REPORT

REPORT REPORT REPORT

From Shelves to Digital Media

On Digital Archiving and Publication

SANGEETHA MENON

NATIONAL INSTITUTE OF ADVANCED STUDIES
Indian Institute of Science Campus
Bangalore 560 012 India

From Shelves to Digital Media
On Digital Archiving and Publication

Sangeetha Menon

NIAS REPORT R1 – 2002

NATIONAL INSTITUTE OF ADVANCED STUDIES
Indian Institute of Science Campus
Bangalore 560 012 India

Published by
National Institute of Advanced Studies
Indian Institute of Science Campus
Bangalore 560 012

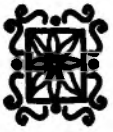
Price : Rs. 70/-

Copies of this report can be ordered from:

The Controller
National Institute of Advanced Studies
Indian Institute of Science Campus
Bangalore 560 012
Phone : 080-3604351
Email : mgp@nias.iisc.ernet.in

ISBN 81-87663-26-X

Typeset & Printed by
Verba Network Services
139, Cozy Apts., 8th Main, 12th Cross
Malleswaram, Bangalore 560 003
Tel.: 334 6692



INTRODUCTION

By the year 2008, it is expected that India would have 20 million personal computers and a 100 million Internet users, compared to the current figures of 4.3 million personal computers and 3.2 million Internet users. This expectation is also vindicated by the fact that in recent times there has been a swifter move towards using Information Technology in private and public sector services as well. Treasure-houses of cultural property such as national museums, archives and private collections, which possess a huge number of manuscripts of antiquity, are also touched by this movement towards an electronic era. What the new digital technologies have to offer is not just the better preservation of the archival material, now restricted to storage on the shelves, but also powerful means for collecting, accessing, distributing and sharing information cutting across spatial and temporal barriers. And all this is made possible essentially due to two breakthroughs in the area of communication. These are (a) the integration of the World Wide Web, photography, computing and electronic publication; and (b) the 'digital transcendence' made possible for the existing

set of electronic devices whether they be photocopiers, telephones, cameras or televisions. With the World Wide Web, which is currently the most interesting spectacle in cyberspace, the concept of localised information shifted to electronic information with new perspectives on communication, storage, retrieval and dissemination. The web provided a new way to link and access large classes of information of various kinds, and to introduce new kinds of applications.

This paper is a summary of the application of a simple technology that is cost-effective and user-friendly, namely digital photography, to professional storage and sharing of archival material such as palm leaf and paper manuscripts, print material, microfiche and microfilm. The project, which was initiated and developed at NIAS in 1998, demonstrated that, fast-deteriorating ancient manuscripts could be digitally preserved and shared with a larger community for many years to come by simpler and faster methods, while maintaining standards of quality.

DIGITAL ARCHIVING

Digital archiving involves converting any information from its analogue form to a digital format of the binary numbers 0 and 1. Once the information (whether it is text, graphics, voice or video) is converted to the digital format it is possible to store it without loss of information and use it for a wide variety of applications such as the World Wide Web.

Conventional means of converting a manuscript to the digital format was by using scanners. But this scene changed with the advent of digital cameras, which changed the concept of traditional photography as well. Photography has changed to a larger extent, although retaining its expression, in terms of technologies and applications. It all started with the advent of digital cameras in 1995 priced under $1000 and marketed for amateur and commercial use. The idea caught the attention of Internet and PC users since it opened up a new pathway connecting PC, photography and Internet publication, enabling cost effective storage, easy access and controlled sharing. Success in the digitization of manuscripts using digital cameras *(see Note 1)* proved that the breakthroughs in communication technologies could influence and integrate many different sources of information for sharing knowledge across a wide spectrum of users. It has also helped to solve, to a large extent, the major problem faced by the users of scanners, namely balancing between compression and quality of the digital image.

DIGITAL CAMERAS

The significant feature of a digital camera, which is also the aspect distinguishing it from a conventional film-based camera and scanner, is that the end product, namely the digital image, is received in the PC with the minimal number of steps in between. Digital cameras are flexible to the extent that the resolution of a picture to be digitised can be pre-set according

to the application it is meant for. This means that we can make choices ranging from print-quality high-resolution graphics to lower resolution Web quality graphics. It is this wide range made possible by digital cameras that has also increased its potential for the digital archiving of information sources like manuscripts.

