Metamorfoze Preservation Imaging Guidelines

Hans van Dormolen; Metamorfoze, Koninklijke Bibliotheek, National Library of The Netherlands; The Hague, The Netherlands

Abstract

Metamorfoze, the Dutch national program for the preservation of the paper heritage, started in 1997. The program is financed by the Ministry of Education, Culture and Science and is jointly coordinated by the Koninklijke Bibliotheek – the National library of the Netherlands – and the Dutch National Archives. The aim of the program is to preserve paper originals that are threatened by autonomous decay.

This autonomous decay is caused by acidification, ink corrosion or copper corrosion. It causes the originals to discolor, become brittle and finally to disintegrate. To protect the originals, analog or digital derivative copies are made that can be used instead of the originals. The originals are withdrawn from use and stored in climate controlled storage rooms. The derivative copies are made following the Metamorfoze Preservation Imaging Guidelines. In this paper I'll focus on the basic principle of the guidelines and the most important technical aspects.

Basic Principle

The essence, the basic principle, of the Metamorfoze guidelines is: everything that can be visually perceptible in the original must also be perceptible in the digital derivative, in the same contrast ratio. The derivative must be as good (as is technically possible within a realistic workflow) as the original. Correct exposure and correct tonal capture are essential in order to realize this goal.

To ensure that the preservation image contains all the visible information of the original a well structured and clearly based relation between original and surrogate must be ascertained.

To do so, guidelines and technical targets are used to assess this relation objectively.

In the guidelines tolerances for technical criteria are incorporated. The accepted tolerance level is established just below a visibly noticeable level. The technical knowledge obtained in the analog preservation world, like sensitometry, provides a good base for the specifications of the tolerance level. All tolerances of the criteria referred to below can be measured objectively with the use of technical targets.

General

Metamorfoze has performed much research in order to formulate these guidelines. We are also still studying how to optimize the guidelines. The techniques that can be used to make derivatives are changing fast. In order to keep the guidelines up to date it is essential to do research permanently. Hence we consider the guidelines a living document that needs to be adapted regularly. All digitization activities are outsourced to digitization companies. We advise and train these companies, explain the guidelines to them and keep them informed of the latest developments.

In the guidelines, many concepts and figures are used. Understanding the guidelines is essential in order to assess the connections between and the importance of the different concepts and figures.

Guidelines can only be effective with adequate monitoring. Therefore the technical monitoring is an important aspect in the guidelines. The correct interpretation and assessment of the fluctuations in quality can only be done with a complete understanding of the guidelines. For working with and monitoring of the guidelines, there are two types of monitoring:

- Calibration of the equipment according to the guidelines
- Monitoring and stabilizing of the production process

Calibration of the equipment

The demands for the calibration of the equipment are stricter than those for the monitoring and stabilizing of the production process. When a tolerance is exceeded in the production process, this is much less damaging than when a tolerance is exceeded during the calibration of the equipment. The quality in the production process will always slightly fluctuate. The equipment must be calibrated daily before starting the production. Or, if there is going to be a lot of production, after a certain number of exposures. After exactly how many exposures depends upon the stability of the equipment.

Calibration of the equipment is an essential but very time consuming job. In order to be able to calibrate faster we are at this moment developing software and special test charts, in cooperation with Image Engineering Dietmar Wueller. The specific technical criteria that are applied for the calibration can differ per machine. Scanners and one shot cameras each have their own character in terms of achievements and deviations. The particular demands for the calibration also depend on the types of originals and the demands for the digitisation. Therefore the tolerances for the technical criteria that are analyzed in the software are separately adjustable.

Technical criteria, such as a gray scale and a color chart, are essential for the calibration of the equipment. Besides the image analysis by software, the exposures must also be visually checked for artifacts.

