Digital Restoration of Erased and Damaged Manuscripts

Roger L. Easton, Jr.
Chester F. Carlson Center for Imaging Science
Rochester Institute of Technology
Rochester, NY 14623-5604
easton@cis.rit.edu

Keith T. Knox
Boeing LTS, Maui, HI
keithknox@compuserve.com

Description: Modern imaging technologies may be applied to the problem of reading "missing" text in manuscripts with historical and cultural significance. This talk describes the various techniques used and results obtained for several documents that have been damaged or degraded by different mechanisms. Among the documents are the Archimedes Palimpsest and a Florentine prayer book.

Roger L. Easton, Jr. is on the faculty of the Chester F. Carlson Center for Imaging Science at the Rochester Institute of Technology. His research interests include application of digital imaging to historical text documents and manuscripts.

Introduction:

Many manuscripts of historic importance are difficult to read because the text has been defaced by accident or even deliberately erased. However, information about the content may still exist even if it is invisible to the human eye. The eye is sensitive only to wavelengths between 400 nanometers (blue light) and 700 nanometers (red light). Fortunately, digital imaging technologies now exist that often allows the text to be recovered. Depending on the condition of the manuscript and the type of information to be recovered, the imaging systems may work in reflection, transmission, or fluorescence. We have demonstrated recovery of significant text from several important manuscripts that were degraded by various means. Among these are the Dead Sea Scrolls, the Archimedes Palimpsest, and a Florentine prayer book. The paper is organized by manuscript, rather than by technique.

This paper is merely the briefest introduction to the methods described here. Further details are available in some of the papers listed in the references or by contacting the authors.

Fragments of the Dead Sea Scrolls:

In 1997, we were invited by Dr. James Charlesworth of the Princeton Theological Seminary to image some fragments of the Dead Sea Scrolls that were in the custody of Father John Peter Meno, the Archdiocesan General Secretary of St. Mark's Syrian Orthodox Cathedral in Teaneck, New Jersey. These fragments had been purchased by Mar Athansius Samuel, who was the Syrian Orthodox Archbishop of Jerusalem in 1947, when the scrolls were discovered. These fragments had been scattered about the floor of one of the caves at Qumran for nearly 2000 years and had suffered greatly from exposure to the elements. Two of the fragments are shown in Figure 1. Note that some parts of the fragments are readable, but much of the text is virtually invisible to the human eye because the contrast between the ink and parchment in the visible spectrum has vanished. We imaged these fragments using a digital camera whose light-sensitive element is a silicon charge-coupled device (CCD) that is

sensitive to wavelengths from about 300 nanometers to 1000 nanometers (one micrometer, also called one "micron") under illumination from a "broadband" tungsten incandescent light source. The range of recorded wavelengths was limited to approximately 850 nanometers to 950 nanometers by an infrared-transmitting filter. The contrast of the text is much enhanced, as shown by the images in Figure 2.[1,2]

Figure 1: Appearance of two fragments of the Dead Sea Scrolls under visible illumination, showing the significant degradation in the contrast between the text and parchment over much of the area.

