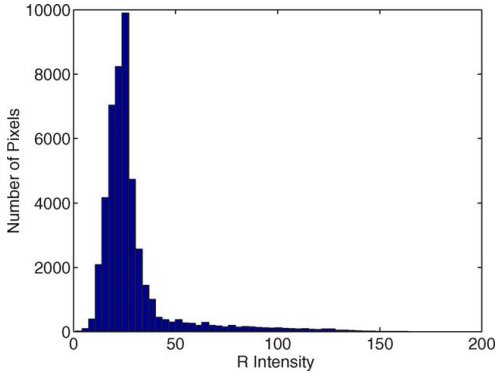

Fig. 3. Intensity histogram of the *Red* band.

due to inaccuracies and some unavoidable factors occurring in the printing process, the actual ink dot size may significantly differ from the nominal one. The actual ink dot size in the halftone area shown in Fig. 6 is much larger than 5%. To estimate the actual ink dot size the image was segmented into '*dots*' and '*background*' parts. In Fig. 6, the segmentation result, obtained employing only the first phase of the segmentation procedure, is provided for the left part of the image. By contrasting the left—segmented and the right—unsegmented parts of the image we can easily notice that the estimated ink dot size is too high. This estimation error occurs due to the fact that the clustering algorithm favours equally populated clusters, while the background area happens to be much more densely represented in the randomly selected training set.

Fig. 7 illustrates the segmentation result of the image shown in Fig. 6 when both phases of the segmentation procedure developed are utilized. One can easily notice that this segmentation result is much closer to the '*ground truth*' than the one shown in Fig. 6.

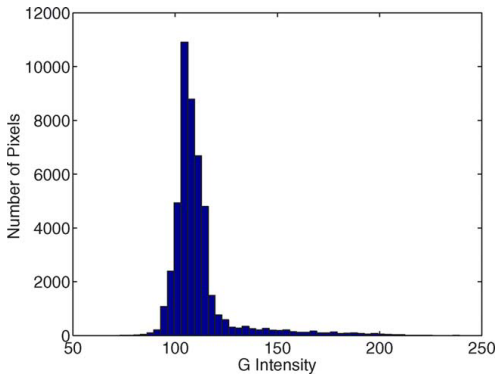

Fig. 4. Intensity histogram of the *Green* band.

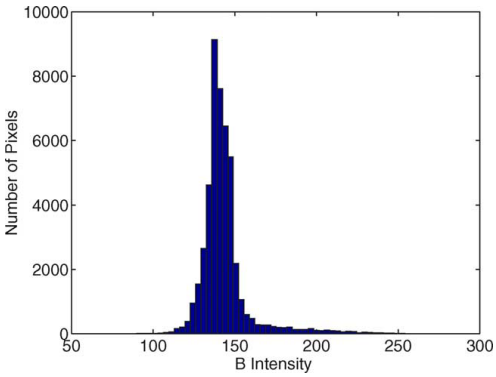

Fig. 5. Intensity histogram of the *Blue* band.

Three factors—equalization of the number of pixels representing the '*dots*' and '*background*' parts of the image in the training set, the use of local information, and elimination of pixels of '*uncertain colour*' from the training set—contribute to improving the accuracy of the estimate of the ink dot size. Amongst them, the first factor is the most significant one.

We experimented with images taken from halftone areas of different nominal ink dot sizes. Images of different colours, including yellow, taken from both printing plates and paper have been used in the experiments.

Table 1 summarizes the results of the experiments when images taken from a printing plate were utilized. The halftone areas used were manufactured in cyan on grey background. For each nominal ink area—column NA % in Table 1—10 images from different physical areas were recorded. In the table, the estimated mean ink dot size—the average percentage of the area covered by ink—and the standard deviation, calculated from these 10 trials are provided. The standard deviations are shown in parentheses.

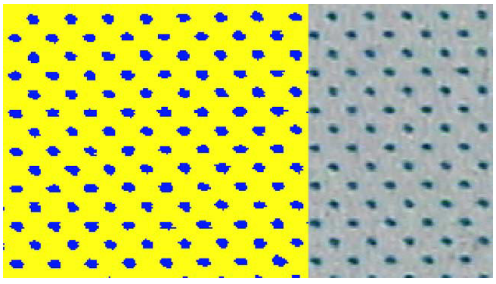

Fig. 6. An image of an enlarged view of a small printing plate area made aiming to get a halftone with 5% ink coverage. The left part of the image is segmented using the 1st phase of the segmentation procedure proposed.

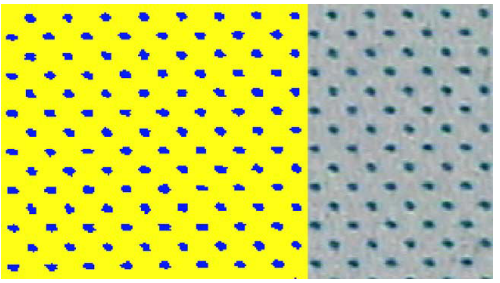

Fig. 7. An image of an enlarged view of a small printing plate area made aiming to get a halftone with 5% ink coverage. The left part of the image is segmented using *both* phases of the sementation procedure proposed.

Four segmentation alternatives have been examined:

(1) segmentation by the SEM algorithm;
(2) segmentation using only the first phase of the procedure proposed—column '`Phase 1`' in Table 1;
(3) segmentation using the first phase of the segmentation procedure when the colour vector representing a pixel is augmented with the local information term—column '`Phase 1+ LI`'
(4) Segmentation using both phases of the procedure and the local information term—column '`Phases 1&2+ LI`'.

As can be seen from the table, the four alternatives give comparable results when segmenting images with approximately equal number of pixels in the '*dots*' and

Table 1
The nominal and estimated average ink dot size for the different image segmentation alternatives

NA (%)	SEM	Phase 1	Phase 1+LI	Phases 1&2+LI
5	14.34 (0.31)	12.94 (0.46)	11.76 (0.43)	09.38 (0.32)
20	25.15 (1.01)	24.63 (0.98)	23.56 (0.92)	22.64 (0.75)
45	45.17 (0.93)	45.27 (1.23)	45.31 (1.13)	44.92 (1.02)
80	78.14 (0.64)	78.84 (0.75)	79.65 (0.67)	81.92 (0.59)
95	91.12 (0.29)	92.07 (0.35)	93.11 (0.31)	95.21 (0.18)

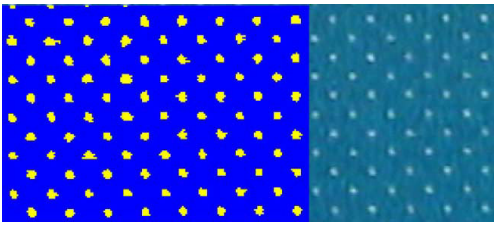

Fig. 8. An image of an enlarged view of a small printing plate area made aiming to get a halftone with 95% ink coverage. The left part of the image is segmented by the SEM algoritim.

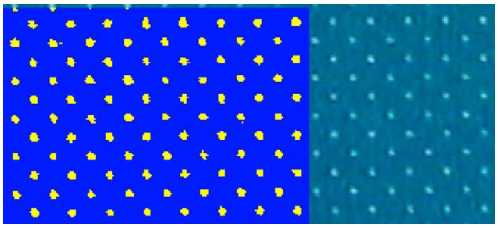

Fig. 9. An image of an enlarged view of a small printing plate area made aiming to get a halftone with 95% ink coverage. The left part of the image is segmented according to the second segmentation alternative.

'*background*' areas of the images. However, when the number of pixels in these areas significantly differ, the procedure proposed is much more accurate than the other segmentation alternatives tested. The equalization of the number of pixels representing the '*dots*' and '*background*' parts of the image in the training set plays the major role in gaining the segmentation accuracy. We recall that the *Nominal Area* (NA) provided is not the '*ground truth*', on the contrary, we attempt to determine the '*ground truth*' through image segmentation.

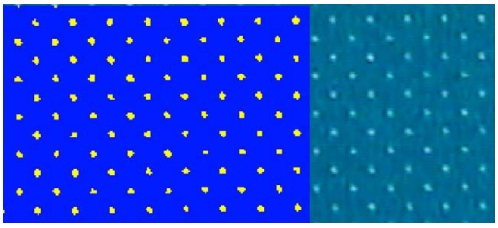

Fig. 10. An image of an enlarged view of a small printing plate area made aiming to get a halftone with 95% ink coverage. The left part of the image is segmented according to the third segmentation alternative.

Figs. 8–10 present one more image segmentation example. The left part of the image shown in Fig. 8 is segmented by the SEM algorithm. The result presented in Fig. 9 is obtained from the third image segmentation alternative. The result obtained using the fourth image segmentation alternative is presented in Fig. 10. Again, by contrasting the left and the right parts of the images, one can clearly see the superiority of the fourth image segmentation alternative.

The developed software was installed in a newspaper printing shop in Sweden and is successfully used on production line to assess the ink dot size on both paper and printing plates.

4.3. Tests using the 'PIXE data'

Six screen dots printed in cyan ink on ordinary newsprint have been used in this experiment. The dots were scanned using both a CCD colour camera and the non-optical PIXE-based method. The same resolution of 2 μm/pixel has been applied in both techniques. A 3-CCD colour camera was used in this experiment [32]. While the CCD colour camera records a 3-band intensity image, the use on the PIXE-based technique results in an image of the amount of ink pigment. The data obtained from the two techniques were preprocessed to bring them into the same coordinate system [32]. The pigment images were thresholded based on knowledge on the composition of the ink and paper used [32]. The thresholding technique is known to be very accurate and robust.

Fig. 11 illustrates a pair of binary images of the same screen dot. The left-hand side image is the thresholded pigment image obtained from the PIXE-based technique. The right-hand side image is the '*optical*' image obtained from the technique proposed. One can easily notice that the optical image does not contain small holes and small isolated '*islands*'. The holes and islands disappear due to light scattering in paper. The dot detected by the image

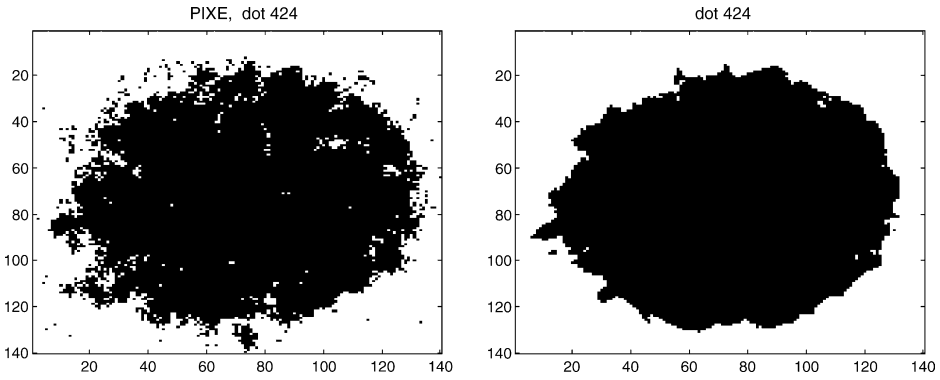

Fig. 11. Left: A screen dot detected by the *PIXE* technique. Right: The same dot detected by the technique proposed.

analysis based approach is 2.28% larger than the one identified by the *PIXE*-based technique. The same range of discrepancy between the optical and non-optical measurements was observed for the other five screen dots. Bearing in mind the light scattering phenomenon in paper, the obtained estimate of the dot size can be considered as very accurate.

5. Conclusions

We presented an option for colour image segmentation applied to printing quality assessment. The segmentation is based on classification of image pixels. Parameters of the decision classes are obtained by clustering a small number of randomly selected pixels—a training data set—represented by a colour vector augmented with one additional component providing information from a local neighbourhood of a pixel being considered. The segmentation is accomplished in two stages. In the first stage, rough image segmentation is performed. The results of the first segmentation stage are then utilized to collect a balanced training data set for learning refined parameters of the decision classes. The final segmentation is obtained by assigning image pixels to the decision classes represented by a set of the refined parameters.

The use of balanced training data sets and the local information significantly improved the image segmentation accuracy if compared to the results obtained when using 'pure' randomly selected training data sets. Small training data sets utilized allowed to significantly seed up the whole image segmentation process. The developed software is successfully used in a printing shop to assess the ink dot size on paper and printing plates.

Acknowledgements

We gratefully acknowledge the support we have received from the Foundation for Knowledge and Competence Development and the TTF research program.

References

[1] A. Verikas, K. Malmqvist, L. Bergman, Neural networks based colour measuring for process monitoring and control in multicoloured newspaper printing, Neural Computing and Applications 9 (3) (2000) 227–242.

[2] A. Verikas, K. Malmqvist, L. Bergman, A new method for colour measurements in multicoloured newspaper printing, in: Proceedings of the Fourth International Conference on Engineering Applications of Neural Networks, Gibraltar, 1998, pp. 189–196.

[3] F. Kurugollu, B. Sankur, A. Harmanci, Color image segmentation using histogram multithresholding and fusion, Image and Vision Computing 19 (2001) 915–928.

[4] R. Ohlander, K. Price, D.R. Reddy, Picture segmentation using a recursive splitting method, Computer Graphics and Image Processing 8 (1978) 313–333.

[5] N. Otsu, A threshold selection method from gray-level histograms, IEEE Transactions on Systems, Man and Cybernetics 9 (1) (1979) 62–66.

[6] C.H. Li, P.C. Yuen, Regularized color clustering in medical image database, IEEE Transactions on Medical Imaging 19 (11) (2000) 1150–1155.

[7] S. Tominaga, Color classification of natural color images, Color Research and Application 17 (4) (1992) 230–239.

[8] L. Xu, M. Jackowski, A. Goshtasby, D. Roseman, S. Bines, C. Yu, A. Dhawan, A. Huntley, Segmentation of skin cancer images, Image and Vision Computing 17 (1999) 65–74.

[9] P.E. Trahanias, A.N. Venetsanopoulos, Vector order statistics operators as color edge detectors, IEEE Transactions on Systems, Man, and Cybernetics—Part B: Cybernetics 26 (1) (1996) 135–143.

[10] H. Huang, Y. Chen, W. Hsu, Color image segmentation using a self-organizing map algorithm, Journal of Electronic Imaging 11 (2) (2002) 136–148.

[11] E. Littmann, H. Ritter, Adaptive color segmentation—a comparison on neural and statistical methods, IEEE Transactions on Neural Networks 8 (1) (1997) 175–185.

[12] N. Papamarkos, C. Strouthopoulos, I. Andreadis, Multithresholding of colour and gray-level images through a neural network technique, Image and Vision Computing 18 (2000) 213–222.

[13] T. Uchiyama, M.A. Arbib, Color image segmentation using competitive learning, IEEE Transactions on Pattern Analysis Machine Intelligence 16 (12) (1994) 1197–1206.

[14] A. Verikas, K. Malmqvist, M. Bachauskene, L. Bergman, K. Nilsson, Hierarchical neural network for color classification, in: Proceedings of the 4th European Conference on Artificial Neural Networks, ICANN-94, vol. 2, Sorrento, Italy, 1994, pp. 847–850.

[15] H.D. Cheng, X.H. Jiang, J. Wang, Color image segmentation based on hamogram thresholding and region merging, Pattern Recognition 35 (2002) 373–393.

[16] A. Tremeau, N. Borel, A region growing and merging algorithm to color segmentation, Pattern Recognition 30 (7) (1997) 1191–1203.

[17] J. Gao, J. Zhang, M. Fleming, Novel technique for multiresolution color image segmentation, Optical Engineering 41 (3) (2002) 608–614.

[18] H. Greenspan, J. Goldberger, I. Eshet, Mixture model for face-color modeling and segmentation, Pattern Recognition Letters 22 (2001) 1525–1536.

[19] C.M. Onyango, J.A. Marchant, Physics-based colour image segmentation for scenes containing vegetation and soil, Image and Vision Computing 19 (2001) 523–538.

[20] T.Q. Chen, Y. Lu, Color image segmentation—an innovative approach, Pattern Recognition 35 (2002) 395–405.

[21] M. Mirmehdi, M. Petrou, Segmentation of color textures, IEEE Transactions on Pattern Analysis Machine Intelligence 22 (2) (2000) 142–159.

[22] D.K. Panjwani, G. Healy, Markov random field models for unsupervised segmentation of textured color images, IEEE Transactions on Pattern Analysis Machine Intelligence 17 (10) (1995) 939–954.

[23] E. Saber, A.M. Tekalp, G. Bozdagi, Fusion of color and edge information for improved segmentation and edge linking, Image and Vision Computing 15 (1997) 769–780.

[24] H.D. Cheng, X.H. Jiang, Y. Sun, J. Wang, Color image segmentation: advances and prospects, Pattern Recognition 34 (2001) 2259–2281.

