Long-term, Permanent Storage of Data

Write Once, Read Forever (WORF)

University of Pennsylvania

New York University

Creative Technology LLC

Contacts:

Richard Solomon <rsolomon@seas.upenn.edu> 413-267-5171

Eric Rosenthal <eric@creative-technology.net> 732-817-1720

Jonathan Smith <jms@central.cis.upenn.edu>

Clark Johnson <clark.johnson@creative-technology.net> 608-467-7835

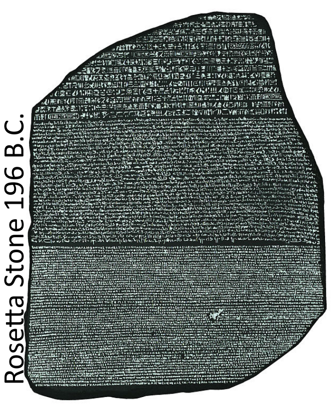
Melitte Buchman <melitte@nyu.edu> 212-998-2668

Write Once, Read Forever

"WORF"— a technical solution for the low-energy storage of information in perpetuity

CTech has IP for re-purposing a tested & proven process for extremely long-term, archival data storage & automatic retrieval

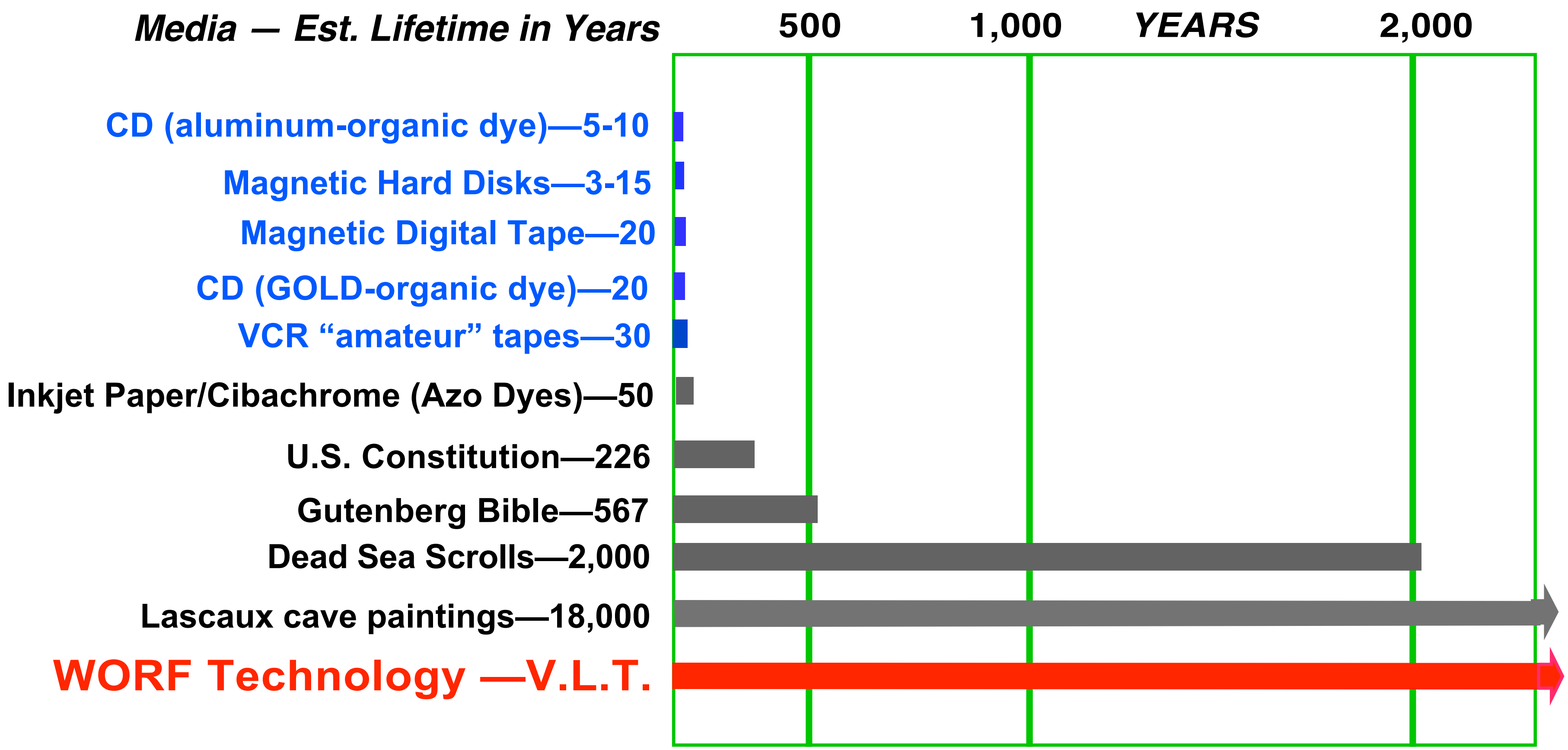
- **Minimal energy for storage** — no need for continual re-writing
- Combines on a stable substrate **digital data clusters, images & human readable documentation** for future decoding
- **Data cannot be altered or hacked**
- **Off-the-shelf, cheap components** for writing & reading digital data + full-color images on same substrate
- **Re-engineering proven, century-old,** silver halide process
  - Applies modern emulsion chemistry — no dyes to fade



Rosetta Stone 196 B.C.

