Long-Term Storage of Digital Data on Cinematographic Film

***Christoph Voges*[1], *Jan Fröhlich*[2]**
[1] ***Technische Universität Braunschweig, Institute for Communications Technology (IfN); Braunschweig, Germany***
[2] ***CinePostproduction GmbH; Berlin, Germany***

Abstract

Cinematographic film is employed for cinema and television productions all over the world. The corresponding exposure and scanning hardware for this kind of film is widely available and installed. Accordingly, it is reasonable to employ this established film material also for 'Bits on Film', i.e., the storage of digital data that has gained a significant interest within the past few years. This promising approach is particularly attractive for archiving of digital film productions and is currently investigated within a research project being described in this paper.

Introduction

Long-term storage of digital data on film has been a frequently-discussed topic in the past few years. As opposed to conventional storage media, such as optical discs, magnetic tapes, or hard drives with relatively limited lifetimes (cf., e.g., [1]), certain films feature an excellent long-term stability (cf., e.g., [2]). A further issue concerning archiving applications is the availability of reading devices in the future. For digital data storage on film it can be reasonably assumed that it is possible to construct reading devices in the future with acceptable effort. As an outstanding feature, film allows hybrid storage of analog and digital data. This can be interesting for certain information, such as texts or images that can be stored both as a digital as well as an analog (i.e., as a photographic image) representation on the same medium. Even a human-readable description of the decoding instructions and parameters can be added to the actual digital data on the film. After the chemical development process that is required for most films, the photographic information cannot be changed. The result is a storage concept that is proof against forgery and manipulation.

The fundamental idea to use photographic film for storage of digital data is not new (cf., e.g., [3]). Today, audio information is digitally stored on cinema film for the Sony SDDS®and the Dolby® Digital systems [4]. A few years ago, a new interest awakened to employ film as a medium for digital long-term archiving applications (see, e.g., [5, 6]). Detailed introductions to this technology that is generally referred to as 'Bits on Film' can be found in [7–9]. In research projects, such as MILLENIUM [10, 11], PEVIAR [12], and DANOK [13] there have already been several contributions to this novel storage concept. Investigated topics include signal and information processing [10], high-density laser exposure [11], channel modeling [14, 15], as well as error correction [7, 16].

Basically, 'Bits on Film' approaches involve digital film exposure devices to record the digital information as data patterns on the film. A major focus has been on microfilm-based approaches in the past (e.g., [10, 13]). Digital film recorders are also widely used to expose cinematographic film after the digital postproduction.[1] Originally these devices have been developed to expose digital image information as analog images on film. Since film recorders (and also film scanners) for cinematographic film are installed all over the world, it is reasonable to use this equipment also for storage of digital information. This approach offers interesting alternatives specifically for archiving of digital film productions. Currently, a research project is conducted by the Technische Universität Braunschweig and the CinePostproduction GmbH in this context. The aim of this contribution is to introduce this project and to present the underlying ideas and technical principles.

This paper is organized as follows: The next section provides a detailed description of the situation for digital production of motion pictures followed by a section describing the specific aims of the project. Subsequently, technical details are described and the paper ends with a set of conclusions.

Digital Production of Motion Pictures

The traditional analog workflow for production of motion pictures consists of shooting the actual scenes, postproduction (i.e., editing, color matching, visual effects), as well as mass duplicating the prints. In Europe, for mainstream productions a fundamental reorientation from analog to digital postproduction took place from 2006 to 2009 and nowadays virtually all motion pictures are postproduced digitally. Similar trends could be observed in many countries all over the world.

The required image data for this digital workflow can be either directly originating from digital cameras or alternatively from scanning an analog original camera negative. Different media are involved for storing the final result of a digital postproduction workflow:

The uncompressed data (about 2 TByte of data for a film in 2K resolution) is normally stored on a set of LTO (Linear Tape Open) tapes either by the production company or the distributor. This image data is intended for direct printing of a negative film, e.g., by means of a laser film recorder. It is normally stored as printing densities (i.e., not linear RGB values) using the DPX (Digital Picture eXchange) file format. The recorded negative film (also referred to as the DUP negative) serves as a basis for mass duplication. Since the DUP negative is suspect to a certain amount of wear during the duplication process (e.g., scratches), it can only be used for a limited number of prints (normally around 500 to 1000). These prints can subsequently be used for projection by means of conventional film projectors. On the other hand, for projection in digital cinemas, the so-called DCP (Digital Cinema Package) is used. The DCP employs visually lossless JPEG 2000 compression with a clearly de-

[1] Due to the perforation of cinematographic film (microfilms are normally unperforated), these devices cannot be used for microfilm (and vice versa). Also, their resolution is usually lower according to the requirements for film productions.

fined color space (see [17] for more details). For a film that is post-produced in 2K resolution, the DCP has a size of about 100 GByte to 300 GByte. Accordingly, the DCP is very attractive for long-term archiving.