Pixel Dimensions and Resolution of a Digital Image

A 'pixel', which stands for 'picture component', is the basic component of any computer graphic. It is the foundational unit of programmable colour in a digital image. A significant component of the camera which is involved in the actual capturing of the image is called the 'charged-couple device' or CCD. The CCD, which can be made up of a large number of sensors, converts light into an electric signal. One pixel is one dot which carries the information from the optical sensor on the CCD of the camera to the image itself. Each pixel displays one colour and all of the pixels together produce the digital image. The specific colour that is described by a pixel is a blend of the three colours of the RGB (red, blue and green) colour spectrum. The colour of one pixel is recorded by three bytes of data corresponding to each unit of the RGB spectrum. The resolution of the digital image is decided by the number of pixels it contains. This means, the more the number of pixels, the better the clarity of the digital image. It is the grid of a vast number of pixels (called the bitmap) which we recognize as the image on the display screen.

The resolution of a digital image is to be understood within two parameters such as the *pixels per inch* (ppi) and the *dots per inch* (dpi) of the print size or the display screen. For a usual 15 inch PC monitor the resolution is set at 800 x 640 *dots per inch*. When an image with pixel dimensions of 640 x 480 fills a screen of a 15 inch monitor on a larger screen each pixel tends to appear distinct and larger (this phenomenon is called *pixelisation*), if the setting of display resolution is not changed accordingly. At the same time, a display screen which can support a maximum resolution of 1280 x 1023 dots can also support 1024 x 768, 800 x 600 and 640 x 480 resolutions. But if the pixel dimensions of the image are lower and the display setting of the screen is larger, the image would be viewed as pixelised. Low resolution images will appear larger (and pixelised) than their print dimension if they are viewed on the display screen set to a high resolution. It is important, therefore, to determine the pixel size of the image according to the intended application.

The input resolution is measured using two numbers, if the unit is 'pixel'; or one number, if the unit is 'dot'. In the first case we can say that the resolution of an image is the number of pixels expressed in terms of the number of pixels on the horizontal axis and on the vertical axis as seen on a display screen. This would correspond to the area resolution of the image. By the second method, we can measure the *dots per inch* of a digital image. The measurement of the digital image

by this unit corresponds to the print size of the image. A higher resolution image does not mean a larger image. The size of the digital image could be viewed at chosen dimensions. However, a higher resolution digital image could be printed at larger sizes without pixelisation when compared to lower resolution images.

These two ways of measuring the resolution as the number of 'dots per inch' (as *dpi*), or as number of pixels in fixed area (such as 1600 x 1200) also influence two different working needs namely:

i) the input from the digital camera, and,
ii) the final output, which is the digital copy of the original.

The input from the digital camera, which is also the identifying feature of the camera, is measured by the unit 'Megapixels'. The term 'Megapixel' refers to one million pixels. A one Megapixel digital camera has an image sensor which is made up of one million pixels. The number of pixels is determined by multiplying the number of pixels on the horizontal by the number of pixels in the vertical. For example the image taken, set to the best resolution, with a one Megapixel camera, is 1024 x 1280, which amounts to 1,310,720 pixels. The individual pixels in the sensor record different colours and together result in the representation of the image. Therefore the larger the number of pixels in the image, the better its resolution is.

The quality of a digital image, apart from resolution, is dependent upon two other factors. The first is the size and physical condition of the original graphic itself, the second is the size of the display screen. An approximate relation between the camera resolution, print resolution and display resolution is shown in *Table I.*

Table I: Resolution, print and display size comparison

Camera resolution	Approx. maximum print size(in.)	Approx. maximum display size(in.)
640 x 480 (0.3 megapixel)	3 x 5	7 x 9
1152 x 864 (1 megapixel)	3 x 5	7 x 9
1280 x 960 (1.2 megapixel)	5 x 7	13 x 18
1600 x 1200 (2/2.1 megapixel)	8 x 10	16 x 22
2048 x 1536 (3/3.3 megapixel)	11 x 14	20 x 28

Features and Prices

In the case of digital archiving of manuscripts the ultimate goal is to get all the information contained in the analogue form. And since it is often difficult to pre-distinguish between the 'noise' (the information which is not part of the original content) and signal in the content of the digital copy, the optimal option would be to record as much information as possible of the original. This would also be helpful in the case of paper and palm leaf manuscripts that contain text which is difficult or cannot be read under normal conditions. If the pixel dimension of the digital image is higher, it could be viewed (enlarged) at higher degrees of zoom value until it begins to pixelise. In some of the digital cameras data interpolation is carried out with complex algorithms producing

enlarged images to some extent. However, since no algorithm can create information (could only 'guess', at the most) that was not there in the original, the interpolated images lose clarity after a certain amount of enlarging. When the digital image starts to pixelise is directly dependent on the resolution of the digital image. Therefore, the more the number of Megapixels, the better is the camera for capturing and comprehending more information.