[25] R.W.G. Hunt, Measuring Colour, Ellis Horwood, 1998.

[26] G. Wyszecki, W.S. Stiles, Color Science. Concepts and Methods, Quantitative Data and Formulae, Second ed, Wiley, New York, 1982.

[27] C.E. Tsao, J.C. Bezdek, N.R. Pal, Fuzzy Kohonen clustering networks, Pattern Recognition 29 (1996) 757–764.

[28] R. Krishnapuram, J. Kim, Clustering algorithms based on volume criteria, IEEE Transactions on Fuzzy Systems 8 (2) (2000) 228–236.
[29] R.O. Duda, P.E. Hart, D.G. Stork, Pattern Classification, Second ed, Wiley, New York, 2001.
[30] A. Baraldi, P. Blonda, F. Parmiggiani, G. Satalino, Contextual clustering for image segmentation, Optical Engineering 39 (4) (2000) 907–923.
[31] S.A.E. Johansson, J.L. Campbell, PIXE—a novel technique for element analysis, Wiley, New York, 1988.
[32] P. Kristiansson, C.M. Nilsoon, H. Busk, L. Malmqvist, M. Elfman, K.G. Malmqvist, R.J.U.J. Pallon, K.A. Sjoland, C. Yang, Optical dot gain on newsprint determined with the lund nuclear microprobe, Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms 130 (1–4) (1997) 303–307.

Paper B – A New Method for Colour Measurements in Graphic Arts

A. Verikas, K. Malmqvist, L. Malmqvist, L. Bergman, A New Method for Colour Measurements in Graphic Arts, Color Research and Application, Vol. 24, No. 3 1999, pp. 185-196.

Lars Bergman

A New Method for Colour Measurements in Graphic Arts

A. Verikas[1,2], **K. Malmqvist**[1], **L. Malmqvist**[3], **L Bergman**[1]

[1] *Centre for Imaging Sciences and Technologies,*
Halmstad University, S 301 18 Halmstad,
Sweden
E-mail:antanas.verikas@cbd.hh.se

[2] *Department of Applied Electronics,*
Kaunas University of Technology, Studentu 50, 3031, Kaunas,
Lithuania

[3] *Department of Atomic Physics,*
University of Lund, Sölvegatan 14, S-22362, Lund,
Sweden

Abstract: This paper presents a method for colour measurements directly on printed half-tone multicoloured pictures. The paper introduces the concept of colour impression. By this concept we mean the CMY or CMYK vector (colour vector), which lives in the three- or four-dimensional space of printing inks. Two factors contribute to values of the vector components, namely, the percentage of the area covered by cyan, magenta, yellow and black inks (tonal values) and ink densities. The colour vector expresses integrated information about the tonal values and ink densities. Values of the colour vector components increase if tonal values or ink densities rise and vice versa. If, for some primary colour, the ink density and tonal value do not change, the corresponding component of the colour vector remains constant. If some reference values of the colour vector components are set from a preprint, then, after an appropriate calibration, the colour vector directly shows how much the operator needs to raise or lower the cyan, magenta, yellow and black ink densities in order to correct colours 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 CMY or CMYK values respectively. Algorithms based on artificial neural networks are used for performing the transformation. During the experimental investigations we have found a good correlation between components of the colour vector and ink densities.

Key words: colour printing; colour classification; neural networks; graphic arts

1 Introduction

Multicoloured pictures in newspapers are most often created by printing dots of cyan (**c**), magenta (**m**), yellow (**y**) and black (**k**) inks upon each other through screens having different raster angles. Fig. 1 illustrates an example of an enlarged view of a small area of a newspaper picture, that contains dots of all of the four inks. A key factor in high quality multicolour printing of pictures is to *measure* and *control* the amount of the four inks transferred to the paper. This paper deals with the measurement of the amount of inks.

Nowadays, the usual way to obtain the quantitative values for controlling the amount of inks in newspaper printing is by measuring the ink densities in specific printed test areas in the newspaper, for example in a colour bar. Such test areas are seldom accepted by the publisher. Besides, such measurements are indirect and hard to handle for the operator controlling the printing process. The operator is faced with a hard task of deciding how to change densities of the four inks in order to improve the quality of pictures. This often leads to over-inked printing. The over-inked printing results in higher costs for inks as well as serious problems of smearing and printing through. Findings of one of the authors show that an operator controlling the printing process can easily come to the zone of over-inked printing [1]. Fig. 2 illustrates the relation between the amount of cyan pigment and the saturation signal measured on the cyan half-tone printed dot [1]. A very accurate method for determining the actual position and amount of ink pigment on paper has recently been established [2]. The measurement technique used in this method is *Particle Induced X-ray Emission (PIXE)* at a nuclear microprobe [3]. The relation shown in Fig. 2 has been obtained by using the *PIXE* technique. As can be seen from the Figure, for small pigment values an almost linear relation holds between the amount of pigment and the saturation signal. However, “when the pigment value further increases a

maximum saturation value seems to be attained" [1]. Printing with an amount of ink larger than that yielding the maximum saturation value we call over-inked printing.

By measuring solid print ink densities an operator can only obtain very fuzzy information about how to control the printing process. This can easily result in attempts to correct the printing process by increasing ink densities in spite of the fact that the process is already running in the over-inked phase. The operator would feel much more secure having an instrument, instead of a densitometer, which, when placed on a multicoloured picture, would show how much he or she needs to increase/decrease the amount of cyan, magenta, yellow and black inks, if compared with some reference preprint. Another important tool printing shops lack today is an instrument providing a possibility to examine the obtained picture at a microscopic level and compare the picture with the desired result. By examination at a microscopic level we mean an inverse colour separation of an image taken from the printed picture and an examination of the obtained printed result separately for each ink used. This would give a better understanding of the interaction of different types of paper, ink and printing devices.

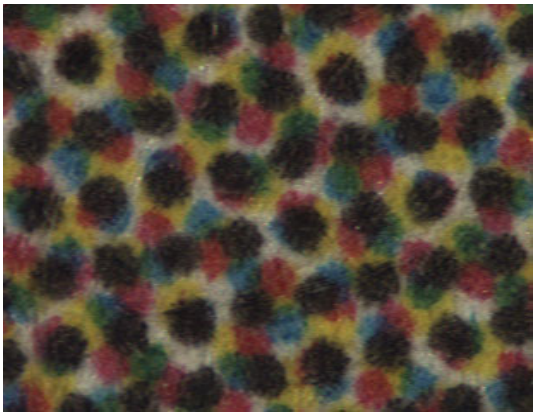

Fig. 1. An enlarged view of a part of a newspaper picture that was created by printing dots of cyan, magenta, yellow and black inks.

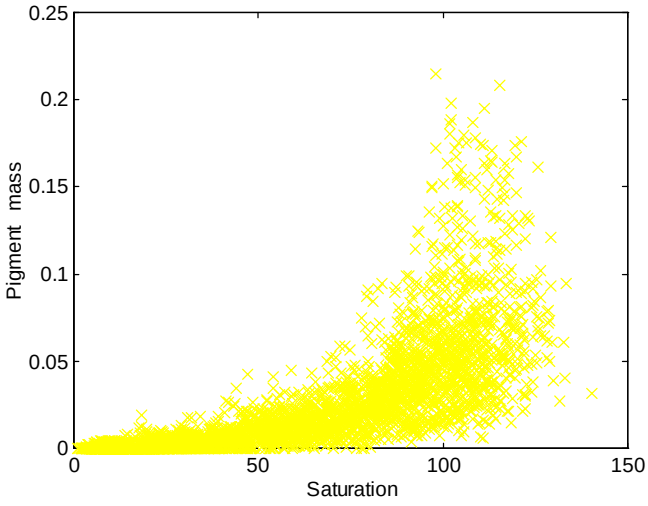

Fig. 2. A correlation plot between the saturation signal and the cyan pigment mass.

In this paper, we present a method and a tool for colour measurements directly on printed half-tone multicoloured pictures as well as for inverse colour separation. We introduce the concept of colour impression [4]. By this we mean the *CMY* or *CMYK* vector (colour vector), which lives in the three- or four-dimensional space of printing inks. Two factors contribute to values of the vector components, namely, the percentage of the area covered by cyan, magenta, yellow and black inks (tonal values) and ink densities. 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 *CMY* or *CMYK* respectively. An intermediate step of the inverse colour separation can be involved when performing the transformation of the set of *RGB* values to the quadruple of *CMYK*. By inverse colour separation we mean the decomposition of the *RGB* image into nine binary images, namely, "white image", "cyan image",…, "green (cy) image" ,…, "cmy image" and "black (k) image". Values of all pixels of the white image are equal to zero, except those corresponding to the areas of the picture that have not got any ink during the printing process. Values of such pixels are equal to one. Accordingly, values of pixels collected into the green (cy) image are set to one if they correspond to the areas of the picture printed with both cyan and yellow inks. The meaning of the other types of binary images is straightforward. We note that such an inverse colour separation is valuable not only as an auxiliary tool for performing the transformation from *RGB* to *CMY*, but also on its own, since it provides a possibility to inspect the printing result. We use algorithms based on artificial neural networks for performing the inverse colour separation and the transformation as well. The algorithms will be described later in more detail.

The colour vector expresses integrated information about the tonal values and ink densities. This is a valuable property, since we never know exactly what actual tonal values we will obtain. Values of the colour vector components increase if tonal values or ink densities rise and vice versa. If, for some primary colour, ink density and tonal value do not change, the corresponding component of the colour vector remains constant. Such a behaviour of the colour vector makes it much easier for the operator to adjust the printing process. If some reference values of the colour vector components are set from a preprint, then the colour vector directly shows how much the operator needs to raise or lower the cyan, magenta, yellow and black ink densities in order to correct colours of the picture being measured. During the experimental investigations we have found a good correlation between components of the colour vector and ink densities.

The rest of the paper is organised as follows. In the next two sections we briefly describe the concept of colour impression and the tools used. The fourth section presents the colour space used. The approach to data normalisation is given in the fifth section. The method for inverse colour separation is briefly described in section six. The seventh section presents a method for accomplishing the transformation from the set of *RGB* values to the colour vector. The eighth section summarises the results of experimental investigations. And finally, section nine presents conclusions of the work.

2 The concept of colour impression

Colour impression of a printed picture depends both on actual tonal values and ink densities. By actual tonal values we mean the nominal tonal values + the mechanical dot gain + the optical dot gain. An operator controlling the printing process never knows the actual tonal values and controls the process by adjusting densities of inks. This is not an easy task for the operator, since it is not clear how big adjustments should be made for cyan, magenta, yellow and black respectively. The degree of adjustments needed can depend on the type of paper and inks, humidity, the duration of the printing process and other factors.

The reason for introducing the concept of the colour vector of colour impression is to integrate information from both actual tonal values and ink densities, and to show how much an operator controlling the printing process needs to raise or lower the cyan, magenta, yellow and black ink densities in order to correct the colours of a picture being measured in reference to some preprint.

Fig. 3 shows the main window of the colour impression software. The components of the colour vector are shown on the right-hand side of the window as colour bars as well as numbers below the bars. Values of the components range between 0 and 1000. The white strips on or above the colour bars indicate values of the colour vector components measured on some reference picture. Having this information the operator can see exactly how the feeds of inks should be changed. For example the figure shows that cyan should be increased by 41 units, magenta increased by 1 unit and yellow decreased by 6 units (see the bottom right corner of the window). The left-hand part of the window displays the image used to calculate the colour vector components.

3 Tools

The equipment we use consists of a one-chip Hamamatsu CCD colour camera C4200, a frame-grabber, a PC and software. The resolution used was such that an image consisting of 512x512 pixels was recorded

from an area of 2.8x4.0 mm^2. All pictures we used were printed on an ordinary newsprint paper of approximately 45 g/m^2 weight.

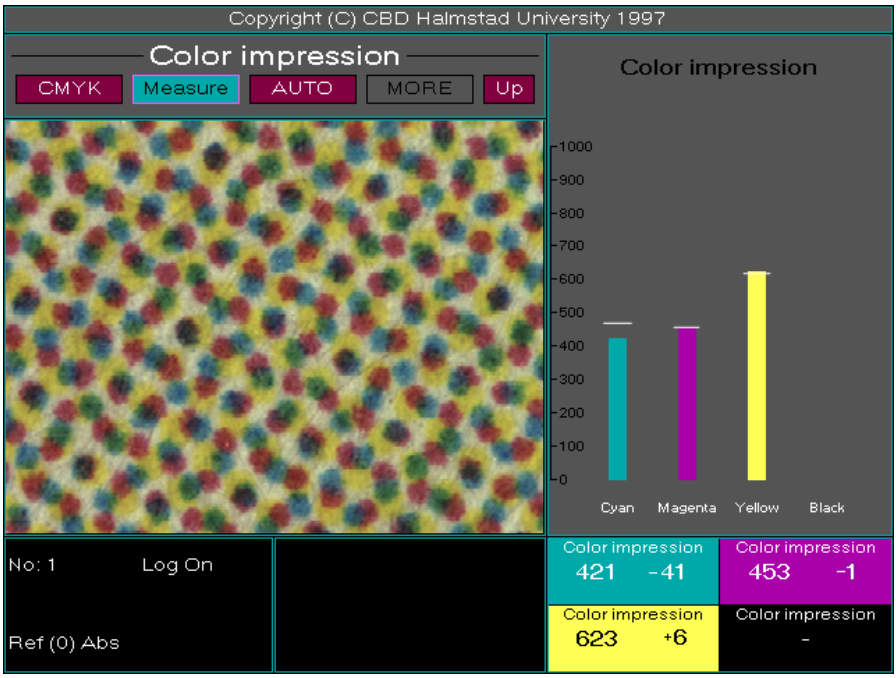

Fig. 3. The main window of the colour impression software.

4 Colour space used

The colour space we adopted uses colour difference signals. Let $\mathbf{r}$ denote the *RGB* triplet given by

$$\mathbf{r} = \int E(\lambda)R(\lambda)\mathbf{F}(\lambda)d\lambda \tag{1}$$

where $E(\lambda)$ expresses spectral properties of the illumination source, $R(\lambda)$ is spectral reflectance function of an object, and $\mathbf{F}(\lambda)$ stands for three spectral sensitivity functions of the colour camera.

If we assume the random variables R, G, and B to be of equal variances (σ^2) and covariances (only variances of the variables have been normalised to be equal to one in our experiments), the covariance matrix of these variables can be written as [6]:

$$\sum = \sigma^2 \begin{bmatrix} 1 & r & r \\ r & 1 & r \\ r & r & 1 \end{bmatrix} \tag{2}$$

where r is the correlation coefficient. The eigensolution of the covariance matrix gives the following eigenvectors ($\mathbf{e}_i$) and the corresponding eigenvalues (λ_i):

$$\mathbf{e}_1 = \{1,1,1\}^t \tag{3}$$

$$\mathbf{e}_2 = \{1,0,-1\}^t \tag{4}$$

$$\mathbf{e}_3 = \{1,-2,1\}^t \tag{5}$$

$$\lambda_1 = \sigma^2(1+2r) \tag{6}$$

$$\lambda_2 = \lambda_3 = \sigma^2(1-r), \tag{7}$$

where *t* means a transpose. The linear transform of the $\{R,G,B\}$ vector by the eigenvectors produces other random variables f_1, f_2, and f_3:

$$f_1 = R+G+B \tag{8}$$

$$f_2 = R-B \tag{9}$$

$$f_3 = R-2G+B \tag{10}$$

which are almost uncorrelated and have zero covariances.

f_1, f_2, and f_3 are the variables of the colour space we use. The choice of a non-uniform device-dependent colour space as well as the use of a *CCD* colour camera for measuring colour prints can be a point of contention. The choice of colour space was based on experimental testing. Five colour spaces, namely, *RGB, HSI, CIELuv*, *CIELab* and $f_1\ f_2\ f_3$ have been tested in the inverse colour separation (colour classification) and compared experimentally. The $f_1\ f_2\ f_3$ colour space provided the highest classification accuracy [5]. Note that we make only relative comparisons when measuring on colour prints. Moreover, the device we developed is tuned to a customer's data set during the training process as it is explained in the following sections. Undoubtedly the *CIELuv* or *CIELab* colour spaces could be used.

5 Data normalisation

We use artificial neural networks to solve the task. In order to accelerate training of the networks, each input variable should be pre-processed so that its mean value, averaged over the entire training set, is close to zero, or else it is small compared to its standard deviation. Therefore, the colour space variables f_i are normalised in the following way.

Let $\bar{f}$ and s^2 be the average and the variance of the colour space variable f. The normalised variable f_n is then given by

$$f_n = \frac{f-\bar{f}}{s}. \tag{11}$$

As it is explained in Section 8, data for training the networks are recorded from a set of colour patches. Instead of averaging over the entire training set when calculating $\bar{f}$ and s^2, we average only over the data recorded from white paper and eight solid print colour patches, namely, cyan, magenta, yellow, cyan+yellow, cyan+magenta, magenta+yellow, cyan+magenta+yellow, and black (k).

