Dot for Dot Proofing

Stefan Livens
Agfa-Gevaert N.V.
Mortsel, Belgium

Abstract

Digital proofing is most successful if it starts from the final rasterised data. When both halftoning and colour properties are conserved in the proof, we call this dot for dot proofing. The match between original and proof should be obtained in three aspects: colorimetry, halftoning and print colorants. The processing includes at least resolution and colour conversion. We propose a general flow that uses an intermediate representation.

An image either represents the colour accurately by giving contone values, or the halftoning by giving high resolution binary data. To reproduce both aspects well in one proof requires methods to handle both types of data together. Solutions for this are presented based on either hybrid or dual representations.

1. Introduction

The goal of proofing is to reproduce images in such a way that they match as closely as possible the appearance of the print, including all its shortcomings. This will generally be more successful if it starts from an image representation that is closer to the one used for the final print.

In classical analog proofing systems, as many processing steps as possible are being shared between proof and print. By using the same films as the plate making, the proofs automatically inherit the exact screening properties of the print. Colour fidelity can be achieved by using colorants for the proof that are very similar to those of the press. The most difficult part is to control the dot gain of the proof so that it also reflects that of the print.[3]

In recent years, analog proofing systems have been replaced rapidly by digital proofing systems. The main reasons for the change are lower cost and faster speed. The fast transition towards computer to plate systems, which eliminate the films upon which analog proofing systems rely, is an additional reason. Regarding quality, both systems have their merits but it is clear that the accuracy of digital proofing solutions is still improving steadily. One direction of improvement is the reproduction of the halftone properties, generally seen as a traditional advantage of analog proofs.

For digital proofing, it is again good practice to have as many components common with the printing. Differences in rendering can be omitted by using the same RIP for proofing. The effect of this is maximised in so called *ROOM* workflows (Rip Once, Output Many): the input data is processed only once and the same output (digital output prepared contone images) is being used for both proofing and printing.

Typical colour proofing solutions start from contone images (in an output colorant space) and use colour management to produce proofs that accurately reproduce the colour of the final print. This usually includes its limitations in gamut, white point and black point, etc.. In some cases, solutions are offered that intend to produce colour accurate proofs with halftoning similar to that of the print (*dot simulation*). Because they start from contone images, and because the colorants of proof and print are different, the resulting dot structure will also be different.

These shortcomings are now being acknowledged and various methods have been developed to remedy them. Their starting point is to have the original rasterisation of the print in common too. Then the proofing starts from the images in their final digital form as rasterised images. When the subsequent processing intends to produce a single digital contract proof that is both colour- accurate and retains the halftoning properties, we classify it as a *dot for dot proofing* method.[6]

Established colour management solutions act on contone images, whereas the halftone appears only after conversion to binary. Clearly, this off-the shelf colour technology does not suffice here and therefore dot for dot proofing poses a technological challenge that calls for new colour management and image processing solutions.

The rest of this text is organised as follows. After outlining some differences between print and proof, we fix three main requirements as aspects in which the proof should match the original. This leads to three basic processing steps which we use to design some processing flows. We propose a general flow which makes use of an intermediate colour space. Throughout, we will emphasize the importance of the image representation. Even more critical is the representation of the dot structure, where we expose the fundamental duality between representations with binary vs. contone data.

2. What is a Good Dot for Dot Proof ?

2.1. Differences between Print and Proof

The basic reproduction properties generally differ from a printing to a proofing device. To convert a printer prepared image for proofing requires specific processing steps for each difference.

2.1.1. Resolution

The proofing resolution is usually lower than the printing (imagesetter or platesetter) resolution. This poses a more stringent limitation on the halftoning that can be (re)produced. In print, the resolution and the minimal printable dot determine the maximal line ruling that can be used for AM screening, and the dot size in FM screening.[5] The same is true for ink jet proofing. The smallest ink droplet determines the minimal halftone dot size. The resolution determines the accuracy of the dot placement.

The line rulings used in AM screening are normally much lower than the absolute highest reproducible frequency, which equals half the (print and proof) resolution. For hybrid screening higher line rulings are typically used,[2] so the limit can become closer.

