

PARADISEC INGESTION WORKFLOW MODEL

1. Pre-ingestion (audio tapes)	2
2. Ingestion Processes.....	5
3. Wavelab processes for generation of derivatives (mp3 and CD-audio quality) (Audio)	10
4. Rimage processes for burning CDs (Admin).....	12
5. Quadriga Processes for BWF generation (Audio).....	13
6. Archive Process (Audio).....	15
7. Restore from archive process (PC).....	18
8. Create a static repository in XML (PM).....	19
9. Check PARADISEC repository reports (Audio)	19
10. Access to Archived files in Paradisec repository	19
11. Appendices	20

Appended to this manual are a diagram depicting Paradisec's Data Flow on p27 and the Austehc Manual of Archival Imaging for Access on p28.

This document reflects practices in the University of Sydney audio ingestion unit of PARADISEC, the Pacific and Regional Archive for Digital Sources in Endangered Cultures. For further information on PARADISEC please visit our website <http://paradisec.org.au>.
This document last updated 15 January 2008.

For questions or comments on this document please contact admin@paradisec.org.au, or write to us at:

PARADISEC
Transient Building F12
Fisher Road
University of Sydney NSW 2006
Tel: +61 2 9036 9557
Fax: +61 2 9351 7572

Role terms used in this document

Admin: project administration officer (2008: Amanda Harris)

Audio: audio preservation officer (2008: Aidan Wilson)

PM: project manager (2008: Nicholas Thieberger)

PC: project coordinator (2008: Tom Honeyman)

1. Pre-ingestion (audio tapes)

- 1.1. assess collection (Admin and PM)
 - 1.1.1. assess whether material is suitable for accession (PM in consultation with steering committee)
 - 1.1.1.1. rights
 - 1.1.1.1.1. does the depositor assert ownership of the material?
 - 1.1.1.1.2. are the rights in the material clearly specified?
 - 1.1.1.2. content criteria (assess against policy) –
 - 1.1.1.2.1. is the material unique?
 - 1.1.1.2.2. is the country and language of the material known, and in the core area (non-Australian)?
 - 1.1.1.2.3. is there an alternative place for deposit of the recording?
 - 1.1.1.2.4. is the content of the recording in a high-risk language?
 - 1.1.1.2.5. is the format or recording medium at high risk?
 - 1.1.1.3. institutional affiliation – give preference to consortium members
 - 1.1.1.4. ease of integration into our systems
 - 1.1.1.4.1. is data in a format we can handle (i.e. cassette, ¼ inch reels)?
 - 1.1.1.4.1.1. if not, direct prospective depositor to suitable alternative facility
 - 1.1.1.4.2. consistency and adequacy of any metadata
 - 1.1.2. if collection assessed as suitable, create a record in the SQL database in the form [collection ID]-ASS
 - 1.1.2.1. check "hide metadata" box on this record
 - 1.1.3. assess tape condition and prepare material for transportation accordingly (see below)
- 1.2. transportation
 - 1.2.1. if the recordings are in bad condition (e.g. mould, rare formats, if tape visibly damaged) send to Sydney (contact Admin)
 - 1.2.1.1. transport any mouldy items in separate ziplock bags and do not pack with non-mouldy items
 - 1.2.2. if in reasonable condition transport to nearest ingestion point
 - 1.2.3. personal transport is preferable; transport in hand luggage if flying
 - 1.2.4. if necessary use registered courier
 - 1.2.5. packing – make sure end fastened on reel tape, pack tightly, wrap in bubble wrap, ensure against unnecessary shocks during transport

- 1.2.6. transport with inventory of contents
- 1.3. receipt of materials at ingestion point (Admin)
 - 1.3.1. check inventory against contents and contact depositor if any discrepancy
 - 1.3.2. enter date in "data received" field of SQL database
 - 1.3.3. assess whether mould-affected and if so implement appropriate procedure (put in ziplock bags and store in 'dirty room')
- 1.4. metadata
 - 1.4.1. where possible get depositors to enter own metadata in SQL database
 - 1.4.1.1. assign login (Admin)
 - 1.4.1.2. associate assessed collection record with user, and change item ID for first record (i.e. change -ASS part of PI to the ID of the first item e.g. LB1-ASS > LB1-001),
 - 1.4.1.3. if appropriate uncheck "hide metadata" box
 - 1.4.1.4. email depositor with instructions
 - 1.4.2. where depositor cannot enter own metadata, create records (Admin)
 - 1.4.3. assign Persistent Identifier (PI) – if depositor already has a consistent naming system, use it, or else use numerical symbols for 'item ID', e.g. 001
 - 1.4.4. enter all metadata available into appropriate fields of catalogue, ensuring that all compulsory fields (marked with a red asterisk) have been filled
 - 1.4.5. label all tapes and tape cases to show PI, ensuring that no written information is covered up
 - 1.4.6. sort all tapes in numerical order in filing cabinet (if to be digitised shortly) or in safe (if filing cabinet is full), or in dirty room if mould-affected
 - 1.4.7. when tapes returned, note date in 'tracking' field in SQL database and tick "Tapes returned to depositor" checkbox
- 1.5. prepare tapes for ingestion (Audio)
 - 1.5.1. assess tape condition – treat mouldy items according to procedure (wear protective gear)
 - 1.5.1.1. dehumidify in vacuum oven if necessary, with separate batches for mould-affected items
 - 1.5.1.1.1. decide duration of treatment and temperature following advice from NFSA <<http://www.screensound.gov.au/screensound/screenso.nsf>>
 - 1.5.1.1.2. if mouldy remove any visible growths first with cotton bud or pad and solution of isopropyl and cetramide
 - 1.5.1.1.3. if sticky shed is present no special prior treatment is needed
 - 1.5.2. clean and prepare for ingestion

1.5.2.1. reel-to-reel

- 1.5.2.1.1. repair or replace lead tape and damaged splices if necessary
- 1.5.2.1.2. clean using slow spool and specialist equipment (cotton pads and isopropyl/cetramide solution); also clean flange (spool) separately
- 1.5.2.1.3. second slow spool to repack tape

1.5.2.2. cassette and DAT

- 1.5.2.2.1. replace splices if necessary (cassette)
- 1.5.2.2.2. fast forward and rewind to minimise print-through (cassette) and to repack tape (cassette and DAT)
- 1.5.2.2.3. if mouldy, bag up separately, to avoid cross-contamination, shelve in dirty room and minimise time in ingestion room, replay on 'dirty' cassette machine.

<i>format</i>	<i>device</i>	<i>comment</i>
Cassette	TASCAM 122 Mk II	tested OK
Mouldy cassette	Yamaha KX-W321	tested OK
Reel-to-reel 15ips	Studer A810	tested OK
Reel-to-reel 7 1/2 ips	Studer A810	tested OK
Reel-to-reel 3 3/4 ips	Studer A810	tested OK
Reel-to-reel 1 7/8 ips	Currently attempting to locate suitable machine	as at January 2008
Reel-to-reel 1/4 track	Studer A810	tested OK
Minidisk	Processed in Melbourne – Minidisk Transfer Editor	tested OK
*.wav files	via network from Melbourne, ANU or APAC	tested OK
DAT	Sony PCM 7040	tested OK
CD	CD inspector on Audiocube 3	tested OK

2. Ingestion Processes

2.1. Wavelab ingestion processes (Audio)

- 2.1.1. ensure the ingestion computer is disconnected from the network (network activity can cause audio interference), turn off any mobile phones or other mobile devices, including close wi-fi transmitters, and as a precaution, close any programs other than WaveLab.
- 2.1.2. reel-to-reel tapes must go through Audiocube-1, cassette tapes may go through either Audiocube 1 or 3. Other sources, such as vinyl records, are to be connected separately as required and may, depending on how the device is connected, go through either Audiocube-1 or 3.
- 2.1.3. open the 'record' dialogue box (the button with the red circle at the bottom of the screen) and type in the name and path of the file that you will create. The destination folder will, on either Audiocube, be "D:/WAV96 ingest" followed by the name of the audio file "[PI].wav". Relevant sub-parts of an item are indicated '-A', '-B' and so on, and are inserted after the PI and before the extension (this is necessary for audio files that will be over 2GB - about 55 minutes at 24-bit/96khz).
- 2.1.4. if further work will be required on a file, such as splitting it into two separate channels, then the file should take a name that is distinct from the files that will ultimately be produced, to preclude any confusion. For instance, if the files to be

archived will eventually be named AW1-001-A.wav and -B.wav and they are two sides of a single ingested recording that will require splitting, then name the file that comprises each of them as AW1-001_1. The underscore as opposed to a dash will prevent the file from being erroneously archived.

- 2.1.5. **#Important#** If the tape is likely to run for more than one hour on 2 channels, or two hours on one channel, then 'auto-split' must be selected, otherwise, the recording will cease when the file exceeds 2GB, the maximum size for any single file in Windows operating systems.
- 2.1.6. it is a good idea to run the signal for a few seconds and adjust the level on the source selector so that it maxes out at about -12.00 dB on the record dialogue (the white stickers on the source selector are a rough approximation of good gain levels for the selected source). After a good level is established, return the tape to the beginning and commence recording, leaving at least 5 seconds of silence before the signal begins.
- 2.1.7. pay attention to the recording and note any analogue errors, such as pops, hisses, buzzing, microphone bumps. These should be noted in the 'ingest notes' in the catalogue. More generally, note the overall quality of the recording; whether it is at all muffled, distorted or good and clean. This should similarly go into 'ingest notes'.
- 2.1.8. analyse the file in Wavelab (control-Y) (takes about 2 minutes for normal length file) - this generates temporary error markers wherever recording faults are apparent
- 2.1.9. check the errors: if they are analogue errors make a note of them in SQL web catalogue; if they are digital errors, re-ingest the recording.
- 2.1.10. enter ingestion metadata into SQL database
 - 2.1.10.1. use 'Ingest Notes' field to record details of recording and audio issues (use '<p>' for linebreaks that need to be kept)
 - 2.1.10.2. tick 'digitised' checkbox (which auto-enters date digitised)
 - 2.1.10.3. add operator info
 - 2.1.10.4. if recorded tape ID mentions date created, tape ID, recordist, or any other info not already in database fill in relevant fields (check with admin as necessary, if unsure, leave blank)
- 2.1.11. if necessary, normalise (takes about 10 minutes). This should not normally be necessary where pre-monitoring has taken place and the tapes are in good enough condition to yield good quality sound
- 2.1.12. after ingestion finished, trim silence at beginning and end of file, leaving margin of 5 seconds' silence