Assume that the networks were trained on some set of colour patches. Let's say that the patches represent a reference colour print and we wish to use the networks, without retraining them, for processing some test colour print that possesses slightly different spectral properties compared to the reference print. We use a least-squares regression technique to make a linear mapping between the colour variables of the reference and the test colour prints. Assume for a while that $\mathbf{f}$ expresses the mean vector of the f_i triplets averaged over some colour patch and that we registered the $\mathbf{f}$ s from the eight solid print colour patches and paper for the reference print as well as for the test print. We then collect the measurements $\mathbf{f}_i$ $(i=1,\ldots,9)$ from the reference print into a 9x3 matrix $\mathbf{F}$. Similarly, we collect the corresponding $\mathbf{f}_i$ from the test print into a 9x3 matrix $\mathbf{X}$. We wish to map $\mathbf{X}$ to $\mathbf{F}$. Assuming a linear mapping, the following objective function minimises the sum of squared errors for such a mapping:

$$E = \sum_{i=1}^{N} \left\| \mathbf{f}_i^{\,t} - \mathbf{x}_i^t \mathbf{M} \right\|^2, \tag{12}$$

where $N=9$ and $\mathbf{M}$ is the unknown transformation matrix sought. The least-squares solution is

$$\mathbf{M} = (\mathbf{X}^t\mathbf{X})^{-1}\mathbf{X}^t\mathbf{F}. \tag{13}$$

The triplet of values $\mathbf{y}_i$ after the mapping is given by

$$\mathbf{y}_i^t = \mathbf{x}_i^t \mathbf{M} . \tag{14}$$

To include a larger number of data points when determining the mapping, a method described in [7] can be used.

6 Inverse colour separation

As mentioned above we use algorithms based on neural networks for performing the transformation from RGB to CMY as well as the inverse colour separation.

6.1 The neural network

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. The information processing that goes on within each processing element can be defined arbitrarily with the restriction that it must be local. Most of known neural networks have their processing elements divided into subsets, called layers. Fig. 4 shows a typical feedforward (no recurrent connections) neural network with explicit division of processing elements into three 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 [8], is called learning or training. See for example [9] for a deeper study of feedforward neural networks. In our work we use neural networks of the type presented in Fig. 4.

Let $o_j^{(q)}$ denote the output signal of the j *th* neuron in the q *th* layer induced by presentation of an input pattern, and $w_{ij}^{(q)}$ the connection weight coming from the i *th* neuron in the $(q-1)$ layer to the j *th* neuron in the q *th* layer.

Then

$$o_j^{(q)} = f(net_j^{(q)}), \tag{15}$$

$$net_j^{(q)} = \sum_{i=0}^{n_{q-1}} w_{ij}^{(q)} o_i^{(q-1)}, \tag{16}$$

where $net_j^{(q)}$ stands for the activation level of the neuron, n_{q-1} is the number of neurons in the $q-1$ layer and $f(net)$ is a sigmoid activation function (neuron's transfer function) given by: $f(net) = 1/(1+\exp(-a\ net))$, where a is a slant parameter.

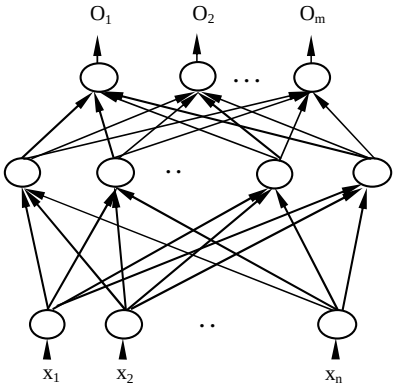

Fig. 4. A feedforward multilayered neural network.

When given an augmented input vector $\mathbf{x} = [1, x_1, x_2, \ldots, x_n]^t$ in the input (0 *th*) layer, the output signal of the j *th* neuron in the output (L *th*) layer is given by

$$o_j^{(L)} = \psi(\mathbf{x}) = f\left(\sum_i w_{ij}^{(L)} f\left(\sum_k w_{ki}^{(L-1)} f\left(\ldots\left(\sum_t w_{tm}^{(1)} x_t\right)\right)\right)\right). \tag{17}$$

When the output values $o_{pj}^{(L)}$ induced by presentation of a particular input pattern $\mathbf{x}_p$ are compared with the desired output values d_{pj}, a mean squared output error cost function is formed as

$$E_o = \sum_{p=1}^{P} E_{op} = \frac{1}{2} \sum_{p=1}^{P} \sum_j (d_{pj} - o_{pj}^{(L)})^2 \tag{18}$$

where P is the number of learning samples. Training the network is done by minimising the error function (18).

When solving the inverse colour separation (colour classification) problem the desired output values d_{pj} are encoded according to the scheme 1-*of*-*m*, i.e. $d_{pj} = 1$, if $c_p = j$ and $d_{pj} = 0$, otherwise; where $c_p \in I = \{1,2,\ldots,m\}$ is the class label of p th input pattern (pixel in our case), and m is the number of colour classes. See the next subsection for a discussion about the number of colour classes.

6.2 Inverse colour separation

We perform the inverse colour separation by determining the colour of every pixel of the image taken from a printed picture. When mixing dots of cyan, magenta and yellow colours, eight combinations are possible for each pixel in the picture. The combination **cmy** produces the black colour. However, in practice, black ink is most often also printed. We assume the black ink to be opaque and therefore **c**+**k**=…=**c**+**m**+**y**+**k**=**k.** Therefore, we have to distinguish between 9 colour classes, namely **c**, **m**, **y**, **w** (white paper), **cy**, **cm**, **my**, **cmy** (black resulting from overlay of cyan, magenta, and yellow) and **k** (black resulting from black ink or black ink plus any other component).

Since we need to distinguish between nine colour classes, our network has nine output nodes (one for each class). The variables f_1, f_2, and f_3 serve as input to the network. In order to achieve an acceptable classification error and a high average classification speed we use a hierarchical-modular neural network [5,10]. Actually, we use a hierarchical committee of neural networks of the type presented in Fig. 4 [10]. The committee solves the classification task in two stages. In the first stage each pixel is assigned to one of six clusters of colours, namely, **c**, **w**, **y**, **cy**, **m-my** and **cm-cmy-k**. In the second stage the clusters **m-my** and **cm-cmy-k** are further divided. We use this approach since the colours **m-my** and **cm-cmy-k** form two clusters of highly similar colours when pictures of newspaper printing quality are considered. In addition to the variables f_1, f_2, and f_3, we extract some features (minimum, maximum, average values of the variables f_1, f_2, and f_3) from the surroundings of the pixel being analysed when classifying pixels from these clusters of similar colours. See [10] for details on the classification procedure.

As a result of the classification procedure an image is decomposed into as many sub-images as there are colour classes in the image being analysed. Fig. 5 shows an example of the classification result. The central part of the image is classified by the developed neural network. An image presented in the figure is taken from a picture printed with two colours, namely, yellow - 50% (50% of the area is covered by yellow dots) and magenta - 100%. Therefore, the image contains two colour classes - **m** and **my**. The figure shows that even if the printing process used for this article is of excellent quality, borders between **m** and **my** areas can hardly be determined. Note, that colour classes **cm, cmy** and **k** overlap even more.

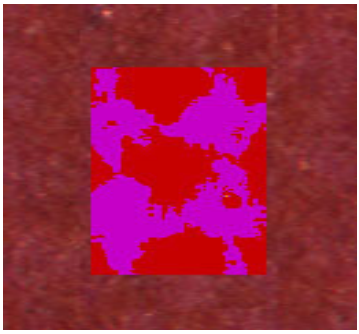

Fig. 5. An image that contains **m** and **my** classes of colours. The central part of the image is classified.

7 Transformation from *RGB* to *CMY*

7.1. Structure of the transformation network and data

In order to reduce variance of the transformation result, we use a committee of neural networks of the type presented in Fig. 4. The architecture of the transformation network is given in Fig. 6. Each member of the committee performs the same task, i.e. estimates the *CMY* output values for the same *RGB* input. The outputs of the members of the committee are then combined to obtain the committee output. It is well-known that a combination of many different estimators can improve the accuracy of the estimate.

Each member of the committee is trained by minimising the following error function

$$E_1 = E_o + \beta \sum_{j=1}^{N_{iW}} (w_{ij})^2 \qquad (19)$$

where N_{iW} is the number of weights in the i th member of the committee, β is a constant and E_o is given by (18). The second term of the error function performs regularisation. If the regularisation parameter β is properly chosen, the regularised network can perform better on unseen data [9,11].

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, when performing transformation from *RGB* to *CMY*. The eight-dimensional vectors serve as inputs to the network, if *CMYK* co-ordinates need to be calculated. The three- as well as the eight-dimensional vectors are measured on a number of test patches. An example of several such patches is given in Fig. 7.

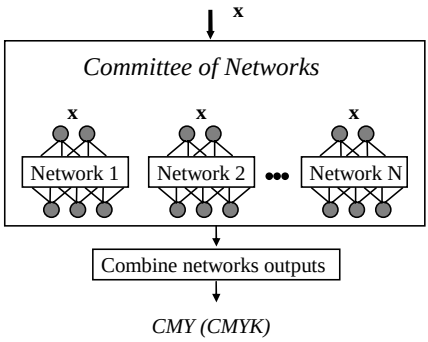

Fig. 6. Architecture of the transformation network.

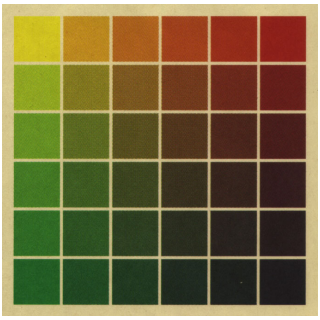

Fig. 7. An example of several test patches used to collect data for training the transformation network.

In order to obtain the eight-dimensional vector, first the inverse colour separation is performed on a test patch. After the separation, the eight-dimensional vector $\mathbf{x} = [x_1, \ldots, x_8]^t$ is obtained in the following way. The components x_1, x_2 and x_3 acquire the mean values of the variables f_1, f_2, and f_3 calculated over a set of pixels coming from the colour classes *{w, c, m, y, cy, my}*. Similarly, the values of x_4, x_5 and x_6 are determined by using a set of pixels from the colour classes *{cm, cmy, k}*. Finally, the values of x_7 and x_8 are set to a fraction of the pixels in the first and the second sets respectively.

In order to train the network each input vector must be paired with the desired output vector, namely, the values of the colour vector components. We recall that values of the colour vector components depend 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. Fig. 8 gives an example of an image taken from one of the test patches as well as the result of the inverse colour separation. The cyan, magenta and yellow areas are highlighted in the corresponding pictures. As can be seen from the figure, after the colour separation the estimate of the actual tonal values can be easily obtained. In what follows, “the actual tonal values” mean the estimate of the actual tonal values.

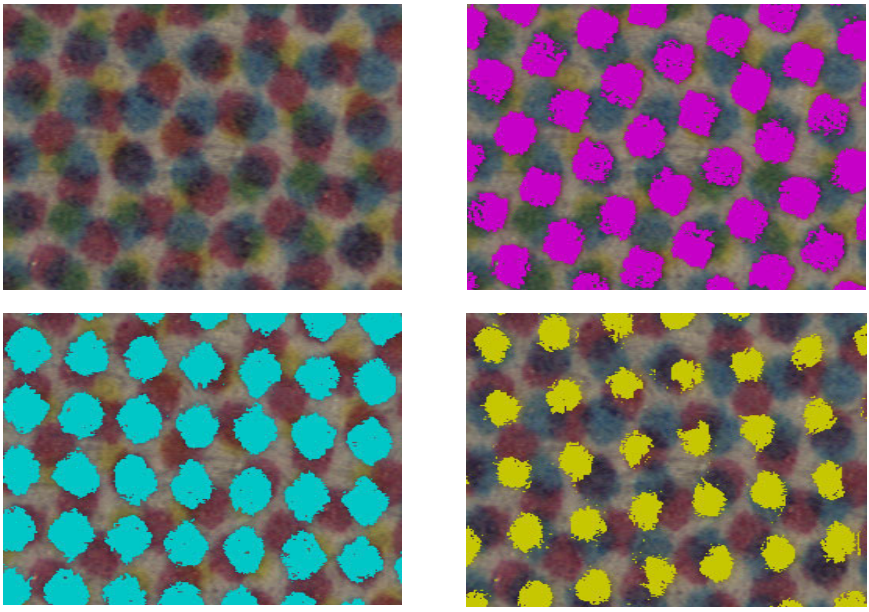

Fig. 8. An example of a colour separation result.

7.2. Combining answers of committee members

It is well-known that a combination of many different estimators can improve the accuracy of the estimate. A variety of schemes have been proposed for combining multiple estimators. The most often used approaches include: weighted average [15-21], average [22], selection of the best in some region of the input space [12], fuzzy integral [13,14].

We can say that a combiner assigns relative weights of importance to estimators in one way or another. The weights can be data dependent [12,17], or can express the worth averaged over the entire space [15,16]. The use of data dependent weights, when properly chosen, provides a higher accuracy of the estimate [12,10].

In this work we integrate two methods, namely, a weighted average and the selection of the best in some region of the input space when combining outputs of the committee members. Let $\mathbf{o}_i(\mathbf{x})$ denote the actual output of the i *th* member of the committee. A weighted combination of outputs of a set of N networks (a committee output) can then be written as

$$\mathbf{o}(\mathbf{x}) = \sum_{i=1}^{N} v_i(\mathbf{x})\mathbf{o}_i(\mathbf{x}) \tag{20}$$

where the weights $v_i(\mathbf{x})$ need to be determined. Note that the weights depend on $\mathbf{x}$. First we choose a number N_{rf} of reference points $\mathbf{m}_i$ in the colour space and determine the prediction error for each member of the committee in the neighbourhood of every reference point. The average squared prediction error for the i *th* member in the neighbourhood of the j *th* reference point can be written as

$$e_{ij} = E\left[\left\|\mathbf{o}_i(\mathbf{x}) - \mathbf{d}(\mathbf{x})\right\|^2\right] \tag{21}$$

where $\mathbf{d}(\mathbf{x})$ is the desired output vector and $E[.]$ denotes the expectation calculated in the neighbourhood of the j *th* reference point. In the experimental tests presented in this paper, we calculated the average empirical error instead of the expected error. The weight of the i *th* committee member in the neighbourhood of the j *th* reference point is given by

$$v_{ij} = \frac{\exp(-\eta e_{ij})}{\sum_{i=1}^{N} \exp(-\eta e_{ij})} \tag{22}$$

where η is a constant. Since the weights v_{ij} can be different in various regions of the colour space, we say that the weights v_i depend on $\mathbf{x}$. Now, given input $\mathbf{x}$, the output of the committee is determined in two steps. First, the nearest reference point k is found:

$$k = \arg\min_{i=1,..,N_{rf}} d(\mathbf{x}, \mathbf{m}_i) \tag{22}$$

where $d(\mathbf{x}, \mathbf{m}_i)$ is the distance between vectors $\mathbf{x}$ and $\mathbf{m}_i$. Then the output of the committee is given by

$$\mathbf{o}(\mathbf{x}) = \sum_{i=1}^{N} v_{ik}(\mathbf{x})\mathbf{o}_i(\mathbf{x}) \,. \tag{23}$$

8 Experimental tests

Experiments we present here mostly concern the transformation from *RGB* to *CMY* and the colour impression. We do not give experiments with the inverse colour separation, since they have been described elsewhere (see e.g. [4,5,10]).

8.1. Training the transformation network

In order to collect data for training the network, 216 test patches, like those presented in Fig. 7, were designed and printed. Two hundred exemplars of such test sheets were printed. The nominal tonal values in the successive test patches were varied in 20% steps, namely 0, 20, 40, 60, 80, 100, for each cyan, magenta, and yellow. For example, nominal tonal values in the test patches presented in Fig. 7 vary as follows. Cyan - from top to bottom: 0, 20, 40, 60, 80, 100%, magenta in the same way from left to right, and yellow is

constant and equals 80%. The test patches measured $1.4x1.4cm^2$ and were printed with approximately the same ink density, namely: cyan - 0.99, magenta - 1.05, and yellow - 0.91.

In order to account for local variations in ink densities and spectral properties of the paper, twelve measurements were made (twelve different colour images taken from 12 different test sheets were recorded) for each of the test patch. Mean values of the variables f_1, f_2, and f_3 were calculated from each of the images and collected into a $Mx3$ (M=216*2=2592) matrix $\mathbf{X}$ to form input data for training the transformation network. Then the actual tonal values were determined for each of the images by performing the inverse colour separation. By collecting the actual tonal values into a $Mx3$ (M=2592) matrix $\mathbf{D}$, we obtained the desired output values for training the network. Five sixths of the data were used for training the network and the rest for validation and testing.

