DISTARNET – A Distributed Archival Network

Lukas Rosenthaler and Rudolf Gschwind
Imaging and Media Lab, University of Basel
Basel, Switzerland

Abstract

The goal of this work is to develop and establish an archival system for digital data with focus on long-term archival as required by archives and museums which preserve the audio-visual heritage of our time. The fundamental idea is that institutions with long-term archival needs of digital data collaborate in order to build a geographically distributed, Internet based archival system. It has been of primary concern in designing this system to minimizing the risk of data loss due to a catastrophic event (on a local or even regional scale such as a devastating earthquake) or false manipulation. In addition the archival network has to cope with the rapidly changing standards of information technology and preserve readability of the data beyond changes of software and hardware technology.

The archival network will be implemented as a distributed, self-organizing and optimized peer-to-peer (P2P) network with a high redundancy. The data to be archived will be distributed in redundant copies among all participants' parties. It will be assumed that the data will be transported over untrusted channels and stored on untrusted nodes. Therefore, strong cryptography will be used to guarantee the integrity and privacy of the archived digital assets.

The goal of this work is to establish an archival system for digital data with focus on long-term archival as required by archives and museums which preserve the audio-visual heritage of our time. The fundamental idea is that institutions with long-term archival needs of digital data (digitized cultural assets such as photographic collections, sound archives etc.) collaborate in order to build a geographically distributed, redundant Internet based archival system. Minimizing the risk of data loss due to a catastrophic event (on a local or even regional scale such as a devastating earthquake) or false manipulation is of primary concern in designing this system. Additionally the archival network has to cope with the rapidly changing standards of information technology and preserve readability of the data beyond changes of software and hardware technology.

In order to achieve both goals, we propose a globally distributed storage network with high redundancy based on a peer-to-peer (P2P) architecture. The integrity and privacy of the data entrusted to the archival system is guaranteed by strong cryptography, especially private/public key encryption schemes. The archival network forms sort of a "closed user group" based on Internet technology where all relevant data is encrypted.

Introduction

This report is printed on paper that has a life expectancy of approximately 50-100 years.

Digital data of all kinds are now for some years a fixed constituent of each archive. Reasons for this are manifold: better use or better access to the collection but also preservation, as the today's audiovisual materials are unstable and are decaying. This new digital cultural heritage contains thus no longer only documents on paper or film materials. Thus new knowledge becomes necessary about the long-term archiving of the electronic data. It is important to hold itself before eyes that a good long-term archiving begins, before the digital data are produced. This has the consequence that digital archives must be structured basically differently than traditional archives.

The digitization of documents and audio-visual material of all kinds the same meaning as the writing and the printing for our language. For the first time it is possible to code these original objects, e.g. a photograph, in a symbolic format (as number bundles). This immaterial code can be transported and replicates without any loss of information! The digital revolution can be regarded as a genuine revolution, since it lets become insignificant in a certain sense place and time and the concept of an "original object" abolishes. The characteristics of each digital code and hence for all multimedia data are thus as follows:

- Digital code is by principle always independent of the medium, onto which it is held. The medium is arbitrarily exchangeable (even if e.g. a digital code "written" in stone is relatively unmanageable).
- Digital information can be replicated, i.e. copied, without information loss. Information loss is - the correct procedure presupposed - in the mathematical sense equal zero. This has the consequence that the term of "original" or "unique" for digital data becomes senseless, since the "original" and the "copy" are identical and indistinguishable. A further important consequence is that by the loss-free copying an in principal unlimited life span of the digital information results.
- Redundancy is important, since thereby a larger security results on damage of the medium, however needs it more to "storage space". Our Latin writing has a

comparatively high redundancy (approx. 64 characters = 6 bits), compared for example with the Chinese writing (approx. 20'000 character = 15 bits).

- The digital code can be transferred over each information channel, and this always with speed of light. Thus for digital data space and time are removed.

The digital revolution has very large effects for archives of any art. When creating the digital data it must however be already clear from the beginning, how the data will be archived. This procedure represents a new work model compared to the conventional, "object oriented" archive.

In order to ensure the long-term availability of digital data, respectively information in digital form, a comparison with the writing is worthwhile. If we regard history, then we see that a large part of our cultural heritage and our knowledge is and was delivered in the form of record. This knowledge survived the centuries:

- It was created in symbolically coded form (text, letter)
- The books and texts were copied regularly, one become independent thereby of the medium. The quality of the medium is of secondary importance, as long as the "code" can be decoded.
- The information was spread. Particularly after invention the printing Gutenberg a mass distribution of the information took place.
- Writing contains much redundancy, so that also with a decay of the medium the text can be still read. Th■s t■xt is■sti■l rea■ab■e.

