

# British Library JPEG 2000 profile

Sean Martin

# Introduction

- The profile will be used for mass digitisation and particularly newspapers
- There is a trade-off between quality and the size and cost of file storage
- It has already been decided that lossy compression is not only acceptable but economically desirable
- It is also desirable that the same master file support the needs for both long term archival and also access

# Principal use cases - 1

- *(a) Navigation*: display of thumbnail images from multiple master files
- *(b) Reading*: display at an intermediate “reading” resolution to a single master with zoom and pan (and occasionally two pages side-by-side)
- *(c) Detailed*: display at full resolution with pan
- *Observation*: (a) and (b) will be much more frequent than (c)

# Principal use cases - 2

Navigation



Reading



Detailed

**A USTRALIA. — WHITE STAR CLIPPERS,**  
from Liverpool to MELBOURNE:—

Ship.	Captain.	Tons Reg.	Tons Bur.	To Sail.
Beechworth ..	G. Frain	1,266	4,000	Dec. 1.
Empire of Peace	T. Baker	1,540	4,600	Dec. 20.
Red Jacket ..	R. Kirby	2,460	5,000	Jan 20.

The magnificent clipper ship Empire of Peace will be despatched punctually on the 20th of December, with a mail, cargo, and passengers. She was built expressly for the Australian passage trade by Messrs. Wright, the builders of the celebrated clippers White Star, Morning Light, &c., which have made some of the fastest passages on record, and it is expected that this noble clipper will fully sustain the high reputation which her owners have earned. She is the largest and finest sailing ship on the berth, and quite new, having only made one voyage from St. John's to Liverpool. Saloons supplied with bedding, and all necessaries. She has excellent accommodation for all classes of passengers. For freight or passage apply to H. T. Wilson and Chambers, 21, Waterstreet, Liverpool; or Seymour, Peacock, and Co., 116, Fenchurchstreet; or to Grindlay and Co., 63, Cornhill, London.

# Aspects of JPEG 2000 which influence performance

- Number of resolution levels: *Influenced by use cases*
- Number of quality layers: *Influenced by use cases*
- Tile or precinct & codeblock size: *2 principal options*
- Progression order: *5 principal options*
- Code/decoder speed-up features: *Beneficial for speed*
- Code-stream markers: *Beneficial for speed*

These lead to 10 potential combinations for investigation

# Performance analysis – 1

<b>Table 1:                      Encoding WITHOUT code stream markers followed by decoding                      Decode Times (mm:ss) for a Test File</b>						
Use Case						
Progression Order	A - thumbnail		B - reading		C - detailed	
	Tiles	Precincts	Tiles	Precincts	Tiles	Precincts
RLCP	00:17	00:07	01:39	01:11	01:22	04:10
RPCL	00:17	00:06	01:40	01:11	01:23	03:59
LRCP	00:16	00:07	01:42	01:22	01:22	04:08
PCRL	02:33	02:50	03:47	04:13	01:22	03:38
CPRL	02:36	03:00	03:50	04:13	01:22	04:24

- There are two poor progression orders: PCRL and CPRL
- No single combination is the best for all use cases
- Precincts are faster than tiles for use cases (a) and (b)
- There is no significant difference between the remaining three progression orders but RPCL is marginally better