There are various models marketed by different companies with a range of Megpixels (*see Note 2*). The prices of most of these brands increase with the increase in the number of pixels. The various models also differ mainly with regard to the storing capacity of the memory card supplied (the range is from a 4 MB to a 128 MB card), degree of zoom, mode of interfaces and transference (either by a serial port or a floppy drive adapter or both), batteries (re-chargeable lithium/nickel hydride battery/4 AA batteries). Most of the cameras come with common features like auto or manual focus, flash, manual and auto settings for adjusting white balance and LCD panel/screen. Almost all models are compatible with both Windows and Macintosh operating systems.

Internal Features, Memory Cards, Image Formats and Data Compression

The images captured by the digital camera are stored in devices called 'memory sticks', 'compact flash cards' or 'smart media

cards'. Smart cards also known as 'solid-state floppy disc cards' are very small, light and thin with average dimensions of 45.0 x 37.0 x 0.76mm, weighing about 2 gm. These cards have small differences in their size, although all perform the same function. Apart from these cards, some models use a 3.5 inch 1.44 MB floppy disk that fits into a normal disk drive on a PC. The advantage of floppy disks is that there is no need of an adapter to transfer files to the PC. The disadvantage is that only 2 images can be taken per floppy disk, if the resolution is set to the highest value. The other professional storage devices are Super Disks which can take information over 80 times greater than a normal disk, Iomega Clik disks which are 40 MB in size, and PC cards which can hold up to 2 GB of information. These three storage devices are not integrated with most of the cameras and are yet to be available for professional use.

A digital camera is different from the conventional camera with respect to its internal features such as the CCD. Each sensor of the CCD records the information for one pixel. Two different settings of the camera determine the quantity of light that strikes the CCD, namely the shutter speed and the aperture setting. The aperture setting determines how much light hits the CCD and the shutter speed determines how long the light stays on the CCD. The digital information in the individual sensors then passes through the internal filters such as white balance, colour etc, and is finally combined by the internal

memory of the camera leading to the image compressed in JPEG format. The images are then either transferred to the memory card one by one after every shot, or a certain number of images are held together if the camera has an internal buffer. In the second case it is possible to capture multiple images in a row with less time in between. Most of the digital cameras have automatic functions for focus, white balance and exposure time. A significant feature of the digital camera, when we talk about digitising manuscripts under different light settings, is that it is possible to adjust the white balance. (*see Note 3*).

The number of images which can be taken at any given time is dependent on two factors, i.e. the memory capacity of the storage device and the resolution at which the image is captured. A comparison of the resolution set, memory capacity of the card and number of images which can be taken is shown in *Table II*.

Table II: Comparison of memory card capacity, resolution and number of images which could be taken

CAPACITY OF THE MEMORY CARD (MB)	STANDARD RESOLUTION (pixels)	NUMBER OF IMAGES
4	1600 x 1200	8-10
4	1024 x 768	14-27
4	640 x 480	27-38

Once the capacity of the storage device is exhausted the digital images are transferred to the PC or the chosen device by a Serial or a USB port using an adapter, so as to re-use the storage device. The most common and preferred adapter is of the size of a 3½ inch floppy which can be recognised by the PC through the floppy disk drive. The used memory card is inserted into the adapter which is then inserted into the floppy disk drive. The images are downloaded to the hard drive of the PC. Thereafter the data in the memory card can be formatted and the card re-used.

The default image format in which most cameras capture the image is the JPEG format, which is named after the committee that designed it, the 'Joint Photographic Experts Group'. The JPEG compression algorithm can achieve a 10:1 or 20:1 compression ratio, depending on the compression settings, without significant loss of information. Some cameras feature apart from the JPEG format, another compression format called TIFF (Tagged-Image File Format) (*see Note 4*) which is a popular image format for scanned images. The one

disadvantage with TIFF files is that there is a significant increase in the time required to open and view those files. But the definite advantage is that being a *non-lossy* compression method, much less information is discarded.