The Dilemma of Digital Archiving

However, experience shows that digital data as a very limited lifespan. On one hand, the life expectancy of digital media seems to be quite limited (e.g. magnetic tape has a life expectancy of about 30 years given proper treatment storage conditions). But to a higher degree the rapid pace of technological advance in the field of information technology results in a much shorter life span of a given technology. For example, while a magnetic tape media may last 30 years, but the tape drive necessary for writing and reading the tape is produced only for 2 years, and service and spare parts are guaranteed for another 2 years. The next generation will read/write a higher density and may be able to treat one generation back, but it is highly probable that after 2 generation, the commercially tape drives are not able to read the original tape because the technology has become incompatible. Therefore, digital data becomes unreadable not because of the limited lifespan of the media, but because a given technology becomes obsolete very fast.

There are different approaches possible in order to tackle this problem:

1. **Hardware Preservation**: Not only the media where the digital data is recorded on is preserved, but also the hardware necessary to read it. In this case, it is necessary to have access to the technical knowledge about the specific hardware, to have access spare parts etc. in order to keep it in an operating state. Therefore this approach is practical only for very special cases.
2. **Emulation**: Sometimes it is possible to emulate old hardware and software on new machines. This approach seems reasonable for software (e.g. games, computer art etc.), but as soon as specific hardware such as tape drives is involved this solution becomes very difficult or impossible (e.g. emulation never allows that a 5 1/4'' floppy can be read on a 3.5'' drive).
3. **Migration**: The digital data is periodically transferred onto new media using the state of the art technology and media formats. If necessary, the data is also reformatted to the latest standardized formats (e.g. file formats etc.)

It seems that, as far as digital archives already dealing with the problem of technology getting obsolete preferred the migration method. However the periodic task (about every 5 years) of migrating all archived data seems prohibitive in cost and effort and therefore prevents many archives and museums using digital data for long term archival.

Data Loss and Redundancy

Every archive which stores items of any kind has the problem that in case an item is destroyed by a mishap, crimes etc., or even that the whole collection is destroyed by a catastrophic event like war, fire, earthquake etc., the originals can not be replaced. There is no redundancy. However, if a library containing printed books is destroyed, it is highly probable that the content of the books will still be available because other libraries contain the same books. The number of books printed for one edition results in a high redundancy. Since digital data can be copied (or “cloned”) with zero loss and transported immaterially over networks, digital data is per se adapted to achieve a high redundancy. The security against data loss is especially high if the redundant data is stored at geographically different places since then the risk is distributed.

However, maintaining such an archive is very complicated and cumbersome and probably beyond the capability of most archives except for the very large ones. And the problem of periodic migration even gets more complicated by introducing this kind of redundancy.

Peer-to-peer Networks

The redundancy of digital data can be achieved by peer-to-peer networks (P2P-networks). This software architecture based on a web of connected peer computers. It is widely used for example for file sharing of music files (e.g. napster, kazaa etc.). There are some projects that are more targeted in the direction of archiving:

- **Mojo Nation http://www.mojonation.net/**
  Mojo Nation describes itself by:

What is Mojo Nation? Mojo Nation is a revolutionary new peer-driven content distribution technology. While simple data distribution architectures like Napster or Gnutella may be sufficient to allow users to trade mp3 files they are unable to scale up to deliver rich-media content while still taking advantages of the cost savings of peer-to-peer systems. Mojo Nation combines the flexibility of the marketplace with a secure "swarm distribution" mechanism to go far beyond any current file sharing system -- providing high-speed downloads that run from multiple peers in parallel. The Mojo Nation technology is an efficient, massively scalable and secure toolkit for distributors and consumers of digital content.

