Scanning Color Negatives

Chris Tuijn
Agfa-Gevaert N.V., GS/EPS/R&D, Mortsel, Belgium

Abstract

The demand for accurate color reproduction has never been as high as it is today. Not only in the high-end electronic prepress market, but also in the desktop publishing and home office markets, the availability of both input and output devices is increasing rapidly.

Most of the input devices today capture positive originals: scanners capture either reflective or transmissive originals; digital cameras are capable of capturing real life scenes as well.

In some market segments (such as, e.g., the newspaper environment), there also is a definite interest in scanning negative originals. Especially with the new emerging APS standard for film (where manual manipulation of the film strips is no longer necessary), the demand for negative scanning will also increase in the home office market.

Scanning negatives, however, is a very delicate process. Not only the input device should be characterised properly, but also the negative film itself is a parameter which needs to be studied carefully. On negative film, the information is stored inverted and due to the color dye layers within the negative film, there also is a density shift between the red, green and blue planes. The main problem, however, is caused by the fact that, due to the variations in the development process, the characteristics of a strip of developed negative film can differ considerably from other strips of the same film type.

In this paper, we first give a brief survey of our approach to scanning negatives presented in the past. Then, we show how the unpredictable properties of negative films can cause this approach to fail and discuss some substantial improvements. In this respect, we show how the adaptive approach taken in the conventional photo-finishing environment can be used electronically. In a following section, we describe how the inverted positive image data can be transformed into a well-known, calibrated color space. In the last section, we briefly discuss the minimal requirements for an ideal negative scanner.

Introduction

Making color pictures using a camera has been a commodity for several decades now, but this does not mean we should underestimate the underlying technology. When making a picture with a conventional camera, we expose a rectangular area of a negative film strip. After exposure of the entire film, the negative film has to be developed, which results in the typical orange-like film strips. From these developed negative strips, prints can be made by photo-finishing labs.

Recently, a new system has been introduced which eases the handling of negative films considerably, called APS: Advanced Photo System. This system has several advantages over the traditional 35 mm film strips:

- There is no need to touch the film anymore, since it is encapsulated in a holder which never should be opened by the photographer. The camera will pull the film strip automatically from of the holder and put it back in when the photographer wants to remove the holder from the camera.
- The camera can position the strip to whatever location is wanted. It is also possible to get partially exposed film strips out of the camera and put them in later. This makes it possible to work with different films using the same camera.
- Other improvements are, e.g., the electronic registration of data (so-called PQI-data) while shooting the pictures (such as averaged incoming light, exposure latitude, shutter speed, geometry etc.).

In Figure 1, you will see a typical APS holder.

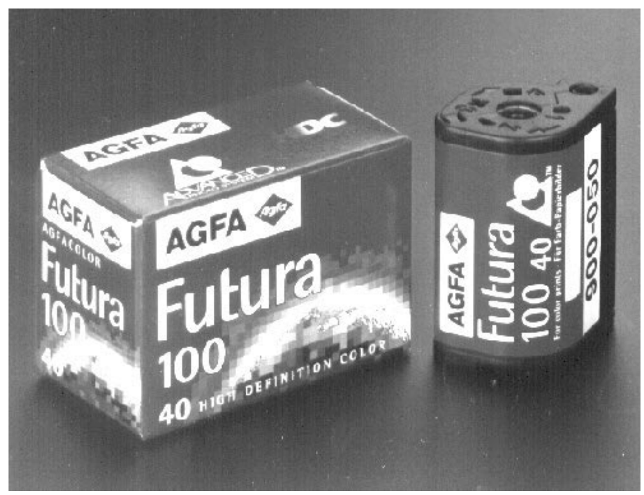

Figure 1. A sample APS film holder

There is no need anymore today to emphasize the importance of digitised information. As pointed out in the abstract, the demand for negative scanning is relatively high in the newspaper market; with the new APS standard for negative film, the home-office market will also become more and more interested in scanning negative (APS) film.

Obviously, correction techniques are necessary to convert the scanned (negative) information into usable positive information. In this paper, we will describe solutions to realise this conversion. For completeness, we will first recall some results presented in the past.[7,8] Then we will present a new approach to solve the problem of the variable characteristics of negative film by analysing the captured scenes adaptively.