Monitoring and stabilizing the production process

In order to be able to monitor as well as stabilize the production process, alongside each original at least 1 technical target has to be digitized, the Kodak Gray Scale. This is a gray scale with which the exposure and contrast transfer can be

analyzed. For originals which contain essential color information, a small color chart must be digitized as well. Besides the analysis of the test charts, visual inspection is also required. This visual inspection is mainly aimed at artifacts that cannot be discerned when analyzing the technical targets, for example artifacts such as stripes, gutter shadow or skew.

The frequency of this inspection depends on the stability of the equipment and the optical quality of the originals. For example, when manuscript material or old newspapers are digitized, correct tonal capture is more important than when purely bitonal originals are digitized, such as modern books. Consequently the frequency of the inspections when digitizing manuscripts or old newspaper will be higher than when purely bitonal originals are digitized.

Technical criteria

For the following technical aspects, tolerances are given in the guidelines, which the derivatives must comply with:

- Tonal capture
- Noise
- Signal to noise ratio
- Uniform illumination
- Colorcast
- Color accuracy
- Detail reproduction and sharpening
- Geometric distortion
- Color misregistration
- Image artifacts and other defects

In this article not all criteria will be described fully, only the aspect of tonal capture. For the complete guidelines see the Metamorfoze website (1).

The first and most important technical aspect is correct exposure. This also holds for digitizing in RAW. A stable and reliable workflow can only be built if there is consistent good exposure. Standardized developing of RAW files to a specific color space can only be done if the exposure is made correctly at all times. If the exposure complies with the tolerance, the contrast transfer in the highlights must be looked at. If these two criteria are met with, the gray scale can be analyzed further. As these first two aspects are so important the complete description of correct exposure and the highlight gamma as specified in the draft version will be described here.

Tonal capture, correct exposure and highlight gamma

The first and most important aspect in the guidelines is correct exposure. The tolerance level is limited to 1/12 F-stop underexposure or overexposure. In general, over exposure is more harmful. Clipping (pixel value 255) over the entire frame must of course be avoided at all times. Therefore, when the equipment is calibrated, no more than 1/12 F stop underexposure is tolerated. Correct exposure is monitored by measuring the pixel value of the first patch of a gray scale, patch A of the Kodak Gray Scale.

The Kodak Gray Scale ranges from near white, which is the first patch, to near black in the last patch. There are twenty patches in all. The first patch (patch A) has an optical density of 0.05, the

next patch is patch 1, with an optical density of 0.15. At the other end of the grayscale patch 19 has an optical density of 1.95. The difference in optical density between the patches (which is called steps) is consistently + 0.10 or –0.10, that is + of - ◌ stop. For assessment of the tonal reproduction, the entire range of the gray scale is measured. All twenty patches are important. However, to gain a quick insight in the quality of the tonal reproduction, our approach is to measure four patches only in the first stage of the assessment process: Patch A, 1, 10 and 19. Once these four patches, have been judged satisfactory, the other sixteen patches are assessed. The density margins (pixel values) are identical for all patches. It is the difference in density (pixel values) between patch A and patch 1 in the image, though that is of essential importance. A norm value is set for the difference between these two patches. This difference is referred to as the highlight gamma.

The term highlight gamma is used by Metamorfoze to indicate the contrast in the highlights. For quality monitoring and description of preservation microfilms commissioned by Metamorfoze, the analog gamma has to comply with explicit norms. For analog film the gamma used to be calculated with the aid of the Kodak Gray Scale at the start of each film. For digital images the contrast plays an equally important role, and therefore has to adhere to specific criteria. To calculate the highlight gamma of digital images one step is used, from patch A to patch 1, or from patch 1 to patch 2, and not, as in the case of microfilms, the first two or three steps together. In other words, the gamma is calculated and described in the 'high lights', hence the term 'highlight gamma.'

Patch A must retain its minimal density in the captured image. Because the optical density of patch A is very low, only 0.05, the exposure margins during capture should also be low. The exposure margin allowed is 1/6 stop in total. This means 1/12 stop underexposure and 1/12 stop overexposure. Expressed in terms of optical density this is: 0.025 and 0.075.