In most European countries, film producers currently have to deposit one print at the particular national film archive. Clearly, due to the printing and duplication processes, the resulting quality is not optimal. Furthermore, standard print film is not intended for long-term-archival (cf., e.g., [18]).

Project Aims

Due to the large amounts of data originating from today's digital film productions, cinematographic film as a digital storage medium is particularly interesting for the film industry. In the past decades, the original analog film negatives could be used as a reliable and quite stable solution for storing a film production as both a cultural heritage and an economic good. However, digital technologies being more and more used for film productions pose a challenge to digital archiving. Current approaches (such as magnetic tapes or hard drives) for storing digital production data require permanent care, maintenance, and thus also financial resources.

The aim of the project is a cost-effective *store-and-ignore* approach for long-term storage of audio and video data originating from today's digital productions. Due to the high stability of the employed film materials, the storage lifetimes are expected to be 100 years or more. Aside from the actual expenses for storing the film rolls, virtually no financial resources are required during the archiving period. By using cinematographic film material which is established worldwide, it is intended to achieve independence of future access and storage technologies.

Technical Aspects

A major focus of the project is digital archiving on black-and-white film since many of these films offer an excellent long-term stability. As suggested in [7, 16] binary modulation is employed, i.e., each data point represents a logical one or zero, respectively.

Exposure devices for cinematographic film, such as the ARRILASER [19] are usually optimized to exposure of digital images as analog photographic images on film. Accordingly, the spot sizes of the modulated laser beams writing the image pixels are chosen to achieve smooth images without single image pixels being recognizable. Writing digital data patterns would result in a strong mutual influence of adjacent data points, also referred to as intersymbol interference (ISI) [7, 16]. Since the spot size of the laser beams cannot be changed by the user for such exposure systems, larger grid spaces d (i.e., distances between the data points) have to be chosen. The ARRILASER uses a grid space of about $d = 5.36\,\mu$m for analog exposures in the *4K Across Academy* exposure format [19]. By omitting image pixels, data points can be written at grid spaces being integer multiples of this standard grid space. A further important issue is the exposure level that has to be adapted for exposing the digital data patterns. If it is chosen too low, not the whole available dynamic range is exploited and the relative impact of noise effects increases. On the other hand, for an exposure level that is chosen too high, the data points may become much larger resulting in increased ISI. Certainly, an optimum exposure level has to be somewhere in between these two cases. Various experiments have been carried out to find an optimum exposure level.

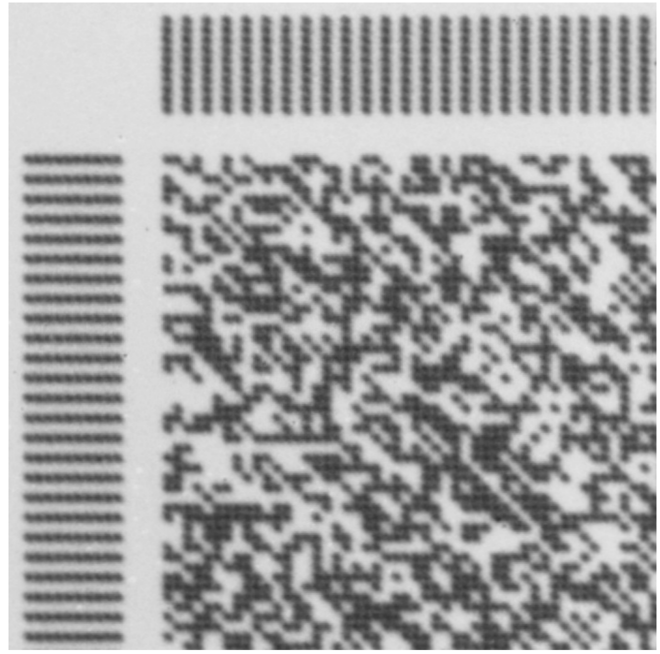

Figure 1. *Microscopic image showing parts of a test pattern (binary modulation, $d = 10.71\,\mu$m grid space).*

An exposed random test pattern with data points at a grid space of $d = 10.71\,\mu$m (two times the standard grid space) can be seen in Figure 1. The pattern has been exposed by the ARRILASER [19] on Fuji Eterna RDS black-and-white film material [20]. Of course, other film materials can be used alternatively, such as Kodak 2238 Separation [21]. It can be observed that two times the standard grid space $d = 10.71\,\mu$m results in very closely spaced but still well-distinguishable data points. Further increasing the grid space, e.g., to $d = 16.08\,\mu$m (three times the standard grid space) would lead to wider spaced data points without changing their actual diameter. Clearly, to achieve a maximum storage capacity, it is desirable to use the $d = 10.71\,\mu$m grid space. For the data points shown in Figure 1 the exposure level has been optimized to digital data patterns as mentioned before. In addition to the actual data points, parts of the synchronization pattern can be seen in the microscopic image. This pattern serves to identify the actual position of each data point (cf. [10]).

