BVCC General Meeting

August 14, 2023

"Scanners"

Joel Ewing, President BVCC

Will Look At A Variety of Scanners

- All-In-One printers
- Flat-bed scanner
- Portable auto-feed scanner
- portable camera-based scanner
- Other Devices

 A PDF of the presentation slides for this program will be posted under Information → Presentations on the BVCC website (bvcomputerclub.org)





Some Generalities

- · Controlling all scanning options and supporting various file formats typically requires software from the vendor
 - Vendor software packages for most retail scanners are Windows only
 - Some more expensive scanners do advertise Mac support
 - Some scanners have driver support available for Linux, but application support tends to be "basic", not full featured support
- There are specialty scanners available, like hi-res auto-feed 35mm slide scanners, but these are hard to cost-justify for casual use
- Vendor software OCR support (for searchable PDF or DOCX formats) has improved in the last decade, but don't expect
 perfection
- There may be problems with out-of-support scanner software if support for a newer scanner from the same vendor is installed on the same Windows system.
- Best scanning parameters depends on objective and quality of original
 - 8x11 document readable at 150 ppi, if want to print reasonable-quality copies, go to 300 ppi minimum
 - standard photo prints: generally 300 ppi adequate
 - Items that are clearly not intended to be in color may scan better as grayscale to avoid color artifacts in image
 - Photo versus document optimization
 - 35mm slides experiment with 1200 4800 ppi. Higher res with marginal slides may just reveal more imperfections (poor focus, camera movement, improper exposure, surface damage) rather than useful details. Stop at point of diminishing returns.

All-In-One Printers (HP6978, \$95 in 2018)



All-In-One Printers

- Many ink-jet printers are print/copy/scan/fax combo
 - Scanner typically used to "align" the print heads
- Advantages
 - No extra desk space, bundled with other printer features
- Disadvantages
 - Slower
 - Scanner Hardware & Software may have fewer features; usually a smaller scanning surface

HP 6978 Office Jet Pro

- ADF (max 8.5"x14") and Scanner Glass (max 8.5"x 11"+) no auto-duplex scan
- Slow ADF of one 8.5x11 page: 28 secs
- Resolution dpi: 75, 150, 200, 600, 1200
- Color vs Grayscale
- Scan area: entire, Letter (8.5x11), A4(210x297mm), 4x6, 5x7. Use of ADF adds Legal (8.5x14) Can adjust guides on ADF, so perhaps can autofeed 4x6 and 5x7 cards, provided stock will bend 180°
- Type: Document, Photo, Multi-Item, Book, ID Card
- Detect Edges option
- Save as: Basic PDF, JPG, Searchable PDF, DOCX, TXT
- · Compression: none, low, medium, high
- High compression on PDF (300 dpi) vs none: size / 4
- Searchable PDF did fairly decent job but not 100%
- DOCX also seemed fairly close, some visual differences
- TXT format seemed to include significant garbage chars
- HP Smart software drawbacks
 - Requires an HP login to fire up software
 - Requires printer/scanner to be on the same sub net, not just reachable. Windows can actually print on the printer from a different sub net if IP address reachable and known, but full function scanning requires HP software.

Epson Perfection V500 Photo Flat-Bed Scanner (\$218 in 2009)

EPOON

EPSON PERFECTION V500 PHOTO

Epson Perfection V500 Photo Flat-Bed Scanner (\$218 in 2009)

Epson Perfection V500 Photo Flat-Bed Scanner (\$218 in 2009)

- Older model, but still works. ADF alternate lid attachment wore out and replacement parts no longer available.
- Scanning area 8.5"x 12" (legal paper required ADF)
- Switches from bottom lighting to back lighting in lid in order to scan slides and film of multiple sizes. Software converts color and b&w negatives to positive images
- Lid can stay open on its own, so flat sides of a 3dimensional object can easily be scanned