Of the different compression algorithms, there are two broad categories called 'lossy' and 'non-lossy'. One of the *non-lossy* compressions called the LZW compression (*see Note 5*) can compress data in the ratio of 2:1. The JPEG is a *lossy* compression method and is the most popular format used by digital cameras. It retains all colour information in an RGB image but compresses file size by selectively discarding data. The JPEG image is automatically decompressed when opened and viewed.

DIGITAL PUBLISHING

New information technologies, in particular the World Wide Web and the Internet, have changed the world of digital or electronic publishing fundamentally. What is involved in 'digital publishing' (popularly called 'electronic publishing') is primarily a network-based publication of text, graphics etc. Most of the programs for digital publishing are developed so as to be compatible with 'hyper text mark-up language', the back-bone language of the Web. The number of applications based on digital publishing is increasing with the increasing popularity of the Web and the rapid advancement in Web technologies. Virtual/Digital Libraries, e-journals, e-texts, e-

prints, e-documents, e-magazines, e-books, e-marks are just some of these applications. There are also many e-catalogues and e-archives supplying users with organised data on various e-journals, e-texts and e-prints world-wide (*see Note 6*).

Online and CD-ROM/DVD publishing

Digital publishing can be classified mainly under two categories: online and CD-ROM. It is not that digital publishing could aim only at Net users. The medium of publication can be chosen depending upon the formulations and understanding that the individual or the institution/s have arrived at about copyrights, nature of content to be published, kind of audience targeted, long term goals etc. In the case of institutions and museums, which possess manuscripts of different kinds of material and content, CD-ROM publication could be the first step. It is also to be noted that the word 'CD-ROM' is used more as a popular acronym for a storage device, which need not be restricted to one kind, in today's electronic world when there is rapid change for catering to new needs and applications. Though CD is still the popular storage device for electronic data, there are also other devices such as, to name a few, Super disks, Iomega Click disks and DVDs (Digital Versatile Disks).

Guidelines for Digital Publishing

The fact that the number of Websites created per day worldwide is increasing in geometrical proportion evidences the popularity of communication and sharing of information through the Net. At the same time, there is rapid realisation that digital publication as well as net-presence could be a successful venture to keep up with ever advancing and changing Web technologies and copyright issues only if the guide lines are worked out well. The major factor that decides what kind of digital publishing is desirable is the content of information.

When we talk about digital publishing of manuscripts, we are at the outset faced with the two issues of information storage and information sharing. Both are equally important. However, the two require different strategies for being effective. This aspect will be discussed in detail under the following subheadings.

NIAS PROJECT FOR THE SETTING UP OF DIGITAL ARCHIVES

Background

In 1998 the National Institute of Advanced Studies began a rare manuscripts preservation project using new digital cameras. Conventional means of preservation used films

(microfilm and microfiche), which were both expensive and had short shelf lives. Called the NiDAC (NIAS Digital Archiving and Computerization) procedure, this method sought to introduce low cost and more permanent storage methods to digital archiving (*see Note 7*).

The project was initiated recognising the urgent need to digitally preserve ancient Indian manuscripts, since the lifetimes of both original manuscripts (palm leaf and paper) and their duplicates in microfilm are limited. The new technology developed in NIAS using still digital cameras and computer interfaces provides a cost-effective and less-time consuming method for storage and dissemination of the archived digital information. It also suggested and demonstrated simple and effective methods for digital publishing of the archived information (text as well as graphics).

Infosys-NIAS Project in 1999

The second year (1999) of the project which was funded by Infosys Ltd., had the following objectives:

(1) to demonstrate that the NiDAC procedure could be used for different kinds of preserved materials such as various kinds of manuscripts (paper and palm leaf), microfiche and microfilm, printed texts, books, paintings, drawings and similar objects,

(2) to assess cost, time and labour involved, and,
(3) to produce a CD-ROM giving examples of different types of preservation.

A CD-ROM was published, as a result of this project, which contained information about five different kinds of preserved materials. The contents of the CD included various selections from paper and palm leaf manuscripts, a digitally copied textbook (monograph) and illustrations from various sources (*see Note 8*).These selections demonstrated some of the many features of digital cameras and their wide range of applications. They also indicated the diversity of the NiDAC procedure and explained how digital cameras can effectively and inexpensively copy a variety of sources and contribute to digital publishing for easy access of information for all users or for registered users.