- 2.1.13. if required, insert track boundary markers - simply press 'insert' (see documentation relating to CD generation for more details)
 - 2.1.14. in the case of stereo reel-to-reel tapes that comprise two simultaneous mono recordings, split the 2 channels
 - 2.1.15. name the files appropriately, for the time being, leave the raw, unedited files (the ones with underscored names instead of dashed ones) in this folder until a later backup can be made
 - 2.1.16. place files in "Azoulay\PDSC Audio Area\Send to Seal" and commence BWF wrapping, (see section 5 for details on the BWF generation process). Otherwise, if more work will be undertaken at a later time, files should be placed on the server in "Azoulay\PDSC Audio Area\Audio Work Files\Workspace\[PI]"
 - 2.1.17. check that mandatory metadata fields (PI, title, date created, country, collector) are in SQL catalogue, if not, add them. If you are unsure of what information to enter into the catalogue, contact Admin
- 2.2. University of Melbourne ingestion unit
- 2.2.1. this suite consists of a purpose-built computer, with a Digital Audio Denmark analogue to digital converter, a cassette player and reel-to-reel player.
 - 2.2.2. digitisation from cassette and reel-to-reel tape
 - 2.2.2.1. check speakers' volume is low
 - 2.2.2.2. turn on computer and tape player
 - 2.2.2.3. check that the AD unit has the correct input (cassette or reel to reel)
 - 2.2.2.4. on desktop, go to 'paradisec login' icon and login
 - 2.2.2.5. open 'Wave Lab'
 - 2.2.2.6. enter the new filename
 - 2.2.2.7. ensure the settings are correct (they should all be standard) 48 khz or 96 khz, 16 bit
 - 2.2.2.8. clean the playback machine
 - 2.2.2.9. for cassette tapes, fast forward once right through both sides before recording
 - 2.2.2.10. on ADDA machine: if reel to reel, analog input is 'XLR'; if cassette analog input is 'jack'
 - 2.2.2.11. play a bit of original tape through to check levels – should not go into red on ADDA, around -12 is a good level.
 - 2.2.2.12. adjust volume for monitoring speakers with right hand knob on blue box

- 2.2.2.13. while the recording is playing, type any audio notes into the "ingest notes" field of the SQL database focusing on the state of the original recording): whether mono or stereo; which channel (right or left); any clipping in original
- 2.2.2.14. when finished, stop tape first then stop recording – leave around 5 seconds (at beginning and end)
- 2.2.2.15. save file
- 2.2.2.16. to send to Azoulay, copy files to hard disk
- 2.2.2.17. take hard disk to a machine that has an ftp connection to the HPC machine called alfred (<sftp://thien@alfred.hpc.unimelb.edu.au//home/thien/az/wav/>) (password is the word for language in South Efate).
- 2.2.2.18. send the most recent files to alfred, then copy the files to the folder 'Sent to Sydney' on the hard disk. There is a (cronjob) script on alfred that will send these files up around 3am the next morning. The cronjob script is located here: alfred.hpc.unimelb.edu.au//home/thien/az/. The person to contact about running this script is Dirk van der Knijff [<dirk@unimelb.edu.au>](mailto:dirk@unimelb.edu.au).
- 2.2.2.19. go to Paradisec catalogue and complete ingest notes see 2.1.10 for instructions

2.2.3. transfer from Minidisk

- 2.2.3.1. the Minidisk Transfer Editor (MDTE) is software that sits on a particular PC that has a dedicated SCSI card installed. The SCSI cable attaches to a Sony MDH-10 minidisk player.
- 2.2.3.2. restart the PC with the MDH-10 connected and powered up.
- 2.2.3.3. insert your minidisk.
- 2.2.3.4. start the MDTE software and you should see the listing of the files on the minidisk. Be sure to name the files for transfer at this point. Click on the relevant files and start the capture program (right click gives the options). This will capture the ATRAC files to the hard disk and then convert them to WAV format.
- 2.2.3.5. when the process is finished, copy the files to a hard disk as per section 2.2.2.16 above.

2.3. Ingestion of photographic slides

- 2.3.1. create a folder inside the directory '\\azoulay\PDSC Audio Area\Non-Audio Work Files' with the name of the collection and item that the slides belong to
- 2.3.2. square slides will need to be processed in the standard scanner (Scanmaker 8700) as the automatic slide feeder will not process them without cropping some of the image. Rectangular slides can be processed in the automatic slide feeder
- 2.3.3. for square slides

- 2.3.3.1. insert slides into a 12-part bracket and place on Scanmaker glass
- 2.3.3.2. open the SF-APP and click 'import'
- 2.3.3.3. once scanning window has opened, verify that the scale is set to 100% and dpi to 4,000
- 2.3.3.4. click 'prescan'
- 2.3.3.5. the resulting scan will show the entire set of slides on the glass, select one slide at a time by drawing a box around it and when selected, click 'scan RGB'
- 2.3.4. for rectangular slides
 - 2.3.4.1. open Nikon Scan 4 on the image ingestion station and, when the slide scanner is turned on, select the 'scanner' menu and open 'Slide Feeder'
 - 2.3.4.2. insert slides into scanner, press enter, in the next window, check that every box is checked, and press enter again
 - 2.3.4.3. enter the collection and item IDs, according to the Paradisec naming conventions, ending with a hyphen (e.g. "AW1-PH01-"). Leave 'suffix' blank
 - 2.3.4.4. select number of digital places for automatic batch numbering; for collections of over one hundred photos, a padding of three digits is necessary. The window will show a preview of how the files will be named, use this preview to double check that the names will be correct.
 - 2.3.4.5. select the directory into which the photos will be saved (defined in step 2.1.16)
 - 2.3.4.6. press enter, the slide feeder will automatically ingest the slides. For hints on using the slide scanner for different types of slides and getting the best possible scan, see 11.1, Appendix 1.
- 2.3.5. when all slides are scanned, rotate those in portrait format using Photoshop (or open source software such as The Gimp). This step can best be done while slides are being ingested as there is sufficient waiting time between each slide being completed to do other tasks.
- 2.3.6. the Paradisec slide ingestion computer has batch process settings to format tiffs to a standard size (5,000 pixels on the longest side) and to produce a jpeg of each tiff. To use these processes
 - 2.3.6.1. open an image in Photoshop
 - 2.3.6.2. select 'portrait-' or 'landscape Photo under A3'
 - 2.3.6.3. press play to start process
- 2.3.7. 'access' .jpg copies should be made of each photo. The access copy should be no smaller than 425 pixels on the longer side. These are for web-preview only and need not be any larger

- 2.3.8. both the archive standard .tif and the access .jpg can then be archived (see 2.1.16 for where to place the file for archiving)
- 2.4. digital image capture
 - 2.4.1. see attached Austehc manual (from page 25) and 11.2 Appendix 2 for tips for Image Capture
- 2.5. digital text
 - 2.5.1. formatted documents, e.g. in MS Word format, need to be saved as rtf and pdf. Check each file for character encodings.
 - 2.5.2. if the file is of a specific type (e.g. a Toolbox file) then it needs to have sufficient documentation to make it readable (e.g. .typ files accompany it).
 - 2.5.3. put the files into the "Send to Archive" folder

3. Wavelab processes for generation of derivatives (mp3 and CD-audio quality) (Audio)

- 3.1. CD-audio quality files
 - 3.1.1. open the file(s) in "\\azoulay\PDSC Audio Area\Audio Work Files\workspace' with WaveLab. It is good practice to keep a backup copy of the files that are to be BWF wrapped inside "\\azoulay\PDSC Audio Area\Audio Work Files\workspace', while the files that get wrapped will be inside "\\azoulay\PDSC Audio Area\Send to Seal' and should be left alone.
 - 3.1.2. if the audio file can be segmented into tracks, do so in WaveLab by pressing the "insert" key while the cursor is at the appropriate point. This will insert a yellow 'generic marker' at that point. It can be moved easily, by holding the yellow, triangular selection point at the top of the waveform and dragging, and it can also be deleted by right clicking it. Move it to an appropriate position for a track boundary. For digital copies of reel-to-reel or cassette tape recordings, there will be audible boundaries where the recorder has been paused or stopped and then re-recorded. These are even visible as a quick high peak followed by silence, then a smaller peak, though this may depend on the recorder. Between these two peaks, in the silence, is a good position to place a track marker. However, be careful to note that the intended track boundaries as specified (if indeed they are specified) by the depositor, on the tape's notes, may not coincide exactly with the stop-starts in the audio file. One 'track' may consist of multiple silence points. Use discretion to evaluate where track boundaries ought to be.
 - 3.1.3. save the file, this will create an *.MRK file that lists the marker positions. The MRK file is then used later on to split the single audio file into separate wav files for each track

- 3.1.4. run WaveLab batch processor from "Tools" > "Batch Process..." on files to downsample the 24bit/96khz '*.wav' file to 16bit/44.1khz. You can either select the files in the current active window using the browse function, or select the entire folder (where you placed the files) in the 'input' tab. Be careful about the path, because if WaveLab cannot find the designated directory, it will create it.
 - 3.1.5. in the 'input' tab, click the 'extra' button to reach the 'processes' window, go to the 'presets' tab and select the preset that says "96@24 to CD 44_1@16" and press 'load'. If this isn't done, the files may be downsampled to 44.1 kHz, but they won't be *dithered* to 16 bit, which means they won't burn onto a CD, and will certainly not play on any audio CD player.
 - 3.1.6. exit the presets window and check that the right processors have been selected, they should be, in order:
 - 3.1.6.1. Normalizer (Untitled)
 - 3.1.6.2. Dither & NoiseShaping (down to 16)
 - 3.1.6.3. Resampler 192kHz(44_1)
 - 3.1.7. if the preset that contains these processors in that order fails to load, the processors can be manually loaded from the 'plugins' folder in the right-hand-side pane of the 'processors' tab.
 - 3.1.8. in the 'output tab' specify "\\azoulay\PDSC Audio Area\CD Work Area\WAV 44-1 for split". This can also be set in 'presets', selecting "96to44_1 for CD" will designate this folder for the output files.
 - 3.1.9. press OK to exit the Processor list dialogue box, and double-check the input and output in the 'input' tab of the batch processing dialogue box. If all is accurate, click 'run'.
 - 3.1.10. for audio files to be burned using the Rimage CD burner, the single file must be split into tracks. From the 'tools' menu, select 'auto-split'.
 - 3.1.11. select the folder already used for the output in step 3.1.8 (by default "\\azoulay\PDSC Audio Area\CD Work Area\WAV 44-1 for split')
 - 3.1.12. in the subsequent window, select to split the file by "markers", then "generic markers" and click next
 - 3.1.13. 'save as separate files' and specify the directory chosen in step 3.1.8
 - 3.1.14. select "create additional folder named as source file".
 - 3.1.15. click "process"
- 3.2. Wavelab batch processor to create 128kbps MP3 files

- 3.2.1. in the batch process window (under 'tools'), select one of the 'mp3' options from the presets tab. This should be set up to make 128 kbps 44.1 kHz stereo mp3 files from whatever input you select, irrespective of the input file format or properties.
- 3.2.2. if not set up, alter the settings of the output files in the output tab
 - 3.2.2.1. confirm that 'encode' is checked and click button underneath
 - 3.2.2.2. select the properties: 44100 Hz, 128 kbps, stereo
 - 3.2.2.3. set the batch processor to normalise if the files require it. If the files were generated from the CD files, they won't need normalising, though they will need the _CD name suffix removed or they won't get harvested by APAC.
 - 3.2.2.4. select 'normaliser -1dB' from the presets tab in the processors dialogue box. No other processors apart from 'normalise' have to be selected, unlike CD generation above.
- 3.2.3. specify the output directory by selecting 'browse' to '\\azoulay\PDSC Audio Area\Send to Archive' and press 'run'. From their output folder, the files will be automatically archived to APAC.