With the aid of the reflection values the pixel values can be calculated. The pixel values are dependent on the color space that is used, on the bit depth, and on the monitor's gamma setting. We assume the use of Adobe RGB 1998 color space, a monitor gamma of 2.2 and a bit depth of 8. (For calculating the pixel value we are using the formula: pixel value = 255 x (Reflection value ^ 1/ monitor gamma)). A density of 0.025 results in a pixel value of 248. A density of 0.075 results in a pixel value of 236. In the captured image the pixel value of patch A on the Kodak Gray Scale must therefore lie between pixel values 248 and 236 (see Table 1). Of course, the pixel values of patch 1 and all other patches must also remain within a margin of 1/6 stop. The pixel value of patch 1 must lie between 224 and 212.

The real or theoretical pixel values (without exposure margin!) of patch A and patch 1 are 242 and 218 respectively.

The highlight gamma is the quotient of the difference in pixel values between patch A and patch 1 measured in the captured imaged, and the real or theoretical pixel values. For example: the difference between measured pixel values in the image between patch A and patch 1 is 248 minus 220 = 28. The difference between the real or theoretical pixel values between patch A and patch 1 is 242 minus 218 = 24. The quotient of the difference in pixel values between patch A and patch 1 is 28:24 = 1.1. So, the

highlight gamma is 1.1. The highlight gamma may vary slightly. This depends on the quality of the capturing device used. For preservation imaging, the acceptable margins of the highlight gamma are set between 0.8 and 1.08.

Uniform Illumination

Illumination can only be assessed if the highlight gamma is within the set tolerances. If the highlight gamma is too low, which is often the case (Over the past months I have done much research into the achievements of scanners, digital backs and SLR camera's. My experience is that without a tonal curve correction, the highlight gamma is generally somewhere between 0.5 and 0.7), assessment of the illumination using a white sheet of paper gives the wrong impression.

To ensure uniform illumination, a frame-filling white sheet of paper with an optical density between 0.05 and 0.15 must be captured. Centered at the bottom of the image a Kodak Gray Scale must be included. To assess illumination, the pixel values are measured in at least five points (with the Photoshop eyedropper tool, 5x5 pixels). Measurements are taken in the centre and at the four corners. In the case of doubt about the evenness of the illumination more points need to be measured. The difference between two random points in the image should never exceed the norm value given below.

The difference in pixel values between the centre and the corners, and between the corners themselves, must not exceed 8 pixel values per color channel. This equals a difference in optical density of 1/10 stop, or 0.03. The maximum pixel value must not exceed 248.

Uniform illumination and mass digitisation

For illumination, a margin of 8 pixel points is rather severe. For the digitisation of unique materials this is certainly justifiable. However, for the large scale digitisation (mass digitisation) of large originals (newspapers) and mass print (books), this demand is not realistic. Therefore the tolerance boundaries for this workflow and this type of originals is stretched to 12 points. When calibrating, it is however important to aim at the most optimal exposure. Therefore we try and stick to the 8 point tolerance as much as possible.

New version of the guidelines

The version of the guidelines that was published in 2007 was a draft version. This year the first official version is expected. In this official version several criteria will be adapted and amended. The most important changes are:
- The color space Adobe RGB 1998 will be replaced by color space eciRGBv2
- The file format that the KB is going to use for long term storage will be JPEG2000 (2).
- The calculation of the signal-to-noise ratio (SNR) must be carried out as described in ISO 21550
- The description of the color accuracy will be described more detailed

eciRGBv2

With a monitor gamma of 2.2, correct tonal capture is not possible. Middle gray in the original, patch 7 on the Kodak Gray Scale (optical density 0.75), shifts to patch 6 (optical density 0.65)

in the digital derivative image. Everyone understands the importance of preserving and transferring the correct contrast ratio when we are talking about digitizing cultural heritage, such as for example the letters of Vincent van Gogh. After all, a pencil drawing by Van Gogh in a letter must look like a pencil drawing in the digital derivative image and not like a drawing made with a pen or a felt-tip.