A “Digital Dark Age”

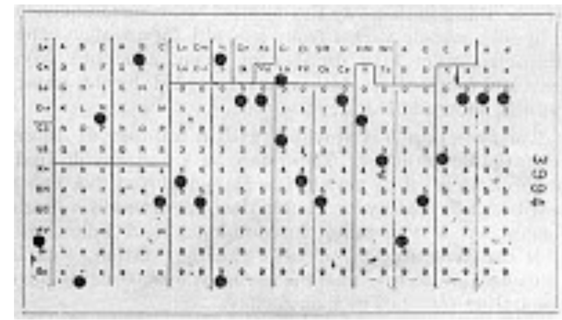
Current Data Storage Technologies Not Suitable For Archiving



A "Digital Dark Age"

Current Data Storage Technologies Not Suitable For Archiving

- **Magnetic hard disks**
  - MTBF really means "mean time before failure"
  - HDs begin data loss right out of the box
  - Energy intensive — requires periodic copying, ~90% energy in data farms to keep magnetic media within thermal limits



Hollerith punched card 1890s

- **CDs, DVDs, solid-state memories**
  - Short-term environmental degradation
- **Magnetic tape**
  - Mechanical deterioration (tapes bind without periodic re-winding)
  - Not random access
- **Complex interfaces for future generations to decode**
  - 1000's of obscure data encoding systems, mechanical linkages, and electrical interfaces since beginning of data processing
- ***Data continues to grow exponentially***

Existing data storage technologies require continuous backup to reduce risk

- **HD MTBF: raw error rates can *exceed 10% from day 1***
  - Digital quantizing errors accumulate over time: bit errors; block errors
- **Natural events destroy data:**
  EMP (lightning), sunspots, earthquakes, gamma rays
- **Hacking can alter data:**
  requiring continual array authentication
- **Tape: mechanical deterioration;** not random access.
- **CDs, DVDs, solid-state memories —**
  short-term environmental degradation
- **Closed, proprietary architectures —**
  obscure encoding, interfaces
  - Documentation not embedded in devices

Limitations of Magnetic Recording

- **Magnetic storage media density *is approaching the thermal limit!***
  - Density is approaching one 1 terabit per square inch —~60%/year increase over 50 years
- **Output signal is compromised as the density of magnetic domains increases**
  - Volume of domains determines strength of magnetic field
  - Higher density decreases voltage and increases system noise
- **Ultimate density depends on statistical fluctuations of domain magnetism**
  - Domain magnetization fluctuates due to *thermally-induced variations* — ie, temperature
    - Decreasing temperature decreases the superparamagnetic size
    - Increasing temperature increases the superparamagnetic size
  - Domain magnetization is subject to its own demagnetizing field
    - Resistance depends upon domain's material, shape, & fields from adjacent domains
  - Smaller domains have larger fluctuations, so:
    at the superparamagnetic limit ***domain magnetization is unable to retain data***
- **Disk drives require error rates of 10^{-12}**
  - At current densities, the error rate (caused mainly by thermal effects)
    is ~0.01 — a correction factor of 10 billion!
  - Achieved by sophisticated error correcting signal processing

Limitations of Magnetic Recording

- **Data stored in magnetic media resides in the polarity of microscopic magnetic regions called domains. The volume of these domains determines the strength of the magnetic field and hence the reproduced voltage during playback. As the density increases this voltage, of course, decreases. Thus the system noise increasingly compromises the reproduced output signal.**

- **A further limitation, and the one that ultimately limits the density, is the statistical fluctuation of the domain magnetism itself. The domain's magnetization is subject to its own demagnetizing field. The resistance to self-demagnetization depends upon the crystallography of the domain's material, the shape of the domain (acicular [needle-like] regions have lower demagnetization) and the fields from adjacent domains.**

- **The domain magnetization (the domain is always saturated) fluctuates as a result of thermally-induced variations. Smaller domains have larger fluctuations. Ultimately, at a critical size, called the superparamagnetic limit, the domain magnetization no longer remains in a fixed direction, and is thereby unable to retain any data. Decreasing the temperature decreases the superparamagnetic size. Increasing temperature does the opposite.**

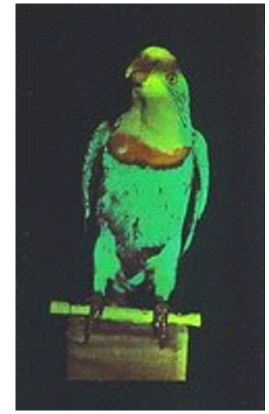
- **Currently, magnetic storage media density is approaching one-terabit per square inch. This density has been increasing at about 60%/year for over 50 years. <u>It is now approaching the thermal limit</u>. At current densities, the error rate (caused mainly by thermal effects—both electrical and magnetic) is about 0.01. To be useful, disk drives require error rates of 10^{-12}, a correction factor of 10 billion! This is achieved by sophisticated error correcting signal processing, made possible by the development of low-cost, ultra-high speed processors**