4. Rimage processes for burning CDs (Admin)

- 4.1. download the files in '\\azoulay\PDSC Audio Area\CD Work Area\Segmented CDs for burn' to burning machine 'D:CDs'
- 4.2. open Rimage System Manager software and check that the Autoloader 1 media settings are set to 'CDR', if not, pause the Production Server and change from DVD to CDR.
- 4.3. for data CDs
 - 4.3.1. start 'new project' in QuickDisc, select 'data CD', drag files into programme
 - 4.3.2. create label in CD Designer and save
 - 4.3.3. select saved label for project in QuickDisc
 - 4.3.4. load CDs into lefthand bin
 - 4.3.5. submit to burn (typically, 2 copies, but programme will handle up to 100 copies)
 - 4.3.6. when finished put in cases
- 4.4. for audio CDs
 - 4.4.1. start 'new project' in QuickDisc, select 'audio CD', drag files into programme
 - 4.4.2. follow remaining steps from 4.3.2
- 4.5. tick 'CD burned' in SQL
- 4.6. if using gold CD stock, enter ID numbers in 'CD ID' field
- 4.7. return CDs to researcher with letter listing inventory of CDs returned and instructions on how to access files online

- 4.8. note return of CDs in 'tracking' field in SQL database
- 4.9. move files from '\\azoulay\PDSC Audio Area\CD Work Area\Segmented CDs for burn' to '\\azoulay\PDSC Audio Area\CD Work Area\Burned CDs' and delete files from burning machine

5. Quadriga Processes for BWF generation (Audio)

5.1. this process commences once the wav file is in the folder '\\azoulay\PDSC Audio Area\Send to Seal' and the necessary metadata has been entered into the catalogue (see section 1.4).

Wrapping includes the necessary metadata and digitally seals the file to prevent it being altered

5.2. keep a backup of the file in a folder named with the collection ID in '\\azoulay\PDSC Audio Area\Audio Work Files\workspace'

5.3. when files are ready to be sealed, run the file_adder.php script.

http://azoulay.arts.usyd.edu.au/paradisec/file_adder.php. This script searches 'Send to Seal' for any newly added wav files and harvests the file name, size and the audio properties and places a reference to them in the catalogue entry for that item

5.4. produce XML files

5.4.1. open Paradisec catalogue and search for the files you are working on:

http://azoulay.arts.usyd.edu.au/paradisec/search_item.php

5.4.2. once this search set is established, run the bwf.php script.

<http://azoulay.arts.usyd.edu.au/paradisec/bwf.php>. This script creates an .imp.xml file for each listed wav file, (the contents of the imp.xml will be identical for each wav file within a single item, but will be named after each file). The imp.xml will be written to the directory: '\\PDSC Audio Area\BWF xml files\XMLImport'

5.5. switch on Audiocube 2, Audio File Inspector (AFI), Quadriga Editor (also called BWF Editor) and BatchProcessor now run automatically

5.5.1. in AFI, press the 'PROCESS' button. This looks to the server, into the directory [\\PDSC Audio Area\Send to Seal](#), as well as the XMLImport folder mentioned in step 3, and copies any wav file that has a corresponding xml file to its local D:\ drive. Once it has finished, it will remove the files from the server.

5.5.2. NOTE: any *mono* file more than 1GB should be divided into two smaller files, named *-A1.wav and *-A2.wav. The reason for this is that AFI converts mono files to stereo files, since BWF files are always stereo. Because of the limit on single files of more than 2GB, if, for instance, a mono file is over 1GB, AFI will double its size. When it reaches 2GB and attempts to continue, AFI will crash.

5.5.3. AFI will also automatically scan the audio characteristics of the file, and produce explicit notes in the BWF header where there are any errors, such as clips, or any

sections of the file that are silent. These will be displayed in BWF Editor when the next stage of the process begins.

- 5.5.4. the third function of AFI is to create a BWF header, comprising the necessary metadata from the catalogue. AFI extracts this metadata indirectly, via the .imp.xml file that was created during 'export', in step 5.4. If AFI cannot find an .imp.xml file with the same persistent identifier as the audio file, it will skip the file and return the error message 'XML Import file not found'.
- 5.5.5. occasionally, audio files can become corrupted in a minor way that means they won't be taken by AFI, even if they are playable files. The error message returned is 'Cannot find Source BWF' and AFI will write a text file named with the audio file's name with the extension .err.txt that contains an error message.
 - 5.5.5.1. the course of action is to move the file out of 'send to seal' and into 'send to seal - rejected'
 - 5.5.5.2. then open the file in WaveLab and to save another copy directly back into 'send to seal'. The newer file will upload without a problem.
 - 5.5.5.3. when this happens for a large number of files in a given collection or batch, there is a preset in WaveLab's batch processor to fix it. Simply move the files to 'Send to Seal - rejected' as per usual,
 - 5.5.5.4. then run the batch processor (from 'tools'), select 'Send to seal - rejected' as the source folder, and load the preset "Cannot find source BWF". This preset just saves a copy back into 'Send to Seal' and deletes the problematic file.
 - 5.5.5.5. running the AFI 'process' again will now work, unless there is a further problem with the file. If this same error occurs, open the file in WaveLab and establish if it has been more seriously corrupted.
 - 5.5.5.6. where more serious corruption has occurred, check the most recent backup (there should be one somewhere). Keep regressing in stages until you find a version of the file's history that is not corrupted. If no such file exists, contact the depositor to establish if the file had been corrupted prior to being received.
- 5.5.6. AFI will deposit the files into a folder on Audiocube-2's D:\TmpBWF'. This folder is the source folder for BWF Editor.
- 5.6. if BWF Editor is already running, it may not have checked its source folder for new files. You can do this by pressing "R" (no ctrl or alt or anything), this will 'refresh' the file list.
- 5.7. click on a file to select it. It will display the quality events that Audio File Inspector recognised and wrote into the BWF header, these are listed on the left side of the screen. Press

"B" to open the dialogue box for the 'bext chunk' or, 'broadcast extension' of the BWF header. It is a good idea to check that all necessary pieces of metadata have been correctly harvested.

5.8. extra metadata must be manually inserted, until such a time as it can be automated.

5.8.1. select the fourth tab in the dialogue box, 'broadcast extension'. This contains information such as the operator, the inspector and the ingest notes. The operator should be filled in already as it is taken from the catalogue. If it is not filled in, ascertain who was in charge of the file (the person who digitised it, for instance) and enter their name. At the same time, enter it into the catalogue so that there isn't any conflicting information. 'Inspector' must be filled in so that the file's history can be tracked. Enter your name, even if you haven't done any 'inspecting' as such.

5.8.2. file quality is a number from 0 to 5 and is meant to represent the audio quality of the file. Entering a default value of 3 for born digital files is sufficient. If you know the file is better or worse, then by all means enter whatever number is appropriate. You can also infer from the ingest notes in the catalogue whether the file was of high quality or otherwise, if the ingest notes are that detailed.

5.8.3. add ingest notes from the Paradisec catalogue at the item level, copy-paste them into 'operator comments' and delete the comments that don't apply to that file. That is, the ingest notes refer to an entire item, while the broadcast extension you are entering them into is file-specific.

5.9. press OK to leave the BWF header dialogue and immediately save the file, otherwise all metadata changes will be deleted without warning when you select another file. Save by pressing ctrl+S.

5.10. now you can safely press the 'Data OK' button, on the bottom-right of the screen, or press ctrl+backspace This moves the file from D:\TempBWF into another folder on Audiocube-2, D:\PresealedBWF.

6. Archive Process (Audio)

6.1. fully automated Quadriga Batch process then begins, to seal the BWF and store it in '\\azoulay\PDSC Audio Area\Send to Archive' folder. Given below is a summary of the processes.

6.1.1. BatchProcessor.exe is a small program on Audiocube-2 that runs continuously and performs one function; to bind the digital seal around BWF files so that no editing, to the audio or the metadata, can occur without first unwrapping the BWF. BatchProcessor.exe automatically starts when Audiocube-2 is booted, and automatically processes any .tmp.bwf file inside 'D:\PreSealedBWF'. It writes the file back onto Azoulay, into '\\azoulay\PDSC Audio Area\Send to Archive'. BatchProcessor.exe contains no 'tools' or 'options' menu, meaning understanding what it does is not easy. It instead runs on a

single .cfg file, which can be viewed in a text editor. The .cfg file is located in C:\Audiocube\Quadriga\BatchProcessor and is called BatchProcessor.cfg. If this .cfg file is moved to any other directory or gets accidentally renamed, it will render BatchProcessor completely useless, as this directory is hard-wired into BatchProcessor's source code.

6.1.2. for reference, here are the contents of the .cfg file:

```

BrowsingCopyProfilePath      = D:\BrowsingCopyProfiles\A.profile.xml
ConcatFileName               = 0
CreateBrowsingCopy          = 0
DeleteGPK                    = 0
DeleteIntermediate           = 1
DeleteMRK                    = 0
DestinationFolder            = I:\Send to Archive
Dither                       = 0
DitherType                   = 1
ErrorOnXML                   = 0
ExpertsOnly                  = 1
Headroom                     = 0.50000000
IsXmlExport                  = 1
LogFileFolder                = D:\LogFiles
NoiseShapeType               = 1
Normalize                    = 0
NormalizeBrowsingCopy        = 0
ProcessPriority               = 0
Reserved                     =
SearchInSubFolders           = 0
Separator                    = _
SourceFolder                  = D:\PreSealedBWF
WriteBwf                     = 1
XmlExportFolder              = I:\BWF xml files\XMLExport

```

The crucial bits here are the source folder, D:\PresealedBWF, and the destination folder, I:\Send to Archive ('\\azoulay\PDSC Audio Area\Send to Archive'). If something doesn't go correctly, check that this .cfg file is correct. Note also, that the .cfg file uses the drive letters, such as I:\ for the server share "PDSC Audio Area" ([\\r0\d0](#)). This drive letter can potentially change on each individual computer; it is the decision of that computer's administrator. For that reason, it is necessary to be aware of which Audiocubes use which drive letters, and to keep them identical.