Archiving a digital production will normally involve multiple film rolls containing the digital data. Accordingly, the data has to be carefully organized on the film. For the file system, we follow the approach suggested in [10], with the file system information being stored separately from the actual data of files. The file system contains the total number of files as well as a header for each file, consisting of filename as well as start and end addresses. Each address serves to uniquely identify a bit within a set of film rolls and specifies film roll number, frame number, as well as subframe number (if a frame is divided into several subframes, cf. [10]). Further metadata and file attributes may be stored in a separate XML (eXtensible Markup Language) structure.

The resulting alignment of frames on a film roll is depicted in Figure 2. As this illustration shows, the film begins with the so-called *punch frame*, i.e., a frame with a hole in its center which

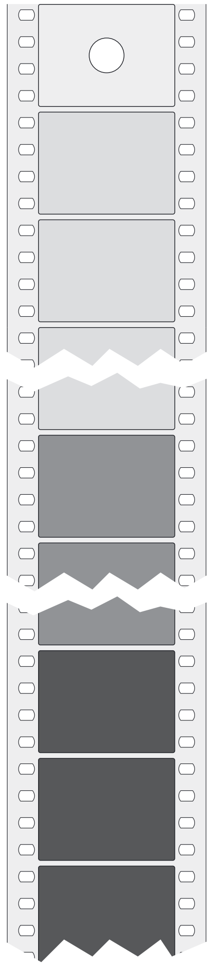

Figure 2. *Schematic alignment of the exposed frames on the film.*

marks the beginning of the role. It is followed by a human-readable description of the content along with decoding instructions. Optionally, more information can be added, such as the file system information in human-readable form or descriptions of the involved data formats. The next set of frames contains the digital file system followed by the data frames containing the actual digital data.

Even if the films are handled with care, a certain amount of dust, scratches, or other minor defects cannot be avoided in practice. Accordingly, FEC (Forward Error Correction) codes are required to ensure a virtually error-free reconstruction of the digital data after the read-out process. For this purpose, redundancy is added by the FEC encoder before writing the data on the film that serves the FEC decoder for error correction within the read-out process. Further information on error correction for 'Bits on Film' can be found in [7, 16, 22].

By neglecting any overhead, e.g., due to error correction, synchronization, and file system, the so-called gross storage capacity can be calculated (cf. [7]). Assuming a grid space of $d = 10.71\,\mu\text{m}$, the gross storage capacity reaches values of approximately $193.8\,\text{Mbit/m}$ for the exposure settings mentioned above.[2]

[2] Assuming $1\,\text{kbit} = 1024\,\text{bit}$ and $1\,\text{Mbit} = 1024\,\text{kbit}$.

Conclusions

In this paper, the fundamental principles of digital long-term storage on cinematographic film have been described. Focus has been a current research project to develop methods for data storage using this kind of film material. Although in principle there are no limitations which data can be stored on the film, the specific focus of this project is long-term archiving of digital productions. The main advantage compared to former 'Bits on Film' approaches is the cinematographic film material which is established worldwide. Corresponding exposure and scanning devices are available and installed almost all over the world. The current situation regarding digital production of motion pictures and the involved storage media has been discussed. Especially the DCPs are interesting for archiving applications since the amount of data is comparatively low due to the visually lossless JPEG 2000 compression. A further advantage is the precise specification including color space and compression techniques.

Regarding the resulting storage capacities, the proposed method is not intended to compete with today's high-density data storage solutions. The objective is merely a *store-and-ignore* offline storage approach which needs virtually no financial resources aside from the costs for storing the film rolls. Whatever data to be stored, cinematographic film offers an attractive alternative for long-term storage. Further results will be presented at the end of the research project.

Acknowledgments

The research project described in this paper is funded by the German Federal Ministry of Economics and Technology (BMWi).

References

[1] M. H. Youket and N. Olson, "Compact Disc Service Life Studies by the Library of Congress," in *Proc. of IS&T Archiving Conference*, Arlington, VA, U.S.A, May 2007, pp. 99–104.

[2] Eastman Kodak Company, "KODAK IMAGELINK HQ, CS, CP and FS Microfilms, Camera Negative Microfilm Data Sheet," Rochester, NY, U.S.A., 1998.

[3] J. D. Kuehler and H. R. Kerby, "A Photo-Digital Mass Storage System," in *Proc. of the Fall Joint Computer Conference*, San Francisco, U.S.A., Nov. 1966, pp. 735–742.

[4] J. Hull, "Surround Sound," in *Handbook for Sound Engineers*, G. Ballou, Elsevier Inc., Oxford, UK, 4th ed., 2008, pp. 1591–1601.

