ISCAS 2000 - IEEE International Symposium on Circuits and Systems, May 28-31, 2000, Geneva, Switzerland

Digital Photography-How Long Will It Last?

*F. Frey, S. Süsstrunk**

Image Permanence Institute, Rochester Institute of Technology
70 Lomb Memorial Drive, Rochester, NY, 14623, USA
*Laboratory for Audiovisual Communication, Communication Systems Department
Swiss Federal Institute of Technology (EPFL), CH-1015 Lausanne, Switzerland

ABSTRACT

Permanence issues for digital photographs arise in three different areas. First, the materials used for digital hard copies should preferably be as permanent as the conventional photographic materials. Longevity of digital hard copy materials is affected by the stability of the material and the storage conditions. Secondly, the digital files that are the counterpart of the conventional negatives need to be readable not only on various systems and platforms today but also in the future. Third, the encoding of digital images should be such that any improvement in processing algorithms and display/output technologies can be applied in future image workflows. The safe keeping of digital data requires an active and regular maintenance of the data. This paper will discuss encoding issues for archival images and strategies to make image preservation happen for digital photographs.

1. INTRODUCTION

The life expectancy of film-based and digital information media is governed by technologically specific standards resulting from extensive testing. Photographic permanence is affected by the stability of the material, the photographic processing, and the storage conditions. Various standards have been published outlining the requirements for each of these areas. Standardization for digital image permanence is more complex. Existing "industry standards" tend to be distillations of vendor responses to a competitive marketplace. As such, preservation is rarely a priority. The same holds for software and hardware.

When developing solutions for the long-term storage of digital photographs, it is important to keep in mind how the users handle their conventional photographs. In most cases, images are not being organized at the moment the pictures were taken. Rather, photographs are put into a shoebox and put on a shelf. When ordering one's memories becomes an important task, usually years later, the photographs are looked at again, sorted and put in albums. As computer files are not human readable, this approach does not work in a digital world: even if the storage media survive, the files written onto them will not be readable anymore 20 years from now. Therefore, new approaches to storing and sorting digital pictures have to be developed.

2. LONGEVITY OF HARD COPY MATERIALS

Various technologies have been developed to print hard copies of digital photographs, including exposure onto conventional photographic materials. However, inkjet imaging is becoming the predominant hard copy output, due to its relative low device and media costs. Speed and image stability are the primary weaknesses of inkjet technology. For photographic applications, it is important that these types of images have at least a permanence similar to conventional photographic materials, which is the current benchmark for stability among image-makers.

In traditional photography, accelerated fading tests have been developed to compare and predict life expectancy of both the image and the support material. For many years the stability of color photographic products was tested differently by various laboratories. The need for standard testing procedures was evident. The resulting ANSI test procedures include methods to determine the stability of color images under dark storage conditions and under a variety of light exposure conditions. [1] Work for similar tests for digital hard copy materials (ink jet, electrophotography, and thermal dye transfer) are now under way. Some of these materials present new and difficult challenges. For example, fingerprints, dye re-transfer to contacting surfaces, and water-fastness is of concern for some materials.

New output technologies, new ink sets for existing devices, and new media are coming to the market daily. With all these new products there is a need to reassure the consumer that they will have a usable life adequate to the chosen application.

3. LONGEVITY OF DIGITAL FILES

3.1 Software and Hardware Issues

Archival issues have never been a prime concern in the computing world. Longevity in the computer industry is measured in months. The rapid pace of development in the computer industry has led to extremely short product cycles. Generations are being counted in months and not years or decades. New versions of software products have a lifetime of at most a year, and computer hardware seems to become outdated in the time between ordering and receiving the hardware.

Hardware compatibility

The longevity of digital recording media is only part of the story. While the coded information may be physically readable for a long time, the device to read it has to exist and the formatting code has to be known. Because of the rapid development in storage technology, the reader/writer of a certain type of digital storage medium is likely to be produced only during a short period. Then, a new enhanced device will be put on the market. For all known media types, about every two years a new

0-7803-5482-6/99/$10.00 ©2000 IEEE

generation can be expected. Generally, a new generation is able to read and write to media of the previous generation, and read only the media of the generation before this. All earlier generations are incompatible. As a result, media has a useful lifetime of about 5 years. After that, there usually exist no more devices to read the media, and/or support for these older devices is no more available.

Software compatibility

Even if the digital data can be transferred to the computer, there is still the problem that the data has to be interpreted to be useful. Thus the longevity of digital data is also determined by the capability of software to read "old" formats. New software releases often change the data storage format. Additionally, metadata has to be stored with digital images in order to interpret the data. The digital representation of an image makes only sense if metadata like the image size in pixels and the meaning of each number is known. By convenience, this metadata is often stored together with the actual image data. Depending on the header structure of the file format, software applications might not be able to interpret that data correctly.

