WHERE WE ARE TODAY: AN UPDATE TO THE UCONN SURVEY ON JPEG 2000 IMPLEMENTATION FOR STILL IMAGES

JPEG 2000 SUMMIT
MAY 12, 2011
LIBRARY OF CONGRESS
WASHINGTON, D.C., USA

David Lowe & Michael J. Bennett
University of Connecticut Libraries, Storrs, CT, USA

Agenda

- The Survey
  - Background & Results
  - Interim Developments
  - Survey Takeaways
- Technical Considerations
  - Tools
  - Format Landscape

Survey Background

- 2003-4: Aware-partnered JPEG2000 project and conference at UConn
- 2007: Adobe product manager's blog queries community re: its JPEG2000 support in Photoshop
- 2008: We develop and post our survey to gauge JPEG2000 acceptance status among cultural heritage institutions

Survey Results

175 responses:

1. **How would you classify your institution?**
   a. **77% Academic/Research Libraries**
   b. **16% Public Libraries**
2. **Breakdown of "other"**
   a. **Corporate**
   b. **Service Bureaus**
   c. **etc.**

Survey Results

3. **Do you use the JPEG2000 file format at all?**
    a. **60% Yes (103** responses)
    b. **40% No (70** responses)
4. **If you do NOT use the JPEG 2000 file format at all, why not?** (59 responses)
    a. **Software shortcomings**
        a. **Creation/Manipulation**
        b. **Access (Browser Support)**
    b. **Lack of staff expertise**
    c. **Patent issues**

Survey Results

5. **Do you use JPEG 2000 as an archival format for new collections?** (142 responses)
    a. **20% Yes**
    b. **80% No**
6. **Do you use JPEG 2000 as an archival format for images converted from legacy formats?** (141)
    a. **16% Yes**
    b. **84% No**

Survey Results

7. **Do you use JPEG 2000 to provide online access images for new collections?**
    a. **53.5% Yes**
    b. **46.5% No**
8. **Do you use JPEG 2000 to provide online access for images converted from legacy formats?**
    a. **46% Yes**
    b. **54% No**

Survey Results

9. **What tools do you use in your JPEG 2000 workflows? (Please indicate all that apply)**
    a. **53% Photoshop**
    b. **37% CONTENTdm**
    c. **19% IrfanView**
    d. **18% Aware**
    e. **17% Kakadu**
10. **Note any tools not listed above.**
    a. (No dark horses.)

Survey Results

11. **If you have migrated sets of files to JPEG2000 from legacy formats, what tools did you use?**
    a. **Photoshop**
    b. **CONTENTdm**
12. **What do you see as the strengths of the available tools?**
    a. **Ease of use**
13. **What do you see as the weaknesses?**
    a. **Slowness**

Survey Results

14. **What do you see as the weaknesses of the available tools?**
    a. **Browser support**
15. **What do you see as viable, lasting alternatives to JPEG 2000 for archival master and/or access derivative copies?**
    a. **TIFF for archival**
    b. **JPG for access derivative**

Survey Results

16. **Do you use or would you consider *mathematically lossless* JPEG 2000 compression?**
    a. **Strong yes for archival**
17. **Do you use or would you consider *visually lossless* JPEG 2000 compression for archival master purposes?**
    a. **Near half yes for archival**
    b. **No and undecided split the rest**

Interim Developments

- Inline Browser support effort
- Djatoka Image Server
  - Local testing promising

Survey Takeaways: Misconceptions

- **Lack of trust in JPEG2000 lossless compression**
  - ***IS** truly lossless*
- **File size savings not seen as significant vs. TIFF**
  - ***Average of 1:2 in size savings vs. TIFF***
- **Lack of awareness of higher bit depth range**
  - ***Includes 48 bit support***
- **Seen as lossy-only and proprietary**
  - ***Neither is the case***
- **Note preference of JPEG2000 as access format**
  - ***Designed to scale from archival through access derivatives***

JPEG 2000: Some Hang-ups Persist

- Among general lack of software support, also notable is the lack of Adobe Lightroom support beyond one known plugin which spoofs the program.
- Limited native DAM software support "out of the box." Dodgy performance once implemented.
- Damaging PR: Nov/Dec 2009 Dib article
- Lossy JPEG 2000 files prone to encoding errors in Adobe Photoshop when created in large batches.

Lightroom JPEG 2000 Plugin

- http://www.lightroom-plugins.com/JP2index.php
- Plugin works by making Lightroom think that it is reading a TIFF instead of a JPEG2000 file.
- Renaming JPEG 2000 filenames using Lightroom subsequently doesn't work entirely correctly.
- Editing embedded metadata first requires the creation of a backup file.
- It is good, however, to see that clever work is being considered in this area.

Limited DAM Support

- Most DAM software packages leverage ImageMagick, http://www.imagemagick.org/script/index.php, to do their heavy image manipulation lifting (see later slides for more on the problems of this with JPEG 2000)
- Many DAMs don't configure JPEG 2000 support out of the box.
- Many DAMs now have browser-based interfaces (both front and back end) and have a hard time displaying JPEG 2000 images as a result of browser JPEG 2000-rendering limitations.