Another event during this project was the organisation of a three-day workshop (11-14 July 2000) for archivists, curators and librarians of museums and other centres in India. The Workshop acquainted archivists with new digitising methods and digital technologies for the preservation of rare manuscripts. Specific and suitable equipment was recommended along with essential software, and proper configuration of equipment was demonstrated from start to finish. The workshop also acquainted the archivists with digital publishing.

Ministry of Culture–Dept. of Science and Technology–NIAS Project 2000–2001

In 2000, a National project was initiated at NIAS jointly with the Ministry of Culture, and the Dept. of Science and Technology (Govt. of India) on setting up of digital archives at the National Museum, Delhi, the National Archives, Delhi, Salarjung Museum, Hyderabad and the Kancheepuram Mutt Library, Sri Chandrasekharendra Sarasvati Visva Mahavidyalaya, Kanchi.

This project had two goals. The first was to help initiate the digital preservation of manuscripts in the four participating institutions. This goal was achieved through three steps:

(1) The equipment bought by and tested at NIAS was installed at the four institutions,

(2) Selected staff in these four institutions were trained for digital copying, image-editing and electronic publishing, through workshops, hands-on-training courses etc.,

(3) A selected number of manuscripts were digitally copied (from the priority list made by each of the four institutions) and published in CD-ROMs.

Towards the fulfillment of this goal, NIAS installed the tested equipment at the four places, offered in-house training courses to the staff, and conducted workshops. A selected sample of digital images, according to the priority list, was published in CD-ROMs.

The second goal of this project was to initiate a systematic research toward image enhancement. When the thresholds of resolution (Megapixels) supplied by the digital camera are crossed, digital images look grainy or pixelized. To correct this problem, a systematic research investigation about how to enhance digitized images has been undertaken. To begin with, one-dimensional digitized data were considered. Four different methods of enhancing such data were implemented and compared. Research related to interpolation and noise reduction in 1D signals is nearing completion. Our goal in the near future is to extend this work to actual 2D images. Preliminary results in this regard are highly encouraging.

The Process and Steps from the Shelf to the Digital Medium

The process of digital archiving and publication through the method adopted by NIAS is simple and easy to implement. The process has five components:

(1) acquisition of equipment (hardware, software and peripherals),

(2) the process of digital archiving,
(3) the process of file maintenance and database management,
(4) the process of digital publishing and
(5) training courses.

(1) equipment constituting hardware, software and peripherals:

As stated earlier the method is based on innovations made in digital photography and tailored to the digital archiving of palm leaf and paper manuscripts, printed texts, books, paintings, drawings and similar objects. The equipment constitutes a digital camera (with recommended configuration), accompanying peripherals like memory card, adapter, cables etc., a PC (with recommended configuration), a CD-Writer, image editing software and electronic publishing software. The recommended configuration for the digital camera is: minimum 2 Megapixels resolution, a memory card (*Smart Media*) with minimum 4 MB storage capacity, adapter which is recognised by the floppy disk drive of the PC, and necessary cables. The recommended configuration for the PC is minimum 64 MB RAM, 200 MHz processor, 4 GB hard drive, CD-Writer with a writing speed of 3x, and Windows OS. Standard image-editing software is recommended for graphic manipulation.

(2) the process of digital archiving

The optimal conditions for the use of the equipment and getting maximum results are a well-lit area (with white light), adjustable stands for holding camera and manuscripts of different sizes, and uninterrupted power supply.

The digital camera can be used with power supply either from the mains or from a battery. The rechargeable battery carried in the camera usually gives power for about one hour. The memory card is inserted in the slot and the settings adjusted for resolution, white balance, macro etc. The camera is mounted on a stand, which can be adjusted according to the height requirement and the manuscript is placed in the appropriate position. The pages of the manuscript are digitised in order. It is recommended that when each page is digitised, the area to be covered could cross a few centimeters beyond the content, so as to facilitate proper editing of the images during the following step. The maximum resolution setting for a 2 Megapixel camera is 1600 x 1200 pixel dimension. A 4 MB memory card can hold from 8 to 10 images at an instance. After the images are digitally copied, they are transferred to the hard drive of the PC using the floppy disk adapter or serial cable. Once the digital images are downloaded to the PC, the memory card is re-formatted and made ready for copying more pages.

