Metamorfoze Preservation Imaging Guidelines "0ne Size fits All"

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, digital derivative copies are made that can replace the originals. In 2007 the draft version of the guidelines was presented. Now, at IS&T Archiving 2009, the first official version of the Metamorfoze guidelines will be presented. A new aspect in this latest version is that Metamorfoze provides three different guidelines for three types of originals. We also present a methodology for calibrating cameras and scanners, using software and technical targets which may simplify the workflow and also reduce time and costs. We started to use color space eciRGBv2 and JPEG2000.

General

Metamorfoze has done 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 keep on doing research. Hence we consider the guidelines a living document, which 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.

These Guidelines are input oriented and put together primarily for the digitization of paper originals such as archives, books, newspapers and magazines that measure between sizes DIN A-4 (21 cm x 29,7 cm) and DIN A-2 (42 cm x 59,4 cm). Guidelines for originals smaller than A-4 and bigger than A-2 are still under construction.

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.

For originals in which color, and more particularly, color accuracy, plays an essential role, such as works of art, we have even accentuated our basic principles: all information that is perceptible in the original must also be perceptible in the digital derivates in, as far as is technically reasonable, the same contrast ratio and, again as far is technically reasonable, the same color as the original.

To ensure that the preservation image contains all the visible information of the original, in other words: to ensure correct exposure and correct tonal capture, a well structured and clear 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 as being 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.

Three types of originals

By dividing the materials into different categories digitizing according to the Guidelines Preservation Imaging becomes more realistic and also easier. Thus, digitization according to the guidelines can be applied on a wider scale. All tolerances for technical criteria cited in the guidelines are listed for each category separately:

1. Originals that can be considered works of art.
2. Unique and general library and archival materials.
3. Non-unique and general library and archival materials.

1. Originals that can be considered works of art. These may be letters with drawings, such as the correspondence of Vincent van Gogh, but also photo collections, special prints or paintings. An important aspect for this type of originals is that the color accuracy of the digital images plays an important role. For these kind of originals it is important that the tonal capture up to the darker areas (up to optical density 1.95 at least, that is L* value 9.97) is performed according to the guidelines.

2. Unique and general library and archival materials. This may be newspapers, books, periodicals and manuscripts with a maximum optical density that does not exceed 1.55 (L* value 19,31). This therefore excludes collections containing photographs. Color accuracy is a less important aspect for this type of originals. For these kinds of originals it is acceptable that the tonal capture is performed according to the guidelines up to density 1.55 (L* value 19,31). The paper heritage that are digitized by Metamorfoze is mostly quite old. The maximum density (D-max) of this type of originals is low. Tests with various scanning systems and one shot systems have repeatedly shown that digitizing of the darker areas (L* values >14,28) according to the guidelines is difficult. This division in categories makes the digitization of the different types of original mentioned above easier and more realistic. When digitizing originals without darker areas there is after all no need to calibrate the scanning or one shot systems in

order to be able to digitize darker areas according to the guidelines.

3. Non-unique and general library or archival materials. These may be books with a maximum optical density that does not exceed 1.50 (L* value 20). An important difference with group 2 is the use of technical targets and the application of technical inspection in the entire work flow. For the calibration of the machines as well as with each exposure technical targets must be used. For the digitization of non-unique library and archive material the use of technical targets for each exposure is optional. Here the technical inspection is aimed particularly at the calibration of the machines.

For the digitization of a collection consisting of various types of materials, for example letters and photographs, it is important to choose a scanner or one shot system that can also digitize the darker areas according to the guidelines. This also holds for the digitization of collections in which color plays an important role, and has to be accurate. Here too a scanning system of one shot system must be chosen that can digitize the darker areas according to the guidelines.

Calibration of the equipment

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

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 with a gray scale. With the gray scale correct tonal capture 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.

Color space eciRGBv2

Color space eciRGBv2 is L* based. This means that the luminance levels, the contrast is built 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, this L* based color space provides slightly more space in the darker area, patches 17, 18 and 19 on the Kodak Gray Scale (densities 1.75, 1.85 and 1.95). In other words, with a correct exposure and conversion of a 16 Bit RAW file to a specific 8 Bit file, using L* the blacks will not fuse so quickly. In the highlights, patches A, 1, 2 and 3 (densities 0.05, 0,15, 0,25 and 0.35), the mutual distances are slightly smaller. In the new guidelines, we have therefore started using this color space.