[5] C. J. Angersbach and K. Sassenscheid, "Long-Term Storage of Digital Data on Microfilm," in *Proc. of IS&T Archiving Conference*, Ottawa, Canada, May 2006, pp. 208–209.

[6] D. Gubler, L. Rosenthaler, and P. Fornaro, "The Obsolescence of Migration: Long-Term Storage of Digital Code on Stable Optical Media," in *Proc. of IS&T Archiving Conference*, Ottawa, Canada, May 2006, pp. 135–139.

[7] C. Voges and T. Fingscheidt, "Technology and Applications of Digital Data Storage on Microfilm," *Journal of Imaging Science and Technology (JIST)*, vol. 53, no. 6, pp. 060 505–1–060 505–8, Nov. 2009.

[8] C. Voges, "An Introduction to Long-Term Archiving of Digital Data on Film Material," *VDT International Convention,* Leipzig, Germany, Nov. 2010, to be published in conference proceedings.

[9] C. Voges, "Bits on Film – Langzeitarchivierung digitaler Daten," *FKT (Fernseh- und Kinotechnik)*, in German language, vol. 65, no. 3, 2011.

[10] C. Voges, V. Märgner, and T. Fingscheidt, "Digital Data Storage on Mi-

crofilm - The MILLENIUM Project: Signal and Information Processing," in *Proc. of IS&T Archiving Conference*, Arlington, VA, U.S.A, May 2009, pp. 187–191.

[11] D. M. Giel, A. Hofmann, W. Salzmann, and C. Voges, "Digital Data Storage on Microfilm - The MILLENIUM Project: Hardware Realization," in *Proc. of IS&T Archiving Conference*, Arlington, VA, U.S.A, May 2009, pp. 80–81.

[12] F. Müller, P. Fornaro, L. Rosenthaler, and R. Gschwind, "PEVIAR: Digital Originals," *ACM Journal on Cultural Heritage*, vol. 3, no. 1, pp. 2:1–2:12, June 2010.

[13] A. Hofmann and D. M. Giel, "DANOK: Long Term Migration Free Storage of Digital Audio Data on Microfilm," in *Proc. of IS&T Archiving Conference*, Bern, Switzerland, June 2008, pp. 184–187.

[14] A. Amir, F. Müller, P. Fornaro, R. Gschwind, J. Rosenthal, and L. Rosenthaler, "Towards a Channel Model for Microfilm," in *Proc. of IS&T Archiving Conference*, Bern, Switzerland, June 2008, pp. 207–211.

[15] C. Voges and T. Fingscheidt, "A Two-Dimensional Channel Model for Digital Data Storage on Microfilm," *IEEE Transactions on Communications*, accepted for publication.

[16] C. Voges, V. Märgner, and T. Fingscheidt, "Digital Data Storage on Microfilm - Error Correction and Storage Capacity Issues," in *Proc. of IS&T Archiving Conference*, Bern, Switzerland, June 2008, pp. 212–215.

[17] Digital Cinema Initiative, "Digital Cinema System Specification, Version 1.2," Mar. 2008.

[18] Eastman Kodak Company, "KODAK VISION Color Print Film 2383, 3383, Techical Data," Rochester, NY, U.S.A., Mar. 2005.

[19] ARRI, "ARRILASER Instruction Manual," Munich, Germany, May 2001.

[20] FUJIFILM Corporation, "FUJIFILM RECORDING FILM for Digital Separation ETERNA RDS," Datasheet, 2010.

[21] Eastman Kodak Company, "KODAK Panchromatic Separation Film 2238, Technical Information Data Sheet," Rochester, NY, U.S.A., 1998.

[22] F. Pflug, C. Voges, and T. Fingscheidt, "Performance Evaluation of Iterative Channel Codes for Digital Data Storage on Microfilm," in *Proc. of IEEE GLOBECOM*, Miami, FL, U.S.A., Dec. 2010.

Author Biography

Christoph Voges received his Diploma degree in Electrical Engineering from Technische Universität Braunschweig, Germany in 2005 with a major in information technology. During his studies he visited the University of Southampton, UK, as an exchange student. After graduating, he joined the Institute for Communications Technology (IfN) in Braunschweig, Germany. His particular research interest is film as a medium for long-term storage of digital data. He is a member of the ITG Technical Committee 3.4 "Film Technology", the AWV Working Committee 6.3 "Data and Storage Management", as well as the AWV Project Group 6.3.2 "Digital Archiving on Film".

Jan Fröhlich is the technical director of CinePostproduction GmbH. He has been involved in a large number of film projects, recently Europe's first animated stereoscopic feature film "Animals United". He also worked with the Fraunhofer Institute for Integrated Circuits (IIS), Erlangen, Germany on the development of new production and archiving systems for digital television and cinema.