Going from Negative to Positive

Some Background on Negative Originals

In order to know how we have to calculate the inversion curves, we first give some background on the which make up a negative film. Negative film is made of three different layers which each are sensitive to one of the three primaries. Cyan, magenta and yellow negative images are formed in the negative films in red-, green- and blue-sensitive layers of the film.[6] To produce a color print, this negative composite is printed on a positive material (the so-called photographic paper) that also has three emulsion layers sensitive to the same red, green and blue primary colors, in order to produce a subtractive color print.

The three emulsion layers incorporate *color couplers* consisting of molecules of compounds that will unite with the reaction products of the silver development to form dyes. The molecules in the blue-sensitive layer form the yellow dye; those in the green-sensitive layer, magenta dye and those in the red-sensitive layer the cyan dye.

In order to compensate for unwanted absorptions of light in the cyan and magenta image dyes, dye masks are also incorporated in the film. This gives the negatives their characteristic overall orange mask. One of the challenges in negative scanning, consists of getting rid of this mask.

How Scanners Perceive Negatives

As already explained above, a developed negative consists of several dye layers which have been activated by a chemical process. It is very complicated to try to model this process mathematically, and therefore we choose for an approach which characterizes the process of:

- exposing the negative film,
- developing it, and
- capturing it with a particular scanner as a whole.

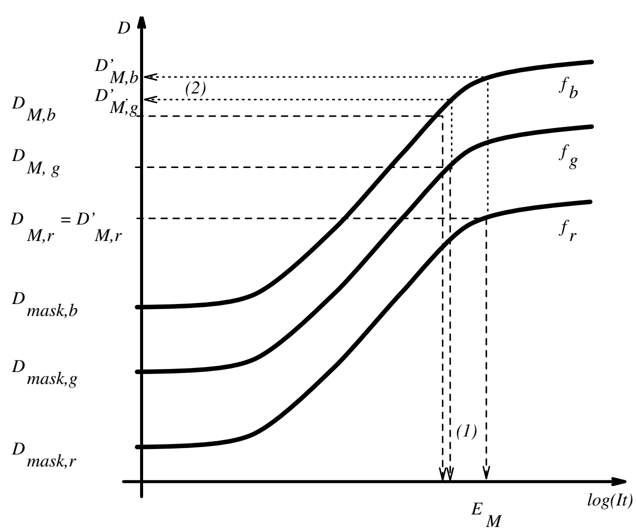

Figure 2. The characteristic curves of a negative film

An experiment which is very useful to characterise this process consists of examining how various intensities of neutral light are recorded on a negative film and how a particular scanner sees these intensities. This experiment can be realized by exposing a negative film in contact with a silver step-wedge and developing it. Figure 2 indicates how a particular scanner perceives this negative step-wedge through it's filters. It consists of three curves (one for each filter); these are usually referred to as the *characteristic film curves*.

On the X-axis, the original intensities are put on a logarithmic scale; the Y-axis contains the densities of the negative (as captured through the scanner's filters). The typical S-shaped characteristics curves are obtained by sampling the different patches of the test strip, and converting them to the right coordinate spaces (log(It),D). In order to eliminate small measurement errors, regression techniques are applied to represent the characteristic curves by an analytical function (as described in[5,7]). This results in the smooth S-shaped characteristic curves of Figure 2.

We can observe that the red, green and blue curves are shifted with respect to each other (the distance is determined by the background mask). The slope in the inflection point of the curves (which is often referred to as the gamma-value) is smaller than 1, which proves that negative film compresses the density ranges. Typical gamma-values for negative film range from 0.65 to 0.80, and can be different for the red, green and blue channels. For our test wedge exposed on Agfa HDC 100* film, e.g., the gamma values (as perceived by the Agfa SelectScan Plus† scanner) are 0.63, 0.71 and 0.73. Pieces of unexposed film generate the mask densities, which for our test strip have values 0.25, 0.58 and 0.70 D. Most of the light will be captured in the linear part of the curve; very powerful light sources will be captured in the shoulder of the curve, which ultimately saturates in the so-called saturating densities (2.21, 2.72 and 3.28 D for our test strip).