- **Freenet**, http://freenet.sourceforge.net/
  Freenet describes itself by
  Freenet is a distributed decentralized information storage and retrieval system. It is designed to allow the free distribution of information on the Internet without fear of censorship. To achieve this it provides anonymity to those placing information into Freenet, and those accessing information from Freenet. It is also totally decentralized, nobody is in control of Freenet, not even its creators. This makes is virtually impossible to force the removal of a piece of information from the system. Further, Freenet provides many efficiencies over more conventional means of information distribution such as the World Wide Web through its dynamic caching and mirroring of content. Freenet was originally designed by Ian Clarke and is being implemented by a number of skilled volunteers.
- **OceanStore**: An Architecture for Global-Scale Persistent Storage, http://oceanstore.cs.berkeley.edu/, developed at the University of California, Berkeley by John Kubiatowicz. OceanStore describes itself by
  OceanStore is a utility infrastructure designed to span the globe and provide continuous access to persistent information. Since this infrastructure is comprised of untrusted servers, data is protected through redundancy and cryptographic techniques. To improve performance, data is allowed to be cached anywhere, anytime. Additionally, monitoring of usage patterns allows adaptation to regional outages and denial of service attacks; monitoring also enhances performance through pro-active movement of data. A prototype implementation is currently under development.
  For a detailed description see: John Kubiatowicz, David Bindel, Yan Chen, Steven Czerwinski, Patrick Eaton, Dennis Geels, Ramakrishna Gummadi, Sean Rhea, Hakim Weatherspoon, Westley Weimer, Chris Wells, and Ben Zhao, Proceedings of the Ninth International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS 2000). This publication and further readings can be found at http://oceanstore.cs.berkeley.edu/ publications/index.html.
- **Eternity Service**
  http://www.kolej.mff.cuni.cz/~eternity/index.html, developed at the Institute of Mathematics and Physics at Charles University in Prague. The project of Eternity service is based on an idea that originally came from Ross Anderson, University of Cambridge Computer Laboratory. Eternity service describes itself by
  Eternity Service is a new attitude to the way of managing data. By using many servers all over the world that are interconnected with Internet we can achieve very high reliability of storage. The purpose of this project is to verify, whether our approach to the problem has a chance to succeed. We want to introduce a new way of creating, storing and handling copies of data of high importance - the Eternity service - with a very high degree of reliability. The Eternity service be resistant not only to usual threats such as natural disasters and vandals, but also to or court decisions, religious leader orders, activities of secret services and so on.
- An other idea is the **MediaLess Archive**, as proposed by Jim Lindner, media matters LLC, New York. It has been communicated by Jim Lindner at various conferences and events including the following:
  - JTS, Joint Technical Symposium of FIAF/FIAT/ IASA , Jan 20th – 22nd 2000, Paris
  - AMIA 2000, annual conference of the Association of Moving Image Archivists, Nov 13-18, 2000, Los Angeles
  - Guggenheim Museum New York, Feb. 2001
  - Tagung der Fachhochschule Bern, Feb. 2001

Distributed Archival Network (DISTARNET)

The goal of the proposed work is to establish an archival system for digital data with focus on long-term archival as required by archives and museums which preserve the audio-visual heritage of our time. The fundamental idea is that institutions with long-term archival needs of digital data (digitized cultural assets such as photographic collections, sound archives etc.) collaborate in order to build a geographically distributed, redundant Internet based archival system. Minimizing the risk of data loss due to a catastrophic event (on a local or even regional scale such as a devastating earthquake) or false manipulation is of primary concern in designing this system. Additionally the archival network has to cope with the rapidly changing standards of information technology and preserve readability of the data beyond changes of software and hardware technology.

In order to achieve both goals, we propose a globally distributed storage network with high redundancy based on a peer-to-peer (P2P) architecture. The integrity and privacy of the data entrusted to the archival system is guaranteed by strong cryptography, especially private/public key encryption schemes. These encryption algorithms are based on the fact that symmetric pairs of keys exist, where each key is able to decrypt data encrypted by the other key. This encryption scheme can be used either to encrypt/decrypt data or to generate digital signatures, which securely identify the originator and guarantee the integrity of data.

Each participating institution provides (n+x) times[*] the amount of storage needed for it's own data to the archival network. The data injected to the system is then cloned n times and distributed optimally among all participants. Each institution has to contribute one ore several storage nodes. All these nodes form a self-organizing peer-to-peer storage network. The archival network will be designed to be self-adapting to changing configurations such as adding or removing nodes. It will react automatically to loss of contact to nodes and guarantee that all data is always stored with the required redundancy. If, due to a catastrophic event, all the data of one of the participants is destroyed, its data will be cloned by the other nodes to regain the redundancy. At any time, all the data of any participant can be retrieved in Toto or partially. The archival network forms sort of a "closed user group" based on Internet technology where all relevant data is encrypted (untrusted transportation channels).