CTech Re-purposing Silver Halide Media For WORF Digital Data Storage

- **Photosensitive metallic emulsions have extremely long lifetimes — to date photographs have lasted >160 years**
  - Accelerated lab testing: 1000+ years for fixed silver halide emulsions
- **Low energy utilization for long-term storage**
  - No need for continual media maintenance or duplication
  - No requirement for environmental controls
- **Data cannot be erased or modified**
  - WORF = **W**rite **O**nce, **R**ead ***Forever***
  - Protected from threat of remote hacking or corruption
  - EMP proof; vibration (earthquake) proof; neutrino & γ-ray proof
- **Low-cost, modern components for writing & reading data & images**
  - CTech's IP re-purposes proven conventional data mechanisms
- **Future proof**
  - Data + human readable text on same substrate

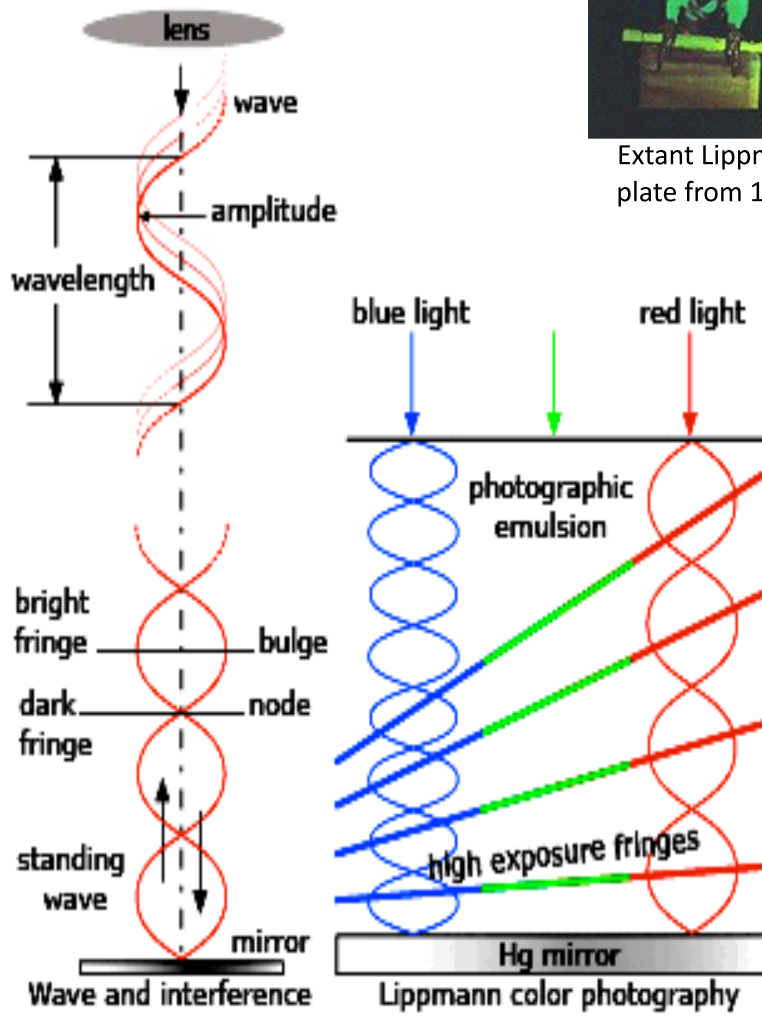
Permanent color imaging __without filters__ demonstrated 120 years ago using silver halide!

CTech's WORF technology *re-purposes* Lippmann media for *21st Century archival DATA storage*



Extant Lippmann plate from 1890s

- **Gabriel Lippmann won 1908 Nobel Prize for capturing color images via photographic interference waves embedded in special fine grain emulsion**
  - Extant, century-old photos at Musée de l'Élysée in Lausanne, Switzerland, today show *perfect* color reproduction
  - no dyes to fade
  - no filters for color
  - too slow then for practical use
  - plates could not be directly reproduced
- **WORF overcomes 19th Century limitations with modern photochemistry, computer controls, & micro-electronics**



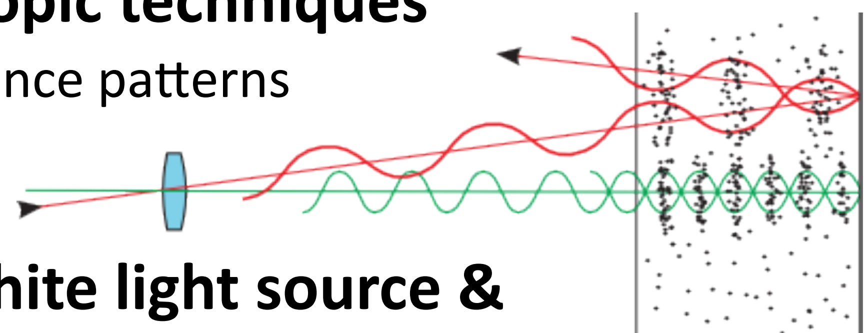
Source: Nobel Prize Committee

Color photographs made by Lippmann in the 1890s on ultra fine-grained silver halide emulsions.