Damaging PR: DLib Article

- http://www.dlib.org/dlib/november09/kulovits/11kulovits.html
- Much-cited by JPEG 2000 skeptics.
- Article states its conclusions on JPEG 2000's weaknesses based primarily upon the performance of two open source tools…

ImageMagick TIFF to JP2000	Version: 6.5.3-5 2009-06-11 Q16	`-compress Lossless -quality 100`
GraphicsMagick TIFF to JP2000	Version: 1.3.6 2009-07-25 Q8	`-compress Lossless -quality 100`

Damaging DLib Article

- Concludes that, “The direct pixels comparison, using both GraphicksMagick's and ImageMagick's compare functionality, indicated that pixels had been changed during migration from TIFF to JPEG 2000…”
- It is worth noting that van der Knijff (2010)* has recently cited that among the tools he tested, ImageMagick was particularly poor in its ability to accurately interpret JPEG 2000 conversions based upon its JasPer JPEG2000 library dependencies.**

* http://www.udfr.org/jp2kwiki/images/4/4f/Jp2kMigrationCharacterisationKBExternal.pdf

** http://studio.imagemagick.org/discourse-server/viewtopic.php?f=1&t=15807

Damaging DLib Article

- van der Knijff goes on to state, “Of all tools, ImageMagick’s ‘Identify’ tool shows the poorest performance: the information it provides on resolution is erroneous and incomplete. It only detects ICC profiles that are of the ‘restricted’ type. Moreover, it reports non-existent ICC profiles when colour is defined using an enumerated (e.g. sRGB) colour space. Because of this, I would advise against the use of ImageMagick for the characterisation of JPEG2000 files.”*

* http://www.udfr.org/jp2kwiki/images/4/4f/Jp2kMigrationCharacterisationKBExternal.pdf

Damaging DLib Article

- On the other hand, independent testing at UConn of TIFF > lossless JPEG 2000 conversion using the Photoshop CS4 JPEG 2000 plugin, confirms that JPEG 2000's lossless compression is **truly lossless at the pixel level** (using stacked TIFF & JPX layers of same image > toggling difference blending mode > histogram check).
- Murray (2007) has also previously done similar direct testing of Kakadu and also reports similar lossless JPEG 2000 compression results.*