For all screening types, the ability to produce small dots is the most important limitation. The smallest dot used in print also needs to be reproducible by the proofer. For both, the dot size needs to take into account dot gain. Consequently, higher dot gains on the proofer aggravate the problem. On the other hand, if additional *light* (lower density) inks are used on the proofer, they can be put to an advantage. Even if dots cannot be reproduced small enough, putting a larger dot of a lighter ink is a viable alternative that can produce a tonally correct result.

For FM screening, the small dot size limitation has to strictly adhered to, as all dots have the same small size. For AM screening, one can choose to relax the condition, since it only affect a small portion of the tonal range (with hybrid screening, this portion becomes substantially larger). For the highlights, one either has to give in on tonal accuracy, or allow selective highlight breakout. This means that some dots will be reproduced, while others are omitted. This yields an effect on the highlights that looks similar to that of hybrid screening.

2.1.2. Colorants

The inks used in printing presses are completely different from those used in most proofers. Typically ink jet proofers are used with standard ink jet inks, which are not optimised for reproducing a certain print process. Almost always, both use a kind of CMYK as their basic ink set. Still, their colorimetric values are generally different so that a pure print ink cannot be reproduced by a pure proof ink. Also, ink jet proofers often use extra inks, such as light cyan and magenta inks, while the print might use one or more spot colours.

2.1.3. Dot Gain

The difference in physical properties of both ink and media, combined with the resolution difference and the alternative processing all contribute to differences in dot gain between proof and print. These differences need to be modeled and compensated for. This is the most challenging part, as it is so interwoven with the processing parts.

2.2. Requirements

By definition, a proof has to represents an original and therefore its properties should always be viewed with respect to a reference print. Such a reference can be an actual print, or a standard contract proof made with a trusted proofing system. If such a proof can be made on the same physical proofing device as the dot for dot proof, the comparison becomes very interesting because the device limitations are then equal.

We identify three distinct properties of a good dot for dot proofing method.

- **colorimetric match**: The proof should match the original output colorimetrically. This applies to a fairly coarse scale. It is the scale at which the colour is achieved in print, which directly corresponds to the line ruling used.
- **halftone match**: The spatial properties, dot placement, size and shape are to be reproduced at a finer scale, preferably that of the print, while a somewhat lower resolution (as usually for the proofing device) is still acceptable.
- **print colorant match**: At an intermediate scale, that of the size of an individual halftone dot, an approximately correct colour for the dot is to be achieved. This last goal is less important than the global colour matching, and should only be enforced to the degree that it does not impede the other goals.

Interestingly, each of these goals relates to a different scale of resolution. This makes it possible to reach all of them simultaneously.

When we compare this with the requirements of a standard colour proof, we see that this also needs to obtain a colorimetric match. The scale at which this needs to be achieved should be small enough to represent all image detail, or at least all detail that also appears in print. The move to dot for dot proofing can be seen as enhancement where the requirements about the rendition of small detail are increased. The halftoning is an additional type of detail that appears on a smaller scale. On that scale, the colorimetric match does not have to be met anymore (neither does the print matches the original contone image on that scale). Still, an approximate colour match between the dots of print and proof is desired, which is the third requirement.

2.2.1. Colorimetric Match

The colorimetric match is a macroscopic property. It can be evaluated with the naked eye, if both a proof and a reference can be viewed side by side using a viewing distance large enough so that the halftoning becomes invisible. Preferably, the colorimetric match should be quantified be reproducing and measuring a colour target both as reference and as proof. The colour difference data should be viewed in line with those of other contract proofing solutions. The general validity of the match is easily tested by trying to match different print standards. A good match should also be pursued for non-process print colours. Again it should be verified by measuring a target that covers the tonal range including overprints.

2.2.2. Print Colorant Match

The match of the proof colorants with the print colorants is most evident for large solid pint colorant patches (primary, secondary. …) . For these, the match is the same as a colorimetric match. Apart from the solids, the print colorant match shows itself as a smaller scale property, namely that of the colour of individual print dots. A visual test with a magnifying glass can suffice because the print colorant match is regarded as less critical. A true colorimetric measurement on that scale on the other hand would be difficult.