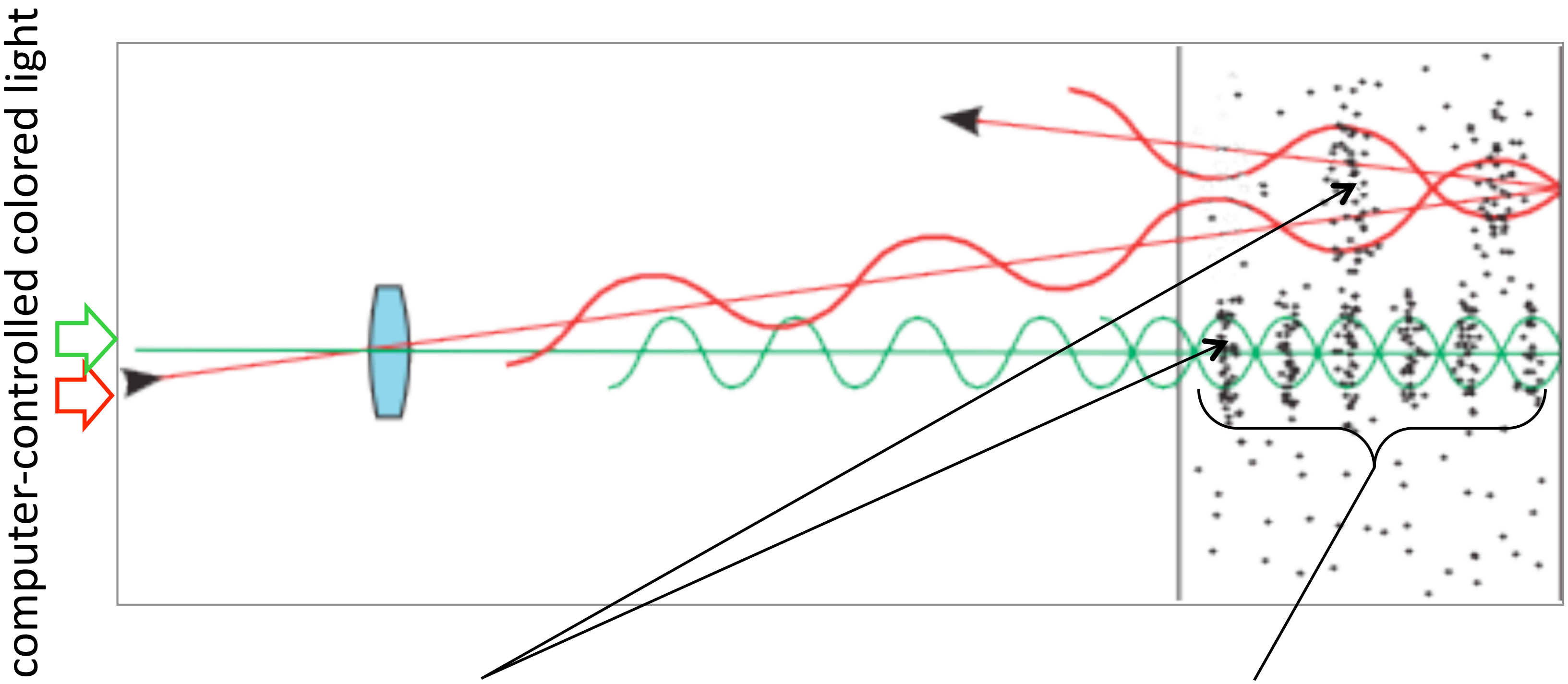
Basic Elements of WORF System

- **Different wavelengths for different numerical states**
  - Practically — 8-16 states in per pixel node (constellation)
  - ≈*1*µm diameter to accommodate light wavelengths
  - Data *cannot be erased or altered,* no dyes to fade
  - Theoretically *>1000 states per pixel node*
- **Data is written using spectroscopic techniques**
  - Embeds different vertical interference patterns in the emulsion representing standing waves



- **Read-out is via conventional white light source & conventional precision mechanisms to locate data**
  - Interference patterns act as diffraction gratings
  - Detection applies conventional spectrometric apparatus
- **High signal-to-noise due to frequency detection**
  - Confined to one octave in visible region = no harmonics