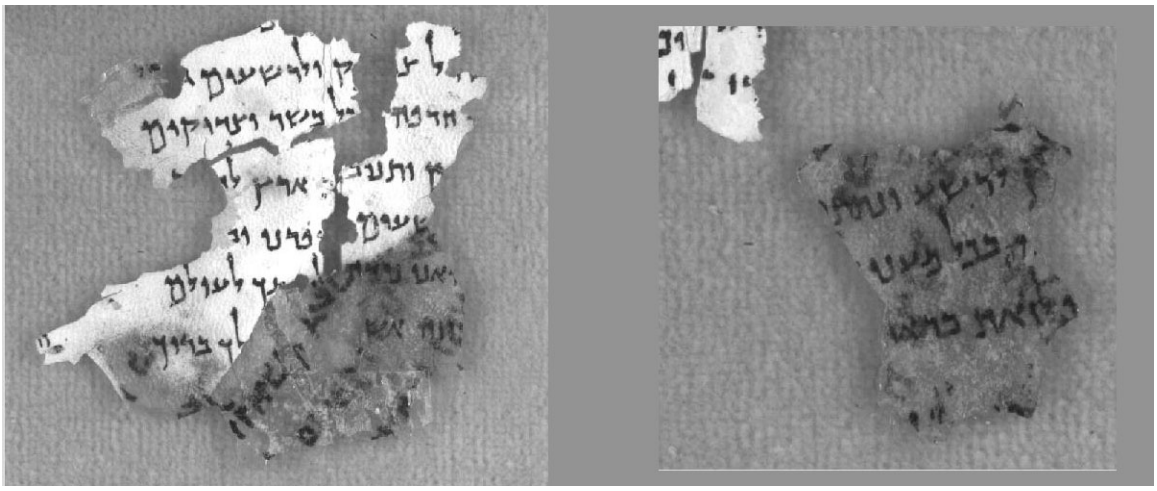

Figure 2: Appearance of the fragments in the near infrared region of the spectrum, showing that virtually all of the text is now readable..

Professor George Brooke of Manchester University in the United Kingdom reconstructed the original text from the digital images and identified the writings as a previously unknown hymn for a harvest festival.

Archimedes Palimpsest:

Archimedes was one of the greatest mathematicians, scientists, and engineers of pre-Renaissance times. He left a legacy of inventions, observations, and mathematical proofs that are used to this day. Our knowledge of Archimedes' work is indirect; his words and diagrams have been passed

down to us via handwritten texts that were copied from other copies by scribes over more than two millennia.

The oldest known link to Archimedes is the so-called *Archimedes Palimpsest*, which is a parchment codex from the 10th century that contains the oldest known (though partial) copies of seven treatises. The manuscript survived in its original form for only about 200 years, and suffered greatly because of its location in Constantinople, which stood at the boundary between the Islamic east and the Christian west. The city thus was an intermediate waypoint of the Crusaders during their journey to the Holy Land. Tragically, Constantinople was sacked during the Fourth Crusade in 1204 and its libraries were looted. Rather than being destroyed outright, texts usually were recycled to provide writing materials for new books because of the expense of making parchment. At that time, the Archimedes manuscript was disbound and erased (by washing off the original ink). The rebound parchment formed an overwritten manuscript, or "palimpsest": the *Euchologion*, which was a Christian prayer book. From the colophon of the Euchologion, we now know that the prayer book was first used on Easter Sunday in 1239. At various times in the next 800 years, the manuscript was in libraries at the monastery at Mar Saba, at a church library in Jerusalem, and finally back in its original home, now renamed Istanbul.

The original Archimedes text was lost to scholarly study until identification by Johan Ludvig Heiberg, a Danish philologist who studied the manuscript in 1906 using tools no more sophisticated than black-and-white photography and a magnifying glass. His discovery was deemed of sufficient import to be reported by the New York Times on July 16, 1907.

During the upheavals in Europe in the early 20th century, the manuscript disappeared again and was feared lost forever. However, it resurfaced into public view in the late 1990s, when the French family that had taken possession after World War I consigned it for sale by Christie's auction house. The authors were approached in the summer of 1998 by Dr. Hope Mayo, a consultant on manuscripts for Christies, with a request to image some pages of the Palimpsest for inclusion in the auction catalog. These images demonstrated the existence of text on pages that had previously not been identified as palimpsested.

The manuscript was auctioned by Christie's in New York on October 28, 1998. The new owner is an anonymous American collector who has pledged to make the manuscript available to scholars., and he has fulfilled this pledge by lending it to the Walters Art Museum (WAM) in Baltimore for study. Dr. William Noel, the curator of manuscripts and rare books at WAM, assembled an international team of conservators, scholars, and imaging technologists to study the manuscript and recover any knowledge that still resides on its pages.

Scholars have long used ultraviolet illumination to read palimpsests because the organic material in parchment fluoresces under such lighting; the parchment absorbs the "short" ultraviolet light and re-emits longer-wavelengths, often in the visible region of the spectrum. The "bluish" visible fluorescence is generated from within the parchment (and thus "beneath" the inks). The faded original text attenuates both the incoming ultraviolet light and the exiting visible light. This "double-pass" attenuation enhances the visibility of the original text in images taken in blue light under ultraviolet illumination, to the point where the visibility of the original text often is comparable to that of the overwriting. This is shown in the images taken at different wavelengths of light shown in Figure 3.