The data itself is organized in "containers" which are encrypted with the public key of the owner. Therefore only the owner of the data is able to decipher the data (using it's private key) and thus unauthorized use is prevented. However, in case of a catastrophic event, not only the data but also the private key of the data owner may be lost. Therefore, it is important that the private keys used to decrypt the data have to be made available (e.g. in a sealed envelope) to trusted institutions in order to guarantee the readability in case the key and it's owner are lost. Such institutions could be some sort of "board of directors" of the archival network, government agencies ("Bundesamt für Kultur", Library of Congress etc.), maybe in future even agencies of the UN (e.g. UNESCO) etc.

Each data container exists in several instances in order to gain the required redundancy. These instances are identical with respect to their data content, but may be identified by a serial number. Each container is tagged with an id, which allows identifying the owning participant.

Architecture / Design Principles

In the description of the architecture, the following terms will be used:

- **Participant**
  An institution, which participates to the archival network by providing storage space. A participant may, but is not required to, inject data into the system. Each participant provides one or several nodes.
- **Node**
  A node is a computer which is able to run the node agent (see below) and controls one or several attached storage units.
- **Storage Unit**
  A storage unit is a logical storage device with direct access capabilities. It can be implemented as a disk farm, a near line storage device etc..

The archival network is designed as a distributed peer-to-peer architecture consisting of three distinct parts:

1. **Node agent**
   It implements the distributed intelligence of the system. It communicates on a peer-to-peer base with other node agents in order to optimally distribute the data containers and controls all it's attached storage units.
2. **Injection/Retrieval Agent**
   This piece of software implements the front-end used by a participant to inject or retrieve data, to query the network and to monitor relevant properties.
3. **Administration Agent**
   This software is used to configure the archival network such as adding or deleting new nodes, adjusting global parameters etc. . It also allows monitoring all relevant system parameters (free space, network activity etc.)

The archival network will be implemented in a platform independent way using JAVA technology.

A node agent knows of and communicates with only a limited number of "direct" neighbor node agents. Requests of a node agent will be transmitted by these direct neighboring members to their respective neighbors until all nodes have been contacted. Responses to requests are transmitted directly to the requesting node agent. The node agents periodically communicate status information in order to be aware of potential dropouts of nodes, addition of new nodes, problems with network connections etc.. Using this architecture, there is no single point of failure and the archival network will be self-adapting to changing topologies of the underlying network. If a node cannot be contacted for a defined period of time, it is considered to be in trouble and all the data containers, which are stored on the failed node, are replicated from redundant copies on other nodes in order to guarantee the required redundancy. These containers again are optimally distributed on all available nodes.

The optimal distribution of the data containers is crucial for the archival network. The algorithm used to locate the storage nodes for a given container has to take into account many criteria, with the following to be of special importance:

- *Geographical distance*
  Each data container instance should be stored on nodes geographically located as far away as possible from the location of the originating node and from the nodes, which store other instances of the same data container.
- *Even distribution*
  The containers of a specific participant should be distributed as evenly as possible over all nodes. This is to prevent the clustering of the data of one participant on

[*] Where *n* denotes the redundancy required. The higher *n*, the higher is security against data loss, but also the higher is the cost of the required storage space. We consider *n = 3* to be a reasonable value, with n = 2 being a minimal value for redundancy. *X* denotes the space which may be temporarily used. This space is also used to guarantee the required redundancy if an institution experiences a catastrophic event and the data stored there has to be cloned. X is expected to be ~1.0, but may get smaller with increasing number of participants.

a specific node or at a specific geographical location (Distribution of the risk).

Other criteria may be the stability of nodes (history of uptime), network bandwidth, storage space, storage media (online, near line,...) etc. . The distribution algorithm will be based on a ranking system, where all criteria are weighted and the node with the highest-ranking "wins". This procedure is repeated for each data container instance, which has to be injected to the system. Such a ranking system is well adapted to the distributed architecture of the archival network.