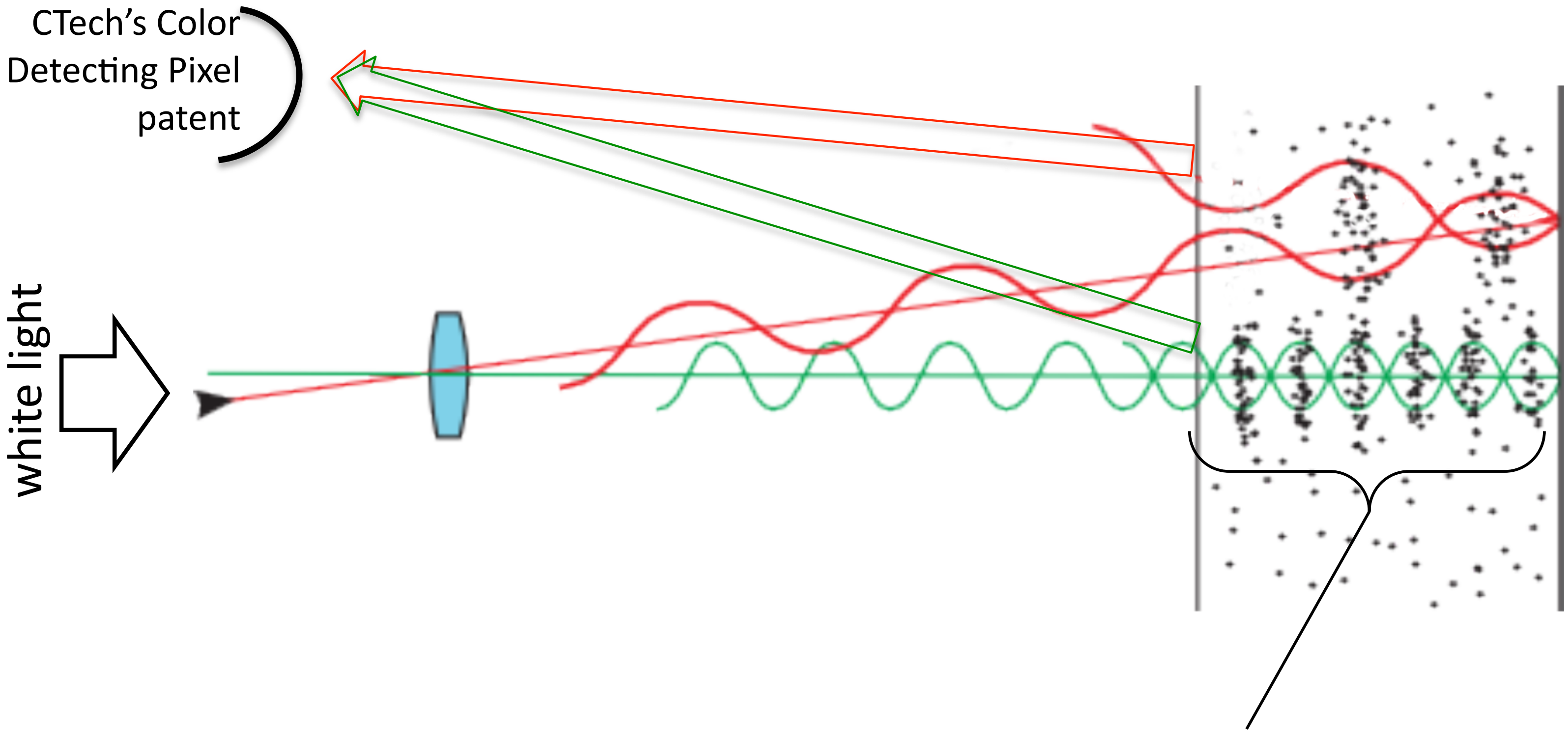
Writing Data via WORF



Interference patterns *are embedded in the emulsion as a pseudo standing wave — representing colors (wavelengths) — using* ***CTech's patented computer controlled spectroscopic techniques***