- 6.1.3. as well as writing the BWF file, BatchProcessor.exe writes an .exp.xml (export XML) and saves it in the directory '\\azoulay\PDSC Audio Area\BWF xml files\XMLExport.'
- 6.1.4. BatchProcessor.exe writes the BWF files to I:\Send to Archive. A script will run nightly on this folder and perform three checks.¹ They check that the file is more than zero bytes, since it is possible that an empty shell of a file could slip through, that the file has a PI of the format *?-*?-*?.???, that is, with exactly two dashes and at least something between them, and that the BWF is listed in the catalogue. If the file passes all three checks, it will be sent to a folder that has three aliases, or links to it. This folder is normally invisible to all users, but visible to APAC as the directory \r0\APAC, and visible to Sydney Uni's local backup archive as the directory \r0\UCC. The third link is visible to the audio person as \r0\Backup Files. The actual file itself is not deleted until all three links to it have been deleted by whoever has write access to their respective directories. If one of the three checks fails, the file will be sent to a separate folder, I:\Send to Archive - Rejected and an email will be sent to everyone involved to inform them.
- 6.2. periodically the APAC store will look at the '\\azoulay\PDSC Audio Area\Send to Archive' folder and copy any new files to the APAC store, including any files it already has but have been since updated, making re-archiving, where necessary, very simple. The backup archive process will store the batch of files on 200GB digital tape under a single archive tag; the tag and the names of files archived are emailed to everyone and are listed in the catalogue in the field "Archives". Clicking on a file in this field will open a message window that displays the tape number that the file is backed up to. The reason for this is file retrieval.
- 6.3. the duplicate tape is returned to the PARADISEC lab to be stored in the safe
- 6.4. files that will not need to be sealed, that is, anything but .wav files that you intend to archive, should simply be sent from '\\azoulay\PDSC Audio Area\Send to Seal' to '\\azoulay\PDSC Audio Area\Send to Archive', thus by-passing steps 5.2 through 5.10. But they will have to go through the same file checking procedures outlined in step 6.1.4 before they can be sent to the archive. Nothing can otherwise be written to the \r0\APAC directory and therefore, nothing can go into the archive unless it is checked by this script.
- 6.5. in case a file has been incorrectly archived and needs to be altered in some way, such as in the BWF header, and re-archived, it can be copied straight from the backup directory - do not delete it - and pasted inside workspace. If you put it through the process again and archive it, the original faulty version that still sits in the backup directory will not be a problem; APAC

¹ This script is located at /home/tomhoney/bin/to_archive.rb and is run nightly at 10pm via cron from the user account "tomhoney" on the server azoulay.

automatically senses that a filename that it has already seen, but with a more recent date, is in the [\\r0\APAC](#) folder and replaces its old version with the newer one. No file then, needs to be manually deleted from APAC, unless the name is different.

6.6. troubleshooting summary:

Error message:

XML import file not found

Action:

Create the XML import file by going back to the catalogue, making sure all the metadata is there, making sure the file has been 'imported', then use the 'export' function to write the .imp.xml file to the server.

Cannot find source BWF

Move all unsuccessful files to 'send to seal - rejected', run WaveLab's batch processor (tools > batch processor), select the send to seal - rejected folder as the input. Choose the preset 'Cannot find source BWF' and click 'run'. Run AFI again.

7. Restore from archive process (PC)

7.1. restoring a file on Azoulay, which was modified or created in the last several days:

- 7.1.1. note: If it is one of your files you can restore it, otherwise this action must be performed by the Super User (either Abed Kassis, or Tom Honeyman). You must be able to use a unix shell to do this.
- 7.1.2. connect to Azoulay by ssh (ssh username@azoulay.arts.usyd.edu.au)
- 7.1.3. navigate to the directory to which you want to recover or wind back a file (cd /r0/... etc)
- 7.1.4. run "recover -s nsrhost"
- 7.1.5. you should get a "recover>" prompt.
- 7.1.6. "help" will list available commands. "ls" will list files that you can recover.
- 7.1.7. note: files older than a month will not be shown. You may need to get an older tape loaded. Abed Kassis (or the ArtsIT server administrator) should handle this process by contacting ICT and getting them to load the correct tape. If you are recovering a file from the backups of "Send to Archive" then the tape number listed in the catalogue is necessary. Note however, that it will probably be easier to recover the file from Store/APAC (see 7.2 below). IMPORTANT: you must monitor the recovery process as it will prompt you if you need to overwrite a file. Leaving your terminal without responding

to the prompt will halt the tape robot, and may result in the automatic backups not working.

7.2. restore from Store/APAC:

7.2.1. if you do not have shell access to APAC then contact Stu Hungerford to request a recovery

7.2.2. if you have shell access and full read access to the collection, you should be able to use an sftp program, or ssh or scp via a unix shell. At present restores can only be done by Abed Kassis. It is advisable to do periodic restore from archive every six months or so.

8. Create a static repository in XML (PM)

(This step will only be necessary until the dynamic repository is working, it is waiting on Tom Murtagh implementing some last items)

8.1. in the 'export' screen of the catalogue (<http://azoulay.arts.usyd.edu.au/paradisec/port.php>), select 'Download static repository'. This will take some time as the XML file is created and presented on the screen.

8.2. when it is done, save the page to your computer.

8.3. this file then needs to be placed by ftp on the <http://paradisec.org.au> website, but first rename the existing PDSCsr.xml to PDSCsr[DATE].xml, where [DATE] is the date on which the earlier file was created

8.4. go to the archive registration site: <http://www.language-archives.org/register/register.php4>

8.5. enter the url (<http://paradisec.org.au/PDSCsr.xml>) into the 'Base URL' box, then click on 'Validate only'. It should be validated and then it will be harvested over the next 24 hours.

8.5.1. if it does not validate you will have to figure out why, edit the file and revalidate it, you can do this in an XML editor or in a text processor. If it still does not work then contact the OLAC administrators.

9. Check PARADISEC repository reports (Audio)

9.1. check PARADISEC repository audit report emailed weekly to all project members from stuart.hungerford@anu.edu.au

9.2. identify any anomalies and rectify, reprocessing relevant files as necessary.

10. Access to Archived files in Paradisec repository

10.1. anyone requiring access to a file in the archive must fill in an access form, available from the website: <http://www.paradisec.org.au/downloads.html> accompanied by a list of the files required

10.2. once the form is received, Admin sends a request to alison.scott@anusf.anu.edu.au to create a username and password for the new user and specifies which collections/items they should have access to and for what period of time

10.3. Admin sends the login details to the new user with instructions of how to access the repository

10.4. the user logs in at <https://auth.apac.edu.au/login> and will be able to see the files available to them

10.5. these files can then be downloaded in full archival or mp3 access versions in the case of audio or full archival .tiff or .jpg in the case of images. For each file, the online/offline status will be listed, any files that are not online will download after a few minutes wait time

11. Appendices

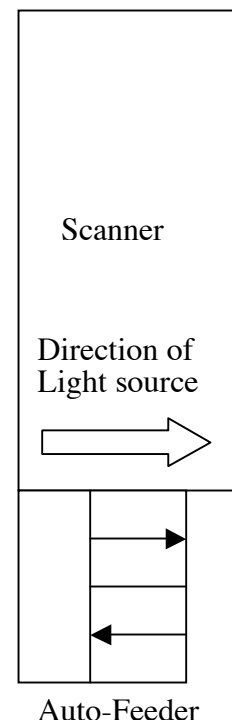
11.1. Appendix 1 Some notes on scanning slides

11.1.1. the slide feeder consistently picks up more than one slide at a time, which will jam the scanner. It is therefore necessary to watch the scanner each time it is at the point of ingestion and hold back all slides except for the first one.

11.1.2. dust on the slide surface may cause visual interference and slides should be clean. Blowing lightly will clear larger dust particles and a *very* light rub with a glasses cleaning cloth will remove the smaller bits, a fine paintbrush is also useful. Don't rub too hard as it can scratch the surface if you rub it with a dust particle.

11.1.3. portrait-oriented photos are fed into the scanner on their side as if they were landscape. Otherwise, only the middle section will be scanned, with black bars on either side. The scanner will scan the images from the side they are fed from. Usually, a slide has the gloss surface as the front, but it is worthwhile making sure, by finding a slide with writing on it to make sure it isn't back-to-front (see diagram; slide should have matt side facing the right, as the scan is taken from the left).

11.1.4. slides are usually numbered already, so ideally this numbering system should be followed. In which case, if some slides are missing, those numbered files should similarly be missing. This means that the auto-number feature of a batch process will have to be stopped and one number skipped.



11.2. **Appendix 2** Notes on Image Capture

11.2.1. equipment:

- 11.2.1.1. Canon Powershot G6
- 11.2.1.2. Powerbook G4
- 11.2.1.3. Camera tripod
- 11.2.1.4. 2 tungsten light bulbs

11.2.2. rotating the photographs

- 11.2.2.1. if the 'Rotation' button is dimmed and not working in the Remote Shooting and Image Browser windows, then the camera might have the 'Auto Rotate' function on. This function automatically rotates the image according to which way the camera is being held (ie. Portrait or landscape), this is useful if the camera is vertical, but not useful if it is being held horizontally. To turn off the Auto Rotate function on the camera:
 - 11.2.2.2. on the camera, press the menu button
 - 11.2.2.3. in the menu, use the arrow keys to scroll down to Auto Rotate
 - 11.2.2.4. press the left/right arrows to change it to 'off'
 - 11.2.2.5. if the Auto Rotate function is off, you can rotate images in the remote capture window before you take the photograph and it will also change the orientation of the view finder. It is also possible to rotate the photograph afterwards in the Image Browser window by selecting the photograph and selecting 'rotate' in the 'edit' menu.

11.2.3. folders

- 11.2.3.1. when starting the Remote Shooting window, you need to choose the folder that the photographs go into, the contents of which are displayed in the image browser. Put the photographs into a temporary folder (named eg. InProgress) first, and then, in 'Finder', to move the photographs to their proper folder afterwards.

There are a few reasons for this:

- 11.2.3.1.1. it is quicker to move images in 'Finder' after they have been taken than to change what folder they initially go in when the photo is taken.
- 11.2.3.1.2. the image browser slows down if it has too many (more than about 100) images to display, so if you are dealing with a large amount of photographs, you can transfer them to their proper folder progressively.
- 11.2.3.1.3. if you accidentally double click 'send to trash' it will not only delete all the photographs in the displayed in the Image Browser, but any other sub-folders that were also in the folder the photographs were saved in. It is all

recoverable from the Trash, but is it still better to have a folder that has all its contents on display.