We included seven members in the committee for the transformation network. In order to make outputs of the members of the committee less positively correlated, the members were trained on partly different training sets. One seventh of the data were different when training each member of the committee.

8.2. Data used to test the colour impression concept

During the training process, the neural network learns the non-linear relationship, for constant ink densities, between the *RGB* and the actual tonal values. However, the *RGB* values measured on an arbitrary colour patch depend on both actual tonal values and ink densities. Therefore, to prove the usefulness of the colour impression concept we need to show that there is a good correlation between the ink densities used and the colour vector components measured. Three series of test-prints, we call them the cyan series, the magenta series and the yellow series in the sequel, were produced where the solid print densities for the cyan, magenta, and yellow inks were systematically varied, one at a time. A sketch of the test-print layout is shown in Fig. 9. Five measuring areas in half-tone pictures were chosen to measure the colour vector components. Solid print ink densities were measured by a densitometer on colour bars situated in the same columns as the measuring areas. Fig. 10 shows the half-tone pictures used to measure the colour vector components. The five measuring points are marked by the white circles.

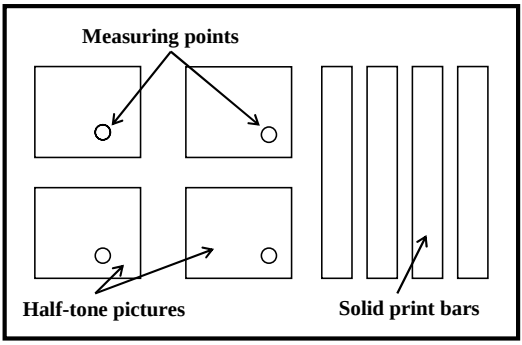

Fig. 9. Test-print layout.

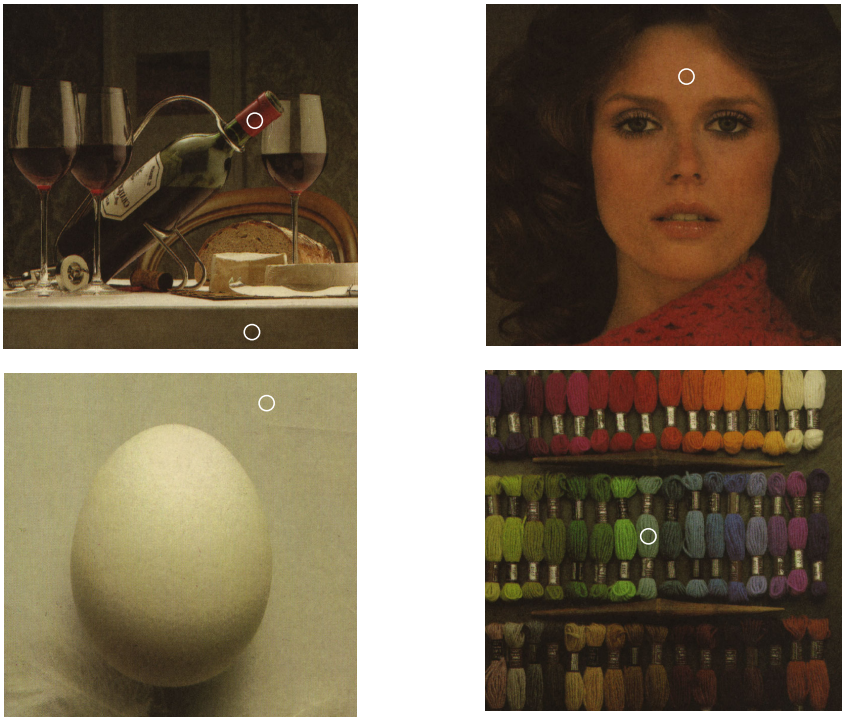

Fig. 10. The half-tone pictures used to measure the colour vector components.

8.3. Results of the tests

We measured the ink densities and colour vector components in all five measuring points and for each series of the test-prints. Fig. 11 and Fig. 12 illustrate the results of the measurements for the cyan series when measuring the colour vector components in point No 4 (the girl picture). In both figures * is used to plot the measurements of cyan, x - the measurements of magenta, and o - the measurements of yellow.

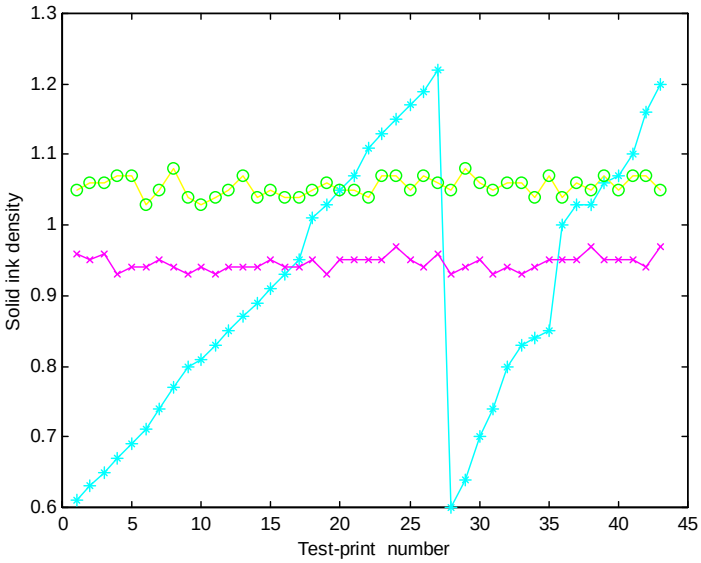

Fig. 11. Densitometer measurements for the cyan series.

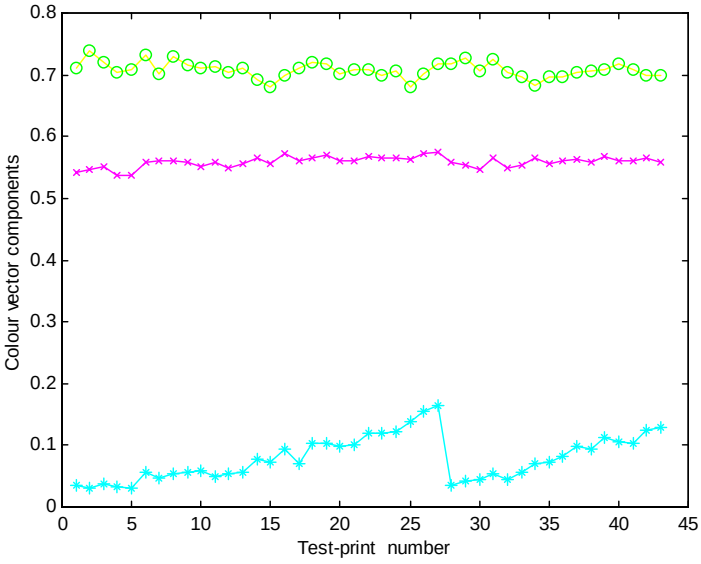

Fig. 12. Colour vector components for the cyan series when measuring in point No 4.

As can be seen from Fig. 11 and Fig. 12, there is a good correspondence between the patterns of variation of the colour vector components and the densitometer measurements. The value of the correlation coefficient between cyan components of the colour vector and solid ink density has been found to be $r_c = 0.971$. The result of the colour separation related to point No 4 is given in Fig. 13. We note that the content of cyan in point No 4 is very low. However, in spite of that, the colour vector follows quite well the

pattern of the cyan solid ink density. It should be noted that the ink density pattern presented in Fig. 11 is a smoothed one, since the values presented are the average values of several measurements. Actually, a variation of 5-10% was observed on the same test-print when ink density was measured.

On the other hand, some discrepancy between measurements of ink densities and the corresponding components of the colour vector is very likely to happen, since these two types of measurements are performed in different environments. The colour vector components are measured directly on half-tone pictures, while ink densities are measured on solid print areas. When measuring the colour vector components we take into account distortions that come from overlay of several inks. We can say, that components of the colour vector reflect the reality of the print, while the solid ink density measurements reflect the greatly simplified reality.

Fig. 14, Fig. 15 and Fig. 17, Fig. 18 illustrate result of the measurements for the magenta and yellow series, respectively. The values of the correlation coefficient for these series have been found to be $r_m = 0.981$ and $r_y = 0.986$. In Fig. 16 we also present a relation between the magenta component of the colour vector measured for the magenta series and magenta solid ink density. As can be seen from the figures presented, there is a good correspondence between the components of the colour vector and solid print ink densities. Similar values of the correlation coefficient have also been obtained for the other measuring points. Fig. 16 shows that an almost linear relation holds between the solid ink density and the colour vector component. The relation shown in Fig. 16 coincides with the observation made about the behaviour of the saturation signal for small pigment values (Fig. 2).

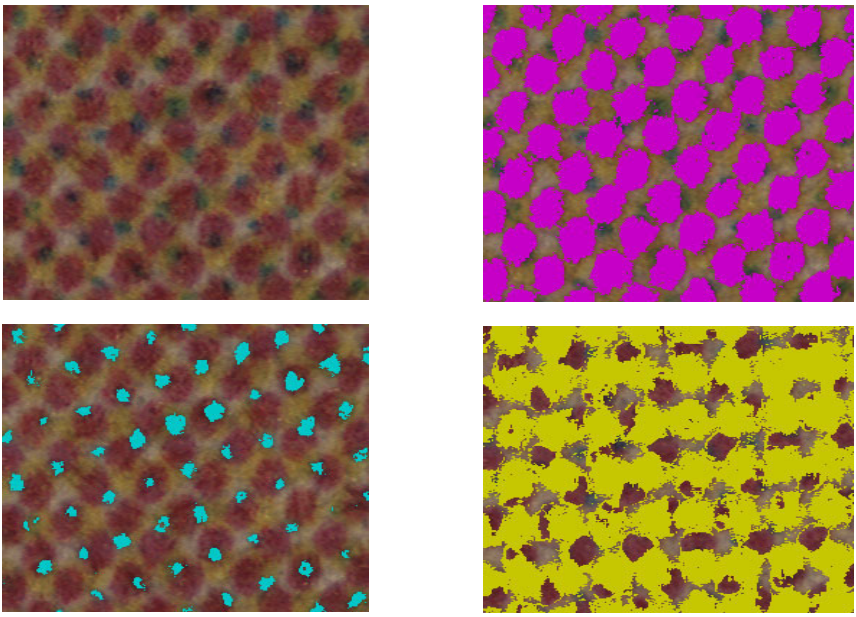

Fig. 13. The result of colour separation related to point No 4.

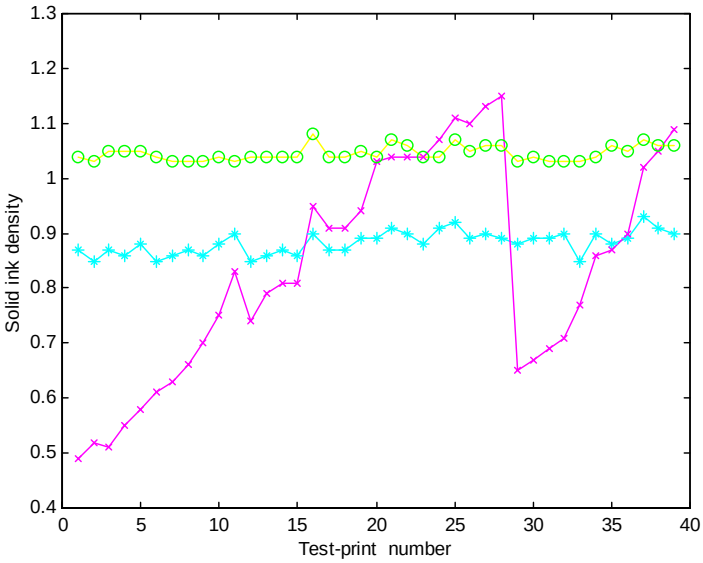

Fig. 14. Densitometer measurements for the magenta series.

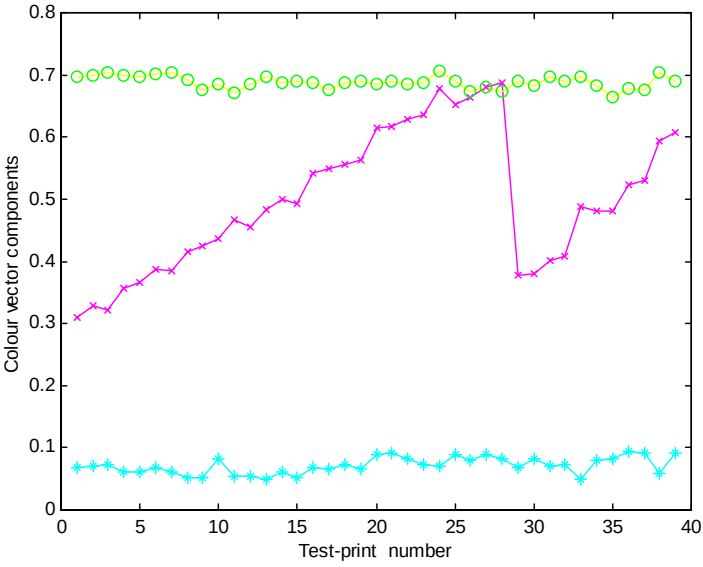

Fig. 15. Colour vector components for the magenta series when measuring in point No 4.

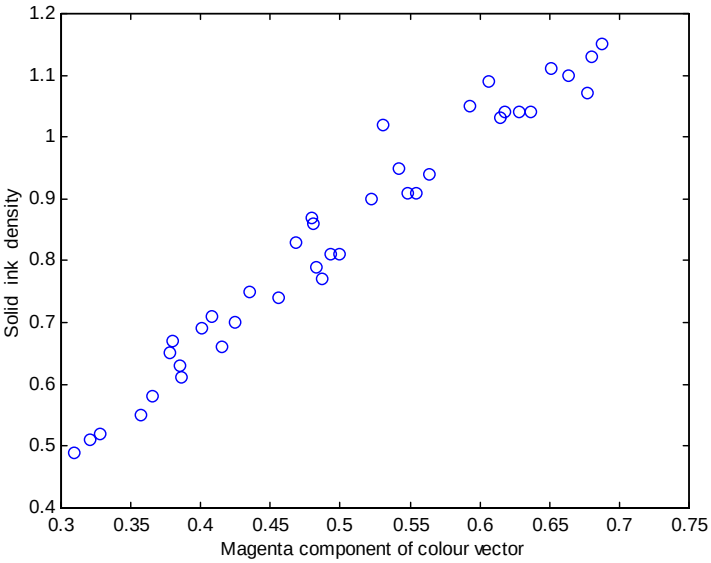

Fig. 16. Relation between magenta component of colour vector and magenta solid ink density.

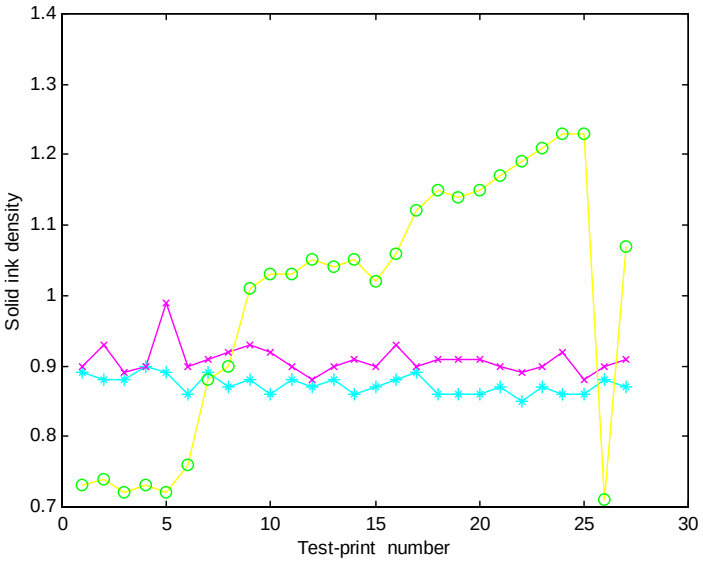

Fig. 17. Densitometer measurements for the yellow series.