Different colors allow ***8-16 bits of data*** *to be stored per pixel*

Reading WORF Data

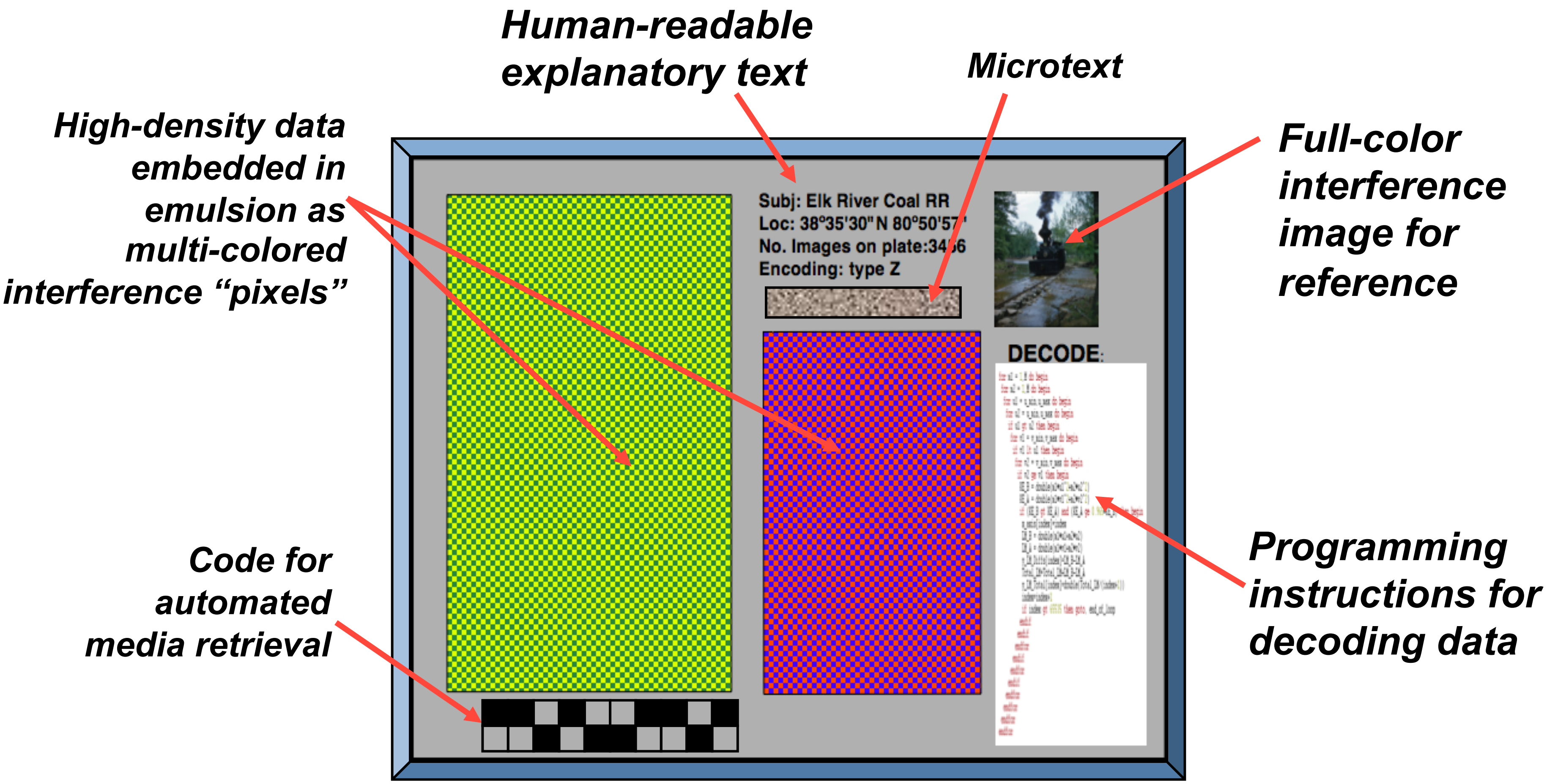


*Read-out uses **conventional light** & **CTech's patented precision mechanisms** to detect & randomly locate data in the diffraction gratings embedded in the emulsion*

*Different colors allow **8-16 bits of data** to be stored per pixel as a standing wave*

One possible WORF media format

(other possibilities include disks, wallet cards, memory sticks, etc.)



Technical Background on the Lippmann Process

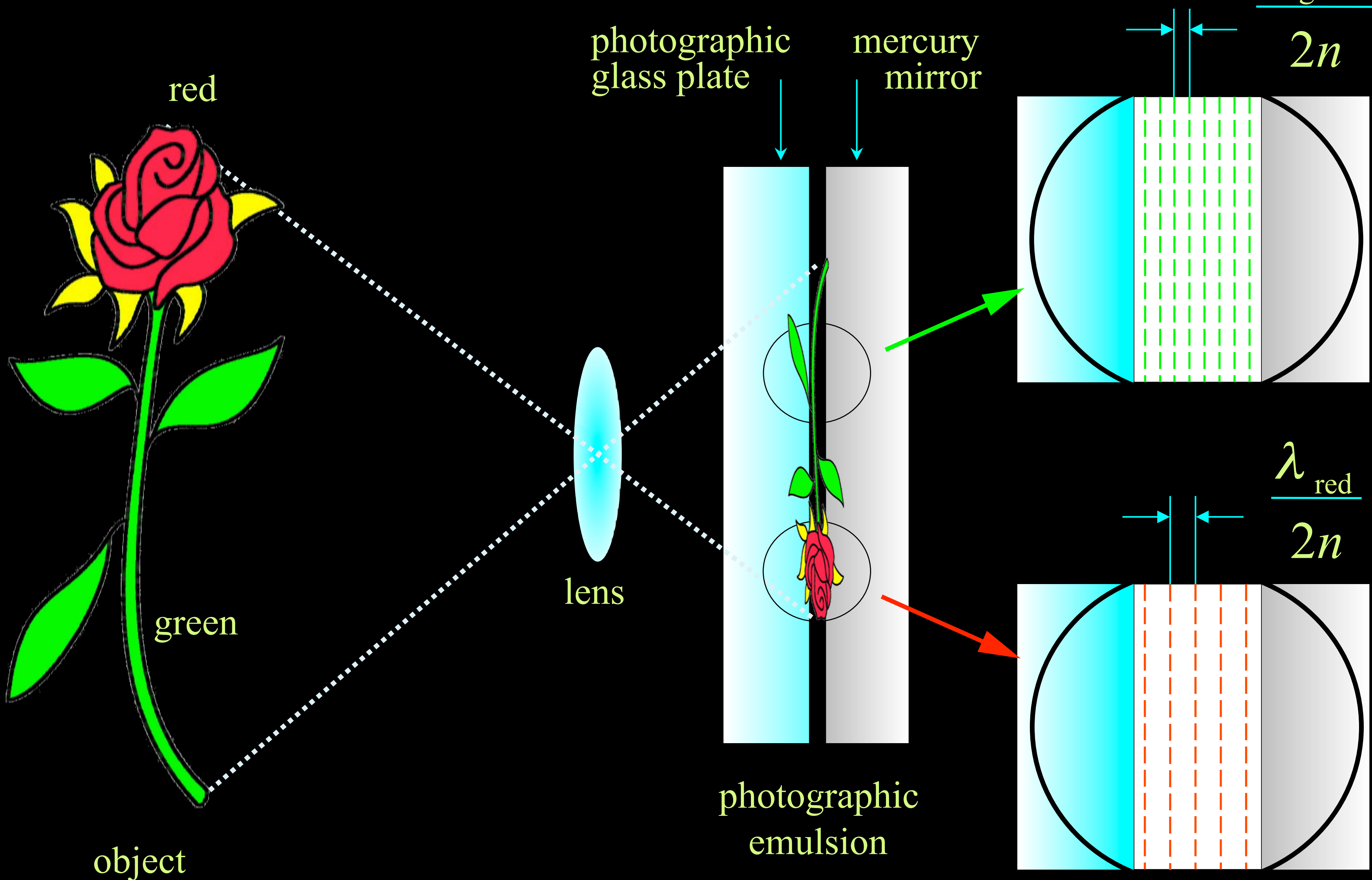
In 1886, a French scientist at the Sorbonne, Prof. Gabriel Lippmann (better known as Mme. Curie's thesis advisor), invented a unique photographic technique based on the phenomenon of light interference: recording color photographs in a black-and-white photographic emulsion. However, the practical execution of the technique at the time was extremely difficult due to slow emulsions and rudimentary light sources. In 1891, after years of experiments, Lippmann presented his invention to the Academy of Sciences [*ref. 1*]. In 1908 his invention won the Nobel Prize for Physics. His 125-year-old photographs are still extant in the Musée de l'Élysée in Lausanne, Switzerland, and show no signs of deterioration or fading as they use no dyes, just silver halide grains.

