Long-Term Storage of Digital Data on Microfilm

Carsten J. Angersbach, MicroArchive Systems GmbH, Germany and Karsten Sassenscheid, Fraunhofer Institute for Physical Measurement, Germany

Abstract

Building on the ArchiveLaser® project, which was presented during the 2005 IS&T Archiving Conference, the further developments of the technology that have taken place in the follow-up NOAH project will be displayed in this paper. It will present a complete system that is able to store digital data as a bit stream on long-term stable optical film (microfilm) and retrieve it afterwards with simple optical scanners. The complete workflow has been demonstrated at the CeBIT fair in Hannover, Germany, in March 2006. One of the core achievements of our research is the density of more than 1 Terabyte per reel of film. This data carrier enjoys all the characteristics and advantages of an optical, analog data carrier: long-term durability of up to 500 years, no manipulation or virus attacks and independence from rapidly changing soft- and hardware generations. This workflow addresses storage providers, whose data necessarily has to be stored for more than 4 years and where the total cost of ownership, including a solid guarantee for future accessibility to this data, is crucial.

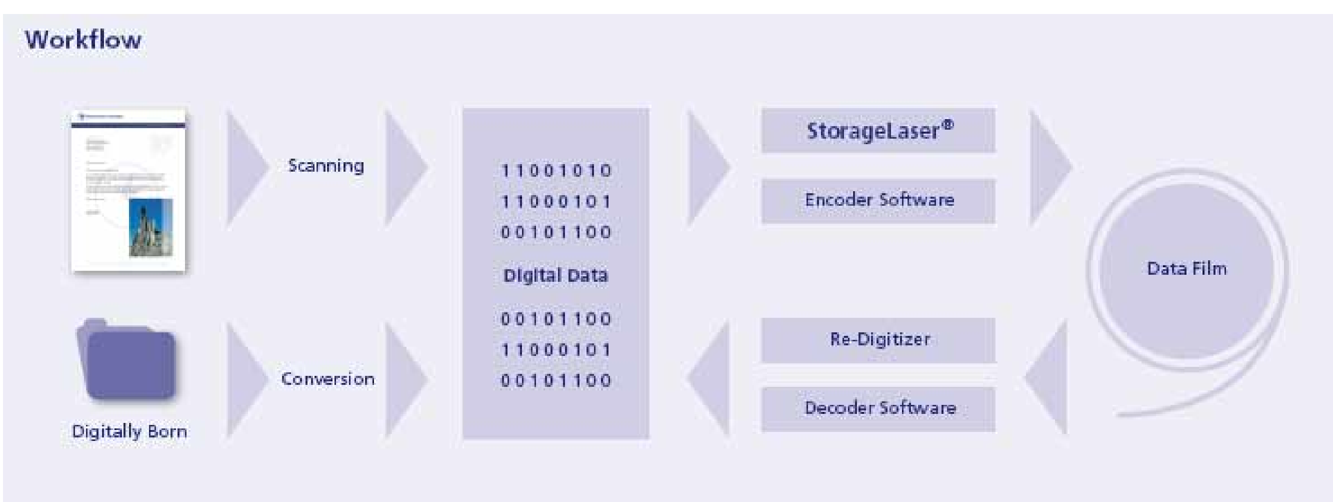
Film and Laser Technology: Blending of Proven Technologies

Various attempts have been made to resolve the problem of safeguarding the ever increasing amount of digital data on a long-term basis in the light of ever decreasing lifetime cycles of up to date computer soft- and hardware. Be it an institution of the public or health sector, the private economy, the music industry or even the movie industry – they all face the challenge of securely saving their digital data for a long period of time and at the same time ensuring immediate accessibility from a digital archive.

Currently, the problem is solved by simply accepting the costs, risks and strains that go with continuously migrating data. New generations of computer processors appear on the market in increasingly shorter time intervals. The same applies for data formats, digital data carriers and reading devices. If you wanted to avoid migration, you would be compelled to archive every generation of data carrier along with its reading device, operating system and software to go with it. This is time intensive, expensive and thus not practicable.

Decades of experience of the cultural heritage sector in the field of archiving have proven that classical microfilm is a widely accepted medium when it comes to long-term durability and data security. Using microfilm enables us to record digital data with special codes. Combining this approach with today's state of the art laser technology, it is herewith possible to achieve a data density of over 1Tera Byte per reel of film and make use of its advantages such as long-term durability of up to 500 years, no manipulation or virus attacks and independences from soft- and hardware changes.

The aim of the NOAH project is to ensure that future access to the information stored on microfilm will under all circumstances be simple and executable using standard technology and no special know-how, following an open source philosophy. Being an optical medium, the microfilm offers two crucial advantages: on the one hand, the media itself, and therefore any reading process in the future is independent of any hardware developments, as long as there will be digital cameras or optical scanners. On the other hand, the interpretation instructions (from simple syntax of file formats or metadata definitions up to complete software emulation) can be attached alongside the stored information in a non proprietary (open source), human and/or machine readable form, such as for instance XML or plain ASCI text.