Digital Storage

All digital data is recorded in form of binary numbers. However, plain binary representation is very rarely used. Rather error correction and data compression are applied, often in combination.

The misinterpretation of an on/off switch always leads to a significant error in the interpretation of data. In order to cope with this inherent error, redundant information is added to the plain digital data. A simple form is the parity bit. This method has been replaced by very elaborate error correction codes like the cyclic redundancy check (CRC). CRC not only allows detection of errors but also allows correcting them, if not too many bits within a group have been misinterpreted. The principle of all error correction schemes is to add redundant information.

Digital information usually contains a certain degree of semantic or syntactical redundancy. In order to achieve an efficient storage as binary data, data compression algorithms are applied. However, compression is mainly an issue for transferring data over networks. Compression in an archival environment has to be evaluated very carefully. Although the user might not have a choice, since pictures from digital cameras are often already compressed in the hardware to allow for faster download. [2]

There are two fundamentally different approaches to data compression. Lossless compression tries to remove redundant information by the detection of patterns in the data and then using optimal coding to represent these patterns. Lossless compression schemes are usually not efficient enough to warrant the time used to compress and decompress the data. In contrary, lossy compression schemes irreversibly removes information that the algorithm determines is not used for the correct interpretation of the data. They are based on psychovisual criteria which in turn are based on output representations. If the output size and device is not known, as is usually the case in an archive environment, lossy compression is not recommended.

Error detection and data compression are very often used in combination. On a hardware level, digital storage devices use error correction in order to guarantee a certain level of data quality. [3] This error correction is performed automatically without the user being aware of it. However, the internal error correction rates (and their development with time) may be an interesting estimate of the quality of a storage medium.

3.2 Digital Media

Much has been written about the stability of digital storage media. [4] As mentioned above, media stability is just one factor in determining permanence. A migration plan that accounts for these issues is ultimately the only way to ensure that data will survive. Another important factor is the hardware/media combination for writing data. Often hardware is optimized for media from a certain manufacturer.

The critical physical properties of magnetic tapes are binder cohesion, binder-base adhesion, friction, clogging of magnetic heads, dropouts, and binder hydrolysis. Magnetic properties of interest are coercity and remanence. However, the consumer is currently left without a recognized specification to compare tape products and the manufacturers without a standardized procedure to evaluate tape life. The only option for the user is to purchase tape with recognizable brand names. [5]

It is a well-accepted fact that good storage conditions will prolong the life of all tapes. The accepted recommendations have been incorporated into an ANSI document, published in 1998, on the storage of polyester base magnetic tape [6]. The lowest recommended temperature is 8°C because subfreezing storage can create problems with the base material. Another storage concern is the avoidance of external magnetic fields. Topics for care and handling of magnetic tapes include cleaning, transportation, use environment, disaster procedures, inspection, and staff training.

Unlike magnetic media, optical disks are manufactured not only in a variety of sizes, are also composed of very different materials. The most common substrates are either polycarbonate or glass. Since they can operate by several different mechanisms, the image-recording layer features various organic or inorganic coatings For example, write-once disks can record information by ablation of either a thin metallic layer or a dye/polymer coating, by phase change, by metal coalescence, or by change in the surface texture. Read-only disks have the surface modulated by molding of the polycarbonate substrate. Erasable disks are based on magneto-optical or phase-change properties. Despite this vast dissimilarity in composition, optical disks have an important advantage over magnetic materials with respect to predicting their life expectancy. Optical disks are recorded and read by light and do not come into contact with moving or stationary parts of equipment. Therefore their useful life is mainly determined by the properties of the material itself and, unlike magnetic tape, physical ware and tear is less of an issue. This has resulted in the use of the Arrhenius method to predict the longevity of disks by several investigative laboratories [7].

Optical disks can fail because of a number of different factors, such as the relaxation of the substrate causing warping, changes in the reflecting layer by corrosion, cracking, or pinholes; changes in the reflection of any dye layers by light, pressure of crystallization, or breakdown of the disk laminate by adhesion failure and layer separation. Of particular interest to the consumer is how long optical disks will last. Various tests reported a very wide range form 5 to over 100 years depending on the product.

As with all materials, improvements in storage conditions will result in improved life. [8] This is regardless of the inherent stability of the material. The standards document also covers magnetic fields, enclosure, labeling, housing, storage rooms, and acclimatization. Particular care should be given to maintaining a low dust and dirt environment. Another important consideration for optical disks is the avoidance of large temperature and humidity variations. Protection from light is vital for many writeable CDs.