It should be noted that, for typical scenes, the original exposure range is about 2 D, and the resulting range on the negative film is about 1.3 D. In order to generate lookup tables which will convert from negative to positive, we have to cut out the piece of the characteristic film curves we need and follow them in the opposite direction, as will be explained in following paragraph.

Calculation of the Inversion Luts

In order to determine the density range of the negative original, a preliminary scan should be carried out. From this scan, the minimum and maximum densities should be determined (denoted as $D_{m,r}$, $D_{m,g}$ and $D_{m,b}$ and $D_{M,r}$, $D_{M,g}$ and $D_{M,b}$ respectively.

Then, the characteristic curves should be followed in the inverse direction (i.e., from D to log(It)) to determine the extreme exposure values. The maximum of the log(It) values (denoted as E_M) coming from the $D_{M,r}$,$D_{M,g}$ and $D_{M,b}$ is used to calculate the corrected maximum densities. More formally: if f_r, f_g and f_b denote the characteristic functions and $D_{M,r}$,$D_{M,g}$ and $D_{M,b}$ the maximum densities, then E_M is defined as:

$$\max_{x\in\{r,\,g,\,b\}} f_x^{-1}\ (D_{M,\,x})$$

The corrected densities $D'_{M,x}$ are then defined as $f_x(E_M)$ (cf. the previous diagram). Analogously, the corrected minimum densities are calculated as the function-value of the minimal exposure value (E_m). Intuitively, E_m and E_M are the minimum and maximum exposure values (in the logarithmic domain) of the original scene.

We now assume that the white point of the scene corresponds to these corrected maximum densities of the negative. (Due to the fact that the exposure conditions are undetermined, this is not necessarily the case. All we can theoretically derive from the above calibration is that the corrected densities correspond to a neutral point of the original scene.)

If we do not want to apply further color corrections, we should convert the log(It) values to actual intensities. We thus obtain a lookup table to be downloaded into the scanner for the final scan. If E is the log(It) value corresponding to a given density and n the number of bits per pixel, then the associated intensity is calculated as follows:

$$I = (2^n - 1).10^{a(E_M - E)}$$

In the formula above, α denotes an additional density scaling to convert the density range of the negative to that of the original scene.

The formula indeed guarantees that a point with $D'_{M,r}$, $D'_{M,g}$ and $D'_{M,b}$ as maximum densities, should have a value $I = 2^n$ 1 i.e., the highest scan value. The factor α makes sure that the minimum densities (or at least one minimum density) correspond to an intensity value of 0.

Refining the Inversion Process

Problems

The technique described above works well under the assumption that a negative film is entirely determined by the characteristic film curve of its test strip. This is, unfortunately, not the case. It turns out that there are many factors that influence these curves and that there can be considerable differences between film batches, even of the same film type. The biggest differences are caused by the development process, which can have a profound impact on the slopes of the curves. Also considerable variations between mask densities of the same film type can occur.

Therefore, it is very difficult to come up with an automatic conversion technique based on the characteristic film curve alone. This suggests that we should modify the curves such that they better reflect the characteristics of the negative film strip under consideration.

Modifying the Characteristic Curves

A first modification assures that the mask densities of the characteristic curves match those of the film to be scanned. This can be realised by moving the characteristics curves of the test strip vertically. The target mask densities can easily be determined by scanning a piece of unexposed film.

This modification already improves the conversion, but still leaves room for improvement. The major problem is that we need more accuracy in the higher densities of the negative (since those will be mapped to the light areas, using an antilog function), whereas adjusting the mask density only fixes the problem in the lower densities (of the negative).

This problem can be overcome by asking the scan operator to indicate a white (or neutral) reference point. This point (which is expressed in original, negative densities) can then be used to make sure that the characteristic curves are adjusted such that red, green and blue report the same original log(It) value for those reference densities. This adjustment can be realised by moving the curves horizontally. It should be clear that, in addition to compensating for the non-standard characteristics of the negative film, we also compensate for the exposure conditions of the original scene. This feature can be used, e.g., to get rid of unwanted orange casts of pictures taken using a flash.

Instead of specifying a neutral point, one might also want to indicate a color point (for instance, to specify skin tones, blue skies etc.). This can be realised by adjusting the horizontal offset which is used to reposition the curves horizontally.