Technical criteria

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

- Correct tonal capture (includes correct exposure and correct high light gamma)
- Noise
- Signal to noise ratio
- Uniform illumination
- Colorcast
- Color accuracy
- Detail reproduction, sampling efficiency and sharpening
- Geometric distortion
- Color misregistration
- Image artifacts and other defects

In this article not all criteria will be described fully, only the aspects 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.

Correct exposure

The first and most important aspect in the guidelines is correct exposure. Correct exposure is monitored by measuring the pixel value (RGB value or count value) in the high light area of the first patch of a gray scale, for example patch A of the Kodak Gray Scale Q-13 or Q-14. Patch A on the Kodak Gray Scale Q-13 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 1/12 F-stop over exposure and 1/12 F-stop underexposure. 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 (RGB values or count values) are dependent on the color space that is used, the gamma and on the bit depth. To calculate the pixel value in color space eciRGBv2 and 8 bit the following formulas are used:

$$L^* = 116 \sqrt[3]{R} - 16$$

R = Reflection value

Pixel value = 2, 55 x L*

Optical density Patch A	Reflection value	L* value	Pixel values in eciRGBv2
+ 1/12e F-stop, 0,025	0,94	97,80	249
0.05	0,89	95,63	244
- 1/12e F-stop, 0,075	0,84	93,51	238

To calculate the pixel value in color space Adobe RGB 1998 and 8 bit the following formula is used:

Pixel value = 255 (R^1/2, 2)

Optical density Patch A	Reflection value	L* value	Pixel values in Adobe RGB 1998
+ 1/12e F-stop, 0,025	0,94	97,80	248
0.05	0,89	95,63	242
- 1/12e F-stop, 0,075	0,84	93,51	236

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. This means in color space eciRGBv2 a limited pixel value of Patch A of 244 to 238.

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 - ⅓ 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.

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.'

The calculated or theoretical pixel values (without exposure margin!) of patch A and patch 1 are 244 and 223 respectively.

Theoretical value

Patch	Density	Pixel value, eciRGBv2
A	0.05	244
1	0.15	223

The difference between Patch A and 1 is 21

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:

Pixel value in the image

Patch	Density	Pixel value, eciRGBv2
A	0.05	248
1	0.15	235

The difference between Patch A and 1 is 13

Highlight gamma is 13:21 = 0, 61

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.

The highlight gamma, therefore, in the example is too low. Over the past years 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. This leads to loss of information in the high light area. This becomes evident in highly inadequate and inaccurate color representation in the high lights. In some cases a too low high light gamma will result in the loss of textual information.

Overexposing Patch A and underexposing Patch 1 in the same image will cause a not acceptable highlight gamma and is therefore not allowed.

Pixel value Patch A	Pixel value Patch 1	Highlight gamma	
249 (+1/12 F)	218 (-1/12 F)	249–218=31, 31:21=1,47	Fail
249 (+1/12 F)	228 (+1/12 F)	249-228=21 21:21=1	Pass
238 (-1/12 F)	228 (+1/12 F)	238-228=10 10:21=0,47	Fail
238 (-1/12 F)	218 (-1/12 F)	238-218=20 20:21=0,95	Pass

The highlight gamma is meant for use in the highlight area only. Flaws in correct exposure are most harmful in the highlights.

Overview of the mentioned technical criteria ordered by group

	Art work	Unique library and archival material	Not unique library and archival material
Color space	eciRGBv2	eciRGBv2	eciRGBv2
Bit depth	16/8 bit	8 bit	8 bit
Correct exposure, Patch A (reflectance 0,89)	249 - 238	249 - 238	249 - 238
Correct exposure Patch 1 (Reflectance 0,70)	228 - 218	228 - 218	228 - 218

Highlight gamma	0,8 -1,08	0,8 -1,08	0,8 -1,08
Analyze correct tonal capture up to	Density1,95 or 2,20 L* 9,97 or L* 5,44	Density1,55 L* 19,31	Density1,55 L* 19,31
Color accuracy	Mean ΔE < 4 Max ΔE < 12	Mean ΔE < 12 Max ΔE < 23	Mean ΔE < 12 Max ΔE < 23

Count values are calculated in 8 bit. Color accuracy is calculated with formula Cie 1976 and a X-rite Colorchecker SG.