Epson Perfection V500 Photo Flat-Bed Scanner

- Scanning requires multiple steps
 - Setting scanning parameters and saving format, Preview scan, selecting the rectangular area for the image, actual scan
 - Selection of scan area via mouse or manual specification of WxH (inches or mm) and moving scan area with mouse
 - When scanning many similar, identical sized, identically positioned items, only the physical placement and "actual scan" needs to be repeated – auto-increment of name of saved file
 - Does auto-adjust for exposure on area selected for scan
 - Time for actual scan depends on size of scan area and resolution
 - Processing options may add time dust removal, Digital ICE (Image Correction and Enhancement) for film/slides (hardware+software)

Scanning Documents

- Use "Document" setting to optimize for text
- Unless extremely fine print, 150dpi is sufficient for a readable document, perhaps go to 300 dpi if object is to be able to print a copy of the original, or if is an important legal document
- If document does not use color, scan as greyscale rather than color saves as a smaller image and no color artifacts from imperfections in document surface
- Best saving format is PDF (retains size info, and supports multiple pages). PDF with OCR text also supported, but not that great with this older software and requires minimum of 400 dpi scan.

Scanning Images From Prints

- Specify auto-exposure "Photo"
- Suggest trying 300 dpi higher res may just reveal defects or texture of print surface
- If non-colored image, scan as greyscale, unless there is unusual coloring from age that you want to preserve.
- Saving format: most likely, jpg (has lossy compression).
 Sometimes when there is important descriptive info on back of print, I will save a front & back image of the print as a 2-page PDF; but one could just save the back as a separate image.

Scanning Images From Film/Slides

- For, consistency may want to specify scan size in mm as film has standard sizes (e.g, 35 mm slides have typical image of 23mm x 34mm or 34mm x 23mm, depending on camera orientation)
- Use an anti-static brush to clean dust from film before scanning
- Film/slides needs to be scanned with emulsion (duller) side up. On slides the side may be marked "this slide toward screen". If you get this wrong, image will be flipped horizontally left-to-right and any text in image will be backward. GIMP can be used to do a horizontal flip to correct, but less work to scan it right in the first place.
- Film potentially has much higher resolution than typical prints, but effective resolution depends on the quality of the camera, skill of its user (accuracy of focus, lack of camera movement), and film speed. Experiment depending on intended use – max resolution supported by scanner probably not the best choice and will result in excessively large image files with no improvement in quality.
- Typical saving format: JPG. If want to preserve image with no compression loss, saving in a lossless-WEBP or PNG format would be better, but my older software only supports TIFF for that purpose, which produces much larger files than either PNG or WEBP – can use GIMP or other image utilities to convert to other image formats.
- Digital ICE enhancement makes sense for slides to "fix" scratches, but adds significant time. I averaged almost 2 mins
 per slide scanning 35mm slides at 3200 dpi (over 33 hours for 1000 slides). Part of that time was overlapped with
 renaming slide images based on content or description written on slide. Each slide was 37 MB as TIFF. Would be 15MB20MB as PNG, slightly less as a lossless WEBP, maybe 5MB as JPG.



- This is from a slide that was scanned at 3200 dpi and illustrates two points:
- How the presence of text can clearly show whether the slide was scanned with the correct side up. Of course any asymmetric object in the picture will serve, provided you know the true left-right appearance of the object.
- How lack of sharpness in the image when individual pixels are not yet obvious can be a clear indication that there's no point in scanning at a higher resolution.

After Thoughts on Flat Bed Scanner

- If I had known in advance how many slides I would inherit in addition to my own, I might have seriously considered spending \$500-\$1,000 for a specialized automated slide scanner.
- Flat bed Scanner is good for scanning small pamphlets or booklets, but scanning large multi-page documents, especially if duplexed, is a pain.
- Items larger than the scanner surface can be scanned as multiple images, even sometimes stitched together successfully into one image with specialized software (like Hugin), but that doesn't always work and can be time-consuming.