(3) File maintenenance and database management

Prior to the digitising of manuscripts, it is necessary that a proper filing and database management system is devised. Naming of folders and individual files and arranging files under proper folders are very important, since, in the long run, the number of manuscripts copied keeps increasing. Effective management of files and folders necessitates proper labelling and allocation. A simple format for naming and labelling is to have a mother folder named after the manuscript and subfolders for the original JPEG images, edited JPEG images and edited PDF files. There are also document management programs which allow for the systematic allocation and access of databases and resources.

(4) the process of digital publishing

Once the images of the manuscript are digitally archived, the next step is to edit them and get them ready for electronic publication. As mentioned earlier, digital data can be made available either online (restricted or universal) or on CD-ROMs. Hence the editing of the images should keep both the alternatives in mind. Since preserving the digital image with maximum amount of information is the first priority, the original digital copy of the manuscript should be kept intact, without any image editing, in one subfolder. This is to avoid losing information when the original

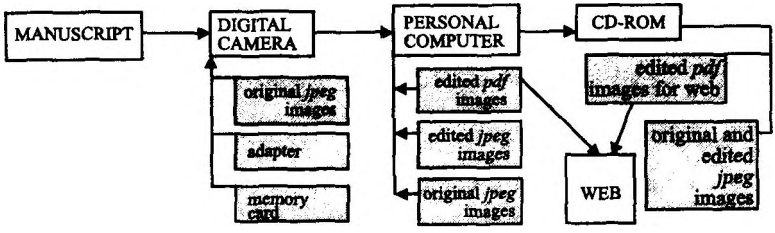
JPEG images are edited and saved. Since JPEG is a *lossy* compression method, every time an image is saved some information is discarded. Preserving a copy of the original digital images in a separate subfolder would prevent such loss.

There are two steps in the editing of the original digital images in order to ready them for digital publishing. The first is 'cropping', that is, exact sizing of the image, and the second is converting the JPEG format to PDF (portable document format) format (*see Note 9*). The conversion to PDF is made so as to lead to the linking of the images with html (hyper text mark up language), the prevalent [and present] language of the Web. If the immediate choice is to publish the images only as CDs the conversion of the edited images to PDF is not needed, though, in any case, PDF provides a very good format for systematic electronic display and presentation. To keep the option for online publishing open, it is advisable that the PDF of the edited files are also separately saved and arranged. A separate subfolder can be created for images which are downsized specifically for the Web. Editing of the images and their conversion to PDF are done using standard image editing programs. The actual allocation, linking, arranging and presentation of PDF files, to be read as an electronic document, are done using an

electronic publishing program. After editing of the images of the manuscript is completed, the mother folder inclusive of the subfolders, is copied ('CD-writing') to the CD-ROM.

The four steps above are outlined and schematically shown in Fig.1:

Fig.1: Processes: from Shelves to Digital Media



(5) modules of training:

The method described above, with all its steps, is easy to implement and can be improved and modified according to specific needs, advances in digital cameras and in Web, design and file management technologies, as per the archivist's imagination. Under the NIAS project, two kinds of meetings and workshops were organised as part of the training module. The first kind helped archivists and curators to review the collection of manuscripts at their centres, study the existing methods of preservation, and make a priority

list of manuscripts for digitisation. The second kind of meeting, which included workshops and in-house hands-on-training, helped the archivists to understand the state-of-the-art in digital archiving and publishing, and the ongoing research in NIAS toward image enhancement, and initiate the project at their own centres. One of the many results of the initiation of the project for digitisation of manuscripts at various centres was the realisation that the method is cost-, time- and labour-effective. Table 3 shows the approximate time to digitise and electronically publish 10 pages of a manuscript. It is also to be noted that once the equipment is installed and the optimal conditions are established in the work place, the time for the entire process involving digital copying, downloading etc. decreases to a considerable extent.

Table.3.: Steps and Approximate Times

Processing Steps	Time (min)
Setting, adjusting and fixing the camera	3-5
Setting the manuscript	3-5
Capturing 8 to 10 images set to a resolution of 1600 x 1200 pixel dimensions using a 4MB memory card	3-5
Downloading 8 to 10 digital images set to a resolution of 1600 x 1200 pixel dimensions using a 4MB memory card	1-2
Minimal file and folder management	2
Formatting of memory card	1
Minimal image editing and *pdf* format conversion for 8-10 images	10-15
Interlinking of 8-10 *pdf* files and their electronic Presentation inclusive of thumbnail and bookmark creation	3-5
Writing 8-10 original and edited images to CD-ROM (this factor depends on the file size and the speed of the CD-Writer and not on the number of images/files.)	3-5
Total time for the digital archiving and publishing of 8-10 images	minimum of 29 min and maximum of 45 min.