The principle of Lippmann photography [*ref. 2*] is presented in the next slide. It involves imaging a colored object, illuminated by white light, on an extremely fine-grained black-and-white photographic emulsion. Lippmann created the standing wave structures in the emulsion by placing the emulsion in contact with liquid mercury serving as a mirror. Incoming and reflected light therefore interferes to form an interference pattern of planes, parallel to the emulsion (following slide).

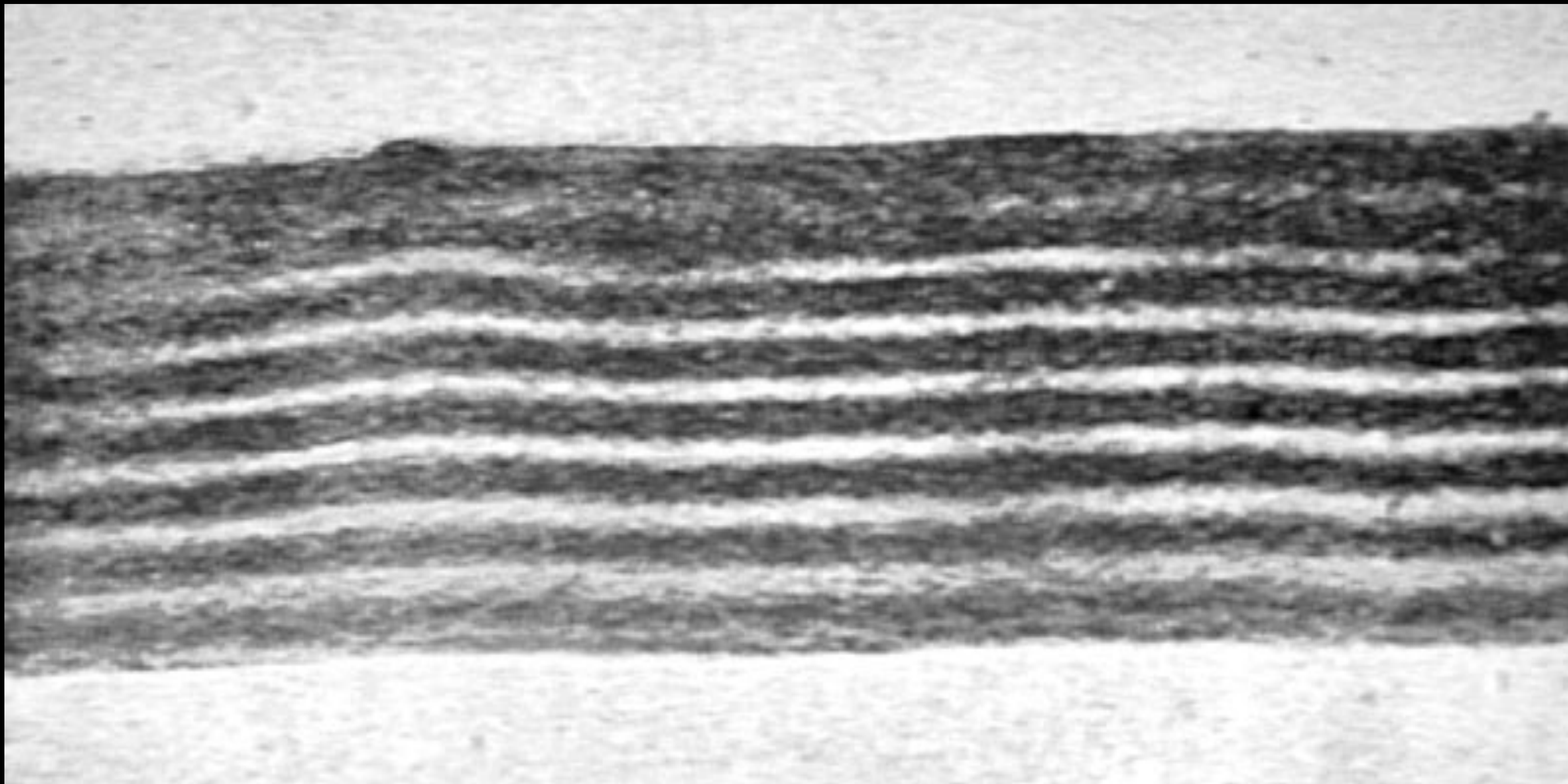
1. Lippmann, G., "Sur la théorie de la photographie des couleurs simples et composeés par la methode interferentielle," J. Phys. (Paris), vol. 3, no. 97, 1894.