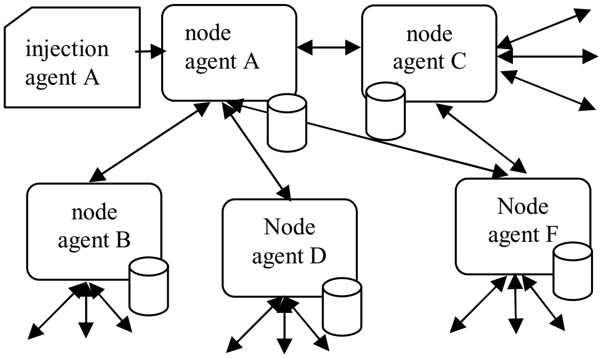

Figure 1. Schema of a distributed archival network

Data Injection

The injection/retrieval agent is launched to inject data. It builds the data container, encrypting the data using the public key of the participant. Then the data container is signed in order to identify the owner and to prove its integrity. The data container is then handed over to the node agent, which clones and distributes the data container on the archival network. The data container may contain many small objects, one large object or part of a very large object. The archival network requires that minimal metadata is associated with each object in order to uniquely identify each object. This metadata may consist of a variable length string, which has to be unique among all objects of a participant. Special data containers will contain the association of these metadata strings with the data objects. Further metadata such as relational databases, descriptive texts (hypertexts) etc. are considered to be ordinary data objects, which can be entrusted to the archival network. It is the responsibility of each participant to choose the appropriate format for these metadata.

Disaster Recovery / Retrieval

There are several levels of disaster recovery. The least severe is after partial loss of data objects. The participant starts the injection/retrieval agent and gives the ids of the objects to be restored. The information is passed to the node agent, which collects the data container(s) with the given objects and passes them to the retrieval agent. The retrieval agent then deciphers the data containers and restores the data objects. Since all transactions are securely identified and encrypted, only the owner is able to restore the data objects.

The most severe case is the total destruction of the site of a participant. Let's assume that all infrastructure of the participant has been destroyed including all metadata and keys (for example by a devastating earthquake). In this case, another institution takes over the responsibility of the content of the destroyed participant and attaches a new node to the archival network. The associated key is handed over by the controlling instance (e.g. by the "board of directors"). Using the appropriate key of the destroyed participant, this new node will reconstruct the entire data object collection of the destroyed participant. The reconstructing node agent will broadcast the request to get all data containers with the given participant id, and it will be able to legitimate this request by signing this request with the appropriate key. It will collect the data containers instances from the "nearest" nodes and the associated injection/retrieval agent will reconstruct all data objects.

Self-Organization

The archival network has to be self-organized in order to keep the network manageable and scaleable, and to avoid any single point of failure. If the intelligence is distributed equally on all nodes, the failure of a limited number of nodes will not affect the archival network as a whole. Self-organization is implemented within the node agents. The first aspect is that the distribution algorithm tries to evenly distribute all data container instances on the network using the ranking criteria mentioned above. The second aspect that each node is monitored by it's neighboring nodes (neighbor nodes are not defined by geographical proximity, but are pure logical neighbors). If a node drops out for a defined time period, the neighboring nodes will broadcast this information through the network and the nodes which store other instances of the data containers of the failing node will negotiate in order to duplicate their instances. Periodically each node will reevaluate it's ranking with respect to the data container instances it is save keeping. If other nodes do have a much better ranking (there will be some hysteresis included in this reevaluating ranking process), it will forward this container instances. This self-organization process can be regarded as a classical control process. Through careful selection of the regulating parameters, oscillations have to be prevented (damping, e.g. through hysteresis).

Administration / Monitoring

Since the archival network is a closed user group, there must be some possibilities to administer and monitor the archival network. All network traffic between the node agents has to be signed by a digital signature. A node accepts only requests and data from nodes with a valid signature. The distribution of keys and certificates follows the well-known procedures of private/public key schemes.

The administration also has the possibilities to monitor the performance of the network and visualize adequate parameters (e.g. distribution of data containers, free/used storage space, state of nodes etc.). These capabilities require some special privileges, which also are protected by private/public key pairs.

Security / Privacy

Security and privacy will depend on public key encryption methods. We will rely on well-known standard encryption methods like PGP (Pretty Good Privacy), GnuPG etc. In a first phase we will evaluate the available standards and choose an appropriate methods using the following criteria:

- Open standard (encryption and signature)
- High security
- Performance
- Flexibility
- Open source

All object data is encrypted (optionally, if the participant requires it for it's data. Encryption can be selected on an object-by-object bases) and signed (required). This guarantees the privacy (encryption) and integrity (signature) of the data. Therefore the archival network can be regarded as the save of a (Swiss) bank, which keeps the belongings of it's customers in a save place, but does not know what it is keeping.

As mentioned above, the private keys of the participants have to be safeguarded at several trusted institutions in order to allow the full reconstruction of a participants archive after a full-scale disaster. This safeguarding has to be negotiated by the participants. This "political" issue is not part of this project and will have to be solved for each participant individually.