* http://dltj.org/article/lossless-jpeg2000/

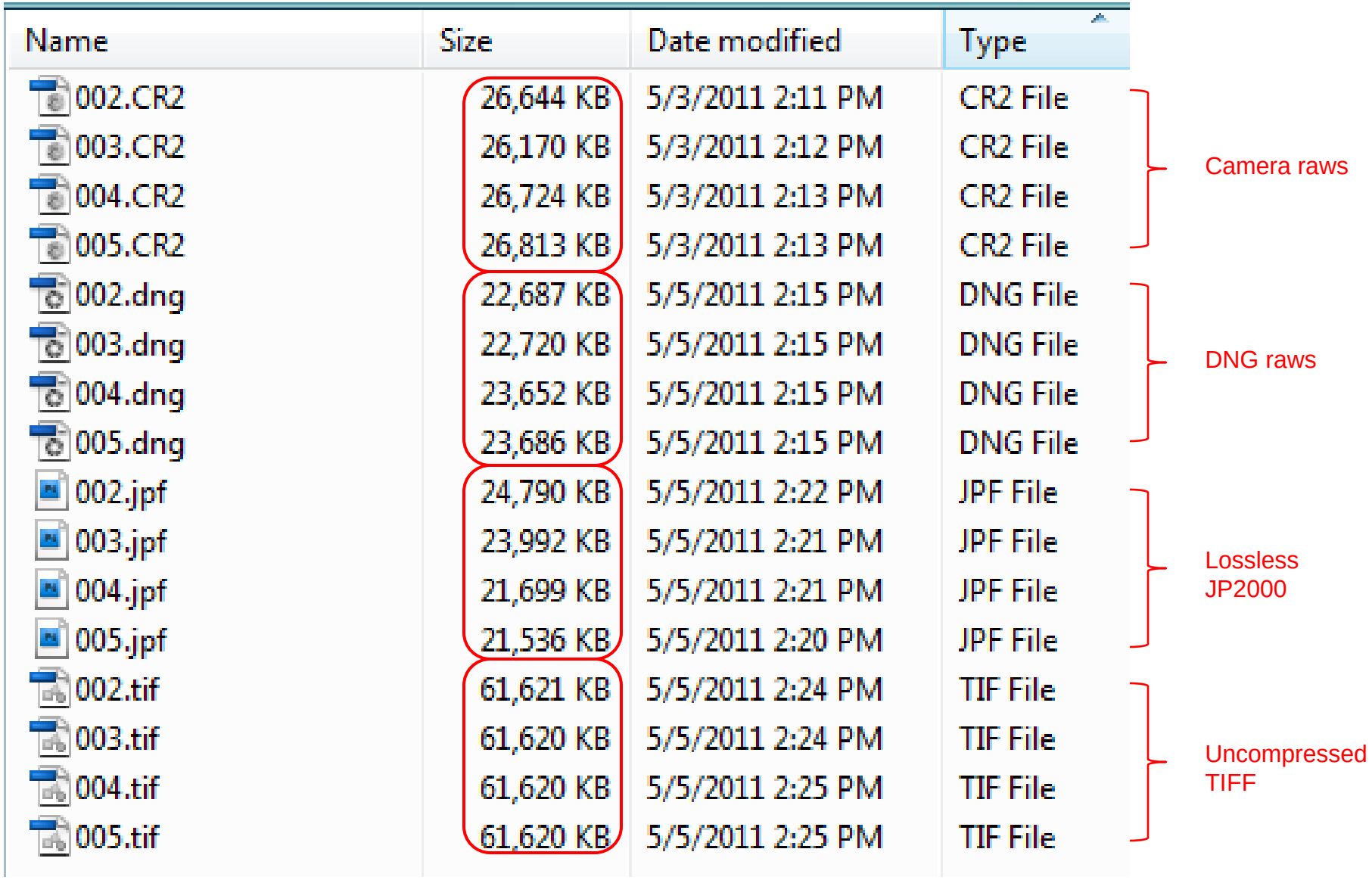
Damaging DLib Article

So, resist the urge to completely judge the specification's attributes based solely on the performance of one of its more inconsistent tools.

JPEG 2000: Leveraging Lossless Compression

- A solid substitute for uncompressed TIFF archival files (for those who need rendered archival files and want to save storage space)
- In so doing, makes it easier to also archive raw DNG "safety masters" along with a rendered format (JPEG 2000). For a given image, storage footprint results in something smaller than a single uncompressed TIFF.

Archival Storage Considerations



Name	Size	Date modified	Type
002.CR2	26,644 KB	5/3/2011 2:11 PM	CR2 File
003.CR2	26,170 KB	5/3/2011 2:12 PM	CR2 File
004.CR2	26,724 KB	5/3/2011 2:13 PM	CR2 File
005.CR2	26,813 KB	5/3/2011 2:13 PM	CR2 File
002.dng	22,687 KB	5/5/2011 2:15 PM	DNG File
003.dng	22,720 KB	5/5/2011 2:15 PM	DNG File
004.dng	23,652 KB	5/5/2011 2:15 PM	DNG File
005.dng	23,686 KB	5/5/2011 2:15 PM	DNG File
002.jpf	24,790 KB	5/5/2011 2:22 PM	JPF File
003.jpf	23,992 KB	5/5/2011 2:21 PM	JPF File
004.jpf	21,699 KB	5/5/2011 2:21 PM	JPF File
005.jpf	21,536 KB	5/5/2011 2:20 PM	JPF File
002.tif	61,621 KB	5/5/2011 2:24 PM	TIF File
003.tif	61,620 KB	5/5/2011 2:24 PM	TIF File
004.tif	61,620 KB	5/5/2011 2:25 PM	TIF File
005.tif	61,620 KB	5/5/2011 2:25 PM	TIF File

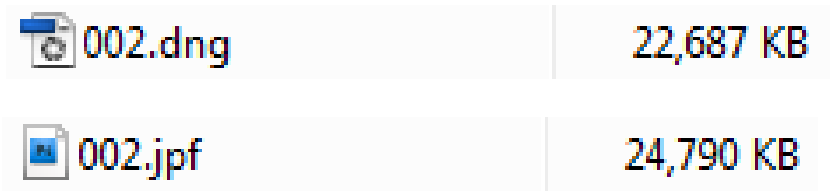
Camera raws

DNG raws

Lossless JP2000

Uncompressed TIFF

Archival Storage Considerations: getting richer data preservation bang for your storage footprint buck



You can archive both the original latent raw image data & a losslessly rendered format…

…all while using less storage space than a single uncompressed TIFF

47,477KB (DNG + JPF) vs. 61,621KB (TIF)

In Turn @ UConn Libraries…

For special collection printed & illustrated texts, and maps we reformat and archive in this manner:

*1) DNG raw "safety masters" (converted either from camera raws or native from scanners running VueScan)**

2) Lossless JPEG 2000 "archival masters" (reversible JPX, Photoshop)

3) Lossy JPEG 2000 "processed masters" (irreversible JP2, Photoshop)

**For additional background on DNG as an archival format, see: http://digitalcommons.uconn.edu/libr_pubs/23/*

Contact

David B. Lowe

Preservation Librarian

University of Connecticut Libraries

david.lowe@uconn.edu

Michael J. Bennett

Digital Projects Librarian & Institutional Repository Manager

University of Connecticut Libraries

michael.bennett@uconn.edu