2. Ives, H.E., "An Experimental Study of the Lippmann Color Photograph," Astrophysical J., Vol. 27, 1908, pp. 325-352.

Principle of Lippmann color photography



http://sechtl-vosecek.ucw.cz/en/expozice5.html



Microphotograph of a cross-section through a Lippmann emulsion showing the laminar layers.

http://sechtl-vosecek.ucw.cz/en/expozice5.html

Development Plans

primarily an engineering & software effort

- Identification of suitable archival photosensitive emulsions & processes
- Re-purposing precision drive & spectroscopic mechanisms for rotating & stationary media
- Re-purposing firmware & software drivers
  - —Develop data storage architecture
  - —Error checking methodology
- Construction of working prototype

Penn Research Directions

- **Develop data storage architecture**
  - Characterize wavelengths/radiance levels of data on substrate
  - Encode bitstrings of data, *e.g.*, as constellations of frequency + amplitude
- **Validating WORF components and prototypes**
  - *E.g.*, design checksums and error-correcting codes for centuries/millenia of reliability

Subsequent Phases for Core WORF Technology

- **Deployment with no moving parts**
  - Credit card size
  - Wearable device
- **Temporal chemical (time stamp)**
  - Proof that data was recorded at a certain time:

    ie: if a portion of the emulsion is not thoroughly fixed its rate of fading could be used to determine date of origination, subsequent alteration, or deliberate expiration

Summary

WORF's Advantages for Archiving

- **Extremely long lifetime**
  - Future proof
  - *human readable decoding information on media*
  - Ease of portability
- **Lower Lifecycle costs**
  - Superior energy footprint
- **Greater data integrity**
  - Not alterable
- **Random access media for automated rapid retrieval**

WORF Archival Features Compared to Conventional Storage

WORF

- **Applies *proven archival media* technology**
  - Scalable over a very long lifetime
  - Stable substrate
- **Lower lifecycle Costs**
  - Room temperature storage
  - Low power consumption
  - Permanently stable — no periodic re-writing
- **Future proof**
  - Media contains *human readable instructions* for future decoding
  - No need for continual array authentication
  - Not dedicated to proprietary hardware
- **Data cannot be erased or modified**
  - **W**rite **O**nce, **R**ead **F**orever
  - Protected from threat of remote hacking/ corruption
  - Error correction significantly reduced

Conventional

- **Magnetic Disk and Tape**
  - Data and Media degrade over time
  - De-magnetization/EMP etc.
  - Consumes energy primarily for thermal limit constraints
  - Reader obsolescence
  - Mechanical failure
- **SSD**
  - Sensitive to gamma rays, neutrinos, EMP, etc.
- **Cloud Services**
  - Reliability and availability
  - Accessibility requires network connectivity
  - Not true archiving according to accepted definitions

Media Comparisons

Media Type	Est. Lifetime	Comparable Lifecycle Costs	Data Corrupt-ibility	Random Access	Archivable	Future-Proof
CD/ DVD	5-10	Medium	No	Yes	Yes	No
SSD	10	Unknown	Yes	Yes	No	Unknown
MAG Tape	20	Very High	Yes	No	Yes	No
MAG HDD	3-15	High	Yes	Yes	Yes	No
Cloud	Contractural	Unknown	Yes	Network Dependent	Unknown	Unknown
WORF	**>1000**	**Very Low**	**No**	**Yes**	**Yes**	**Yes**

WORF – Write Once, Read Forever

WORF intellectual property and patents

- 3 US Patents issued applicable to WORF's write and read technology:
  - 6,985,294 (full spectrum color projector)
  - 7,437,000 (full spectrum color detecting pixel camera)
  - 7,521,680 (electromagnetic spectral-based imaging devices and methods)

- Patent pending on WORF archival data storage technology:

  "The invention combines several techniques applying high-resolution photosensitive emulsions for the long-term, archival storage of data, images and text. Data is stored as vertical interference patterns of multiple frequencies in a photographic emulsion. Read-out of the stored data uses a precision mechanism to locate and decode stored data."

Long-term, Permanent Storage of Data

Write Once, Read <u>Forever</u> (WORF)

University of Pennsylvania

New York University

Creative Technology LLC

Contacts:

Richard Solomon <rsolomon@seas.upenn.edu> 413-267-5171

Eric Rosenthal <eric@creative-technology.net> 732-817-1720

Jonathan Smith <jms@central.cis.upenn.edu>

Clark Johnson <clark.johnson@creative-technology.net> 608-467-7835