We originally sought to segment the Archimedes text from the prayer book text, i.e., we wanted to "strip off" the overtext and leave the original text. To accomplish this, we used a scientific digital camera (a SenSys™ camera from Roper Scientific) that is sensitive over wavelengths from about 300 to 1100 nanometers. We collected images of each section using a filter that passed bands of wavelengths only 10 nanometers wide over the full visible region. We then processed those images to identify pixels that "belonged" to the original Archimedes text, the later prayer book text, and parchment. We could then create individual images of each "class" of object. An example is shown in Figure 4.

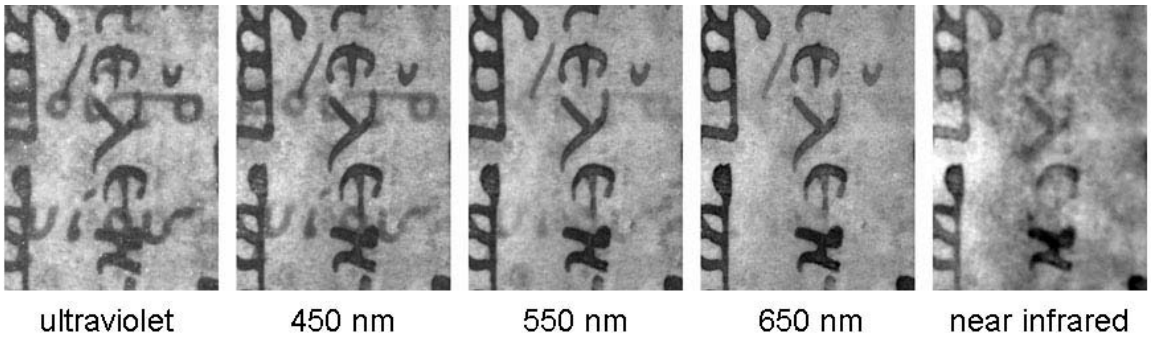

Figure 3: Appearance of a small section of the Archimedes Palimpsest under illuminations at different wavelengths. The original Archimedes text runs horizontally. Note the decrease in the visibility of the Archimedes text with increasing wavelength. (*Image produced by the Rochester Institute of Technology and Johns Hopkins University. Copyright resides with the owner of the Archimedes Palimpsest.*)

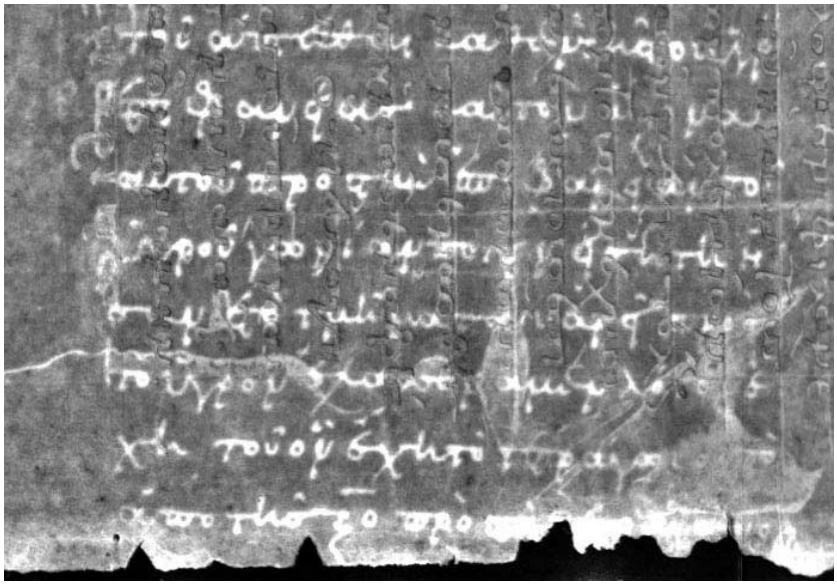

Figure 4: Section of one folio of the Archimedes Palimpsest after image processing to segment the original text (running horizontally). (*Image produced by the Rochester Institute of Technology and Johns Hopkins University. Copyright resides with the owner of the Archimedes Palimpsest.*)

When submitted to scholarly review by Dr. Reviel Netz of Stanford University and Dr. Natalie Tchernetska of Cambridge University, the results of this rigorous segmentation were judged to be in some ways "too good". The characters revealed by the processed images exhibit "gaps" that

may be due to any of several reasons. The gap may be a true feature of the character, or the ink may have flaked off or otherwise been removed, or it may be due to obscuration of the ink by an overwritten character. When reading the actual manuscript, the scholars can often tell which of these situations applies in individual cases, which they could not as easily do from the processed images.