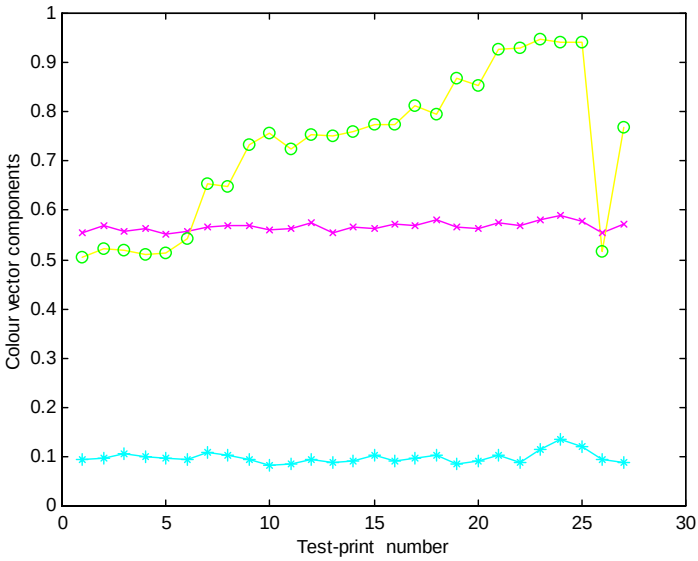

Fig. 18. Colour vector components for the yellow series when measuring in point No 4.

9 Discussion and conclusions

Nowadays, the usual way to obtain the quantitative values for controlling the amount of inks in newspaper printing is by measuring the ink densities in specific printed test areas in the newspaper, for example in a colour bar. By performing such indirect measurements an operator is faced with the hard task of deciding how to change densities of four inks in order to improve the quality of pictures. Besides, such test areas are seldom accepted by the publisher.

In this paper, we have presented a method for colour measurements directly on printed half-tone multicoloured pictures. We introduced the concept of the *CMY* or *CMYK* colour vector, which lives in the three- or four-dimensional space of printing inks. Two factors contribute to the values of the vector components, namely, tonal values and ink densities. The colour vector expresses integrated information about the tonal values and ink densities. If some reference values of the colour vector components are set from a preprint, the colour 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 colours 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 *CMY* or *CMYK* values respectively. We presented a neural networks based approach for performing such a transformation as well as for inverse colour separation.

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

The ability of the tool developed to perform the inverse colour 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 coloured paper. The use of the system can result in reduced consumption of inks and, therefore less severe problems of smearing and printing through.

Acknowledgements

We gratefully acknowledge the support we have received from The Swedish National Board for Industrial and Technical Development as well as from the MoDo and STORA concerns. We also wish to thank two anonymous reviewers for their valuable comments on the manuscript.

10 References

[1] P. Kristiansson, C.M. Nilsson, H. Busk, L. Malmqvist, M. Elfman, K.G. Malmqvist, J. Pallon, K. A. Sjöland, R.J. Utui, and C. Yang, Optical dot gain on newsprint determined with the Lund nuclear microprobe, *Nuclear Instruments and Methods in Physics Research B.* **130**, 303-307 (1997).

[2] P. Kristiansson, K. Malmqvist, L. Malmqvist, C.M. Nilsson and W. Persson, Distribution of ink in solid prints related to the structure of the paper, in *Proceedings of the International conference "Printing and Graphic Arts"*, Mineapolis, USA, 1996, pp. 39-43.

[3] S.A.E. Johansson, and J.L. Campbell, *PIXE - a novel technique for element analysis*, John Wiley & Sons, 1988.

[4] K. Malmqvist, L. Malmqvist, A. Verikas, and L. Bergman, A new method for colour measurements in multicolour newspaper pictures and its use for ink feed control, in *Proceedings of the International Conference "Printing and Graphic Arts*", Mineapolis, USA, 1996, pp.181-185.

[5] A. Verikas, K. Malmqvist, and L. Bergman, Colour image segmentation by modular neural network, *Pattern Recognition Letters.* **18**, 173-185 (1997).

[6] T. S. C. Tan, and J. Kittler, Colour texture classification using features from colour histogram, in *Proceedings of the 8th Scandinavian conference on image analysis, SCIA-93*, Tromso, Norway, 1993, pp. 807-813.

[7] R. Lenz, P. Meer, Spectral based illumination estimation and color correction, in *Proceedings of the 10th Scandinavian conference on image analysis, SCIA'97*, Lappeenranta, Finland, 1997, vol. 2, pp. 885-892.

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

[9] C. Bishop, *Neural networks for pattern recognition*, Oxford University Press, 1995.

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

[11]M. Taniguchi, and V. Tresp, Averaging regularised estimators, *Neural Computation*. **9**, 1163-1178 (1997).

[12]K. Woods, W. P. Kegelmeyer, and K. Bowyer, Combination of multiple classifiers using local accuracy estimates, *IEEE Trans. Pattern Analysis and Machine Intelligence*. **19**, 405-410 (1997).

[13]S. B. Cho, and J. H. Kim, Combining multiple neural networks by fuzzy integral for robust classification, *IEEE Trans. Systems, Man, and Cybernetics*. **25**, 380-384 (1995).

[14]P. D. Gader, M. A. Mohamed, and J. M. Keller, Fusion of handwritten word classifiers, *Pattern Recognition Letters*. **17**, 577-584 (1996).

[15]P. Sollich, and A. Krogh, Learning with ensembles: How over-fitting can be useful, in D. S. Touretzky, M. C. Mozer, and M. E. Hasselmo (eds), *Advances in Neural Information Processing Systems 8*, MIT Press, 1996, pp. 190-197.

[16]A. Krogh, and J. Vedelsby, Neural network ensembles, cross validation, and active learning, in: G. Tesauro, D. S. Touretzky, and T. K. Leen (eds), *Advances in Neural Information Processing Systems 7*, MIT Press, 1995, pp. 231-238.

[17]V. Tresp, and M. Taniguchi, Combining estimators using non-constant weighting functions, in: G. Tesauro, D. S. Touretzky, and T. K. Leen (eds), *Advances in Neural Information Processing Systems 7*, MIT Press, 1995.

[18]A. Verikas, M. Signahl, K. Malmqvist, and M. Bacauskiene, Fuzzy committee of experts for segmentation of colour images, in: *Proceedings of 5th European Congress on Intelligent Techniques and Soft Computing*, Aachen, Germany, vol. 3, 1997, pp. 1902-1906.

[19]T. Heskes, Balancing between bagging and bumping, in: M. C. Mozer, M. I. Jordan, and T. Petsche (eds), *Advances in Neural Information Processing Systems 9*, MIT Press, 1997, pp. 466-472.

[20]C. J. Merz, and M. J. Pazzani, Combining neural network regresion estimates with regularized linear weights, in: M. C. Mozer, M. I. Jordan, and T. Petsche (eds), *Advances in Neural Information Processing Systems 9*, MIT Press, 1997, pp. 564-570.

[21]S. Waterhouse, and G. Cook, Ensemble methods for phoneme classification, in: M. C. Mozer, M. I. Jordan, and T. Petsche (eds), *Advances in Neural Information Processing Systems 9*, MIT Press, 1997, pp. 800-806.

[22] P. W. Munro, B. Parmanto, Competition among networks improves committee performance, in: M. C. Mozer, M. I. Jordan, and T. Petsche (eds), *Advances in Neural Information Processing Systems 9*, MIT Press, 1997, 592-598.

Paper C – Intelligent Monitoring of the Offset Printing Process

L. Bergman, A. Verikas, Intelligent Monitoring of the Offset Printing Process, Neural Networks and Computational Intelligence proceedings 2004, pp. 173-178.

Lars Bergman

Proceedings of the 2nd IASTED International Conference
NEURAL NETWORKS AND COMPUTATIONAL INTELLIGENCE
February 23-25, 2004, Grindelwald, Switzerland

INTELLIGENT MONITORING OF THE OFFSET PRINTING PROCESS

L. Bergman
Intelligent Systems Laboratory, IDE-section
Halmstad University
S-30118, Halmstad, Sweden
email: lars.bergman@ide.hh.se

A. Verikas
Department of Applied Electronics
Kaunas University of Technology
LT-3031, Studentu 50, Kaunas, Lithuania
antanas.verikas@ide.hh.se

Abstract—**In this paper, we present a neural networks and image analysis based approach to assessing colour deviations in an offset printing process from direct measurements on halftone multicoloured pictures—there are no measuring areas printed solely to assess the deviations. A committee of neural networks is trained to assess the ink proportions in a small image area. From only one measurement the trained committee is capable of estimating the actual amount of printing inks dispersed on paper in the measuring area. To match the measured image area of the printed picture with the corresponding area of the original image, when comparing the actual ink proportions with the targeted ones, properties of the 2-D Fourier transform are exploited.**

Keywords—**Neural modelling, Neural network committee, Fourier transform, Offset lithographic printing.**

1 Introduction

Offset lithographic printing is the most widely used commercial printing process. It is used to produce high quality pictures in the production of magazines, catalogs, newspapers, etc. The pictures are represented by cyan (*C*), magenta (*M*), yellow (*Y*), and black (*K*) dots of varying sizes on thin metal plates. These plates are mounted on press cylinders. Since both the empty areas and areas to be printed are on the same plane on the plates, they are distinguished from each other by one being water receptive and the other being ink receptive. As the press runs, a thin film of water is applied to the plate followed by an application of the corresponding ink. The inked picture is transferred from a plate onto the blanket cylinder, and then onto the paper. An example of an image taken from such a picture is shown in Fig. 1.

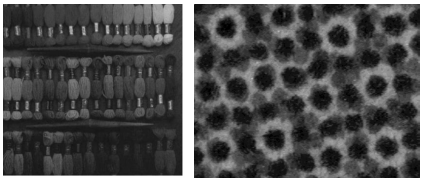

Fig. 1. **Left:** A multicoloured picture. **Right:** An enlarged view of a small part of the picture shown on the left.

Throughout the job run, the press operator repeatedly samples the prints to assess print quality. The samples are compared to the approved *reference print*. The initial comparison is visual. If deviations discernible to the human eye are detected, then a *densitometer* or *spectrophotometer* is used to obtain a numerical assessment of the deviations. Deviations in print consistency from the *reference print* occur as a result of changes in operating conditions and press variability. The types of adjustments that are frequently required during a job run are often attributed to adjusting the proportions of each ink dispersed on the corresponding plate. It is not uncommon for a press operator to alter the amount of ink by nearly 20% during a job run to maintain the same printing quality [1]. Each individual operator performs inking adjustments based on his perception of the relationship between the proportions of the different inks and their overlap that results in the *full-colour* printed picture. The perception is very subjective and so is the printed result.

To measure the ink proportions, usually small test areas are printed. Fig. 2 presents an example of two types of such areas. Full tone test areas of cyan, magenta, yellow, and black colours, as those shown on the upper part of Fig. 2, are usually used for densitometer or spectrophotometer based ink density measurements.

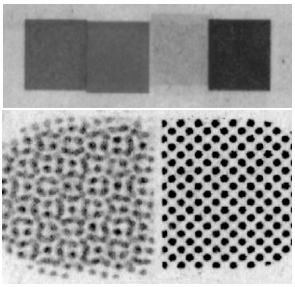

Fig. 2. **Above:** An example of full tone test areas. **Below:** An example of a double grey-bar.

However, print quality of a printed picture depends on both ink densities and dot sizes [2]. Thus, to take both these factors into consideration, one need to measure the ink proportions on halftone areas, for example, on the so called double grey-bars, shown in the lower part of Fig. 2. One

half of the double grey-bar is printed with cyan, magenta, and yellow inks while the other half is printed as a black halftone screen. Obviously, the most adequate estimate of the ink proportions would be obtained if a press operator, instead of using test areas, could measure the proportions directly on multicoloured halftone pictures.

The estimated ink proportions, the press operator has to relate, to the appropriate adjustments of the amount of ink dispersed on paper necessary to maintain print quality. Since each individual press responds differently to on-line adjustments, then it is the experience that a press operator has gained from monitoring the behaviour of the same press over time that allows him to apply appropriate adjustments that can maintain print consistency. Obviously, this is quite a difficult task.

In this paper, we propose an approach to using artificial neural networks for estimating the ink proportions directly on arbitrary areas of multicoloured halftone pictures. The obtained ink proportions are then automatically compared with corresponding areas of printing plates containing information about the targeted proportions of the four printing inks. The comparison result—the deviation—is further used by the press operator/control system to compensate for the colour deviation.

2 Estimating Ink Proportions by Neural Networks

To estimate the ink proportions, an RGB image from a measuring area is first recorded by the image acquisition equipment used. Since the RGB colour space is device dependent and rather nonuniform, the recorded RGB image is transformed to the $L^*a^*b^*$ counterpart [3]. The device independent $L^*a^*b^*$ colour space, which will be briefly introduced in Section 2.2, is known to be much more uniform. Then the mean $\widehat{L^*a^*b^*}$ values are calculated for the image. Now, the ink proportions—the *CMYK* values—are easily obtained for the measuring area if we know the mapping $L^*a^*b^* \rightarrow CMYK$. The mapping is learned by an artificial neural network using a predetermined number of training data measured on specially designed test patches. An example of several test patches used is shown in Fig. 3.

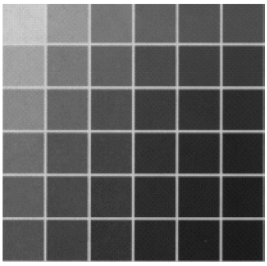

Fig. 3. An example of several test patches used to collect data for training the transformation network.

In this paper, we consider mappings $L^*a^*b^* \rightarrow CMY$—it is assumed that K results from the CMY overlap. However, mappings $L^*a^*b^* \rightarrow CMYK$ can also be learned. *CMY* values for a test patch show how large part of the area of the patch is covered by dots of cyan, magenta, and yellow ink, respectively. For example, *CMY* values equal to 0.2, 0.4, and 1.0 indicate that dots of cyan ink cover 20% of the area, dots of magenta ink cover 40%, and yellow ink covers the entire area of the patch.

To train the network, the mean $\widehat{L^*a^*b^*}$ values measured on the test patches are used as the input data and the ink proportions—*CMY* values—estimated for the corresponding patches, serve for the target data. The way used to estimate the target values is outlined in Section 2.2.

A feedforward two hidden layer perceptron with three input and three output nodes is the neural network used to learn the mapping. The number of hidden nodes is found by cross-validation. To mitigate the data registration burden, it is highly desirable to use small neural network training data sets. It is known, however, that neural networks are unstable to perturbations in a learning data set. Hence, the use of small training data sets may cause generalization problems. Therefore, to cope with the problem, we use neural network committees to learn the mapping.

2.1 Neural Network Committee Used

A variety of schemes have been proposed for combining multiple estimators [4]. For the purpose of the application described here, the averaging approach proved to be the most reliable. This approach simply averages the individual neural network outputs:

$$y_C(\mathbf{x}) = \frac{1}{L}\sum_{i=1}^{L} y_i(\mathbf{x}) \tag{1}$$

where L is the number of neural networks and $y_i(\mathbf{x})$ represents the output of the *i*th neural network given an input data vector **x**.

We apply the *Bayesian* framework for training members of the committee, using a simple *Gaussian* prior for the weights [5]. In the *Bayesian* framework, the assumption is that the network weights are normally distributed and give rise to normally distributed outputs, where each output y_i is effectively the mean of the distribution and σ_i the associated standard deviation. In the committee, therefore, we assume that each network makes an approximation to the "true" distribution of outputs, which has mean y_C and variance given by

$$\sigma_C^2 = \frac{1}{L-1}\sum_{i=1}^{L}(y_i - y_C)^2 \tag{2}$$

With some simple manipulations we get

$$\sigma_C^2 = \frac{1}{L}\sum_{i=1}^{L}\sigma_i^2 + \frac{1}{L}\sum_{i=1}^{L}y_i^2 - \left(\frac{1}{L}\sum_{i=1}^{L}y_i\right)^2 \tag{3}$$

Note that this equation assumes that the networks in the committee all approximate the same *Gaussian* distribution of the outputs.

2.2 Estimating Target Values

In our experiments, six nominal CMY values of 0, 0.2, 0.4, 0.6, 0.8, and 1.0 are used to print the test patches. Thus, there are 216 test patches when using three printing inks. Though the nominal *CMY* values used to print the test patches are known, the actual values are unknown, since dots grow during the printing process. To estimate the actual *CMY* values—the target values for training the neural network—the Neugebauer model [6] is used.

According to the Neugebauer spectral equations [6], the predicted average spectral reflectance of the printed test patch $\widehat{R}_S(\lambda)$ is given by

$$\widehat{R}_S(\lambda) = \left[\sum_{i=1}^{P} w_i R_i(\lambda)^{1/n} \right]^n \tag{4}$$

where $R_i(\lambda)$ is the known spectral reflectance of the ith Neugebauer primary colour—measured on a corresponding test patch, w_i is the area covered by the ith primary, P is the number of the primaries, and n is the Yule-Nielsen correction factor, which accounts for light scatering in paper [6]. There are $P = 8$ Neugebauer primaries when using cyan, magenta, and yellow inks, namely, white paper, cyan, magenta, yellow, green—overlap of cyan and yellow, blue—overlap of cyan and magenta, red—overlap of magenta and yellow, and black—overlap of cyan, magenta, and yellow. Accordingly, there are $P = 16$ primaries when using four printing inks. It is clear, that knowing w_is for each of the test patches, the CMY values are easily calculated. The parameters w_i are obtained from Eq.(4) through an optimization process using data from test patches. Rather than performing the optimization in the spectral space—Eq.(4)—we fulfill the search in the approximately uniform $L^*a^*b^*$ colour space. The $L^*a^*b^*$ coordinates are calculated as follows.