2.2.3. Halftone Match

The halftone match is the property that undoubtedly enjoys the most interest, as it is the factor that makes a dot for dot proof stand out. This match is situated on the smallest scale. In the print the dot size, shape and position are defined at the print resolution. For the typical case of proofer with a lower resolution, at least some of the accuracy has to be given up, even with an ideal dot for dot method.

Therefore the halftone match can preferably be interpreted as: *how well does the halftone match so that the visually important properties, derived from the halftoning are retained.* In this sense the exact dot placement, dot overlap and dot shape are no longer the ultimate goals. Still, individual dots can be checked visually against a reference with a magnifying glass. Doing so is actually investigating the limits of two different printing techniques, so we cannot expect them to look exactly the same and neither is this required.

It makes much more sense to check other properties like the screen angles and line ruling. Practical comparisons should include the reproduction of small image detail, of smooth vignettes and flat tones. In those the overall sharpness, the possible tonal jumps and the apparent amount of noise can be compared with the print. The black channel should be preserved so that the GCR and UCR settings from the print remain visible, Similarly the difference between dot centered and clear centered rosettes should be clear.

2.2.4. Moiré Prediction

One very important derived property of the halftoning is the creation of moiré as an interaction with the screen raster. This is of particular interest since the capability to predict unwanted moiré patterns is one of the major advantages a dot for dot system can offer. There exist different types of moiré and aliasing, which originate in the diverse interactions between object patterns, image grid, halftone raster, printer grid. ..,[4,1]

Unless the print and proof resolutions are the same, it is to be expected that not all these artifacts can be reproduced by dot for dot proofing (most of them can be reproduced however). A detailed study of moiré types is usually overkill from a practical point of view. However, testing with a small set of moiré-sensitive images can be very valuable. Test images involving patterns in neutral grey areas are particularly useful, as here the moiré can also be steered by changes in GCR.

3. How to Make a Good Dot for Dot Proof ?

3.1. Processing Steps

Given the differences between print and proof, the following processing components will be needed when a dot for dot proofing method is designed:

- **resolution conversion** (RC), whenever the proofer resolution differs from the printer resolution.
- **colorant conversion** (CC) to achieve a colorimetric match.
- **dot gain correction** or colour correction, in order to obtain the correct tonal and colorimetric result for the whole tone scale.

3.1.1. Resolution Conversion (RC)

The resolution of the proofs R_{proof} is usually lower than that of the prints R_{print}. Whenever they differ, resolution conversion will be an indispensable processing step. Only when $R_{proof} = n.R_{print}$ (n integer), the exact pattern of the print can be reproduced.

In all other cases resampling techniques are needed. Minimally, the average needs to be preserved. This rules out nearest neighbour resampling in favour of averaging resamplers(e.g. bilinear resampler). They converts binary data onto data containing intermediate values and inevitably lead to loss of sharpness. Unless $n.R_{proof} = R_{print}$ ($n > 1$, n integer), single print pixels are spread out to several proof pixels too, which can cause aliasing unless the resampler is well chosen.

3.1.2. Colorant Conversion (CC)

The solid primary print colorants can be mapped accurately to mixtures of the proofer colorants using conventional colour management techniques. The same is true for the solid overlaps of the colorants. Together they form the Neugebauer primaries. For CMYK process colours, there are 16 different combinations. This very simple CC method can be applied to the image with the dot structure, which is then converted from binary (expressed in print colorants) to contone (proofer olorants). If it is to be applied on data that is already been through RC, the occurrence of intermediate values make the CC more complex. This is mainly due to the fact that the RC creates additional mixing of values, after which the original amount of overlap cannot be fully recovered.

3.1.3. Correction

CC methods based on primary conversion are capable of producing approximate dot for dot proofs, but in the general the results will not be colour correct over the whole tonal range. The physical reason for this is found in the effects of the halftoning themselves, such as differences in dot gain. Precisely because the CC acts on binary data, it cannot take into account the dot gain of intermediate contone values. Solving this requires knowledge about the contone values from which the halftoned image originates. It leads to more fundamental issues of image representation that will be discussed further on.