Based on the assessment of the scholars, we adopted a simpler processing scheme that enhances the contrast of the Archimedes text and the color difference between the two texts. This method is based on the observation that the faded ink of the Archimedes text is "redder" than the overwritten text, and often is virtually invisible in red light, or in the red "channel" of an image taken under white-light illumination. In other words, the gray values of pixels that contain Archimedes text are approximately the same as the surrounding parchment under red light. The color of the overwritten prayer book text is much more neutral and therefore it is quite visible in all channels of the visible-light images. We take advantage of this observation by using a high-quality professional digital camera (Kodak DCS-760, 3040×2008 pixels) that takes color images (red, green, blue, or RGB) under visible and ultraviolet illuminations. Using this camera, we collect images with a spatial resolution of approximately 600 pixels per inch (25 pixels per mm). The images are processed by segmenting the color "channels": the red channel under tungsten illumination and the blue channel under ultraviolet illumination and balancing the contrast (mean and variance of the image histogram) across the frame. We then put the tungsten red image in the red channel and the ultraviolet blue image in the green and blue channels of a new "pseudocolor" image. The Archimedes text is "bright" (looks like parchment) in the red channel and darker in the green and blue channels, and thus appears with a reddish tint. The prayer book text is dark in all three channels and appears "neutral." An example is shown in Figure 5.

Figure 5: Pseudocolor image processing of f.48r. The original appearance of the section is shown at top (Archimedes text running horizontally). The pseudocolor image (bottom) exhibits enhanced contrast of the Archimedes text and the color difference that gives a "cue" of the source of the text (*Image produced by the Rochester Institute of Technology and Johns Hopkins University. Copyright resides with the owner of the Archimedes Palimpsest.*)

The scholars have successfully read significant original text from the pseudocolor images of the palimpsest created by this method.[3,4,5]

Florentine Prayer Book:

We were contacted early in 2001 by Evelyn Cohen, who wanted to recover information from the colophon of an illuminated fifteenth-century prayer book from Italy (Ms. 8224 in the Library of the Jewish Theological Seminary in New York). The colophon had been erased, probably by a subsequent owner of the book. As shown in the left side of Figure 6, the text is virtually invisible under visible illumination.

Ms. Cohen and Nellie Stavisky brought the book to our laboratory at the Rochester Institute of Technology (RIT) in February 2001 for imaging. During our first attempts, we were surprised to see that the text was no more visible by imaging the visible light fluorescing from the parchment. During this visit, two subsequent imaging sessions at the Seminary, and one session at RIT in May 2004, we applied many other combinations of illuminating and imaging wavelengths. Fortunately, we found that the text became at least somewhat visible when using ultraviolet illumination and imaging through an ultraviolet-transmitting filter. In other words, we view an "ultraviolet reflectance" image (right side of Figure 6). Ms. Cohen has been able to read most (but not all) of the text of the colophon from these images, so we assess the technique as helpful, but not completely successful.

We are currently investigating the use of optical character recognition[6] to try to recover the remaining text and to increase our confidence in the text thus far determined.