Epson ES-200 (~\$200)

- Needed a better, faster way to scan: legal-sized documents, duplexeddocuments, unusual-sized receipts.
- Found Epson-200
 - Small, easily portable
 - Auto-feeds documents from 2" to 8.5" wide up to 36" long and automatically sizes the image or PDF appropriately, including postcards and envelopes
 - Scans both sides of each sheet at the same time, auto-detecting blank sides cuts scanning time for duplexed documents in half (5-6 sec/page).
 - Special feed for scanning business cards
 - Can save as JPEG, PNG, TIFF, multi-TIFF, PDF, Searchable PDF (does a better job on OCR than older software)



Epson ES-200 Scanning Multi-Page Document

EPSON

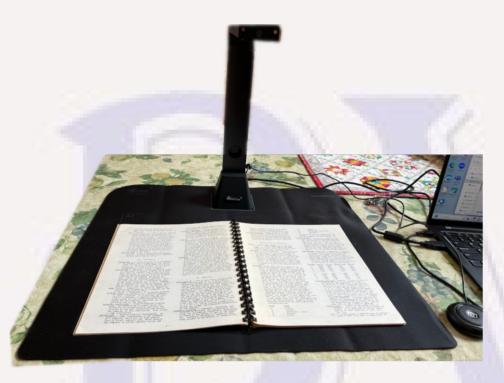
A Scan Problem

- How to scan a delicate oversized book
 - Battle of Waterloo, pub 1817
 - Large (2.5" x 11.5" x 13")
 - Not in pristine condition, brittle pages, smoke damage on some page edges (better copies online at French museum)
 - Of family historical interest. Handwritten comments by ancestors in 1880's

A Solution

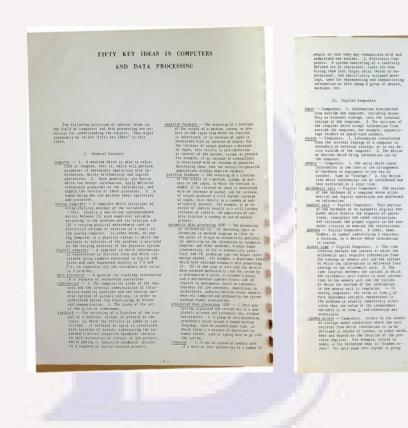
- An practical solution was finally found in 2022: IRIS ScanDesk 5 Pro (\$165)
 - Designed to "scan" two pages at a time from smaller books using a camera embedded in a pedestal, but also meant it could handle the size of a single 11.5"x13" page
 - Not a perfect solution but could produce a readable copy with much less effort and minimal risk of book damage.
 - Most useful output formats: PDF, searchable-PDF, JPG, also supports EPUB OCR text, but quality depends on original font quality.
 - Can take of picture of two pages in two seconds, but scan speed limited by how fast you can turn pages

IRIS ScanDesk 5 Pro



- Hi-res camera & LED light on metal arm & base
- Black pad with position markings for scanner base and books
- USB connection to PC for data and power
- Push switch connected to scanner to initiate scan - about 2 secs for both pages
- Software partially compensates for camera parallax, senses division between pages
- Can save as PDF with each page a separate page or with two pages per PDF page
- Searchable PDF or EPUB formats require significant processing time for OCR, but makes use of multiple CPUs if available. Can also save as one JPG per page
- Was able to successfully scan one single large page at a time, but a lot of extra work because software not really designed for that

IRIS ScanDesk Results



to look up is the book may be almost anywhere in the book with roughly equal probability. <u>pulse</u> — Circuits. In general, a sharp difference ustally over a relatively short period of time, between the normal level of some physitime, between the normal level of some physic city variable corresponding to the average level of a wave or waves and a high or low level of their physical variable corresponding to the cresi or trough of the wave; often, a sharp voltage change. For eximple, if the voltage at a terminal changes free -10 to 400 volts and remains themes for a period of