11.2.4. shooting Settings in the Remote Capture Window

11.2.4.1. photo Effect

11.2.4.2. changing the Photo Effect from 'Effect Off' to 'Vivid' makes light pencil and faded ink stand out better.

11.2.4.3. focusing Point

11.2.4.4. the 'Manual Selection – Center' function means the process of taking the photograph and uploading it to the computer is slightly quicker because the camera does not refocus for each picture, because it automatically focuses on the centre of the image. However, if there is nothing in the centre of the page (for example the text is only at the top of the page, or in a column down the side of the page) or the writing is very pale, the image will be out of focus and it is better to change the Focusing Point to 'Automatic Selection'. Generally, leaving it on 'General Selection' lessens the chance of a fuzzy picture because the camera re-focuses for each photograph.

11.2.5. macro

11.2.5.1. if the 'Automatic Selection' focusing point is on and the image is still fuzzy, then change the Macro to 'On'.

11.2.6. lighting

11.2.6.1. having the lighting rig further away from the page makes the light more even. A higher lighting rig will also make it less likely to appear in large images where the camera has zoomed out, so it will not have to be moved very often.

11.2.6.2. if the paper is shiny or there is pencil which reflects the light, putting paper light shades on the light bulbs or turning them off altogether (if there is adequate lighting from the room light or a window) reduces the reflection.

11.2.7. general things

11.2.7.1. you can also release the shutter button by pressing the spacebar.

11.2.7.2. you can change the name of an individual in the Image Browser window by double clicking on it.

11.2.7.3. for images with little contrast (eg. Light pencil, faded ink), turning off the Tungsten light bulbs and, if there is daylight from a window, the room lights, as well as changing the Photo Effect in the Shooting Settings to 'Vivid' is a good way to get a clearer image.

- 11.2.7.4. if taking a lot of photographs of papers of a uniform size, drawing pencil dots on the backing cardboard to line up the paper speeds up the process.
- 11.2.7.5. folded-over post-it notes are a good way to stick down curling paper.
- 11.2.7.6. if you want to make a very small adjustment to the zoom, it is easier to adjust the height of the camera tripod than adjusting the zoom on the computer.
- 11.2.7.7. a suggestion for making copying books easier: if it were possible to make the automatic page numbering 'odd' or 'even', then every second page could be copied. This would be useful when photographing books because you could photograph all the left-hand pages and then all the right-hand pages, and you would only have to turn the page while keeping the book in place, rather than having to move it after taking each photograph.

Appendix 3: Web Delivery of Digital Files

Often the easiest method of returning digital files to depositors or other people who request access to individual files, is via the Azoulay web space. This document details how to make files accessible, and also how to protect them from being accessed by others.

1. Directories

The directory that forms the Azoulay webspace is located at `/r2/d0/webapp`, and is softlinked by the directory `/home/webapp`. It is also shown as a server share alongside PDSC Audio Area, Contract Audio Area, and other shares visible (though usually not accessible) to everyone on the network. This directory represents the root level of the Azoulay website, and any directories created here, show up on the website as sub-directories. That is, the folder `/r2/d0/webapp/AW1` will be accessible and viewable via the url `http://azoulay.arts.usyd.edu.au/AW1`. Without any `index.html` file inside this directory, the url will simply display a list of contained files and folders.

2. Softlinking

For ease of space and file management, softlinks will be used instead of placing files directly in the webapp folder. To provide web access to a complete directory of files that is already sitting somewhere on Azoulay, all that is required is to create a link inside `/r2/d0/webapp` that points to the desired directory on Azoulay. The benefit of doing this is that copies of files need not be created, saving plenty of time, bandwidth and space on the server, and also, completely blocking access to the files (once they have been downloaded by the desired individuals, for instance) is simply a matter of severing the link rather than making sure backups are kept and deleting them from the webapp space. However, a few precautions must be taken.

3. Permissions

The target directory must allow, at every step along the way, read and execute access to 'everyone'. In unix terminology, the modes must be something like `drwxrwxr-x` or, from left to right, it is a directory, it is readable, writable and executable for the creator, it is readable, writable and executable for the group, and it is readable and executable for everyone else. If at any stage there is a directory in the hierarchy that prevents read access to everyone else, the web page will simply display 'forbidden'. Contract Audio Area (`/r0/d1`) is set up to allow read access to everyone, so creating a folder in this directory to link to via the webapp, will work, provided your created directory also allows read access from everyone.

4. Procedure

The procedure for creating a directory to house web-accessible files and linking to them, therefore, is as follows:

1. Create a directory somewhere on Azoulay, such as, for instance, inside the Contract space:
`mkdir /r0/d1/example/`
2. Make sure the permissions are correct (`775` will give the permissions detailed above):
`chmod 775 /r0/d1/example`
3. Change directory to the webapp:
`cd /r2/d0/webapp`

4. Create a link to the target directory (this will name the link, and the directory on the web “example”):

```
ln -s /r0/d1/example example
```

5. Place the files that you wish to make accessible in the target directory.

The above procedure will allow all users to access the files inside /r0/d1/example simply by navigating to <http://azoulay.arts.usyd.edu.au/example>, but it will be more secure if you can password-protect the directory, so that you can control who can access those files.

5. Password protection

There is a website at <http://www.webmaster-toolkit.com/htaccess-generator.shtml> that will generate all the files that you need to restrict access to web pages. The required files to password-protect a page are .htaccess and .htpasswd. The prefixed periods ensure that these files are invisible to users for security reasons.

The .htaccess file looks like the following:

```
AuthUserFile /home/webapp/example/.htpasswd
```

```
AuthGroupFile /dev/null
```

```
AuthName "Password Protected Area"
```

```
AuthType Basic
```

```
<limit GET POST>
```

```
require valid-user
```

```
</limit>
```

and the .htpasswd file looks like this:

```
guest:guP41be8JGvIg
```

The coded line after the username 'guest' is an encrypted password that is created by the htaccess generator. When a user navigates to this page, in this case <http://azoulay.arts.usyd.edu.au/example>, they are prompted for a username and a password. The username is stipulated in the .htpasswd file, and the password they enter gets encrypted. If and only if the encrypted password matches the .htpasswd file will the user be allowed access to the directory.

The procedure for creating these files is as follows.

1. Enter the desired username and password into the htaccess generator, as well as the path to the directory you want to protect. This path will be something like /home/webapp/example, or to use the actual location rather than merely the links, /r2/d0/webapp/example. Click the 'generate' button.
2. Create two files in the directory that you want to protect. This directory must be the same as the directory that you specified in the htaccess generator. The files will be .htaccess and .htpasswd. Importantly, these absolutely must be raw text files with no formatting.
3. Copy the contents of the first box into the .htaccess file and save it.
4. Copy the contents of the second box into the .htpasswd file and save it.
5. Keep a record of which passwords you specified for each username.
7. Make sure that all users are able to read the .htaccess and .htpasswd files, otherwise they will be blocked.

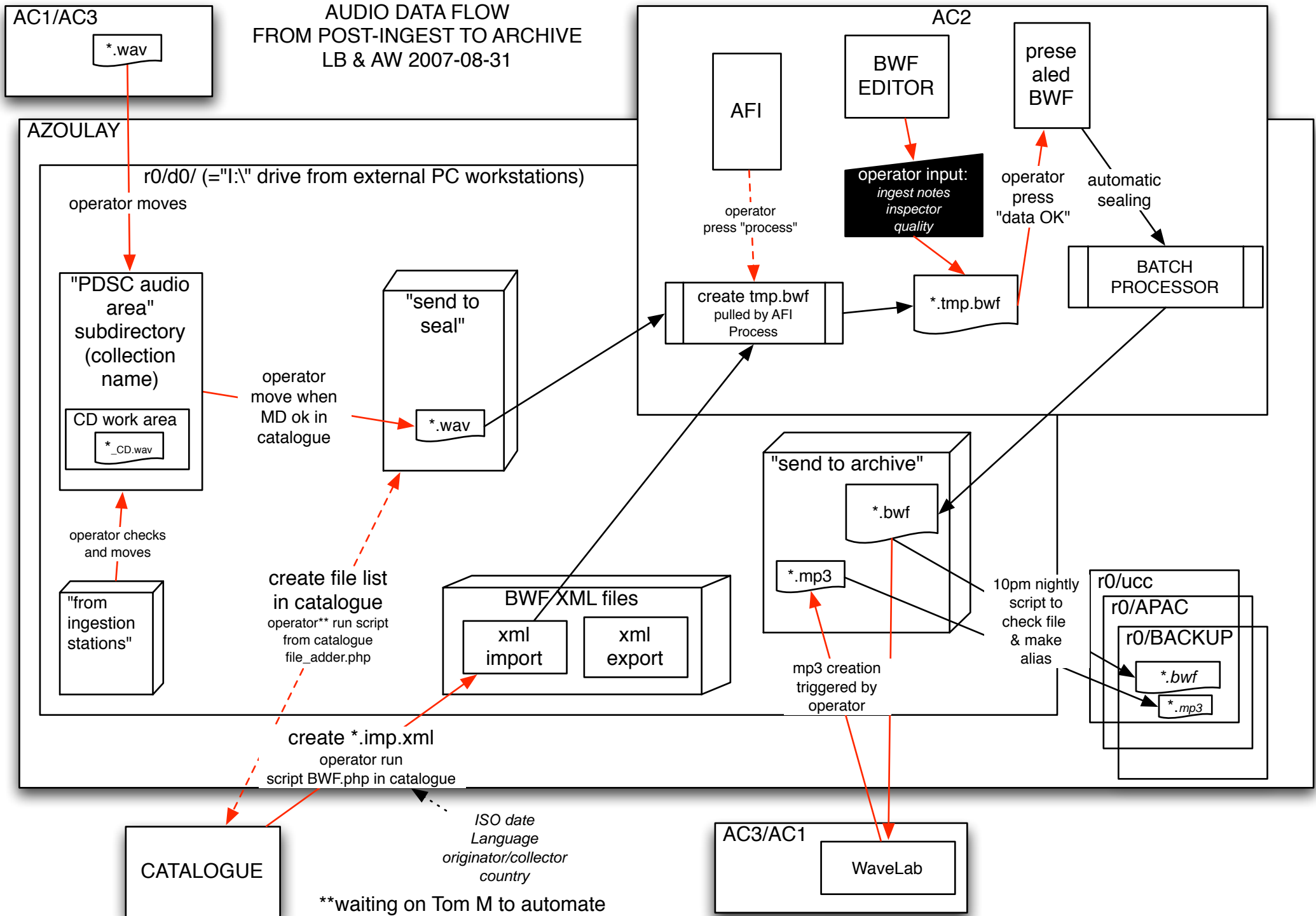
All that's then left to do is to email the user and inform them of the url of the directory, and tell them the username and password used to protect the page.