Technical Components of the NOAH System

The film recorder (StorageLaser®)

The data code, represented by a pattern of pixels, is transformed into a TIFF image which is later transformed by acusto-optic modulators into laser beam intensities. Every pixel on the film is then exposed by a laser beam writing 160 line pairs/mm.

The film

The challenges in writing an optical code onto photographic media are as follows [1]:

- photographic film is highly non-linear
- characteristic curves
- noise (film grain)
- grayscale reproduction (inter-image effects, side absorptions)
- chemical diffusion & scattering
- optical distortions (film recorder & scanner)

The encoder

In order to keep the reading process as simple as possible, the intelligence of the archiving system is located within the writing process. This includes error correction, redundancy, delocalisation, sampling and other aspects of data security technology.

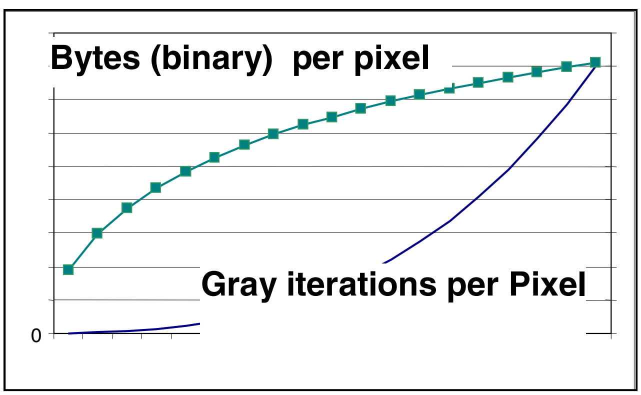


***Figure 1**: Storage capacity determined by gray iterations and bytes per pixel*

The re-digitizer (scanner)

For the reading process of the optical archival media only an optical scanner is necessary. The technical requirements not only depend on physical aspects as over-sampling, but also on economic issues: Depending on the service level, the storage- or asset management provider demands faster or slower reading times for a given entity of data.

The decoder

The instruction set for the reconstruction of the encoder preferably is deposited various times within an archival unit. Foremost, it must be written down in an easily readable way (machine readable, open-source code).

Crucial Points

From a technical point of view the characteristics of the film have a crucial impact on the storage capacity. The processes of encoding, writing, re-interpreting of the information patterns (scanning) and decoding must not be looked at separately [2]. Therefore the aim of locating the complexity of the storage system into the writing process and simultaneously guaranteeing an open source and simple reading process, an interdisciplinary expertise of photochemistry, optics, encrypting and information technology is crucial.

The data density determines the marginal costs of large scale storage systems. It is therefore suitable to collaborate with film manufacturers in order to optimize costs.

The Fraunhofer Institute for Physical Measurement has numerous years of experience in this field. The ARRI-Laser, which was developed for the movie industry, and which has won the "Scientific Engineering Award" from the Academy of Motion Picture Arts & Sciences in 2002, was invented by the same team of experts who developed the ArchiveLaser®. This device was primarily designed for the cultural heritage sector. It was developed within the ARCHE project, which was supported and funded by the German Federal Ministry of Economics and Labor. The results of the ARCHE project can be inspected as they will be presented as a case study at the current IS&T Archiving Conference. The subsequent NOAH project aims at iteratively broadening the applicability of this technology initially designed for the cultural heritage community to serve the needs of long-term storage by writing any digital code on a stable media (bits-on-film) with the innovative StorageLaser® System.

References

[1] Rudolf Gschwind, Lukas Rosenthaler, PEVIAR: Long term storage for digital data on photographic film material, Draft to the Swiss National Science Foundation, 3001 Berne, Switzerland, 2005

[2] Brunner, Raimund, Fraunhofer Institute für Physikalische Messtechnik, The ARCHE Project, Freiburg i.Br., 2006

Author Biography

__Dr. Carsten J. Angersbach__, LL.M., received his education at the Universities of Freiburg, Berlin and Munich (Germany) as well as at the Université de Lausanne (Switzerland) and the University of Chicago (USA). He holds academic degrees from the University of Munich and the University of Chicago.

__Dr. Karsten Sassenscheid__ is Head of Department for Optical Spectroscopy and Optical Systems (OSS) at the Fraunhofer Institute for Physical Measurement Techniques (IPM) in Freiburg i.Br. He earned his PhD in Physics at the University of Würzburg. He is author of various publications in scientific journals. He is Chairman of the "Arbeitskreis Optische Messtechnik von Photonics BW" and member of various scientific associations (VDI-Ausschuss OPTAM et al)