rections is to value any collective forms of positive pairs is characterized by ring of interest free eer will get the unial preservation is and correctly determined the preservation black to the original main-depreservation black to the original main-ter and the state of the state of the characterized by the state of the characterized by the state of the distance states is and of the state of the distance states is and of the state of the distance states is a state of the state distance states is a state of the distance states is a state of the state of the distance states is a state of the state of the distance states is a state of the state of the distance states is a state of the state of the distance states is a state of the state of the distance states is a state of the state of the state of the distance states is a state of the state of the state of the distance states is a state of the state of the state of the distance states is a state of the state of the state of the distance state of the state of the state of the state of the distance state of distance distance of the state of the state of the distance state of distance distance of the state of the stat or absence of polarized spots, singly or in sets. 3. Delay-Line Memory such as a Quartz Rod. A circular path forward through the delay-line memory and back through electrical circuits along which a pattern of pulses rec circuits along which a pattern of pulses rep-rementing information pulse stored. Links — Digital Computers. In a synchronia com-pulses at equal times which schedule the operations of the computer. ... In a synchro-mout computer, there is no used for such a clock, since the closing or completion of ome circuit initiats the operation of a subsequent

circuit. <u>transducer</u> — A device which converts energy from one form to another. For example, a slab of quartz crystal embedded in mercary can change electrical energy to sound energy (land vice warsa), as is done in ponic delay lines in versa), as is done in sonic delay lines in certain computer smoory systems. "Any device which emables the conversion of any physical, chemical, robiological phenomenon let an electrical, hydraulic, pneumatic, methanical, potical or other signal for transmission, re-cording, measurement, malysis, actuation, output or control. Transducters have been referred to as pick-ups, sensing elements, prinary elements, command devices, seasors, de ectors and probes as examples, transducer exist which convert pressure, temperature, acceleration, force, muss, time, distance, rediation and a multiplicity of other phenom-ena into electrical form. These signals are than readily fed to amplifiers, data proces-sing systems, or output controllers. Many

nstrumentation problems may be resolved int

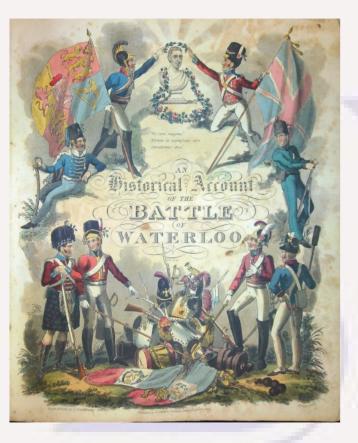
the choice of input and output transducers with an interconnecting information-processing link. Transducers extend, refine, supplement, or replace man's senses in remote and heardous

- Two pages from 50 page booklet scanned in under 2 mins, + 2 mins post processing for OCR scannable PDF
- Some minor differences in page • rotation, but pages fairly well cropped with minimal parallax distortion

Foreword to the Reader

in working with the computer and information sciences from day to day. It is useful to have at one's elbow certain reference information. Included among the reference information which one might like to consult from time to time, or even become informed about for the first time, are the selections contained in this reference diary. Not every selection will be useful to every reader of course, but we do hope that many of our salections will be useful or informative to many of our readers. For lack of space this publication could not include all the reference information which might desirably be provided.

IRIS ScanDesk 5 Pro on Large Book



THE BATTLE OF WATERLOO.

an object of reverence even to her victors, and rejuced in the possession of that power, which was still formidable enough to intimidate the insolence and enthusame of conquest. Sustained by the one, and elated by the other, she screetly cherished the idea that the day could not be far distant, when she should again shine forth in all the luster of aniable tyranawy, and again decorate her present oppressors with the interesting badges of imperial ascendancy. The fascinating experiment was soon tried indeed, but the result differed a little from what was expected, perhaps hoped, by the enlightened advocates of that benigrant system which Napoleon invented and matured, for the safety of personal liberly, the freedom of public opinion, and the general happiness and independence of nations.