The Adaptive Approach

Manual selection of a reference point to indicate either a neutral point or a point with a particular color cast turns out to be quite difficult. If the scan operator has no print at his disposal, it is very difficult to indicate a reference point with associated color. Although the result of a default correction without reference point is often acceptable, experiments have shown that in at least 20% of the cases additional corrections might be required. This is not very productive and asks for a more sophisticated approach. One might argue that no or little manual corrections are required in the photo-finishing or so-called mini-labs, and those are, after all, generating positive prints out of the negative.

The approach taken in those labs is based on a combination of using statistical information on the entire film strip and using adaptive information of the original to be printed. Often, the entire film is scanned to get a rough idea on the negative film's characteristics. The next step consists of coming up with three exposure parameters, which are then used to expose the film paper. Basically, these exposure parameters determine the right mix of red, green and blue.

In order to generate the exposure parameters, the mini-labs analyze each scene separately and come up with special adjustments for, e.g., landscapes, originals with a lot of grass or snow, portraits etc. The approach taken in the photofinishing world can also be used for scanning the images electronically. Indeed, the 3 exposure parameters can be interpreted as a virtual point on the negative which should be mapped to a neutral reference point. Agfa's solution in this respect is called TFS, which stands for Total Film Scanning.[‡]

In Agfa's new FotoLook 2.5.2[§] scanning application, this adaptive approach will be taken. Based on information produced by the TFS algorithm, a virtual point of the negative will be derived which will be mapped to a neutral (positive) value. The scan operator still can modify the cast of the virtual point if he wants to.

Empirical studies have proven that changing the cast of the virtual point is needed in less than 1 percent of the originals. It is needless to say that this kind of editing is also subject to the taste of the scan operator. Some operators like *warmer* neutrals, some want *colder, bluer* neutrals.

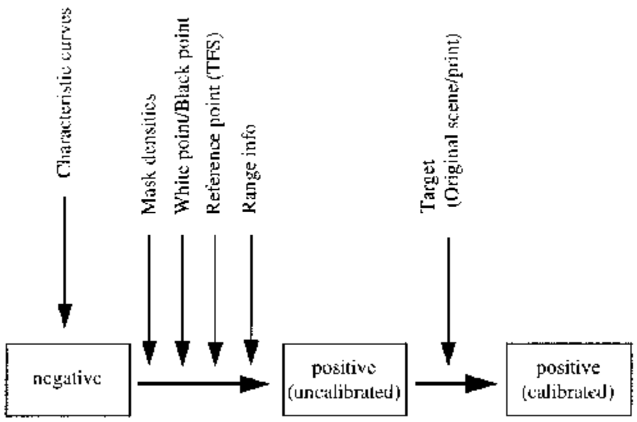

Figure 3. Overview of the negative to positive conversion

In Figure 3, we give a compact overview of the negative to positive conversion. The figure shows that the characteristic film curves of the test strip of the film type to be scanned are taken as a starting point. These default curves can be adjusted to fit the real curves by adjusting the mask densities, by applying a white point/black point setting or by setting a reference point (possibly with a color cast). Ideally, the reference point will be calculated adaptively by a TFS algorithm. The range information of the original to be converted is used to cut the right piece out of the global characteristic curve. After conversion, a color transformation can be applied to map the positive RGB signal into a calibrated RGB space. The color transformation is, as will be explained in the next section, dependent on the chosen target.

Color Characterisation

The process described above ultimately generates an image in a positive RGB space. It is obvious that this positive space is not calibrated and depends on a number of components in the process chain leading towards the end result, such as the negative film type, the scanner, the development process etc. The main idea of the techniques described above is that the result is a positive image where the neutrals have equal R, G and B.

We now want to bring the positive images to a color space which can be used as a starting point for further color management. It is not so straightforward to apply a color transform of any kind to a well-known, calibrated color space here since we do not have a reference. Different options can be taken as ideal reference space:

- **The Original Scene:** this is undoubtedly the reference space which is best defined (even though the light source is very variable: daylight, Tungsten light, artificial light etc.). A drawback of this reference space is the fact that it does not always generate pleasing colors. Often, the correct tonal mapping generates images which are too pale.
- **The Print:** the print on photographic paper is also a possibility. A definite advantage to this reference consists of the fact that the colors are quite saturated, the contrast is high enough and the overall result is quite pleasing. A serious drawback is the fact that there is no such thing as *the standard color paper* and, on the other hand, that basing the colorimetric correction on the final print is also depending on the previous steps in the chain (such as development, mini-lab processing and so on.) Another drawback of this approach is the fact that due to the very high gamma of the film paper (which causes the wanted, high contrast and accounts for the very wide, safe exposure latitude), some detail is lost in the dark parts of the original scene. In pre-press applications, this is often considered to be an unwanted side-effect of the high contrast.
- **A Combination** of the print on paper with less contrast.