FUTURE OF DIGITAL ARCHIVING AND PUBLISHING

The significant feature of digital photography is that it is simple in concept and provides easy means of transfer between a wide range of electronic devices and applications. It also has a Net-friendly image format. The advent of digital photography is certainly a breakthrough for faster communication and easy storage. The first of its advantages relates to the fewer number of intermediate processes between the capturing and processing of images. Unlike traditional photography, digital photography is instant and flexible. The image of the object photographed can be seen on a small LCD screen at the back of the camera. There are no film costs. The memory card or memory stick can be reused after the digital images are downloaded to the PC.

These digital images are flexible since they can be stored in universal image formats (most of the cameras use the 'JPEG' format) which can be edited using standard image editing software and transferred between many different kinds of devices ranging from a Palmtop to a PC. Digital images also allow a wide variety of applications. They can be sent to another destination by e-mail, posted on a Website or copied to a CD-ROM.

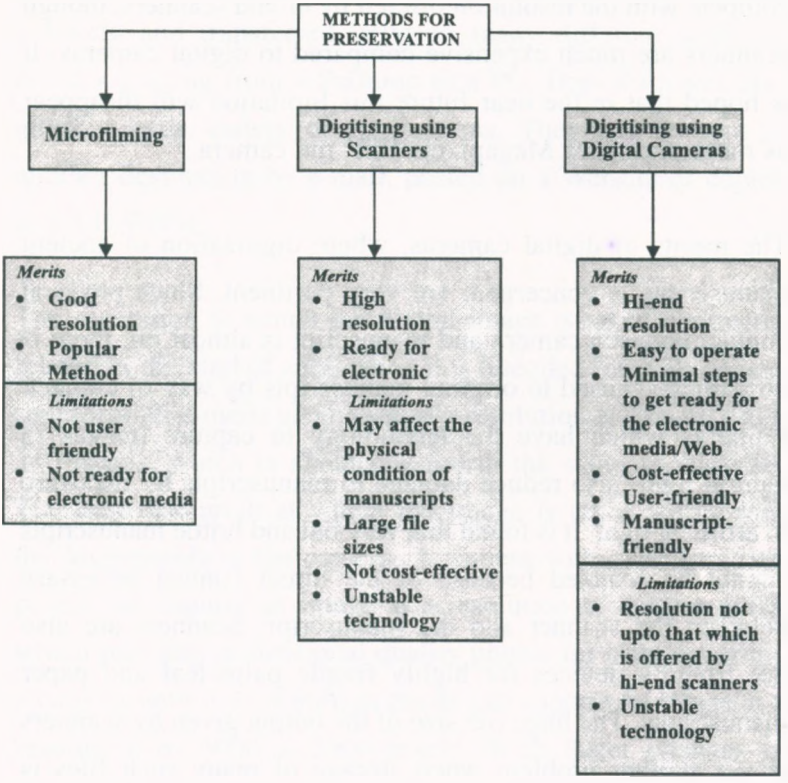
The resolution at which the digital image is taken is directly related to the kind of application it is intended for. Still images sent through e-mails can be set to a resolution as low as 320 x 240 pixels, which is about one-fourth the standard file size. The best resolution at which the image is taken varies with the Megapixels of the camera. A camera with over 2 million pixels can capture an image at a resolution of 1600 x 1200, which provides professional quality photos for printing, while a camera with over 6 million pixels can capture an image at a resolution of 3000 x 2000 pixels. With faster changes in technologies and emerging applications, the market for digital cameras is today witnessing the entry of newer cameras with more and more Megapixels.

A frequent question which is asked when the use of digital cameras is recommended for digital archiving is related to the use of scanners. Digital cameras have advantages as well as disadvantages over scanners (*see Fig.2*). The major limitation

of digital cameras with the present technology (though digital cameras can offer print quality resolution) is that they cannot compete with the resolution offered by hi-end scanners, though scanners are much expensive compared to digital cameras. It is hoped that in the near future this limitation will disappear as more and more Megapixels 'fill' the camera.