Let us assume that $R_S(\lambda)$ is the measured spectral reflectance of a test patch illuminated by a light source with the spectral power distribution $S(\lambda)$. The predicted spectral reflectance of the patch is denoted by $\widehat{R}_S(\lambda)$. Having the spectral reflectance and the colour matching functions $\overline{x}(\lambda)$, $\overline{y}(\lambda)$, $\overline{z}(\lambda)$ [3], we can calculate the tristimulus values XYZ characterizing the patch:

$$X = k \int R_S(\lambda) S(\lambda) \overline{x}(\lambda) d\lambda \tag{5}$$

where

$$k = \frac{100}{\int S(\lambda) \overline{y}(\lambda) d\lambda} \tag{6}$$

The variables Y and Z are obtained likewise. Having the XYZ tristimulus values, the $L^*a^*b^*$ colour space is defined as follows [3]:

$$L^* = 116(Y/Y_n)^{1/3} - 16 \tag{7}$$

$$a^* = 500[(X/X_n)^{1/3} - (Y/Y_n)^{1/3}] \tag{8}$$

$$b^* = 200[(Y/Y_n)^{1/3} - (Z/Z_n)^{1/3}] \tag{9}$$

where X_n, Y_n, Z_n are the tristimulus values of X, Y, and Z for the appropriately chosen reference white. The Euclidean distance measure can be used to measure the distance ($\triangle E$) between the two points representing the colours in the colour space:

$$\triangle E = [(\triangle L^*)^2 + (\triangle a^*)^2 + (\triangle b^*)^2]^{1/2} \tag{10}$$

The cost function minimized during the optimization process when estimating the parameters w_i is given by

$$J = \Sigma_{k=1}^{K} \Delta E_k \tag{11}$$

where K is the number of the test patches used and ΔE_k is the difference between the kth patch colour (according to Eq.(10)) evaluated using the predicted $\widehat{R}_{Sk}(\lambda)$ and the measured $R_{Sk}(\lambda)$ spectra of the patch. The optimal values of parameters w_i are found by stochastic optimization. Knowing the parameters the target values are easily calculated for all the test patches.

2.3 Training Committee Members

Bootstrapping [7], Boosting [8], and AdaBoosting [9] are the most often used approaches for data sampling when training members of neural network committees. Some studies show that boosting may create committees that are less accurate than a single network [10]. Boosting may suffer from overfitting in the presence of noise [10], [11]. In our application, data sets are rather noisy. Therefore, we have chosen the bootstrapping sampling technique.

We train each member of the committee by minimising the following objective function

$$E = \beta E_D + \alpha E_W = \frac{\beta}{2} \sum_{n=1}^{N_D} \sum_{k=1}^{Q} \left[\mathbf{y}_k(\mathbf{x}^n; \mathbf{w}_i) - t_k^n \right]^2 + \frac{\alpha}{2} \sum_{j=1}^{N_{iW}} (w_{ij})^2 \tag{12}$$

where $\mathbf{x}^n$ stands for the nth input data point, N_D is the number of input data, Q is the number of outputs in the network, $\mathbf{w}_i$ is the weight vector of the ith member, N_{iW} is the number of weights in the ith member of the committee, α and β are hyper-parameters of the objective function. The second term of the objective function performs regularization. We use the *Levenberg-Marquardt* algorithm for neural network training [12] and *Bayesian* regularization techniques [5]. In the Bayesian approach, the posterior probability distribution for the weights, $p(\mathbf{w}|D, \alpha, \beta, M)$, is given by

$$p(\mathbf{w}|D, \alpha, \beta, M) = \frac{p(D|\mathbf{w}, \beta, M) p(\mathbf{w}|\alpha, M)}{p(D|\alpha, \beta, M)} \tag{13}$$

where M is the particular neural network model used, $p(\mathbf{w}|\alpha, M)$ is the prior probability distribution for the weights, $p(D|\mathbf{w}, \beta, M)$ is the data likelihood function, and $p(D|\alpha, \beta, M)$ is the normalization factor. Assuming Gaussian prior distribution of the weights, additive zero-mean Gaussian noise for the target data, and provided the data points are drawn independently, the probability densities are written as

$$p(D|\mathbf{w}, \beta, M) = \frac{1}{Z_D(\beta)} \exp(-\beta E_D) \tag{14}$$

$$p(\mathbf{w}|\alpha, M) = \frac{1}{Z_W(\alpha)} \exp(-\alpha E_W) \tag{15}$$

where $Z_D(\beta) = (2\pi/\beta)^{N_D/2}$ and $Z_W(\alpha) = (2\pi/\alpha)^{N_W/2}$. The optimal weights maximize the posterior probability $p(\mathbf{w}|D, \alpha, \beta, M)$.

3 Comparing Printed Result and the Target

To assess the actual ink proportions on an arbitrary area of a multicoloured picture, first an $L^*a^*b^*$ image of the area is recorded. The average $L^*a^*b^*$ values for the area are then calculated and presented to the neural network. The network predicts the actual ink proportions, which have to be compared with the targeted ones. The targeted ink proportions are known as a digital copy of the printing plates. The original image to be printed is also known, usually in an RGB or $L^*a^*b^*$ format.

To compare the actual ink proportions with the targeted ones one has to match the recorded image area of the printed picture with the corresponding area of the original image. To perform the match we exploit properties of the Fourier transform.

Let $f_1(x, y)$ and $f_2(x, y)$ be two images, and $F_1(u, v)$ and $F_2(u, v)$ their corresponding Fourier transforms. Let also $f_2(x, y) = f_1(x + x_0, y + y_0)$, as illustrated in Fig. 4. Then according to the Fourier shift theorem:

$$\frac{F_2(u, v)F_1^*(u, v)}{|F_2(u, v)F_1^*(u, v)|} = e^{j2\pi(ux_0+vy_0)} \tag{16}$$

where $*$ denotes the complex conjugate. The shift parameters (x_0, y_0) can now be determined from the inverse Fourier transform:

$$F^{-1}\left(e^{j2\pi(ux_0+vy_0)}\right) = \delta(x_0, y_0) \tag{17}$$

Fig. 5 visualizes the inverse Fourier transform calculated for the images shown in Fig. 4. The translation parameters are given by the position of the peak.

The same technique can be applied to identify the rotational displacement and scaling. The rotation is represented as a shift in polar coordinates, while scaling as a shift in logarithmic coordinates [13].

If, for example, $f_1(x, y)$ is $f_2(x, y)$ scaled with factors (a, b), then the Fourier transforms of the images are related as:

$$F_2(u, v) = \frac{1}{|ab|} F_1\left(\frac{u}{a}, \frac{v}{b}\right) \tag{18}$$

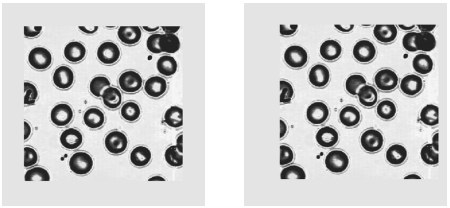

Fig. 4. **Left:** An original image. **Right:** The translated original image.

Fig. 5. Visualization of the result, given by Eq.(17), for the images shown in Fig.4.

Ignoring the factor $\frac{1}{|ab|}$ and using the logarithmic scale we get:

$$F_2(\log u, \log v) = F_1(\log u - \log a, \log v - \log b) \tag{19}$$

or

$$F_2(s, t) = F_1(s - c, t - d) \tag{20}$$

where $s = \log u$, $t = \log v$, $c = \log a$, and $d = \log b$,

The translation (c, d) is found by applying the shift theorem and the scaling (a, b) is given by

$$a = e^c, \quad b = e^d \tag{21}$$

Having the shift, rotation, and scaling parameters, matching of the image areas can be performed.

4 Experimental Tests

The experimental tests performed concern an offset newspaper printing process. First, to test the ability of neural networks to learn the mapping, several test sheets, each containing 216 test patches, were printed keeping the same ink density. An example of thirty-six such pathes is shown in Fig. 3. For each cyan, magenta, and yellow inks, the average ink coverage of a patch area—the dot size—was varied in 20% steps, namely, 0.0, 0.2, 0.4, 0.6, 0.8, and 1.0. Thus, from one test sheet we get one set of data containing 216 data points.

4.1 Target Values and Colour Mapping

To train the mapping networks, one test sheet—216 data points—were used. Fig. 6 presents the result of estimating the target vales for the training data. The target values are given by the average actual ink coverage—average dot size—of the patches. The difference between the nominal and actual dot sizes is known as dot gain.

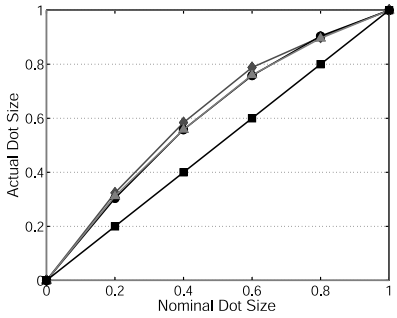

Fig. 6. Actual versus nominal dot sizes.

The optimal structure of the mapping network was found by cross-validation—it turned to be a structure of 3-8-5-3 nodes. The number of networks included into the committee was found to be seven. In Fig. 7, the evaluation result

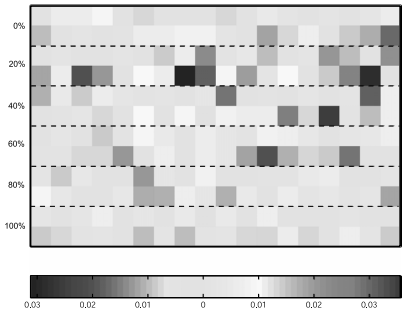

Fig. 7. The error map for the yellow ink coverage, with the percentage of the nominal coverage shown on the left.

of the mapping networks is shown. The Figure visualizes the prediction errors of the yellow ink coverage for 216 colour patches of the test data set. One pixel of the image corresponds to one colour patch. When scanning the image from left to right the nominal cyan ink coverage varies with the highest frequency—20% each step—from 0 to 1 and then repeats it self. The magenta ink coverage varies with 6 times and yellow ink with 36 times lower frequency. The percentage shown on the left of Fig. 7 is for the yellow ink coverage. The maximum error observed is 0.035, while most of the errors are below 0.01. The maximum as well as the average errors obtained for cyan and magenta inks were about 30% lower than those observed for yellow ink. Thus, we can conclude, that the trained committee of neural networks is capable of predicting the actual ink coverage with very good accuracy.

It is worth noting that the exact ink coverage value is not known for each colour patch, since the estimated target values are the average ink coverage values. Colour patches printed with the same nominal ink coverage values may exhibit slightly different the actual ink coverage values. Thus, the ink coverage prediction errors may be even lower than those exemplified in Fig. 7.

4.2 Comparing Target and Result Images

The target and result images used in this test origin from different sources and are of rather different quality. The difference may cause problems in the matching process.

The target colour image is obtained from the digital images used to produce the printing plates. The digital images are high resolution monochrome images, one for each printing ink. From these monochrome images we calculate an RGB image, which is transformed to the $L^*a^*b^*$ counterpart. Fig. 8 presents an example of the target image used.

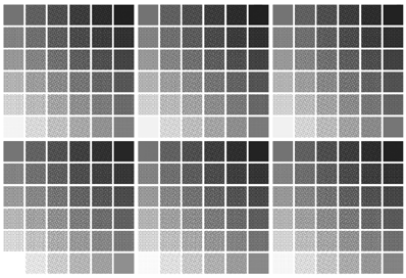

Fig. 8. The target image calculated from the high resolution digital images used to produce the printing plates.

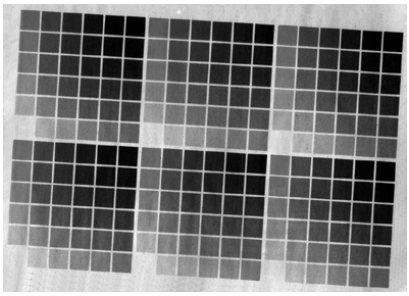

Fig. 9. The result image obtained by scanning the offset printed picture.

To obtain the result image we scan the offset print using a HP ScanjetIIc scanner. The scanner produces an RGB image, which is then transformed to the $L^*a^*b^*$ counterpart. Fig. 9 shows a part of the scanned image to be matched against the target image shown in Fig. 8.

To make the matching process less expensive in computing time and memory, we down-sampled the images used in the matching process from 1500×1500 to 512×512 pixels.

The quite large discrepancy between the target and result images caused by the printing process and effects of

the log-polar conversion, introduces undesirable peeks in the matching result. However, we know a priory that the result image is only moderately translated, rotated, and scaled compared to the target image. Thus, we can avoid selecting one of these false peeks simply by limiting our peek search area. Fig. 10 visualizes the matching result characterizing the scaling and rotation steps. The square shown in Fig. 10 indicates the peek search area while the arrow marks the peek identifying the angle and scale of the result image as compared to the target.

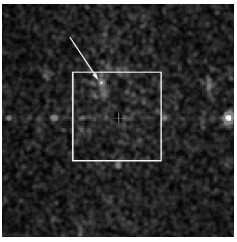

Fig. 10. Visualization of the matching result characterizing the scaling and rotation steps. The square shows our peek search area while the arrow marks the peek identifying the angle and scale of the result image as compared to the target.

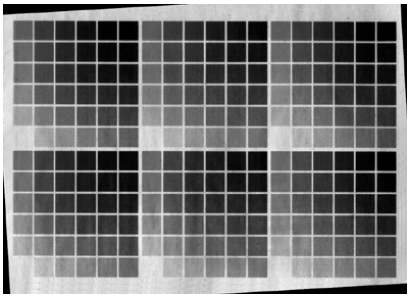

Fig. 11. The adjusted—scaled, rotated, and translated—result image.

5 Conclusions

To maintain an acceptable quality of multicoloured pictures, the offset printing process requires the press operator to make appropriate and timely on-line ink feed adjustments to compensate for colour deviations from the *reference print*. To assess the deviations and monitor the printing process is quite a difficult task for the printing press operator.

In this paper, we presented an approach to assessing colour deviations from direct measurements on halftone multicoloured pictures. From only one measurement a trained neural network committee is capable of estimating the actual relative amount of each cyan, magenta, yellow, and black inks dispersed on paper in the measuring area. The target vales used to train the networks are obtained from the spectral Neugebauer equations based reflectance modelling. A quite high estimation accuracy of the amount of inks was observed in the tests performed. Registration of the measured image area of the printed picture against the corresponding area of the original image is achieved by exploiting the shift, rotation, and scaling properties of the 2-D Fourier transform.

References

[1] S. Almutawa, Y. Moon, Process drift control in lithographic printing: issues and connectionist expert system approach, *Computers in Industry* 21 (1993) 295–306.

[2] A. Verikas, K. Malmqvist, L. Malmqvist, L. Bergman, A new method for colour measurements in graphic arts, *Color Research & Application* 24 (3) (1999) 185–196.

[3] G. Wyszecki, W. S. Stiles, *Color Science. Concepts and Methods, Quantitative Data and Formulae*, 2nd Edition, John Wiley & Sons, 1982.

[4] A. Verikas, A. Lipnickas, K. Malmqvist, M. Bacauskiene, A. Gelzinis, Soft combination of neural classifiers: A comparative study, *Pattern Recognition Letters* 20 (1999) 429–444.

[5] D. J. MacKay, Bayesian interpolation, *Neural Computation* 4 (1992) 415–447.

[6] W. Rhodes, Fifty years of the Neugebauer equations, *Proceedings SPIE* , Vol. 1184, 1989, pp. 7–18.

[7] B. Efron, R. Tibshirani, *An introduction to the bootstrap*, Chapman and Hall, London, 1993.

[8] R. Avnimelech, N. Intrator, Boosting regression estimators, *Neural Computation* 11 (1999) 499–520.

[9] Y. Freund, R. E. Schapire, A decision-theoretic generalization of on-line learning and an application to boosting, *Journal of Computer and System Sciences* 55 (1997) 119–139.

[10] D. Opitz, R. Maclin, Popular ensemble methods: An empirical study, *Journal of Artificial Intelligence Research* 11 (1999) 169–198.

[11] T. G. Dietterich, An experimental comparison of three methods for constructing ensembles of decision trees: Bagging, Boosting, and Randomization, *Machine Learning* 36 (1/2) (1999) 1–22.