Redundancy / Distribution of Risk

The archival network will be designed in a way that guarantees the optimal distribution of the data containers with respect to the risk of data loss. The archival network must adapt itself automatically to changes in the environment (e.g. failure of a storage node or a communication path) by replication and redistribution of data containers in order to achieve at any time the minimal required redundancy with optimal use of available distributed storage space.

The algorithm we will implement follows closely the thermodynamic principal of minimizing the **free enthalpy**: minimizing the free energy while maximizing the entropy of the system. Let's consider the following analogies:

- Each data container is considered to be an electron: The repulsive forces between electrons augment the energy of the total system if the electrons are close together. Minimizing the energy will therefore maximize the average (geographical) distance between the data containers, which also maximizes the entropy.
- Each archival node acts like an electronegative electron attractor. The more free storage is available, the more electro negativity (and therefore attraction) is assumed.
- Each data container has an equivalent of spin that has the same value for all redundant (identical) copies of the data container. In analogy to quantum chemistry, data containers with the same "spin" may not be located at one storage node (as electrons with the same spin my not occupy the same energy level). This will enforce that the identical copies of a data container will never be located on the same storage node.
- The archive has a "constant mass", i.e., if one archive is no longer available, the missing parts are automatically generated until the same amount of data containers is reached

The system will react like a quantum electron gas to external influences and find an optimal dynamic equilibrium by minimizing the "free enthalpy" of the system. We consider the implementation based on quantum-chemical and thermodynamically models in the context of the question of long-term archival of digital data to be fundamental research in computer science. Further optimization conditions are, e.g., the limited bandwidth of communication channels therefore leading to a minimization of network transfers, statistical stability of individual nodes etc.

Further properties are:

- All data has to be encrypted using private/public key methods. Since the amount of data is very large (Terabytes), very efficient methods have to be used. Existing algorithms have to be evaluated, optimized and modified for this task.
- The whole "distributed archival network" has to be designed in a way to guarantee almost 100% security with regard to a) data loss and b) privacy in an insecure environment (Internet) with instable nodes and communication paths.

Data Migration

Since both hardware and software will change very often with respect to the expected lifetime of a real world implementation of a distributed digital archive, the migration of data to new hardware and new software standards (formats, programming languages etc.) has to be considered from the beginning and will be one of the principal design criteria. This will be achieved by a careful design of open and flexible communication protocols (on the level of communication protocols we consider XML as one the currently dominant standards), modular, platform independent, object oriented software design and an open source approach. Migration processes must be aware and make use of the distributed nature of the long-term archive.

Due to the properties of distarnet, migration becomes almost a non-issue: new hardware can be integrated by simply switching of the old hardware and attaching the new hardware to the network. The archival network will react to this event and automatically rebuild the redundancy and integrate the new hardware. If new file formats are introduced, small programs which perform the conversion of the data can be distributed on the network and perform the conversion also in a distributed manner.

Current State of the DISTARNET

A proof of concept has successfully been implemented in a platform independent way using JAVA technology. Currently we are working on the definition of a network protocol to be used by distarnet. We expect starting the implementation using several (redundant) technologies within the next few months. First results are expected within one year after implementation started.

References

1. Mojo Nation http://www.mojonation.net/
2. Freenet, http://freenet.sourceforge.net/
3. OceanStore: An Architecture for Global-Scale Persistent Storage, http://oceanstore.cs.berkeley.edu/
4. Eternity service,http://www.kolej.mff.cuni.cz/~eternity/index.html, developed at the Institute of Mathematics and Physics at Charles University in Prague

Biography

Dr. Lukas Rosenthaler, born 1960, studied Physics, Mathematics and Astronomy at the University of Basel, and got 1987 a Ph.D. in Applied Physics in the field of Nanotechnology building an early Scanning Tunneling Microscope. During his Ph.D. thesis he got involved with image processing and image analysis. From 1988 to 1992 he worked as postdoc at the Swiss Institute of Technology in Zürich in an interdisciplinary project about the understanding and the computational simulation of the vision system of human beings. From 1992 to 2001 he worked in the field of computer graphics and visualization within Cadwork Corp., the leading software manufacturer for CAD-software in Switzerland. During this period he also developed new methods for the restoration of damaged movie films, in affiliation with the Scientific Photography Lab of the University of Basel. Since 2001 Lukas Rosenthaler is a full time staff member of the Imaging and Media Lab of the University of Basel, Switzerland. The main research topics are the restoration of movie films and the long-term preservation of digital images.