When the allies permitted the restoration of the Bourbons, they contemated that act as a service done to France, and a security obtained for Europe. upon her political condition, there could be little reason to doubt, though it was obvious that the immediate effect could not be productive of general tranquillity or cordial satisfaction among all classes of society. They who had acquired fortunes from the common wreck of property, during the most convulsed and disastrous periods of the revolution, would naturally feel some alarm, lest the dangerous pretensions of the ancient possessors should disturb their claims : they, also, who justly rejoiced at the extinction of many feudal and ecclesiastical privileges, which circumscribed or destroyed the rights of the citizen, would look with anxious forebodings at the return of a dynasty, under whose sway those privileges had been conceded and maintained ; while they, in whom all sense of morals was annihilated by the profligate maxims, and equally profligate practice, which sprung from the rank soil of democracy, who regarded the dictates of virtue and decorum as a restraint, and restraint as oppression, could not anticipate, with much delight, the restoration of a sovereign, who ventured to respect not only the forms but the precepts of religion, and whose example was likely to enforce the observances of piety.* If, however, the apprehensions of the first and second class were

Ben Jonson has accurately described the inducate which courts and sovereigns ought to have up the morals of the people :---

⁴⁰ Princes that would their people should do well,
 ⁴⁰ Must at themselves begin, as at the head ;
 ⁴¹ For man, by their sample, pattern out,
 ⁴² Their initiations, and regard of laws ;
 ⁴⁴ A virtuous contra a world to virtue draws,⁴¹

Was France to be the only country denied the possibility of such a blessing? - Country denied the possibility of such a blessing? - Country denied the possibility of such a blessing?

administration of the 5 fronts The langest capacity for the Stemis to furning the higher or to dation to have deliver

- Two pages from 1817 book Battle of Waterloo
- Many hand-colored pics
- Hand-written note from abt 1880 by gggf Vaughan
- I have read Victor Hugo's long description of the Battle of Waterloo in Les Misérables, Volume II, Book I.
- In my grade school days I remember riding West from Lexington, MO on old US 24 through little towns Wellington, Waterloo, and Napoleon

Why not Use Smartphone Camera as Scanner?

• Pros

- widely available, portable, hi-resolution on newer models

Cons

- Usability optimized for pictures
- cropping is a manual operation
- Parallax distortion rectangular pages no longer rectangles. Parallax distortion can be significantly fixed with GIMP perspective transformation, but is labor intensive.
- Document image lacks meta data of actual document size
- Quality variation: resolution of actual document, focus, rotation, lighting, shadows all depend on skill of user, which means not consistent. Difficult to judge quality on a small smartphone screen.
- Transfer to PC platforms for organization and archival adds additional steps

A Typical Smartphone Camera "Scan"

THE REFERENCE DIARY OF THE LIBBARY OF COMPUTER AND INFORMATION SCIENCES

Edmund C. Berkeley

Copyright (c) 1966

The Library of Computer and Information Sciences

All rights reserved

This book or any part thereof must not be reproduced in any form without written permission of the publisher

1967 Edition -- Published 1966 by Ibrary of Computer and Information Sciences 866 Third Avenue New York, N. Y. 10022

Printed in the United States of America

pose Digital Computers".

Newtonville, Mass. 02160 August, 1964

Table of Contents Foreword to the Reader In working with the computer and information in working with the tis useful to have at one's ellow certain reference information. Included among the reference information which one might like to consult Fifty Key Ideas in Computers and Data Processing reference information which one might like to consult from time to time, or even become informed about for the first time, are the selections contained in this refer-ence diary. Not every selection will be useful to every Summary of Binary Arithmetic and Related Number Systems reader of course, but we do hope that many of our selec Some Binary, Octal, and Decimal tions will be useful or informative to many of our readers. For lack of space this publication could not include all the reference information which might desirably be ration Tables Summary of Boolean Algebra (Contrasted with Elementary Algebra) We gratefully acknowledge the permission of "Computers and Automation" magazine to reprint here "Over 1000 Areas of Application of Computers", the Ranges of Computer Speeds of Commercial Computers "Computer Census", and "Descriptions of General Pur Over 1000 Areas of Application of Computers All suggestions, comments, corrections, and criticisms will be most welcome, and will help to make the next edition of this reference diary more serviceable Computer Censu Descriptions of General Purpose Digital Computers Edmund C. Berkeley

Page

24

36

Foreword to the Reader

In working with the computer and information sciences from day to day, it is useful to have at one's elbow certain reference information. Included among the reference information which one might like to consult from time to time, or even become informed about for the first time, are the selections contained in this reference diary. Not every selection will be useful to every reader of course, but we do hope that many of our selections will be useful or informative to many of our readers. For lack of space this publication could not include all the reference/information which might desirably be provided.