[12] M. T. Hagan, M. Menhaj, Training multilayer networks with the Marquardt algorithm, *IEEE Trans Neural Networks* 6 (5) (1994) 989–993.

[13] B. S. Reddy, B. N. Chatterji, An FFT-based technique for translation, rotation, and scale-invariant image registration, *IEEE Trans Image Processing* 5 (8) (1996) 1266–1271.

Paper D – Modelling and Control of the Web-Fed Offset Newspaper Printing Press

L. Bergman, A. Verikas, C. Englund, J. Kindberg, J. Olsson, B. Sjögren, Modelling and Control of the Web-Fed Offset Newspaper Printing Press, TAGA Proceedings 2003, pp. 179-194.

Lars Bergman

Modelling and Control of the Web-Fed Offset Newspaper Printing Press

Lars Bergman[1] Antanas Verikas[1,2] Cristofer Englund[1] Josef Kindberg[1]
Joakim Olsson[1] Björn Sjögren[1]

Keywords: colour control, neural network, newsprint

Abstract: We present an approach to modelling and controlling the web-fed offset printing process. An image processing and artificial neural networks based device is used to measure the printing process output – the *observable variables.* The *observable variables* are measured on halftone areas and integrate information about both ink densities and dot sizes. From only one measurement the device is capable of estimating the actual relative amount of each cyan, magenta, yellow, and black ink dispersed on paper in the measuring area. We build and test linear and non-linear printing press models using the measured variables and other parameters characterising the press. The *observable variables* measured and the press model developed are then further used by a control unit for generating control signals – signals for controlling the ink keys – to compensate for colour deviation. The experimental investigations performed have shown that the non-linear model developed is accurate enough to be used in a control loop for controlling the printing process. The control accuracy – the tracking accuracy of the desired ink level – obtained from the controller was higher than that observed when controlling the press by the operator.

1. Introduction

In lithographic printing a key issue to produce a high quality print is to control the amount of ink dispersed on the paper. In web-fed offset, such as newspaper print, most parts of the process is highly automated. However, one step in the process, namely the control of the inking level is still mostly done manually.

The operator controls the amount of ink dispersed on the paper by controlling the flow of ink to the printing plates. The ink demand is determined by the inked area of the printing plate. The ink is fed from an ink tray by a special roller, ink ductor, through a small adjustable gap controlled by a metal plate – the ink key. These ink keys are divided into narrow areas – ink zones – across the printing press.

[1]*Halmstad University, Halmstad, Sweden*
[2]*Kaunas University of Technology, Kaunas, Lithuania*

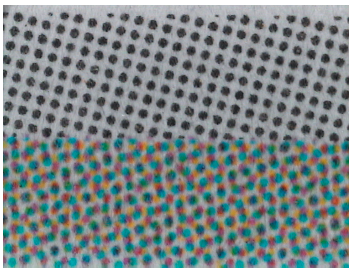

Fig 1. Sample of a grey-bar.

The operator controls the flow of ink, our *adjustable variables*, by adjusting the settings of the ink keys.

Three basic choices have to be made when attempting to automate the ink feed control. More specifically, one must choose the way *to quantify* the degree of deviation in the inking level from the desired one, the way *to model* the process, and the way *to control* the process.

To quantify colour deviation, traditionally solid ink density has been used in the graphic arts industry. However solid ink density, when used as a colour deviation measure, have some drawbacks, for example, it does not reflect the influence of water and other factors on halftone screens.

An ideal instrument would, when placed on a four-coloured print, show how the amount of cyan, magenta, yellow, and black ink need to be adjusted. To be able to determine the adjustment levels, the measured values have to be compared to desired ones. Two major problems then arise. One of determining the measuring position and one of determining the desired value.

Controlling colour by using grey balance have become popular. To do this, test-areas – grey-bars – composed of two halftone screen parts are used, as shown in figure 1. One part of such a bar is printed as a black halftone screen, and the other one as a balanced mix of cyan, magenta and yellow halftone screens to produce a neutral grey print of the same darkness. Grey-bars, originally developed for the naked eye, can also be used to measure on. Measuring on grey-bars does not pose the positioning problem.

We have recently developed an instrument called Malcolm for measuring the amount of ink on grey-bars (Malmqvist et al. 1999). The Malcolm instrument is a multi-functional instrument for print quality control in newspaper printing. The Malcolm instrument is equipped with four main tools for Measuring *Colour Impression*, *Dot Size*, *Density*, and *Registration.*

The *Colour Impression* is a quantity integrating information from both dot sizes and ink densities and yields the amount of ink in the measured area separately for cyan, magenta, yellow, and black.

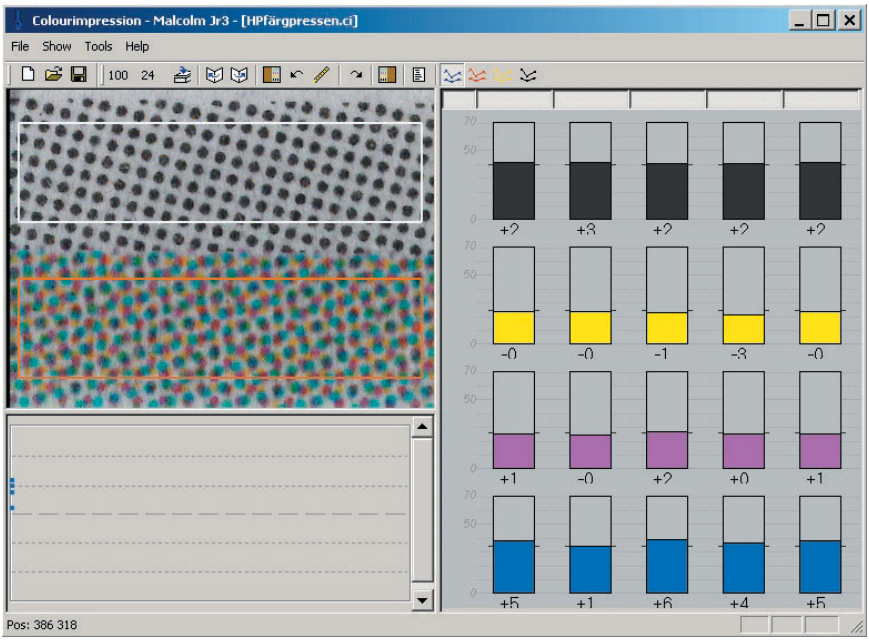

Fig 2. Malcolm screen showing the Color Impression values.

The Malcolm Instrument is optimized to measure *Colour Impression* on grey-bars. The *Colour Impression* values are provided separately for each ink. The values are independent of registration and overprint. *Colour Impression* values are normally presented to the operator as a difference between the value measured for the actual print and the reference print. Thus the values indicate if the operator should increase or decrease the amount of ink. Figure 2 illustrates the main window of the *Colour Impression* tool.

To model the printing process we use both linear and neural network based non linear models.

To control the process we build and test four different controllers: the press operator model based controller, the model predictive controller, the fuzzy logic controller, and the printing press inverse model based controller.

2. Related work

By measuring the amount of ink using a densitometer or the Malcolm Instrument the operator gets a hint on how to control the printing process. The operator has to do all the interpretation of the measured values and ”translate” them into what actions have to be done.

There have been several attempts to develop systems for controlling the ink flow

process. The systems can be categorised into two groups, namely support systems and control systems.

A support system is an expert system used to guide the press operator in the decision process when minimizing the colour deviation in the printed result from the desired one. An example of such an operator support system is the system called CONES (Almutawa S. et al. 1999). The CONES system is a neural-network based expert system that was developed in order to model the behaviour of an experienced press operator taking actions to compensate for colour deviations. The CONES system has been designed to capture the expertise of an operator obtained on a specific machine. The CONES system consists of a neural network, and a rule based expert system.

The CONES system operates in Expert Mode or Novice Mode. In Expert Mode, the system is trained by following the on-line remedial control actions taken by an experienced operator. The purpose of training is to extract the relationship between the *observable* and *adjustable variables.* As an *observable variable* the CONES system uses solid ink density measured on control bars. Density is sampled from the match-print and the actual prints. In the Novice Mode the operator is given advice from the CONES how to adjust key settings to obtain recommended density.

The major conclusion drawn from the development of the CONES system is that the operator's knowledge was specific to the particular machine, where all tests had been made.

A control system is a system that is able to directly control the press. In this case, the measure of the inking level must be obtained on-line. In a semiautomatic control system, first, the press operator adjusts the inking level to the desired one, thereafter the control system keeps the level throughout the job run.

A fully automatic control system can control the inking level without the press operator have to intervene. An example of a control system was presented at TAGA 2000 by Pope (Pope B et al. 2000). The paper describes the results of an implementation of an on-line colour control device. The device was tested in a real-time environment and compared with an open-loop system.

The paper proposed three hypotheses. First the on-line closed-loop colour control system should reduce long-term drift in ink densities. Secondly, short-term variations should be reduced as well. The third hypothesis is that the system should be capable of running the press to programmed target densities during the printing process.

Experiments were run on a heatset press in a commercial printing shop. The system used a spectrophotometer and a video camera to measure solid ink density in bars on line.

The first hypotheses appears to be ”semi-true”, meaning that though long-term drifts are corrected by the closed-loop system, operators also manage to correct the drifts, though not as quickly as the closed-loop system. The second hypothesis is true. Short-term variations are strongly reduced by the closed-loop system. Even the third hypothesis has shown to be true. The closed-loop system drives the press, during make ready, to the desired densities.

Both the CONES and the closed loop system use solid ink density as *observable variables.* By contrast, to model and control the printing process we use the *Colour Impression* values.

3. Modelling the printing process

Building mathematical models of dynamical systems by observing input and output data (system identification), is a necessary step in investigating the system behaviour and constructing a controller.

3.1 Process variables

Numerous variables need to be measured or estimated for modelling the printing process. We categorize the process variables into three groups, namely *observable, adjustable*, and *additional variables.*

Our *observable variables*, the *Colour Impression* values, are measured on greybars.

The *adjustable variables* are the variables used to control the inking level. The only *adjustable variables* we use in this work is the ink key levels. In future work it may be necessary to control other press parameters such as ink ductor speed, dampening degree and others.

In addition to the *observable* and *adjustable variables* we use a set of *additional variables* characterising the printing process. Some of them are determined in advance and do not change during the job run, others constantly changes during the job run.

To ensure an even film of ink to the printing plate, ink is fed to the plate through a number of ink rollers. One or more of these rollers oscillates sideways to even out banks or traces of scratches. This causes ink to flow not only in the machine direction but also locally in the cross direction. In our models, we take this cross flow into consideration.

We obtain all our *observable variables* off-line after the job run. By taking samples both prior and after the operator changes of the *adjustable variables* we can take into account the press dynamics.

The whole set of variables needed for modelling the printing process is listed below:

x_1, x_2 – the copy number, and the printing speed in copies per hour, respectively.

x_3, x_4 – the ink ductor speed, and the ink tray level, respectively. In most presses the ink ductor speed is changed when print speed is changed in order to compensate for the change in ink flow due to the speed. The level of ink in the ink tray influences the pressure at the outlet and therefore it does influence the ink flow in the press. We have chosen not to use x_3 and x_4 in the tests we present here, since our tests were done during the relatively short job runs at a constant speed.

x_5 – the CTP dot size error compensation. The parameter used to compensate for errors in dot sizes on the printing plate depending on the behaviour of the CTP. We have chosen not to use x_5 in the tests we present here, because we used the same CTP to produce all the plates, and the behaviour of the CTP did not change during the relatively short period of the tests.

x_6, x_7, x_8 – the estimated *ink demand* – the desired amount of ink for a given area (ink zone). *The parameters* x_6 and x_8 are *ink demands* for the left-adjacent and the right-adjacent ink zone, respectively. The estimated *ink demand* depends on the job and does not change during the job run.

x_9, x_{11}, x_{13} – the ink key settings prior to the change of an *adjustable variable*. The parameters x_9 and x_{13} are obtained from the left-adjacent and the right-adjacent ink key, respectively.

x_{15} – the *Colour Impression*, our *observable variable,* from the previous time step (prior to the change of an *adjustable variable*).

x_{10}, x_{12}, x_{14}, and x_{16} – are given by x_9, x_{11}, x_{13}, and x_{15} after the changes of the *adjustable variables.*

3.2 Linear model

In the linear model, the output variable y – the *observable variable* (the *Colour Impression* in our case) – linearly depends on the input variables x_i characterizing the printing process:

$$y = w_0 + w_1x_1 + w_2x_2 + w_3x_3 + \ldots + w_nx_n \qquad (1)$$

were w_i are the parameters of the model and n is the number of model parameters. We find the parameters of the model by minimizing the sum of squared errors.

To examine the importance of each variable and to get information if any variable can be excluded from the model, we use the so called Z-score test (Hastie T. 2001).

3.3 Non linear neural network based model

If the relationship between the process input and output is not linear we need a

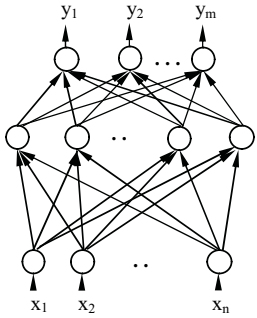

Fig 3. Example of a feedforward neural network.

non linear model for modelling the process. To build non linear models, we use neural network based methods which have proven themselves to be very good at modelling non linear systems. Neural network based methods do not require any apriori knowledge about the data distributions.

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. Most known neural networks have their processing elements divided into subsets called layers. Fig. 4 shows a typical feed forward (no recurrent connections) neural network with explicit division of processing elements into three layers. The layer related to the input, $x, x_2 \ldots x_n$, is called an input layer and that related to the output, $y, y_2 \ldots y_m$, 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 back-propagation algorithm (Bishop 1997), is called learning or training. See for example (Bishop 1997) for a deeper study of feed forward neural networks. In our work, we use feed-forward neural networks with one hidden layer, such as the one shown in Figure 3.

4. Controlling the printing process

We can use different approaches when building our controller. The controller can be based on a forward model which predicts the *observable variables*, or an inverse model which predicts the *adjustable variables.*

An open loop controller predicts a control signal independent of the process current value. A closed loop controller acts upon changes in the process and therefore depend on the status of the process. We can choose to build either an open loop or closed loop controller because we have values both before and after changes of our *adjustable variables.*

We have built and tested four different controllers. The model predictive controller – controller which has been recognized as being an effective tool for tackling some of the difficult control problems in industry, the fuzzy logic based controller, the inverse press model based controller, and the press operator model based controller.

We have chosen to present here the test results obtained from two controllers. The press operator model based controller was chosen because it is interesting to model the behaviour of the printing press operator. We can assume that a press operator models, to some extent, the inverse behaviour of the printing press. Therefore the inverse press model based controller was included to. We have tested these two controllers using both linear and neural network based non-linear models.

4.1 The press operator model based controller

The press operator model based controller is an example of an inverse model based controller. The model predicts the required control signal. Based on the *observable variables* and other variables influencing the printed result, the operator takes a decision how to adjust the inking level. The press operator model based controller is trained to extract the relationship between the *observable*, *adjustable*, and *additional variables* by observing the actions performed by an experienced operator when adjusting for colour deviations. Both linear and non linear models have been implemented and tested. We can expect that the behaviour of the press operator model based controller will be highly dependent on the operator/operators involved and their skills. Such a controller can hardly perform better than the operator. It is also known that operator's knowledge of an offset printing press is specific to the particular press.

The variables x_1, x_2, $x_6 - x_9$, x_{11}, x_{13}, and x_{15} are the input variables used for modelling the operator behaviour. The variable x_{12} is the output variable of the operator model and the controller.

4.2 Inverse model based controller

The press operator model based controller models the behaviour of the press operator, while the inverse model based controller models the inverse dynamics of

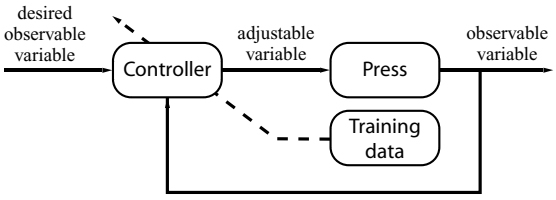

Fig 4. Flowchart of an inverse model based controler.

the press. There is an analogy between these two controllers. An ideal operator would behave as a good inverse model based controller. Figure 4 shows the flowchart of the inverse model based control configuration.

The variables x_1, x_2, $x_6 - x_9$, x_{11}, x_{13}, x_{15}, and x_{16} are the input variables to the controller, and the variable x_{12} is the output variable from the inverse printing press model and the inverse model based controller.

5. Experiments

5.1 Experimental setup

Before our experiments started, the press was investigated and adjusted. All ink keys where adjusted according to the manufacturers instructions to ensure that all ink key settings are comparable.