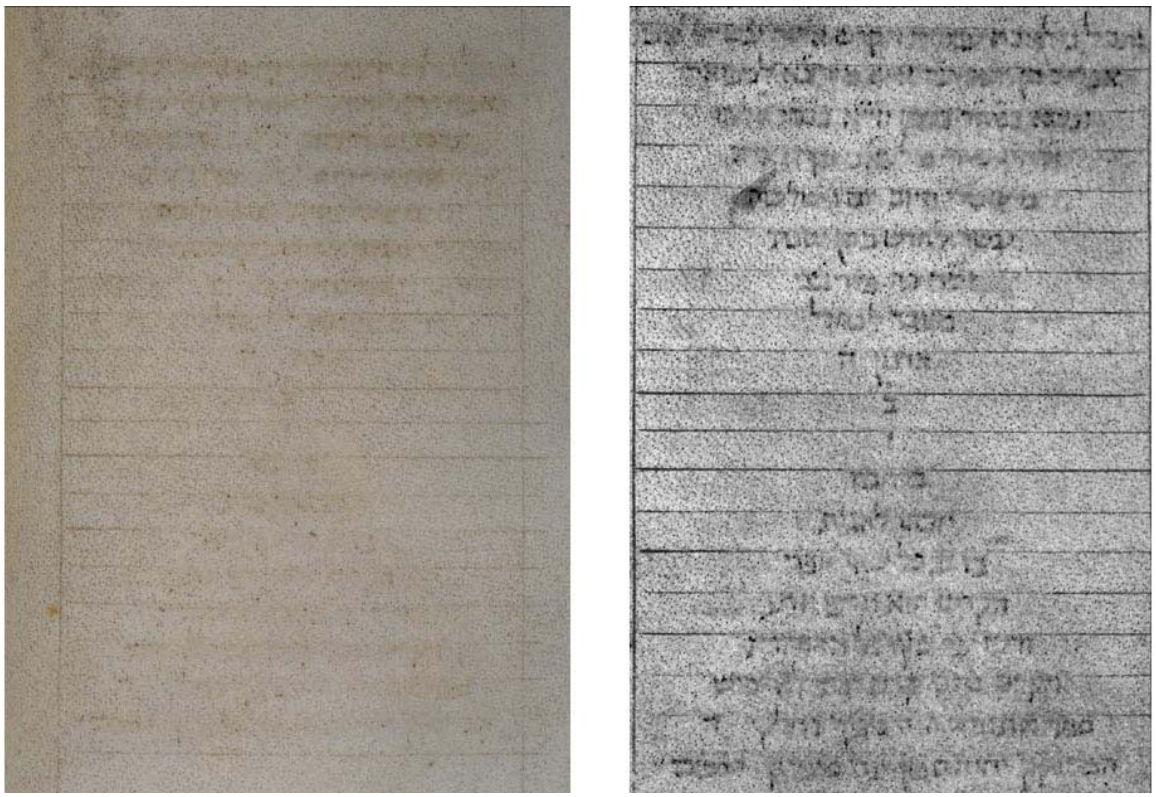

Figure 6: Appearance of the colophon of the Florentine prayer book under visible illumination (left) and using ultraviolet illumination through an ultraviolet-transmitting filter (right). Much of the text can be read from this image, though some significant words are still obscure.

Acknowledgements:
The authors are indebted to the owner of the Archimedes Palimpsest for supporting some of this work, and wish to thank the staff of the Walters Art Museum, especially Dr. William Noel, Dr. Abigail Quandt, and Dr. Cathleen Fleck for their invaluable contributions. We also thank Dr. Robert Johnston, formerly of the Rochester Institute of Technhology, Dr. William A. Christens-Barry of Equipoise Imaging, LLC, Mr. Michael Toth of R.B. Toth, Associates, Dr. Reviel Netz, Dr. Natalie Tchernetska, and Dr. Nigel Wilson, Evelyn Cohen, Nellie Stavisky We also want to thank Charles Dickinson, Derek Walvoord, Alvin Spivey, Pamela Kingsbury, and Matt Heimbueger for processing the images.

References:

[1] "Imaging the Dead Sea Scrolls", K. Knox, R. Johnston, and R.L. Easton, Jr., **Optics and Photonics News 8**(8), pp.30-34, August 1997.

[2]Roger L. Easton, Jr. Keith T. Knox, and William A. Christens-Barry, *Multispectral Imaging of the Archimedes Palimpsest*, **Proc. 32nd Annual Conference on Applied Imagery Pattern Recognition (AIPR)**, pp. 111-118, 2003.

[3]"*A New Reading of Method Proposition 14: Preliminary Evidence from the Archimedes Palimpsest*". Reviel Netz, Ken Saito, and Natalie Tchernetska, ***SCIAMVS* 2,** 9-29, April 2001.

[4]"*A New Reading of Method Proposition 14: Preliminary Evidence from the Archimedes Palimpsest (Part 2)*". Reviel Netz, Ken Saito, and Natalie Tchernetska, ***SCIAMVS* 3,** 109-125, August 2002.

[5]**The Works of Archimedes, Translation and Commentary, Volume 1: The Two Books *On the Sphere and the Cylinder***, Reviel Netz, Cambridge University Press, Cambridge, 2004.

[6] "*Variations of matched filtering for reduced noise amplification"*, Derek J. Walvoord and Roger L. Easton, Jr., **Proc. SPIE 5298**, pp. 59-69, 2004.