6. Removal

After the user has successfully downloaded all files they need, you might want to remove the files from being web-accessible. Due to the softlinking option, this can be safely done without risking the files, simply by removing the link. The target directory will still exist and will still contain all files, but without the link there will be no way to navigate to them from the web.

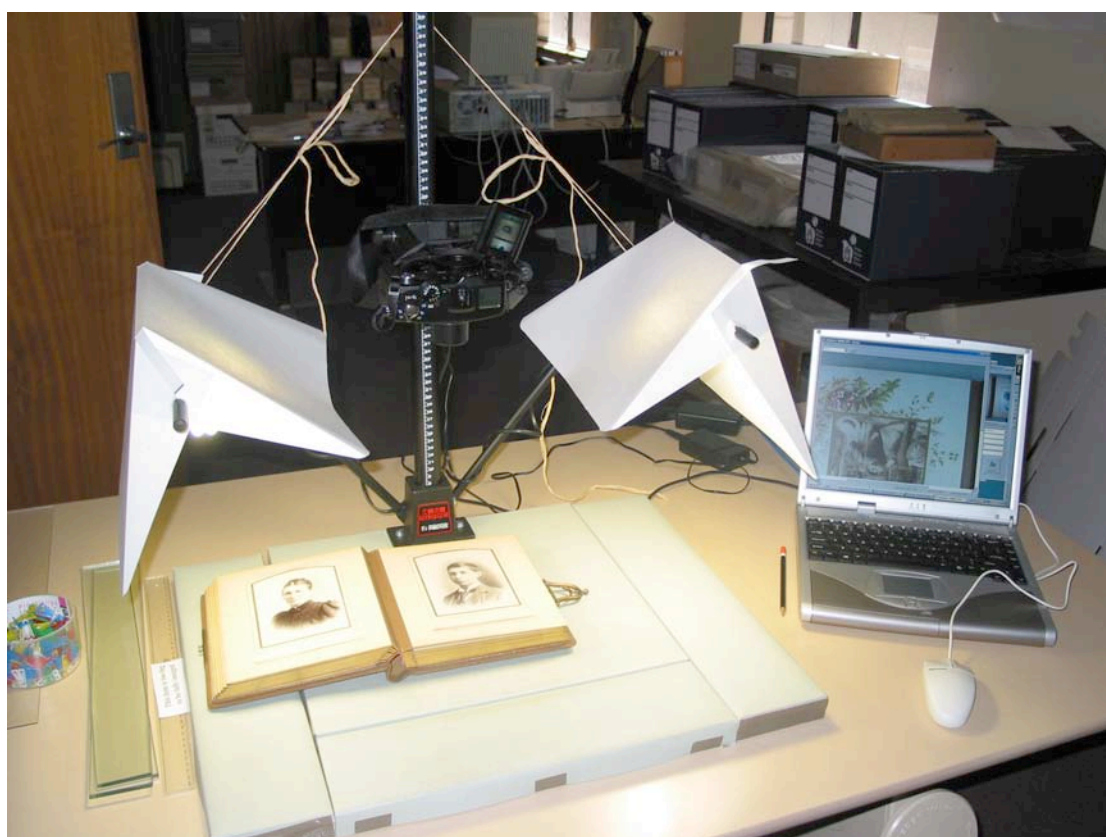
Remember to delete the link, and not the target. Navigate to the link via the webapp directory and delete the link from there. Remember also that the target directory will contain both the .htaccess and the .htpasswd files. You can safely delete these too, though they will be hidden from view.

**AUDIO DATA FLOW
FROM POST-INGEST TO ARCHIVE
LB & AW 2007-08-31**



Archival Imaging for Access

Manual 2004



Australian Science and Technology Heritage Centre

Contents

General	30
Overview	30
System Requirements	30
Hardware Requirements	30
Installation.....	30
Management.....	30
Capturing the Images.....	31
Setting Up.....	31
Settings	32
Basic Settings	32
Advanced Settings	33
Shooting	33
Reshooting or adding images	34
Reshooting as you go	34
Adding or replacing images in an unprocessed group.....	34
Adding or replacing images in a processed group.....	34
Tips and Trouble shooting.....	34
Positioning	34
Reducing reflection	34
Focus.....	34
Colour.....	34
Imaging Slides.....	34
Processing the Images	36
Setting Up.....	36
Editing	36
Rotating.....	36
Cropping.....	37
Reordering.....	38
Deleting.....	38
Item Title	38
Processing	38
Reprocessing	39
Editing existing images	39
Adding or replacing images.....	39
Creating an Index page.....	39
Using Images with the HDMS.....	40
Import Inventory Images Data.....	40
Page through Images associated with an Inventory Item	41
Include Images in Finding Aid	41

General

Overview

In September 2002, as part of an ARC funded project involving the University of Wollongong and the National Archives of Australia, Austehc developed a prototype system for the digital imaging of archival records and the automation of access to those images via finding aids (such as those created by the HDMS). This is not digitisation for preservation. It is digitisation for access and use. The primary goal of the system is to get as much imaged as possible for the lowest cost and have images web deliverable. This system, AIFA (Archival Imaging for Access) is not currently available as an add-on to HDMS users.

This manual documents the basic functions of AIFA (Archival Imaging for Access).

System Requirements

Hardware Requirements

- Computer
- Camera
- Stand
- Lights

Installation

- 1 Install Canon software that comes with the camera
- 2 Install ImgPro [more detail]

Management

- 1 If you are using a HDMS make sure it is the latest version with imaging capability, and set it up BEFORE starting the imaging process.


Check that you have an images folder under the HDMS web folder, i.e. something like this: c:\projects\rern\hdms\web\images.

- 2 If you are NOT using a HDMS you will still need to set up a folder in which to store images, i.e. something like this: c:\slides\web\images or c:\images.
- 3 In both cases you should also set up a folder under the c:\g5 (the Canon G5 folder) for each distinct project you are imaging for, i.e. c:\g5\brow, c:\g5\perr.

Capturing the Images

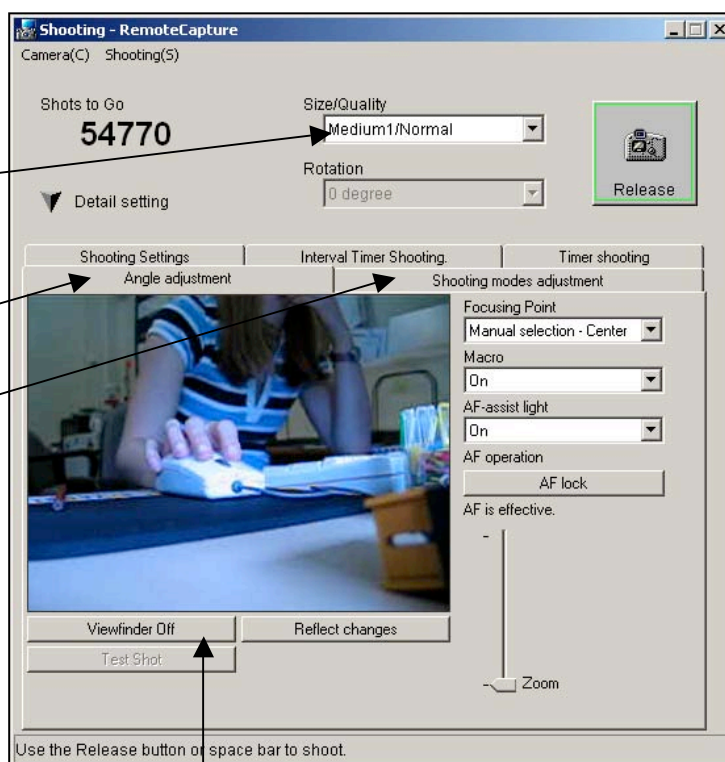
NOTE: the following instructions refer to the Canon G5, locations of features may be different for other camera models.

Setting Up

- 1 Turn on the computer first then attach the camera. The relevant software will appear after the camera is attached via the USB cable to the computer.
- 2 Select 'Canon ZoomBrowser EX' and click OK.
- 3 Close 'Camera Window'. ZoomBrowser will automatically open after this.
- 4 In ZoomBrowser, select 'Camera & Memory Card'.
- 5 Select 'Remote Shooting'.
- 6 Click , to connect to the camera. There will now be two windows open: 'Save – RemoteCapture' and 'Shooting – RemoteCapture'.

- 7 Change the settings in the following locations:
(See **Settings** for details)

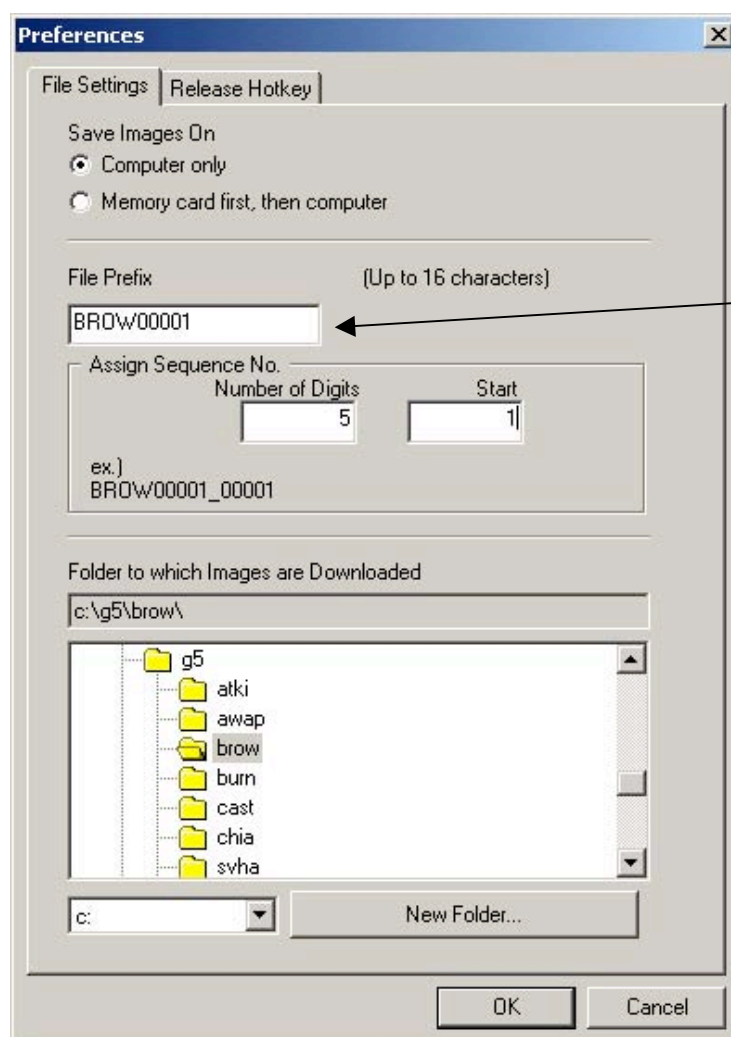
- In the 'Shooting - RemoteCapture' window
- In 'Angle adjustment' window
- In 'Shooting modes adjustment' window



Click VIEWFINDER ON to see what you are shooting on screen. [n.b. sometimes this darkens up as you shoot, which makes it harder to see if your alignment is correct; just click VIEWFINDER OFF and then VIEWFINDER ON again – it's a toggle button – which should brighten it up].