3.2. Building a Flow

The CC and RC components can be put in a sequence and form a basic dot for dot flow. Two straightforward flows are shown in fig. 1 a and b. Both start from the binary separated and halftoned image. Its colour is expressed in the input colorant space, usually a print standard CMYK space, possibly with additional spot colours.

- RC-CC flow(fig. 1a): The input is first resampled to the proofer resolution. The RC introduces intermediate values. which makes the CC that follows more complex. On the other hand the intermediate

values also offers possibilities to refine the conversion in order to make it more colour accurate.
- CC-RC flow(fig. 1b): The input is first transformed to the proofer colorant space. This CC acts on the print resolution data, which is usually high resolution, but it can be very simple since the binary input has only a small number of different possible values. After this the image is rescaled to the proofer resolution.

Both flows have advantages regarding both output quality and processing complexity. Fact is that neither of them produces perfect results. A good implementation of the third processing component is essential in order to achieve a good dot for dot proof. The correction is not necessarily a separate module, but can take the form of an addition to a basic flow in many different ways.

The output of the flows is a contone proofer image. This image can be transformed further to make it printable on the proofing device. For this, the same techniques can be used as for proofing without dot simulation. FM screening methods are particularly well suited, as they do not interfere with the screening inherited from the print.

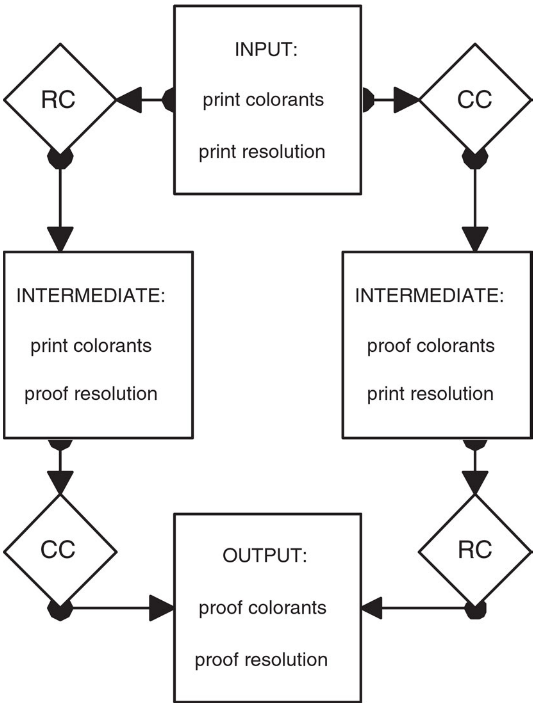

Figure 1. Dot for dot flows: a) RC-CC flow, b) CC-RC flow.

3.3. Image Representation: Intermediate Colour Spaces

While putting together a flow, we stressed importance of the image representation, as it makes the fundamental difference between the flows considered. Logically, any processing (such as RC) yields different results if performed in a different representation.

3.3.1. Generalised Flow

The two basic flows presented above are not the only possibilities. We propose a more general flow (see fig. 2) where the colour conversion is split into two subprocesses:
- CC1: conversion from press colorant space into an intermediate colour space.
- CC2: conversion from an intermediate colour space into the proofer colorant space.