The merits of digital cameras, where digitization of ancient manuscripts is concerned, are very pertinent. Since physical contact between camera and manuscript is almost nil, there is no damage caused to original manuscripts by way of contact. Cameras which have the technology to capture images in natural light, also reduce damage to manuscripts by exposure to artificial light. It is found that very old and brittle manuscripts cannot be scanned because of the direct contact necessary between the scanner and the manuscript. Scanners are also not friendly devices for highly fragile palm leaf and paper manuscripts. The huge file size of the output given by scanners is yet another problem when storage of many such files is concerned. Three other merits of cameras over scanners are that they occupy less installation space, are portable and also produce the output (digital images) in much less time.

Fig.2: Comparison: Microfilming, Digitising using Scanners and Cameras



A comparison of digital archiving, using cameras, is made with the conventional method of microfilming. Though microfilming has a long-standing record as a stable means for preservation, it has limitations relating to access and sharing. The fear about microfilming is about its short shelf-life (about 30 years), and the physical damage that may be caused by a microfilming camera to a fragile manuscript. Microfilm readers are also not user-friendly, since their continuous use could be

strenuous and tiring for the eyes. The attractive feature about digital cameras is that by using them it is not only easy to copy and share information, but also that the process is cost effective and interferes the least with fragile and ancient manuscripts.

The highlight of the future of digital technology will be the increasing integration of information resources that will lead to new and creative ways of preserving, sharing and using information. The method suggested by NIAS for digital archiving and publishing is a stepping stone toward a revolution not only in the field of archiving and display of ancient manuscripts but also in research and application relating to India's lesser known contributions in the fields of art, science and humanities.

Acknowledgements

I am grateful to Prof.R.Narasimha, who has continously encouraged us in this work. I thank Prof. P. G. Vaidya and Prof. B. V. Sreekantan for many interactions. I thank Ms. K. R. Suchitra, Ms. Shilpa Prakash, Ms. V. B. Mariyammal and Mr. H. L.Vasu for their hard work and assistance.

The project on the 'setting up of digital archives' was funded by the Ministry of Culture and the Dept. of Science and Technology, Govt. of India. We acknowledge the support of Dr. Ramamurthy, Dr. Vaidyanatha Iyer, Mr. V. B. Pyarelal and Dr. V.K. Mishra. We also acknowledge the guidance of the Technical Advisory Committee to the project.

NOTES

1. See under 'NIAS PROJECT FOR SETTING UP OF DIGITAL ARCHIVES' and *Note 7.*
2. Visit *www.fujifilm.com www.canon.com www.sony.com www.nikon.com www.olympus.com www.gadgetguru.com www.cnet.com* for a review of the latest models of digital cameras.
3. 'White Balance' refers to what the digital camera determines as white colour in the image. According to different light conditions the colour white would appear differently. Most of the digital cameras have the option to pre-set the white colour.
4. Visit *www.fujifilm.com www.cnet.com www.adobe.com* for a comparison of graphic file formats.
5. LZW is an acronym for the names of its inventors: Lempel, Ziv and Welch.
6. Visit http://citd.scar.utoronto.ca/capa/e-publishing.html for a list of e-journals, other e-documents and e-catalogues.
7. A detailed report about the NiDAC project was published in *Current Science* 76:1299-1301.This article can be read at the URL *http://www.iisc.ernet.in/currsci/may25/articles12.htm*
8. The five kinds of preserved materials are the following: (1) The *Vakyarthachandrika* , a 529 page paper manuscript that is more than 300 years old. It is written in Devanagari, and is a commentary on a commentary of the *Brahmasutras* from the perspective of *Dvaita Vedanta*; (2) A paper manuscript on the *Bhagavad Gita* from Osmania University Library; (3) Selections of text, written in various scripts, and illustrations from manuscripts of the *Bhagavad Gita*, (4) A printed text on *Ramanuja Vedanta*; (5) Illustrations from various sources.
9. PDF (Portable Document Format) files can be read by Acrobat Reader which is the software freely distributed by Adobe, and is downloadable from the Web.

Dr. Sangeetha Menon is an Associate Fellow in NIAS. She has a Ph.D in Philosophy. Her research interests include new methods in digital archiving, Indic approaches to 'consciousness' and 'mind' studies in the context of current multidisciplinary discussions. She is an avid photographer and web-designer.
Her contact email is prajnana@yahoo.com and Homepage address is www.geocities.com/prajnana.

ISBN 81-87663-26-X