The dampening system of the press consists of spray ramps with eight valves and nozzles each. The amount of water emitted from each individual valve/nozzle was measured for different settings. In this way it was possible to give each valve an individual setting to ensure that the same amount of water was emitted from each valve.

The reaction of the printing press to changes in ink key settings was investigated in order to ensure that samples are taken as soon as possible after adjustments in ink key settings, but not before the changes of the *observable variables* caused by the adjustments of ink key settings fade away. The investigations showed that the *ink demand* has a big influence on the reaction time. The reaction time is short at high or medium *ink demand* and increases at low *ink demand.* Problems occur when decreasing the ink key settings at very low *ink demands.* In such a case, it can take a very long time to "get rid of the ink in the press". If we exclude cases with very low *ink demand,* two minutes is a sufficient, but not to long waiting time to take a sample after a change of ink key setting.

The experimental tests were done in ordinary production on a four high offset newspaper printing press. The press has 38 ink zones for each cylinder. Each ink zone is approximately 42 mm wide. Samples were collected at 12 occasions during the first part of the job runs, during approximately one hour of production each. During the 12 job runs the same type of newspaper was printed, and the same type of paper was used.

All the print runs started off with a "cold" press. The press used, uses a preset system to adjust the ink key settings prior to production. This reduces the workload of the press operator and decreases the number of adjustments the press operator has to do during the start-up. This was unfortunate for us, because we got fewer data samples per job run, and less variations in our data set.

5.2 Data collection

We used a log system to record all our variables during the 12 job runs. The log system recorded the *ink demand* for each ink zone and ink, and all the changes of the variables x_1, x_2, $x_6 - x_{13}$ for each ink and ink zone during the job runs.

All adjustments made by the press operator during the job runs were followed by the log system and the page number where the adjustments have occurred was recorded. Every time the log system recorded a "new" change for a page, a change of setting of any ink key for any ink, one sample – one newspaper copy – was collected. When no more changes were recorded for that page by the log system during the two minutes delay time a new sample was taken. In this way, it was possible to obtain the *Colour Impression values* before and after the changes.

The sampled newspapers were measured off-line with the Malcolm Instrument after the job runs. The data collected consisted of the *observable variables* – CMYK values – measured on the grey-bars printed along the edge of each newspaper page, and the *adjustable* and *additional variables* recorded by the log system.

In total 672, 621, 678, and 618 data points were collected for cyan, magenta, yellow, and black ink, respectively.

The data collected were normalized to ZERO mean and variance ONE, according to the following equations:

$$x_{norm} = \frac{(x_i - \overline{x})}{\sigma} \tag{2}$$

$$\sigma^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left(x_i - \overline{x}\right)^2 \tag{3}$$

were N is the number of data points, x_i is the variable before normalization, $\overline{x}$ is the mean value of x_i, and σ^2 is the variance of x.

6. Results

Our data sets contain data with rather small variations, so we have to take special care when removing outliers. Outliers are erroneous data values due to noise and press operator errors. We have chosen to define and remove outliers from our collected data set in two different ways, and therefore we got two different data sets. In the first data set – SET I we define the outliers in a more restrictive way and get a data set with low noise and small variations. In the second data set – SET II we define the outliers in a less restrictive way and get a data set with larger variations and higher noise.

TABLE 1. LINEAR PRINTING PRESS MODEL TEST RESULTS FOR THE *SET I* DATA SET.

Data/Training set #	*Maximum error*	*MSE error*	*Mean error*
1	*5.58*	*2.10*	*1.11*
2	*4.80*	*2.18*	*1.12*
3	*4.91*	*2.46*	*1.22*
4	*8.91*	*2.82*	*1.13*
5	*7.51*	*2.81*	*1.13*
6	*7.25*	*2.05*	*1.03*
7	*6.85*	*2.48*	*1.21*
8	*6.71*	*2.25*	*1.16*
9	*7.82*	*3.95*	*1.30*
10	*6.90*	*2.79*	*1.19*
Mean	*6.72*	*2.59*	*1.16*

Our data sets were randomly dived into learning, validation, and test data sets according to the following proportions: 0.7 – 0.1 – 0.2. The learning data set was used to estimate the parameters of the linear and non linear models. The validation set was used to compare the different models. The test data set was used to test the models chosen.

We tested our models using both SET I and SET II data sets. Table 1 shows prediction error for the test set data for the linear model, and Table 2 shows the test data set prediction error for the non linear neural network based model. In both cases, the SET I data set was used. The experiment was repeated 10 times with different

TABLE 2. NEURAL NETWORK BASED PRINTING PRESS MODEL TEST RESULTS FOR THE *SET I* DATA SET.

Data/Training set #	*Maximum error*	*MSE error*	*Mean error*
1	*5.77*	*1.95*	*1.08*
2	*6.30*	*2.73*	*1.25*
3	*5.79*	*2.77*	*1.24*
4	*8.21*	*2.77*	*1.16*
5	*7.02*	*2.72*	*1.17*
6	*4.93*	*1.86*	*1.02*
7	*5.01*	*2.26*	*1.19*
8	*5.26*	*2.03*	*1.14*
9	*7.03*	*2.76*	*1.12*
10	*4.17*	*2.08*	*1.17*
Mean	*5.95*	*2.39*	*1.15*

random divisions of the data set available into the learning, validation, and the test data sets. A press operator normally accepts errors up to 2 or 3 units, thus the errors obtained are quite acceptable.

When we used the SET II data set, the linear model showed errors twice as big as the ones obtained from the non linear neural network based model. Our tests have also shown that there is a significant difference in the results obtained for the different colours.

TABLE 3. TEST DATA SET PREDICTION ERROR FOR THE LINEAR OPERATOR MODEL.

Data/Training set #	*Maximum error*	*MSE error*	*Mean error*
1	*3.79*	*1.39*	*0.80*
2	*5.70*	*1.75*	*0.86*
3	*4.06*	*1.34*	*0.77*
4	*6.62*	*1.92*	*0.89*
5	*5.13*	*1.66*	*0.82*
6	*5,93*	*1.78*	*0.82*
7	*5.96*	*1.31*	*0.79*
8	*8.37*	*1.74*	*0.79*
9	*3.69*	*1.23*	*0.78*
10	*5.68*	*1.33*	*0.79*
Mean	*5.60*	*1.55*	*0.81*

TABLE 4. TEST DATA SET PREDICTION ERROR FOR THE NEURAL NETWORK OPERATOR MODEL.

Data/Training set #	*Maximum error*	*MSE error*	*Mean error*
1	*3.92*	*1.46*	*0.85*
2	*5.58*	*1.88*	*0.93*
3	*4.26*	*1.43*	*0.79*
4	*6.97*	*2.15*	*0.96*
5	*5.53*	*1.78*	*0.94*
6	*5,98*	*1.84*	*0.86*
7	*5.58*	*1.28*	*0.78*
8	*11.11*	*2.37*	*0.82*
9	*3.88*	*1.63*	*0.93*
10	*6.01*	*1.67*	*0.87*
Mean	*5.89*	*1.75*	*0.87*

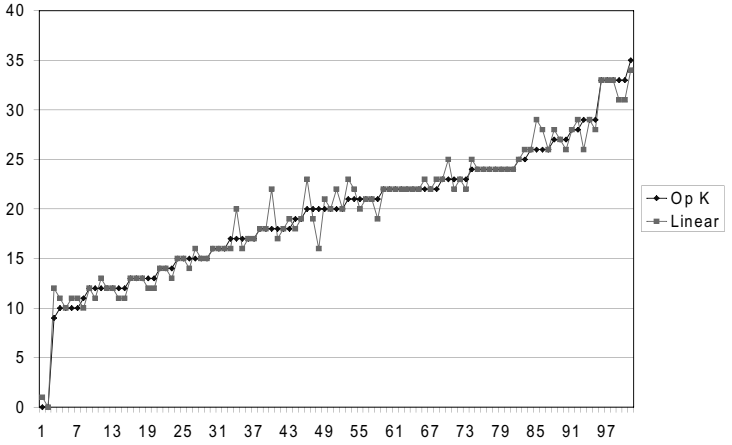

Fig 5. Ink-key levels set by the operator (Op K) and the linear operator model based controller (Linear).

6.1 The press operator model based controller

If the operator modelling task is a non-linear problem, the neural network based press operator model should perform better than the linear one. Tables 3 and 4 presenting the test data set prediction error for the linear an non linear operator model, show that there is no benefit in choosing a neural network for designing the operator model.

The press operator model based controller performed well on the test set data. However such a controller can hardly perform better than the operator. The results show that the operator is moderate in his adjustments to avoid over-inked printing. This is an obstacle for increasing the operator model based controller performance.

Figure 5 shows one hundred different ink-key levels set by the operator and the linear press operator model based controller. As can be seen in Figure 5, the con-

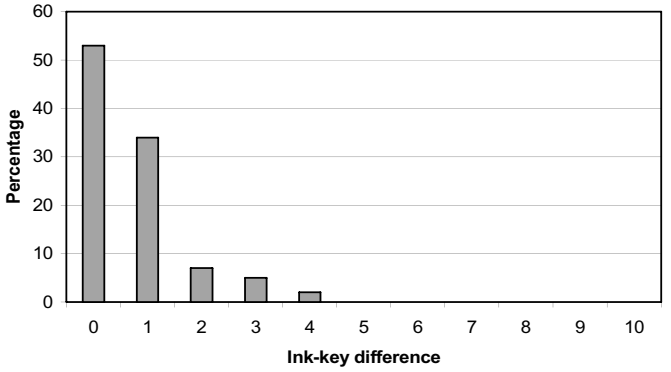

Fig 6. Distribution of the difference between the operator and the linear operator model predictions of ink-key settings.

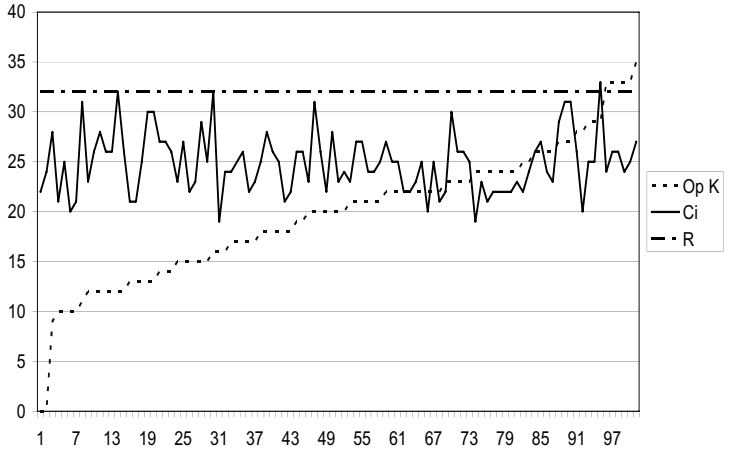

Fig 7. Ink-key level (Op K) as set by the operator, the colour impression value (Ci) and the reference colour impression (R). Data are sorted to obtaion a set of increasing ink key values.

troller predictions follow the operator settings quite well. The histogram of the difference between the operator ink-key settings and the corresponding linear press operator model based controller predictions, shown in Figure 6, indicates that 87 % of the predictions have an error of not more than 1 unit.

Ink key settings are dependent of the *ink demand.* In Figure 7, *Colour Impression* values (Ci) for the test data set and the ink key levels set by the operator (Op K) are shown. The figure shows that there is no clear correlation between the *Colour Impression* values and the ink-key settings. This is due to the fact that the *ink demand* is different for the different data points. As can be seen in Figure 7, the *Colour Impression* (Ci) has a negative bias to the desired *Colour Impression*

TABLE 5. TEST DATA SET PREDICTION ERROR FOR THE INVERSE PRINTING PRESS MODEL.

Data/Training set #	*Maximum error*	*MSE error*	*Mean error*
1	*3.24*	*1.30*	*0.82*
2	*4.03*	*1.21*	*0.78*
3	*3.49*	*1.16*	*0.79*
4	*4.88*	*1.29*	*0.80*
5	*5.43*	*1.91*	*0.97*
6	*4.00*	*1.10*	*0.70*
7	*4.68*	*1.11*	*0.76*
8	*6.10*	*1.20*	*0.70*
9	*3.77*	*1.50*	*0.93*
10	*5.85*	*1.61*	*0.82*
Mean	*4.55*	*1.34*	*0.81*

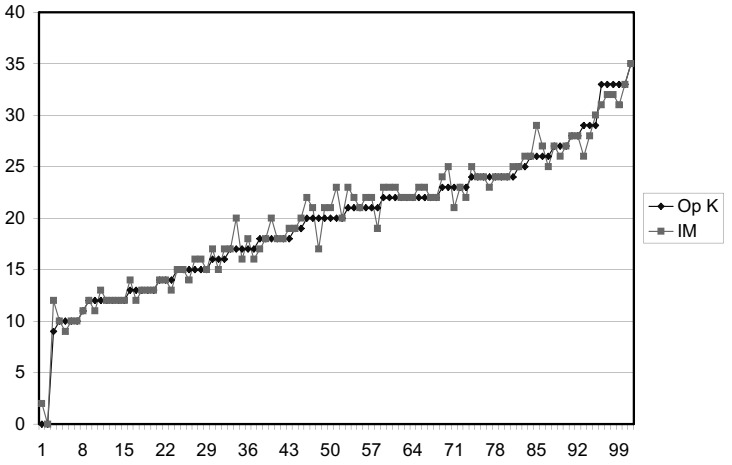

Fig 8. Correct ink-key level (OpK) and the level predicted by the inverse model (IM) based controller.

value (R). This negative bias makes it difficult for the press operator model based controller to perform better than the operator.

6.2 Inverse printing press model based controller

The inverse model predicts the ink-key level required. Table 5 illustrates the prediction accuracy of the model. The predicted ink key values follow the correct ink-key levels in the test data set quite well, as it can be seen from Figure 8. The histogram of the difference between the ink-key levels set by the operator and the corresponding inverse model based controller predictions, shown in Figure 9, indicates that 44% of the predictions have no error, and 42% of the predictions have an error of 1 unit. The maximum error of 3 units and the mean error of 0.77 units have been observed.

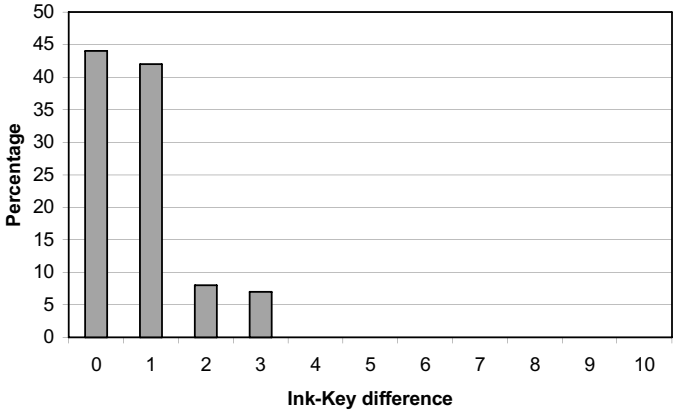

Fig 9. Distribution of the difference between the correct ink-key level and the level predicted by the inverse model.

7. Conclusions and discussion

The printing press modelling results have shown that it is possible to build models to be used for controlling ink flow in a newspaper printing process. The non linear neural network based models have shown better performance than the linear ones.

Improving the prediction accuracy of our models can be done by averaging predictions from multiple models and by averaging measurements from several samples.

The performance of the press operator model based controller depends of the skill and experience of the press operators participating in the training process.

Inverse printing press model based controller has shown the best performance among all the four controllers tested. The obtained controlling accuracy of the ink flow process was higher than that obtained from the experienced printing press operator.

In future work we are going to take steps toward implementing the colour control system on line.

Acknowledgments

The authors wish to thank the staff of the VTAB printing plant in Halmstad, Stora Enso Newsprint Hylte Mill, Hallandsposten AB, and the KK-foundation of Sweden for founding this project. Special thanks to EAE (Ewert Ahrensburg Electronic) GmbH.

Literature cited

Almutawa S., and Moon Y. B.
1999, “The development of a connectionist expert system for compensation of color deviation in offset lithographic printing”, Artificial Intelligence in Engineering 1999:13, pp. 427–434.

Bishop C. M.
1997, “Neural Networks for Pattern Recognition.”, Oxford University Press.

Hastie T., Tibshirani R., Friedman J.
2001, “The Elements of Statistical Learning”, Springer.

Malmqvist K., Verikas A., Bergman L., Malmqvist L
1999, “Malcolm – a new partner in graphic art quality control”, TAGA Proceedings 1999, pp. 473–484.

Pope B., and Sweeney J.
2000, “Performance of an on-line closed-loop color control system.”, TAGA Proceedings 2000, pp. 417–431.