Mapping the converted negative to the chosen reference space can be realised using standard mathematical techniques. Theoretically, it can be proven that applying a linear transformation in density space (i.e., a 3×3 matrix) provides a good mathematical framework to model the required color transformation, which also has been verified empirically. The coefficients of the 3×3 matrix can be obtained by using non-linear regression techniques[3,4] where the errors are measured in intensity space but the values are calculated in density space.

In order to calculate the matrix, we need target values to map to. Depending on the goal, there are two different alternatives to obtain these target values.

- **Going back to the Original Scene:** in this case, we must measure our target originals colorimetrically (using, e.g., a spectro-photometer) to obtain the CIE Lab[1] values of the different patches. These patches now provide our target. Common reference charts are the IT8.7/2[2] original and the MacBeth color checker.
- **Mapping to the Print:** in this case, we first have to a make a picture of the target original, develop it and make a print of it. Then, we must measure the patches of the target original on the print colorimetrically and use those as target values.

In Figure 4, we visualize a matrix transform in intensity space for an Agfa HDC 100 film. Since the matrix is being applied in density space, it may be considered as a generalised, multi-dimensional gamma function with cross terms.

Once the converted negative is mapped into a well-known, calibrated RGB space, the images can be transferred to any output color space using standard ICC profiles.¶

Scanner-Related Quality Issues

In this section, we briefly discuss the requirements for a good negative scanner. Most CCD scanners on the market are mainly used for positive, reflective or transmissive originals. These originals have density ranges starting at 0 D (i.e., absolute white). The density ranges for positive originals are typically quite big and for transmissive originals bigger than for reflective originals. This can be explained by the fact that positive film has a built-in gamma of about 1.8, whereas reflective originals tend to be tonally linear. Negative originals, however, have a much smaller density range but are much darker.

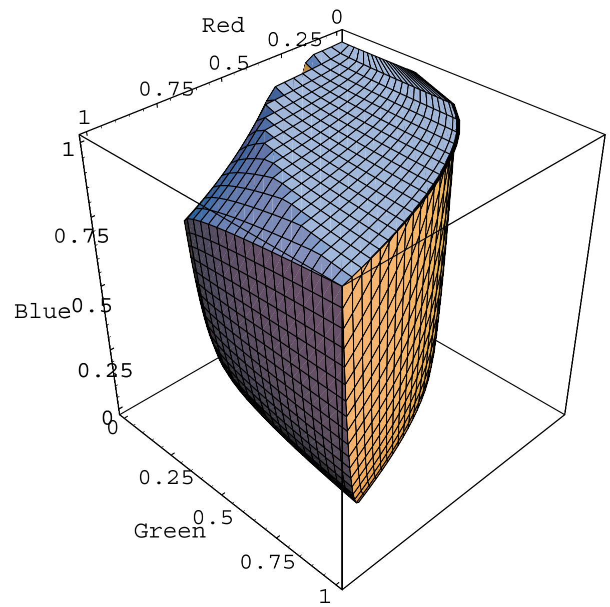

Figure 4. A 3-dimensional color transformation for negatives

The smaller density range can be seen as an advantage, but the fact that negatives are substantially darker than positives poses a number of problems. The problems are twofold: on one hand, CCD scanners tend to report more accurate information in the highlights of an original (which is a mere signal/noise issue); on the other hand, the darker areas of the original become the (important) highlights of the converted negative, and therefore emphasise all the shortcomings. A related problem also consists of potential dust particles or scratches present on the negative. These come out as light fringes on the converted negative, and are often very disturbing on a dark background. Applying image processing techniques to get rid of scratches and dust is quite difficult; therefore, it is better to clean the negative film before scanning. APS provides an even more fundamental solution, since manipulation of the negative strips is not necessary anymore.