- 8 Still in 'Shooting', select Camera > Camera Display Settings and click Camera LCD monitor and OK if you want to view images through the LCD screen on the camera.

- 9 Now in the 'Save' screen, select File > Preferences. **This is important!** Type in a file prefix; if you are using the HDMS then this *must* be the Inventory ID, i.e. BROW00001. This will become the name of the folder under your 'images' folder (or folder as named and set up above under Set up folders/HDMS). You must also select the folder to which images will be downloaded from the camera (that is, the folder you set up under c:\g5).



The file prefix must be **unique**, have **no spaces** and **no slashes (/)**.

If you are using an item ID created pre-computing days such as series sequence with slashes, you *must* replace the slash with a dash, i.e. 1/146 must be entered as 1-146.

If using a HDMS, the HDMS software is programmed to find the dashes and replace them with slashes.

Note: You will always need to check and change the preferences in the 'Shooting' and 'Save' software on re-opening.

Settings

Basic Settings

The following settings will produce images of a usable quality. For further options, or for trouble shooting, see *Advanced Settings*.

	Type of Globes	
	Cool white globes	Warm white globes
Shooting – RemoteCapture window		
Size/Quality	Medium1/Normal	Medium1/Normal
Angle adjustment window		
AF Lock	On	On
Shooting modes adjustment window		
Flash	Off	Off

White Balance	Tungsten	Fluorescent
Exposure Compensation*	0	0
Flash Exposure*	0	0
Metering Mode*	Evaluative	Evaluative
ISO Speed*	50	50
Photo Effect*	Effect Off	Effect Off
Tv/Av setting*	Program AE	Program AE

*These settings should not need to be changed

Advanced Settings

The following settings will need to be tested until you are achieving satisfactory results.

Size/Quality: Medium1/Normal is the setting used for capturing images of reasonable quality and size. Other settings can be used if images are required for purposes other than as part of an archival collection. See the Canon G5 manual.

AF Lock: This locks the focus on the centre of the viewing area, which speeds up image capturing because the camera doesn't have to adjust the focus for every shot. For items with fine detail or close range photography, better results can be achieved by unlocking the AF and turning Macro on.

White Balance: The white balance is generally set depending on the type of globes being used with the imaging rig (see basic settings). However other lighting in a room, such as overhead lights or natural light from windows, can effect this setting. If you are not happy with the results you are achieving, try other White Balance settings until your images are accurately reproducing the colours of the original item.

Tv/Av setting: Select 'Manual' to control both the Aperture Priority (Av) and Shutter Speed Priority (Tv).

Av: the higher the number, the small the lens aperture. This brings more of the image into focus. Try setting at **8.0**

Tv: faster shutter speed allows for moving objects, slower shutter speed allows for shooting without a flash in dark areas. Try setting at **1/8**

Shooting

- 1 Set camera height so that the bottom of the rig's bracket lines up with 34.5cm (the standard height for optimal capture of standard sized images).

NOTE: Raise the height as appropriate to capture oversized objects, such as maps and foldouts, or lower the height for slides to 11.5cm.

- 2 Position item under camera and when satisfied it is in frame, click the RELEASE button in 'Shooting'. See Positioning below for tips.



- 3 Do not remove the item until you hear the camera click.
- 4 You can view an enlarged version of an image in the 'Save' screen. You can also delete and reshoot.

Reshooting or adding images

Reshooting as you go

If you are not happy with an image, make sure the image is selected (in blue) and click Delete. Reshoot the image. It is preferable that you do this as you go so that the images are in the correct order. However they can be reordered at the processing stage.

Adding or replacing images in an unprocessed group

It is possible to add or reshoot images to an existing group at a later stage. As long as the existing images with the same file prefix remain in the same folder and have not been processed in ImgPro, ZoomBrowser will allocate images with the same prefix the next number. So even if image 340_00001.jpg was taken three weeks ago, it would still give any more images with the prefix 340 the next numbers. If it is not possible to image the numbers in sequential order the images will have to be carefully reordered at the processing stage.

Adding or replacing images in a processed group

See Reprocessing below.

Tips and Trouble shooting

Positioning

Always line up items in the centre and try to align the edge of the item parallel with the edge of the viewfinder. The images will be rotated (and can be cropped) at the processing stage. If landscape items are positioned facing the camera rig they will not require rotating.

Reducing reflection

Glossy photographs and documents can cause problems with lights reflecting on them. Some ways to reduce reflection are:

- Lower light arms on rig as far as possible
- Angle shades on the lights away from the item
- Turn off room lighting or block light from windows
- Place tissue paper between the lights and the item in order to filter the light.

Focus

If an item has very little contrast, such as a document with little or faded text, the camera can have difficulty focussing. If this happens, carefully place something with contrast over the document and remove it after you have hit release but before the camera captures the image.

Colour

When an item is very similar in colour to the background surface, the camera can have trouble differentiating between them. This can result in the colour of the item being altered as the camera attempts to focus on it. If you are unhappy with the results you can try using different colours as a background (e.g. black or white) to stop the background colour fighting with the item colour.

Imaging Slides

- 1** Lower the camera to 11.5cm.

2 Turn off lights on rig and room lights.

3 Use the following settings:

Size/Quality	Medium1/normal
Flash	Off
Macro	On
White Balance	Auto
Exposure compensation	-1

4 Place slide, with labelled side down, on a light box with a mask around the slide.

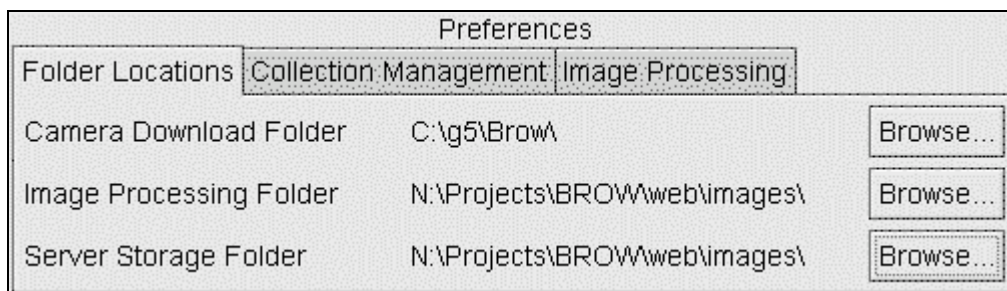
5 Capture the image.

Processing the Images

Setting Up

1 Open ImgPro

Select File > Preferences. **This is important!** There are three tabbed options under preferences (most data can be entered using the BROWSE buttons). You **must** set your preferences in Folder Locations and Collection Management.



Folder Locations:

Camera Download Folder – select folder set up at Management 3, i.e. c:\g5\Brow
Image Processing Folder – select folder set up at Management 1 or 2, i.e. c:\images
Server Storage Folder – select folder set up at Management 1 or 2, i.e. c:\images

NOTE: The camera initially downloads images to the Camera Download Folder. The AIFA software processes the images into the Image Processing Folder. You may wish to store the processed images to another location, such as the final location where the images will be 'served' from (i.e. a network server). Select this location as the Server Storage Folder.

Collection Management:

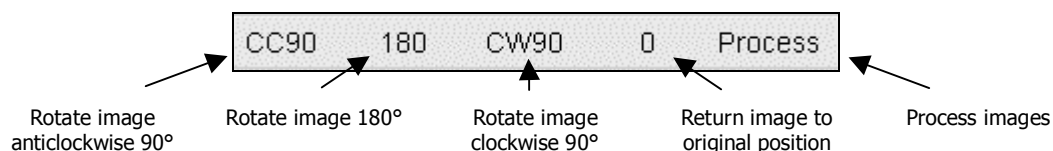
Click on the Collection Management tab to enter the Collection ID Code and the Collection Title and to browse for the HDMS location (leave HDMS location blank if not using HDMS).

- 2 Select File > Open. Select the Item ID that you wish to process from the list and click OK. A thumbnail version of each image will appear down the left side of the screen and a large version of the selected image will appear on the right.

Editing

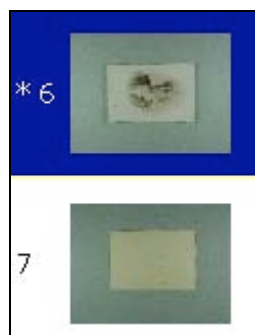
Rotating

Across the top of the screen is a bar with five buttons on it. They perform the following functions:

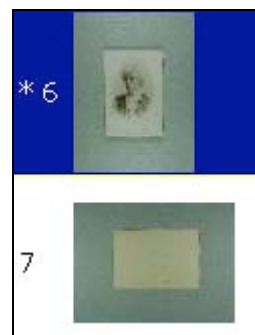


-
- 1 Select image by clicking on it. The selected image will be highlighted (blue) and will have an asterisk (*) beside the number.

NOTE: To select more than one image hold the ctrl button and click on each required image.



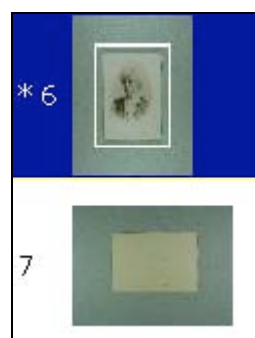
- 2 Click on the button that carries out the required function (see above).



Cropping

Autocropping:

Select one or more images for cropping, then click the right mouse button and select 'Autocrop'. This will place a white box around the area to be cropped on each image. This is not 100% accurate but it is a quick way of cropping large numbers of images. It is possible to go back to images you are not happy with and manually crop them.



Manual Cropping:

This needs to take place before an image is rotated. Select one image. In the large version of that image, click in the top left-hand corner of the required area. Hold the mouse button down and drag across the image to the bottom right-hand corner of the required area. Release the mouse button. A white box will appear 50 pixels outside the area.