L* is a monitor gamma which is based on the way the human eye experiences contrast. Middle gray in the original remains middle gray in the digital derivative image. Besides this great advantage, L* provides slightly more space in the darker area, patches 17, 18 and 19 on the Kodak Gray Scale. In other words, with a correct exposure and conversion of a 16 Bit RAW file to a specific 8 Bit file, using the L* monitor gamma, the blacks will not fuse so quickly. In the highlights, patches A, 1, 2 and 3, the mutual distances are slightly smaller. eciRGBv2 is a L* based color space.

Color accuracy

In order to guarantee color reliability a color chart is recorded: the Gretagmacbeth ColorChecker Rendition chart (with 24 colored patches) or the Gretagmacbeth ColorChecker SG. Next, the produced colors are compared with the reference values of these colors. The comparison is done by software. Metamorfoze uses IE-Analyzer 3.0 (Image Engineering Dietmar Wueller). The comparison takes place in CIE-L*A*B*. The color specifications for which the color is compared and judged are: luminance (Δ L), chrominance (Δ C) and hue (Δ H). For these three specifications together, Delta E (Δ E) tolerances have been formulated in the guidelines. The acceptable scope of the tolerances in color deviations depends on the importance that is attached to the color information in the original. In general, for paper library and archive materials, color information does not form an essential part of the entire information transfer. In other words, although red must be red in the digital derivatives, the color red in the derivative does not have to be identical, within the most rigorous margins, to the color red in the original.

This type of originals is referred to as 'General archive and library materials'. For originals that are considered works of art the correct color transfer does have a vital role. The acceptable margins within the color deviation, Delta E, are minimal for these scans. This type of originals is referred to as 'works of art'. When these types of originals are scanned, an ICC color profile must be made. Metamorfoze hopes to be able to provide an ICC profile on a small scale from September 2008. At this moment, research is carried out into the tolerances of the color deviations that will be introduced for the digitization of works of art. For the digitization of general archive and library materials the following tolerance applies:
- An average delta E of 12 is acceptable. The maximum delta E per color on a color chart cannot be more than 21.

These values are based on measurements of color deviations of camera systems without the use of ICC profiles. The formula that is used to measure the color deviations is formula CIE 1976. The white of the color space in which the files were supplied (Adobe RGB 1998) was used as a reference for the white point.

In general, we can say for color deviation:

- Excellent and unreservedly acceptable: color deviation in Delta E of 1 – 6.
- Acceptable, but of mediocre quality: color deviation in Delta E of 6 – 10.
- Very mediocre quality: color deviation in Delta E higher than 10.

Working according to the Guidelines Preservation Imaging Metamorfoze requires technically skilled operators. In contrast, most archives, libraries and commercial parties aim at mass production and high speed. The amount of exposures is given priority over quality. This policy allows no time for studying manuals of digital equipment, color management or Photoshop, which in its turn leads to inadequate use of expensive digital equipment and inadequate workflows. This can, of course, only result in insufficient images.

I would therefore like to stress that calibrating digital equipment according to the Guidelines Preservation Imaging Metamorfoze is a one-time adjustment which does not stand in the way of mass production and high speed. The guidelines will bring forth more insight into the technical performance of the equipment and the technical quality of the images. This facilitates the construction of a reliable and efficient workflow for mass digitization.

References

[1] Website Metamorfoze http://www.metamorfoze.nl
      Metamorfoze Preservation Imaging Guidelines
      http://www.metamorfoze.nl/publicaties/richtlijnen/richtlijnen.htm
[2] JPEG 2000
      http://www.kb.Nl/hrd/dd/dd_links_en_publicaties/links_en_publicaties_intro-en.html

Author Biography

After working for 15 years as a professional photographer Hans van Dormolen started in 1999 a career as quality manager Metamorfoze for the National Library of the Netherlands. He is responsible for the technical quality of preservation substitutes, analog and digital. He is the author of the Metamorfoze Preservation Microfilming Guidelines and co author of the Metamorfoze Preservation Imaging Guidelines.