Theoretically, one can show that a typical negative film needs very steep inversion luts in order to invert negatives to positives.[9] Under the assumption that we are carrying out a full-range scan of a negative film with gamma γ_f, the look-up table (from I to I') to be downloaded into the scanner shows the following property:

$$\frac{dI'}{dI} = -\frac{I'}{\gamma_f . I}$$

If we assume, for example, that $\gamma_f = 0.65$, the slope of the inversion lut is incredibly steep:

$$\frac{dI'}{dI} < -300$$

This explains why we either need a scanner which internally works with at least 12 bits per pixel, or a scanner which is capable of adjusting the density ranges (usually realized by increasing the line-times of the CCD's). Since the density ranges for red, green and blue differ, we even might want a scanner which is capable of adjusting density ranges on a channel per channel basis. Most of the scanners today apply one pass only (using dichroic prisms or coated CCD's), and therefore are not capable of applying different line times per channel. A better analog to digital conversion (i.e., enough bits per pixel) is even more important there.

Conclusions

In this document, we described a global approach to scanning negative originals. We proposed a technique to characterize a negative film to be scanned on a particular scanner. In order to account for the variations which exist between negative film strips of the same film type, we discussed a number of techniques to customize the characteristic film curves, such as matching the mask densities, indicating white or neutral points or indicating reference points with a specific color cast. In order to minimize the actions to be taken by the scan operator, we showed how a virtual (neutral) reference point can be calculated based on a statistical analysis of the entire film strip and an adaptive analysis of the scene to be converted. In addition to the one dimensional inversion, we also showed how multi-dimensional color corrections can be applied to map the converted negative to a well-defined, calibrated color space. In the last section, we discussed some of the requirements for an ideal negative scanner.

Problems we are still working on relate to originals which have extreme light sources (often called specular highlights). Such specular highlights can be detected after analysis of either one- or multi-dimensional histograms. Also the new emerging APS standard brings in new challenges, such as interpreting the digitally encoded data.

References

1. CIE, “Colorimetry”, *CIE PUBLICATION*, vol. **15,** 2, 2nd Edition, pp. 29-30, Commission Internationale de l'Eclairage, Vienna, 1986.
2. ANSI IT8.7/2-1993, *American National Standard*, “Graphic Technology-Color Transmission Target for Input Scanner Calibration”, NPES, 1993.
3. C. L. Lawson and R. Hanson, “Solving Least Squares Problems”, Prentice Hall, Englewood Cliffs, New Jersey, 1974.
4. D. W. Marquardt, *Journal of the Society for Industrial and Applied Mathematics*, vol. **11**, pp. 431-441, 1963.
5. O. Pytela and J. Majer, “Quantitative Interpretation of Sensitometric Curves of Photographic Materials”, *Journal of Imaging Science,* vol. **35,** 6, pp. 362-364, 1991.
6. L. Stroebel, J. Compton, I. Current and R. Zakia, “Photographic Materials and Processes”, Focal Press, Boston, 1986.
7. C. Tuijn, “Input Calibration for Negative Originals”, *Proceedings of IS&T/SPIE's Symposium on Electronic Imag-*

ing: Science and Technology, Device-Independent Color Imaging II, Volume **2414**, San Jose, 1995, pp. 76-83.

8. C. Tuijn, "The Influence of Precomponsation Curves on Multidimensional Color Modeling", *Proceedings of IS&T/SPIE's Symposium on Electronic Imaging: Science and Technology*, *Color Imaging: Device-Independent Color, Color Hard Copy, and Graphic Arts,* Volume **2658**, San Jose, 1996, pp. 326-333.
9. C. Tuijn, "Hardware and Firmware Requirements for a the Ideal Negative Scanner", Agfa-Gevaert N.V., Internal Memo, Mortsel, 1996.

[*] Agfa HDC 100 is a registered trademark of Agfa-Gevaert.

[†] Agfa SelectScan Plus is a registered trademark of Agfa-Gevaert.

[‡] TFS is a registered trademark of Agfa-Gevaert.

[§] Agfa FotoLook is a registered trademark of Agfa-Gevaert N. V.

[¶] ICC stands or the International Color Consortium, an organization which pursues the establishment of a color management standard in the electronic imaging world.