It is also important to keep a number of back-ups in different places if images are to survive. While this seems like a common sense issue, experience has shown that users, but also professional archives often neglect to back-up their data properly.

4. LONGEVITY OF ENCODING

4.1 Spatial Resolution

There is no magic formula to determine how much resolution is enough, it all depends on the application, input and output devices. In case of digital photography, the spatial resolution of digital images being archived is usually determined by the capture device. The majority of semi-professional and consumer cameras are still in the 1-2 million pixel range. This resolution is adequate for current displays and hardcopy output technologies of up to 8 by 10 inch. The 35 mm film format for these applications has an equivalent information content of 18 to 37 million pixels, depending on the film, camera, and shooting conditions. [9] These current spatial resolutions are still a limiting factor for high quality hardcopy reproduction, as well as for future display devices that will have a resolution of up to 200 dots per inch.

Scanners or scan back cameras that use linear array charge-coupled devices of 8K pixel can capture at much higher resolution, sometimes exceeding the resolution of the photographic material scanned. In that case, the spatial resolution should be chosen to resolve the smallest detail, which can be the smallest image detail or the photographic grain, depending on the application. While high spatial resolution images and therefore large files are not that costly anymore to archive, there is a performance hit of writing, reading, processing and delivering these files.

4.2 Image Representation

Sensor Space Images and Unrendered Images

When a scene or original is captured, either by a scanner or by a digital camera, its first representation is device and scene specific, defined by illumination, sensor, and filters. In the case of scanners, the illumination should be constant for each image. With digital cameras, the illumination can vary from scene to scene, and even within a scene. When images are archived in sensor space, camera or scanner characterization data, such as device spectral sensitivities, illumination, system MTF and linearization data (opto-electronic conversion function OECF) have to be maintained so that further color and image processing is possible. [10]

Recent findings indicate that there is still a lot of research to be done to define color reproduction models for complex images. [11][12][13][14] Consequently, a high-quality digital master image should be archived in sensor space. Any further transformations are dependent on current engineering practices and knowledge, which might be improved in the future. If this is not possible, an unrendered image space should be chosen to contain a colorimetric estimate of the original.

The purpose of an unrendered image is to represent an estimate of the scene's or the original's colorimetry in a standard color space. An unrendered space maintains the relative dynamic range and gamut of the scene or original. The advantage of unrendered image spaces, especially if the images are encoded in higher bit-depth, is that they can always be tone and color processed for all kinds of different rendering intents and output devices at a later date. The quality of the colorimetric estimate depends on the ability to choose the correct scene adopted white point, and the correct transformations. There are several color spaces that can accommodate unrendered data: CIEXYZ, CIELAB, Photo YCC, ISO RGB, and RIMM RGB.

Unrendered images will need to go through additional transforms to make them viewable or printable. Appearance modeling can be applied when an equivalent or corresponding reproduction is desired, and the output medium supports the dynamic range and gamut of the original. In most applications, the goal is to create a preferred reproduction, meaning the image is optimized to look good on a specific medium with a different dynamic range and gamut than the original. In that case, a digital photography reproduction model is applied.

Rendered Images

Rendered image spaces are color spaces based on the colorimetry of real or virtual output characteristics. The complexity of these transforms varies: they can range from a simple video-based approach to complicated image dependent algorithms. The transforms are usually image specific and non-reversible, as some information of the original scene encoding is discarded or compressed to fit the dynamic range and gamut of the output. For example, an image that has been pictorially rendered for preferred reproduction cannot be re-transformed into a colorimetric reproduction of the original without knowledge of the rendering transform used.

There are several color spaces that can be used to encode rendered images in. The choice is usually device and/or application dependent. From an archival point of view, the color space needs to be well defined to allow for future output rendering. The current method is to include an ICC (International Color Consortium) profile with the image that contains gamut mapping information from the sensor color space to the profile connection space. [15] There are still some inconsistencies in the ICC profile specification, and backward compatibility is not guaranteed. From an archival point of view, it is therefore preferable to encode images in a well defined color space such as sRGB, CIELAB, or ROMM RGB and create profiles "on-the-fly" when needed. [16]

4.3 Encoding

Linear encoding (in intensity) is acceptable when high-bit depth information can be retained, and file size doesn't matter. In most cases, a nonlinear, perceptually compact encoding (nonlinear in intensity, but linear in lightness or brightness) is preferable, since the visual artifacts due to image processing would be equally visible across the tone scale.

Depending on the color space used to archive the image in, 8-bit encoding might not be enough. Banding effects can appear, depending on image color distribution, editing, and/or color space conversion, especially if the color space is large or unlimited. However, 16 bit/component RGB is not widely supported yet in either applications or file formats. Images archived in sensor or unrendered representations will go through extensive image processing and color space conversions and should be encoded in higher bit-depth. Rendered images archived in large gamut spaces should also be encoded in higher than 8-bit per channel. If that is not possible, the size of the gamut should be reduced.