Readable, but with perspective (parallax) distortion

But, Current iPhone Magic

- Create a Note
- Select the camera icon
- Select "Scan Documents"
- Position camera to get entire document in view will take picture when it sees entire document
- Repeat for any additional pages, then press save to save all pages
- All scanned pages are automatically cropped, corrected for perspective distortion and saved as multi-page PDF. There are even some editing options for manual rotation or manual cropping. Even does some OCR to attempt to come up with meaningful name for PDF.
- Use the send icon (rectangle w up arrow) to email or to save PDF in cloud storage that is accessible to your computer.
- Results are at least as good as a cheap scanner!

Example of Note "Scanned Document" Page

THE REFERENCE DIARY OF THE LIBRARY OF COMPUTER AND INFORMATION SCIENCES

Edmund C. Berkeley

Copyright (c) 1966

The Library of Computer and Information Sciences

All rights reserved

This book or any part thereof must not be reproduced in any form without written permission of the publisher

1967 Edition -- Published 1966 by The Library of Computer and Information Sciences 866 Third Avenue New York, N. Y. 10022

Printed in the United States of America

Digital Computers

Foreword to the Reader

provided.	In working with the computer and information sciences from day to day, it is useful to have at one's ellow cortain references information. Included among \oplus references informations which one might like to consult from times to time, or even become informed house for most days. Note every alsoritors will be useful to every reader of course, but we do hope that many of our selec- tions will be useful or informative to many of our reader into a will be useful or informative to many of our selec- tions will be useful or informative to many of our selec- tions will be useful or informative to many of our selec- tion and the reference on information which might destrably be
	all the reference information which might desirably be provided.

We gratefully acknowledge the permission of "Computers and Automation" magazine to reprint here "Over 1000 Areas of Application of Computers", the "Computer Census", and "Descriptions of General Pur-pose Digital Computers".

All suggestions, comments, corrections, an criticiams will be most welcome, and will help to make the next edition of this reference diary more serviceable Edmand C. Berkeley

Newtonville, Mass. 02160 August 1966

Table of Contents	
	Page
Fifty Key Ideas in Computers and Data Processing	2
Summary of Binary Arithmetic and Related Number Systems	7
Some Binary, Octal, and Decimal Conversion Tables	11
Summary of Boolean Algebra (Contrasted with Elementary Algebra)	13
Ranges of Computer Speeds of Commercial Computers	19
Over 1000 Areas of Application of Computers	20
Computer Census	24
Descriptions of General Purpose	26

Foreword to the Reader

In working with the computer and information sciences from day to day, it is useful to have at one's elbow certain reference information. Included among the reference information which one might like to consult from time to time, or even become informed about for the first time, are the selections contained in this reference diary. Not every selection will be useful to every reader of course, but we do hope that many of our selections will be useful or informative to many of our readers. For lack of space this publication could not include all the reference information which might desirably be provided.

Hand-Held Scanner

- In search for ways to scan problem books, tried out a rechargeable, hand-held bar scanner
 - Bar must be held horizontal to page while moving vertically down the page. Multiple pages could be scanned, and then bar connected to computer via USB to transfer images and recharge.
 - Nice concept, very light, but didn't function reliably for me:
 - Difficult to move bar at an even speed while keeping it parallel with top&bottom of page and keeping
 movement precisely parallel with sides of page. Failure to maintain the required alignment produced "waves"
 in scanned pages.
 - Difficult to keep correct consistent pressure of bar against page, needed for tracking wheel to accurately
 measure speed of vertical movement failure resulted in parts of scanned image being compressed or
 stretched vertically.
 - As a result, I found after downloading images from scanner to computer that many pages needed to be rescanned before an acceptable scan of all pages was obtained. Even with utilities to make it easy to merge together PDF pages from multiple attempts, generating a "good" PDF for a document with even a modest number of pages was a very frustrating process.