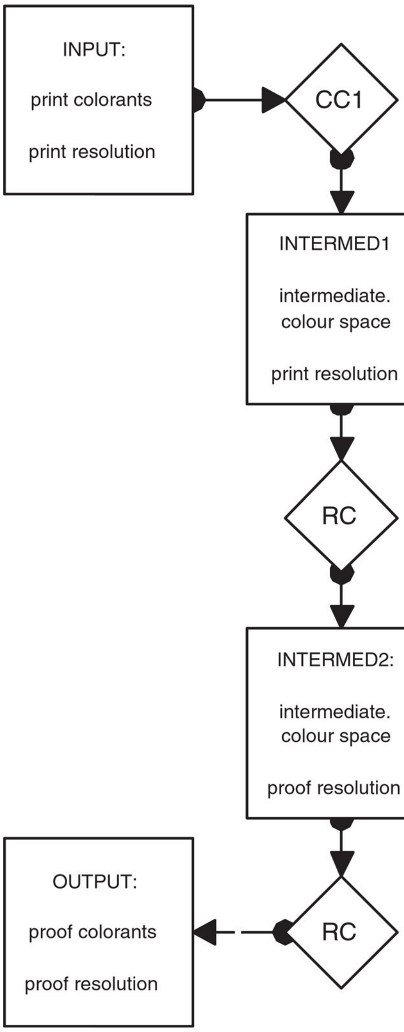

Figure 2. Dot for dot flow with intermediate colour space.

The RC in between is computed using the intermediate colour space. With this flow, the previous flows are contained as special cases. If the intermediate space is equal to the press colorant space, CC1 is an identical operation and the flow reduces to a RC-CC flow. Similarly, the intermediate space equal the proofer colorant space, CC2 is trivial and a CC-RC flow is obtained.

The general flow offers however far more possibilities because the intermediate representation can be freely chosen. An interesting choice is the space of the Neugebauer primaries. An advantage of this is that the amount of overlap is preserved through the processing (such as the RC), since all overlaps are stored in separate channels. A practical disadvantage is that the number of channels increases.

3.3.2. Device Independent Colour Space

Very useful are intermediate representations in a device independent colour space. Approximately perceptually uniform spaces such as CIELab have the advantage that averaging results in a correct visual average. Computational limitations such as rounding errors also have the least visual influence. For CIE XYZ, which directly relates to the amount of light, the averaging is physically justified because of the additivity of light. They also allow a practical way to deal with spot channels without extensive overhead: the spot channels are converted at the start just like the process colour channels.

3.3.3. Creating Custom Representations

The choice of intermediate representation does not need to be limited to standard colour spaces. In fact any representation can be created. A complete definition of the colour at every position is a necessary part of any usable representation. This requires at least three channels (for a colour description), but can also be given with more channels as is the case with CMYK.

In addition to the colour definition, extra channels can be extracted from the original input and added to the representation. These can contain specific information about the input image, such as the ink amount of a specific ink (e.g. black or a spot colour), or a derived property, e.g. the total ink amount. This information can be used to enhance further processing.

A good example of this is the following. Print CMYK (4 channels) is converted to Lab (3 channel colour definition). The original K channel is added as additional information, creating an overcomplete representation. In the CC2 conversion (to proof CMYK), the K channel is used as a key to steer the usage of proofer black. With this scheme the usage of black ink in printer is mapped to the proofer in addition to the colour mapping.

4. Colour Correction

4.1. The Dot for Dot Paradox

The basic flows will in general not yield colour accurate results. The physical reason for this to be found in the effects of the halftoning themselves, such as differences in dot gain and dot overlaps. When trying to correct for this problem, one expects to know both the colour and the halftoning at every position of the image. Then, one inevitably encounters the so-called *dot for dot paradox*:

An image either represents the colour accurately by giving contone values, or the halftoning by giving high resolution binary data.

The contradiction is only apparent as the contone information is in fact contained in the rasterised image. After all we perceive rasterised prints as if they were contone. The contone information in a rasterised image is spread out over an area of pixels, and not present at every individual pixel. When only a pixel by pixel processing is performed, one cannot profit from the contone information necessary for generating a colour correct proof.

The embedded contone information however can become available when some spatial processing is performed. This processing can be a full descreening method, but it can also be revealed through morphological filtering, or even by a simple resolution conversion.

4.2. Solving the Paradox with Contone Data

The key issue in the design of a dot for dot method is the way to present the data so that both the binary and the contone aspects can be taken into account by the processing in a convenient way.

The binary aspect is best presented by staying as close as possible to the original (binary) input data. The contone aspect must be derived by performing extra processing. The result of this can be expressed in two different ways. Explicit contone is obtained by adding extra contone channels and create a double representation. Implicit contone information is obtained by altering the existing channels so that they contain contone information at the individual pixel level without completely erasing the binary information (hybrid representation).