Improvement work flow Metamorfoze

A too low highlight gamma must be corrected when a 16 Bit RAW file is converted to a certain color space. This can be done by making a correction curve. In most cases when making an icc profile, using for example Profile Maker 5.0, the just made correction curve will be ruined. Almost always a second correction curve needs to be made in such cases. The first time before an icc profile is made, and the second time after the icc profile has been made and is effective.

Making a correction curve, an input icc profile and subsequently another correction curve, and measuring and analyzing is a complex and very time consuming job. With various manufacturers (amongst them Zeutschel, 12S, Leaf and Hasselblad) and digitization companies Metamorfoze is researching the possibility of working more efficiently. Metamorfoze also collaborates with Scott Geffert (Center of Digital Imaging) and the Van Gogh Museum Amsterdam. Together with Scott Geffert and Hasselblad Bureau Metamorfoze is testing new software and a work flow based on the updated version of the Metamorfoze guidelines which include color space eciRGBv2.

In order to establish whether a scanning system or one shot system has been calibrated according to the Metamorfoze Guidelines four test scans have to be made now with different test targets. Subsequently these test targets must be analyzed. This is a time consuming job which needs to be performed regularly. According to the guidelines these scans and analyses must be done for each shift or after every 500 exposures. In order to work more efficiently, Bureau Metamorfoze is developing a new technical target in cooperation with Dietmar Wueller. The starting point for this target is that all technical criteria that are referred to in the Guidelines must be able to undergo analysis. Apart from this, the existing IE analyzer software is converted to be able to analyze this target in accordance with the Metamorfoze Guidelines (and therefore following various tolerances).

The German organization MFI (Fachverband für Multimediale Informationsverarbeitung e.V) has approached Metamorfoze in order to be part of the design of the new target and to make it applicable for a wider use. Metamorfoze has expressed its interest in this cooperation and hopes that together with Dietmar Wueller and MFI it will come to a good and solid target.

Using the new target and the software will save much time in the daily fork flow. Particularly so since for the technical analysis

we now use three different software kits. After technical analysis the software will deliver a go or no go. Obviously, scanning can only start after all criteria have been met with. All tolerances of the technical criteria that are described in the guidelines can be adjusted freely. Thus every possible preferred tolerance can be measured with this software. All data obtained by analysis and also the go or no go signal will be stored in the metadata of the image. The target will be developed for sizes DIN A-4, A-3 and A2 and will probably become available up to A0-format. In this large target a smaller target will be incorporated that must also be scanned together with the originals. The target will be presented during IS&T Archival 2009. The Beta version of the software will be demonstrated here as well.This software can also be used in digitisation projects meant for access rather than preservation. As an example, the guidelines show all specifications that an access digitisation project done by the KB, using bitonal books as source materials, has to comply with.

Thus, the guidelines, the software and the target can be applied to any type of digitisation project: from bulk feed through scanning to the letters of Vincent van Gogh.

Thanks to

Dennis Schouten, Robert Gillesse, Henriette Reerink, Bureau Metamorfoze, Ronnie Mampaey, Joop Korswagen, Don Williams, Peter Burns, Dietmar Weuller, Scott Geffert, Daniel Johnston, Volker Jansen, Cedric Muscat, Martin van der Veen and all the companies and institutions I've worked with.

References

[1]   Bureau Metamorfoze http://www.metamorfoze.nl
[2]   Koninklijke Bibliotheek, Den Haag http://www.kb.nl/
[3]   European Color Initiative, www.eci.org
[4]   Image engineering http://digitalkamera.image-engineering.de/
[5]   Image Science Associates http://imagescienceassociates.com/
[6]   Center for digital Imaging, www.cdiny.com
[7]   Fachverband für Multimediale Informationsverarbeitung e.V. http://www.fmi-ev.de/
[8]   Daniel L. Johnston, *A Simplified Standard Method of Digital Image Tonal Capture for Archival Projects*, IS&T's 2002 PICS
[9]   Scott Geffert, *Adopting ISO Standards for Museum Imaging,* 1/2008. Center for Digital Imaging, inc
[10] Roy S. Berns and Franziska S. Frey, Rochester Institute of Technology, *Direct Digital Capture of Cultural Heritage – Benchmarking American Museum Practices and Defining Future Needs – Final report 2005.*

Author Biography

After working for 15 years as a professional photographer he started in 1999 a career as imaging specialist 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. Hans is a member of IS&T and nominated for an IS&T award.