# Performance analysis - 2

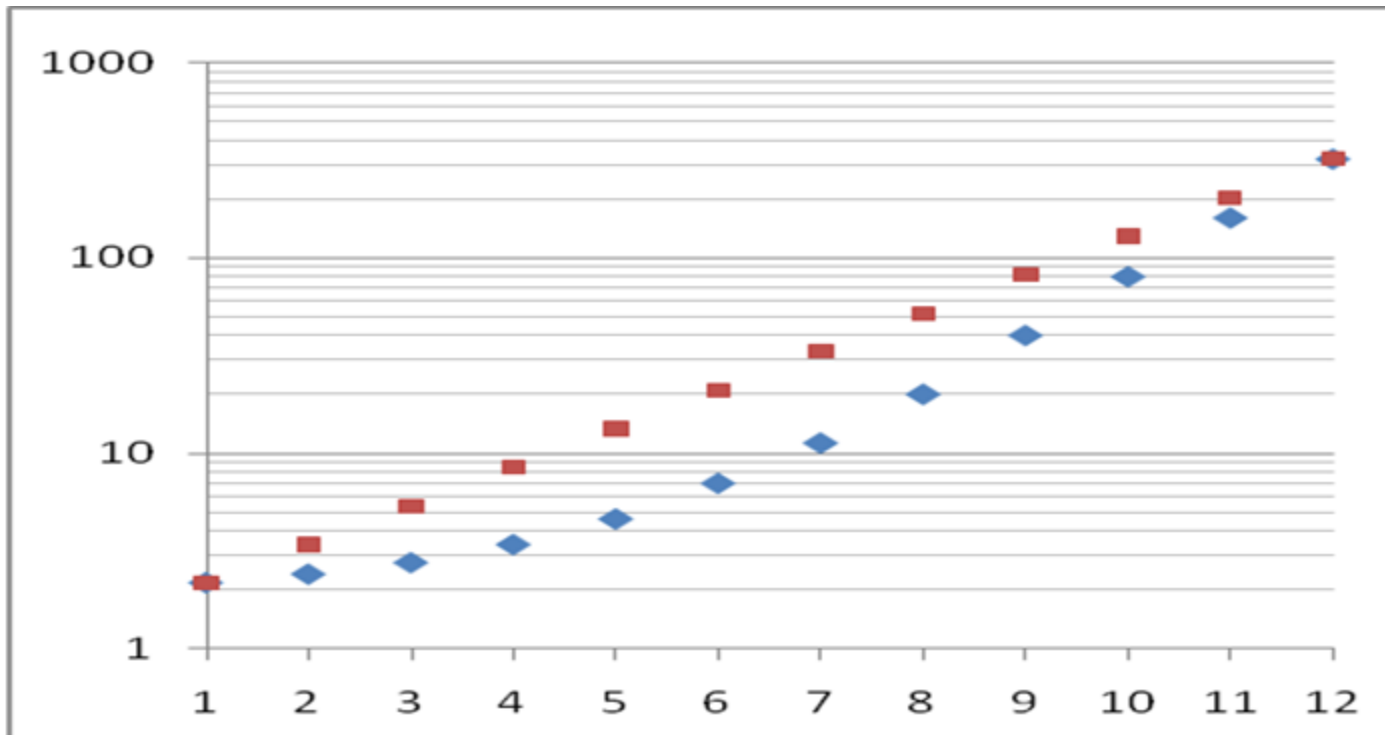
Table 2: Encoding WITH code stream markers followed by decoding Decode Times (mm:ss) for the same test file						
Use Case						
Progression Order	A - thumbnail		B - reading		C - detailed	
	Tiles	Precincts	Tiles	Precincts	Tiles	Precincts
RLCP	00:15	00:06	01:39	01:11	01:23	04:12
RPCL	00:15	00:07	01:42	01:10	01:24	01:41
LRCP	00:16	00:06	01:47	01:23	01:25	04:08
PCRL	02:37	00:08	03:39	01:11	01:25	01:41
CPRL	02:25	00:07	03:39	01:10	01:24	01:41

Table 3: Extended Decode Times for selected files from Test File Set 2	
Test File	Use Case a
RPCL with precincts	1:03
PCRL with precincts	1:14
CPRL with precincts	1:15

These also show that precincts and RPCL are best for the anticipated use

# Specifying quality layers

- Two choices: (a) adopt linear logarithmic spacing or (b) choose your own
- The selection of quality layers was chosen to give better coverage in the areas of greatest expected use





# Recommended JPEG 2000 profile

Parameter/Field	Value
Compression	Lossy
Number of components	3
Component Transform	Yes (irreversible)
Tile size	One tile for entire image
Wavelet Filter	9-7 irreversible
Number of levels	Variable; 6 used for test image
Number of layers	Multiple
Progression order	RPCL
Codestream markers	Packet-length markers
Precincts	256x256, 256x256, 128x128
Codeblock size	64x64
Coder Bypass	Yes

Example Kakadu command line (for minimally lossless):

```
kdu_compress -i test.tif -o test.jp2  
-rate -,10,8.7,7,5.2,3.4,2.1,1.2,0.6,0.3,0.15,0.075  
Creversible=no Clevels=6 Cmodes=BYPASS Corder=RPCL  
Cblk={64,64} Cprecincts={256,256},{256,256},{128,128}  
ORGgen_plt=yes
```

# Concluding comments

- The profile does not specify a specific degree to which compression is applied
- It supports a range of “degrees” of compression progressively relaxed from the “minimally lossless”
- This means that it can be adapted for other types of content, and where ....
  
- A final choice on the degree of compression can be based on:
  - Quantitative measures: *such as PSNR*
  - Qualitative measures: *can you tell the difference?*
  - Comparison: with the *variability and noise inherent in the imaging process*
  - Affordability

## Annex – conjecture about djabatoka

- The decode tests used different quality levels for the three use cases:
  - `kdu_expand -i test.jp2 -reduce 5 -layers 1`
  - `kdu_expand -i test.jp2 -reduce 3 -layers 4`
  - `kdu_expand -i test.jp2 -layers 10  
-region {0.5571,0.5569},{0.0707,0.1661}`
- Extracting higher quality layers for lower resolution images is not visually discernable but slows down the access
- However, its possible that a typical browser client and djabatoka extracts all quality layers even for thumbnails?