4.4 File Format

Ideally, archived images should be saved in a standard file format whose source code is readily available. For sensor space and unrendered image representations, TIFF/EP, a developing ISO international standard, has predefined tags to store all sensor information necessary to further process the image. [17] It also allows for 8 or 16 bit/channel encoding. Some other well known file formats such as PDF, JPEG, PICT, EPS, and PNG have ways to associate profiles to define the color spaces, usually RGB, but do not define other system parameters for future processing mentioned before. Kodak's PhotoCD format stores color information in luminance-chrominance form (YCC). FlashPix, MNG (Multi-image Network Graphics) and HTML 4.0 support only the sRGB space—RGB color data in HTML 4.0 is defined in sRGB color space, which is effectively the default color space. These file formats are only recommended when storing rendered images.

5. DIGITAL ARCHIVING

The principles of secure preservation for digital data is fundamentally different from those for traditional analogue data. First, while in traditional preservation there is a more or less slow decay of image quality, the digital image can either be read accurately or cannot be read at all. Secondly, every analogue duplication process results in a slight deterioration of the quality of the copy. The duplication of the digital image data is possible without any loss at all. In a traditional image archive the images should be stored under optimal climatic conditions and ideally never be touched again. As a consequence, access to the images is severely hindered while the decay is only slowed down. A digital archive has to follow fundamentally different strategy. The safe keeping of digital information requires an active and regular maintenance of the data. The data have to be copied to new media before they become unreadable. Since information technology is evolving rapidly, the lifetime of both software and hardware formats is generally less than the lifetime of the recording media.

6. REFERENCES

[1] Wilhelm H. and Brower C., *The Permanence and Care of Color Photographs: Traditional and Digital Color Prints, Color Negatives, Slides, and Motion Pictures*, Preservation Publishing Company, Grinell, Iowa, 1993.

[2] R. Dale, "Lossy or Lossless? File Compression Strategies Discussion at ALA," RLG DigiNews, 3(1), <info.wgbh.org/upf/www.rlg.org/preserv/diginews/diginews3-1.html>.

[3] Rosenthaler L. and Gschwind R., "Long Term Preservation and Computers—a Contradiction?" *Proceedings ICPS International Congress on Imaging Science*, Antwerp, Belgium, September 7-11, 1998.

[4] Adelstein P., and Frey F. "New Developments on Standards for the Permanence of Electronic Media," Proceedings IS&T's 1st PICS Conference, Portland, OR, May 1998, pp. 127-132.

[5] Van Bogart, J.W.: "Magnetic Tape Storage and Handling", Report by The Commission of Preservation and Access, 1785 Massachusetts Ave. N.W., Suite 313, Washington, DC 20036, June 1995.

[6] ANSI/PIMA IT9.23-1998, Polyester Base Magnetic Tape—Storage Practices.

[7] ANSI/NAPM IT9.21-1996, Life Expectancy of Compact Discs (CD-ROM)—Method for Estimating, Based on Effects of Temperature and Relative Humidity.

[8] ANSI/PIMA IT9.25-1998, Optical Disc Media—Storage.

[9] Holm, J. "The Evaluation of Digital Photography Systems," Short Course Notes, IS&T 49th Annual Conference, May 1996.

[10] Süsstrunk, S. and Holm, J. "Camera Data Sheet for Pictorial Electronic Still Cameras." SPIE Vol. 2416, pp. 5-16, 1995.

[11] Holm, J. "Issues Relating to the Transformation of Sensor Data into Standard Color Spaces," Proceedings IS&T/SID 5th Color Imaging Conference, pp. 290-295, 1997.

[12] Brill, M.H. "Color Management: New Roles for Color Transforms," Proceedings IS&T/SID 5th Color Imaging Conference, pp. 78-82, 1997.

[13] Kohler, T. "ICC Achievements and Challenges," Proceedings International Colour Management Forum, University of Derby, pp. 1-6, 1999.

[14] Holm, J. "Integrating New Color Image Processing Techniques with Color Management," Proceedings IS&T/SID 7th Color Imaging Conference, pp. 127-134 November 1999.

[15] International Color Consortium Specification ICC.1:1998-09, <www.color.org>.

[16] Süsstrunk, S. Buckley, R. and Swen, S. "Standard RGB Color Spaces," Proceedings IS&T/SID 7th Color Imaging Conference, pp. 127-136, 1999.

[17] ISO 12234-2 DIS Photography - Electronic still picture cameras – Removable memory - Part 2: Image data format - TIFF/EP. <www.pima.net/standards/it10a.htm>.