Uncropping:

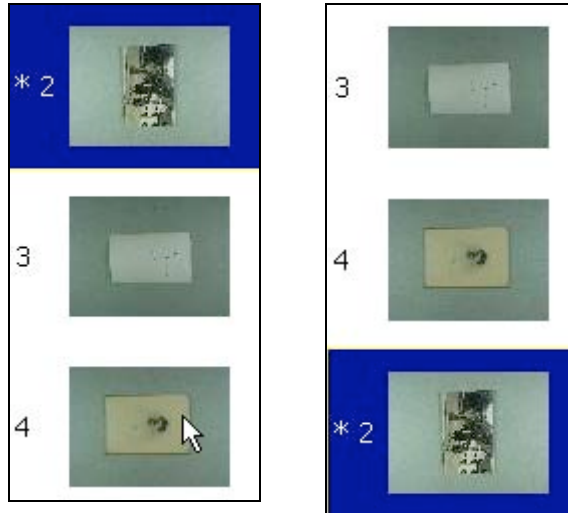
Select image that you no longer wish to crop. Click the right mouse button and select 'Uncrop'. This will remove the white box from the image

Reordering

Place an image **later** in list

Select image to be reordered. Click the right-hand mouse button on image that you want selected image to be **after**. Select 'Move Selected'. The image will retain its original number but will be renumbered accordingly during processing.

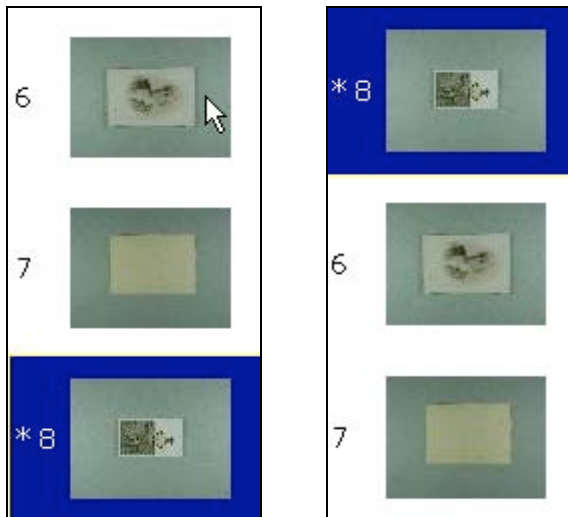
i.e. to reorder image 2 **after** image 4, click the right-hand mouse button on image 4.



Place an image **earlier** in list

Select image to be reordered. Click the right-hand mouse button on image that you want selected image to be **before**. Select 'Move Selected'. The image will retain its original number but will be renumbered accordingly during processing.

i.e. to reorder image 8 **before** image 6, click the right-hand mouse button on image six.



Deleting

Select one or more images to be deleted. Click the right-hand mouse button and select 'Delete'.

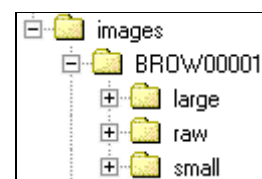
Item Title

Select Edit > Item Title. Enter a title for the item. Hit OK. If using the HDMS the item title should be the same as the inventory title or the inventory identifier.

Processing

When you are satisfied that all the images are correctly rotated and cropped, click on 'Process'. Processing takes a short period of time, longer for larger groups of images.

When the images have been processed a folder named with the file prefix are created in the Image Processing Folder. In the folder are three folders labelled large, raw and small. The folders contain large, small or raw (images created by the camera) copies of each of the images.



Reprocessing

Editing existing images

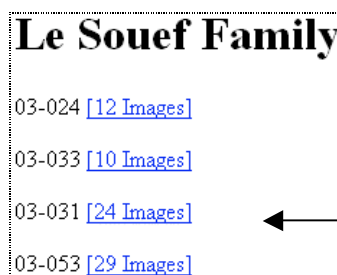
- 1 Select File > Preferences. In Folder Locations select the Image Processing Folder that contains the processed images you wish to revisit. Set all the Preferences as you would when processing a batch of images (see Setting Up).
- 2 Select File > Reprocess. Select the Item ID that you wish to reprocess from the list. Thumbnails of the images will appear down the left-hand side of the screen. Copies of the original images will be returned to the Camera Download Folder.
- 3 Edit **All** images.
- 4 Process.

Adding or replacing images

- 1 Select File > Preferences. In Folder Locations select the Image Processing Folder that contains the processed images you wish to revisit. Set all the Preferences as you would when processing a batch of images (see setup above).
- 2 Select File > Reprocess. Select the Item ID that you wish to reprocess from the list. Thumbnails of the images will appear down the left-hand side of the screen. Copies of the original images will be returned to the Camera Download Folder.
- 3 In ZoomBrowser, make sure that you change the file prefix in File > Preferences so that it is the same as the group of images you are revisiting. You **MUST** also reset the start sequence number to the next number, ie. if you already have 20 images, start at 21.
- 4 Shoot additional images, or reshoot images
- 5 Return to ImgPro and select File > Refresh. The thumbnails of new images will be added to the end of the existing list of thumbnails
- 6 Edit and reorder **ALL** of the images
- 7 Process

Creating an Index page

- 1 In ImgPro, select File > Write Index Page.
- 2 Select the collection to be indexed from the list, hit OK.
- 3 Copy the file `\\bart\general\prjects\ai\test_image_viewer.html` into the image processing/server storage folder, alongside the newly created PROJ-index.html file.
- 4 Open file PROJ-index.html in your web browser.



This creates an index page with the name of the collection, a list of all of the item titles and a link to the images for each item.

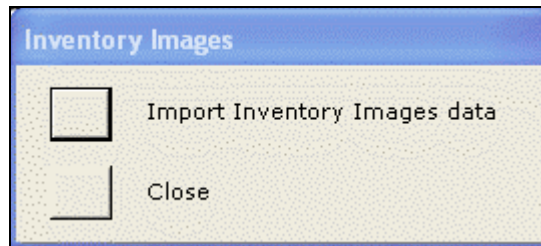
Page through the set of images associated with this item

Using Images with the HDMS

Import Inventory Images Data

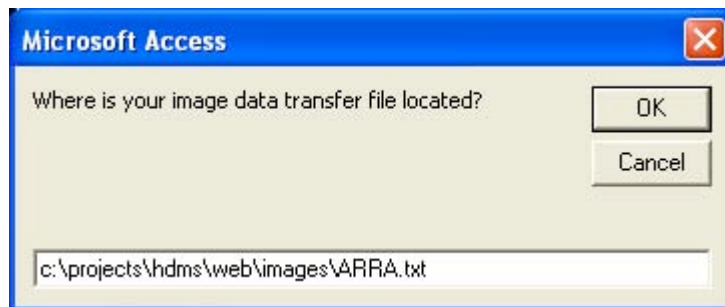
After using AIFA to digitise records documented in HDMS you may associate those images with the HDMS collection.

- 1 Select INVENTORY IMAGES from the MANAGE drop down menu in the Main Menu screen.

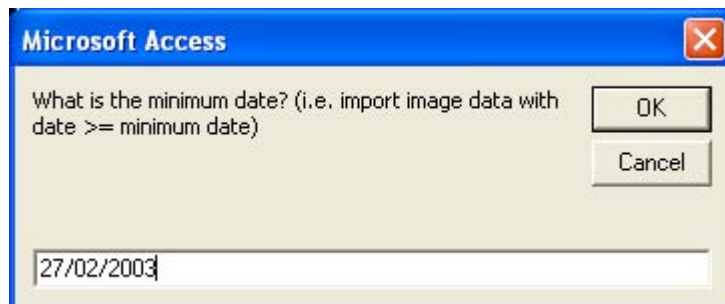


- 2 Click 'Import Inventory Images Data'.

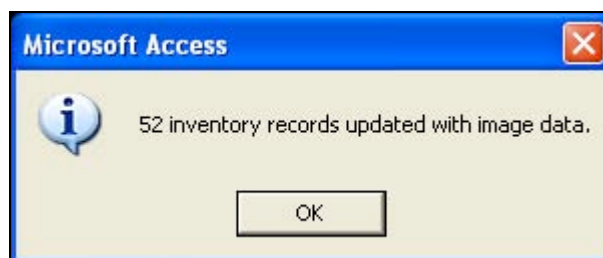
- 3 Type in the location of the image data transfer file (a txt file).



- 4 The txt file holds cumulative data from the imaging process. Input the minimum date of data you wish to import. The default is today's date, and this may be changed.



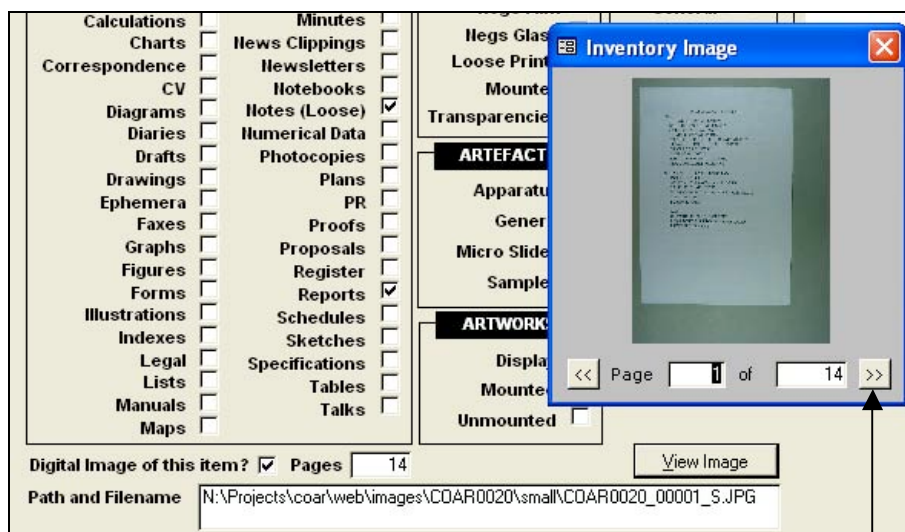
- 5 If successfully updated, you will get a message telling you how many inventory records have been updated with image data.



Page through Images associated with an Inventory Item

Once inventory images prepared using AIFA have been imported into the HDMS, you can view and page through the images per inventory item.

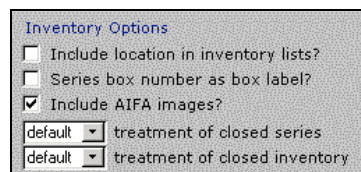
- 1 In the Inventory table, scroll to the bottom of the screen.
- 2 'Digital Image of this item?' will already be selected and the path and filename will be filled in.
- 3 Click VIEW IMAGE to page through the images.



Page through images associated with this item

Include Images in Finding Aid

At step 5 of the HTML Wizard you have the option to include images in the finding aid, which refers only to those images prepared using AIFA which have been imported into the HDMS. If you check this box then images will be included in the finding aid:



Inventory Listing

5-1 Brown Family - Photographs

Envelope inscribed 'Brown Family', containing: 42 photographs of Gilbert, Marie and Ian Brown and members of their families, some labelled.

Provenance [Ian Brown Family](#)
Date range c1900 - c1940 **Quantity** 2.3 cm, 1 file
Formats Loose Photographic Prints and Mounted Photographic Prints
Inventory ID BROW00007 Box Number 1 [Series 5](#)

[Image >>](#)

Click Image>> to page through the set of images associated with this inventory item