4.2.1. Explicit Contone: Dual Representations

The most straightforward way to add contone information is to complement the existing binary image with a contone version. Some advantages of this approach are: Then, the binary data can be kept in a state optimal for the dot reproduction, without the need for making compromises needed for the contone aspect. One can also rely maximally on proven contone colour management techniques, while the processing can be reduced by using a contone image at a lower resolution. Still the amount of data to be processed will be larger. Unless original contone is available, the descreening itself is a time-consuming process in itself, while it inevitably introduces artifacts (especially important if line art or text is reproduced).

4.2.2. Generating Contone Images

The addition of contone information usually implies that this data must be regenerated from the raster data. Of course if original contone data is available alongside the binary data, it is preferred to use this. Otherwise, the contone data can be can be estimated with well-established techniques known as descreening or inverse halftoning.

They generally produce better results if more is known about the characteristics of the halftoning. It is known that many descreeners produce artifacts or fail to completely invert the halftoning. This does not necessarily generates quality problems in a dot for dot flow. After all, the descreened images are not used for proofing by themselves, but only to correct the halftoned images. In this sense, the requirements for the descreener are less strict.

4.2.3. Implicit Contone: Hybrid Representations

Alternatively, the binary input data can be transformed in such a way that it represents both the binary dot and the contone colour information . Such a hybrid representation seems to be more complex and potentially less accurate. It can however be a natural part of a dot for dot workflow. The processing in such a flow usually includes some spatial processing, so that the pixels in the resulting image already contain some neigbhourhood information.

As an example, low pass filtering is often performed together with the resolution conversion, in order to prevent aliasing effects. That filtering is a spatial processing and generates intermediate values near the edges of every halftone dot. This is also true for the resolution conversion itself, if the interpolation generates intermediate values.

In fact, an implicit method will generally be faster and consume less memory than an explicit method. However, the extra processing of the image can result loss of accuracy regarding the dot structure, more than what is needed for the spatial printing only.

The main difficulty of the hybrid method lies in the colour correction. It is not evident to make good use of the intermediate values, since they do not represent contone data in the conventional sense. For this, one as to bear in mind the colour accuracy does not need to be achieved at the pixel scale, but at a larger scale. Therefore it suffices to correct colours to become correct *on the average*.

5. Conclusion

We introduced the concept of dot for dot proofing as a logical step towards more faithful print reproduction. We outlined the requirements of dot for dot proofs and the basic processes that are needed to produce it. We presented a general flow that incorporates conversion to an intermediate representation.

The fundamental issue turned out to be the representation of binary halftone and contone colour data. The latter needs to be made visible through processing of the first. It can then be stored either explicitly as additional channels or implicitly with a hybrid representation. Both methods offer the possibility for subsequent processing that delivers correct colour without losing the dot structure.

References

1. I. Amidror. The theory of the Moiré Phenomenon. Kluwer, 2000.
2. R. Bartels. Reducing patterns in the fm part of tile-based hybrid screens. In Proc. IS T PICS Conference, pages 241–244, Portland, Oregon, April 2002.
3. M. H. Bruno. Principles of Color Proofing. Gama Communications, 1986.
4. P. Delabastita. Screening techniques, moiré in four color printing. In Proc. of TAGA, 1992.
5. X. Kang. Digital Color Halftoning. Wiley, 2002.
6. S. Livens. Dot for dot proofing: how to zoom into the dots without losing the big picture. In Proc. SPIE 5293: Color Imaging, Device- Independent Color, Color Hardcopy, and Graphic Arts IX, San Jose, CA, January 2004.

Biography

Stefan Livens obtained a Ph.D. in Physics from the University of Antwerp in 1998,for which he applied colour image analysis to pictures of silver halide crystals.

When working at Barco Graphics, he was involved in halftoning and developed a dot gain compensation software called "Intellicurve".

Since 1999, he is with AGFA working in their color technology centre. He is responsible for the calibration technology used for Agfa's Sherpa proofers. In the last two years, he has developed a novel dot for dot contract proofing technology which is now part of Agfa's ApogeeX workflow system.