Experimental Study on Lossless Compression of Biometric Iris Data

Kurt Horvath, Herbert Stögner, Andreas Uhl, Georg Weinhandel

Presenter: Rade Kutil

Department of Computer Sciences
University of Salzburg, Austria

uhl@cosy.sbg.ac.at

http://www.wavelab.at/

Outline

- Introduction & Motivation
- Compression in Biometric Systems
- Aim of this Work
- Experimental Setup
- Results
- Conclusion

Introduction

In biometric systems, the compression of acquired sample data may become imperative under certain circumstances, due to the amount of data involved and potentially weak network links between sensor and feature extraction / matching module.

Lossy compression techniques maximize the benefit in terms of data reduction. However, the distortions introduced by compression artifacts may interfere with subsequent feature extraction and may degrade the matching results.

Lossless compression avoids any impact on recognition performance but is generally known to deliver much lower compression rates. An additional advantage of lossless compression algorithms is that these are often less demanding in terms of required computations as compared to lossy compression technology (which is beneficial for the sketched target-scenario often involving weak or low-power sensing devices).

Compression in Biometric Systems: Standards

- ISO/IEC 19794 standard on “Biometric Data Interchange Formats”: current version supports JPEG and JPEG2000 (and WSQ for fingerprints) for lossy compression and JPEG-LS for lossless compression. The most recent (draft) version (ISO/IEC FDIS 19794-6 as of August 2010) supports only JPEG2000 for lossy compression and PNG for lossless compression. The latter draft is mostly based on the NIST Iris Exchange (IREX) program recommendations.

- ANSI/NIST-ITL 1-2011 standard on “Data Format for the Interchange of Fingerprint, Facial & Other Biometric Information”: for lossy compression JPEG2000 is supported, and JPEG2000 as well as PNG for the lossless case.

Aim of this Work

Focus: Lossless compression of rectilinear iris sample imagery (corresponding to IREX KIND1 or KIND3 records).

Methods: Application of various lossless compression algorithms to iris images available from public iris biometric databases (experimental study on achieved compression ratio).

Questions:

1. Is PNG a sound solution with respect to achieved compression ratio (as compared to JPEG-LS and JPEG2000) ?

2. Do we find identical ranking among compression algorithms for different datasets ?

3. Are general purpose file compression algorithms competitive ?

Compression Algorithms

- Dedicated lossless image compression algorithms: Lossless JPEG, JPEG-LS, GIF, and PNG

- Lossy image compression algorithms in lossless mode: JPEG2000, SPIHT, and JPEG XR

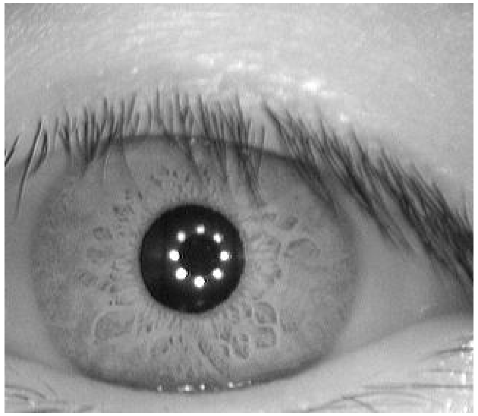
- General purpose file compression algorithms: 7z, BZip2, Gzip, ZIP, and UHA

Sample Data

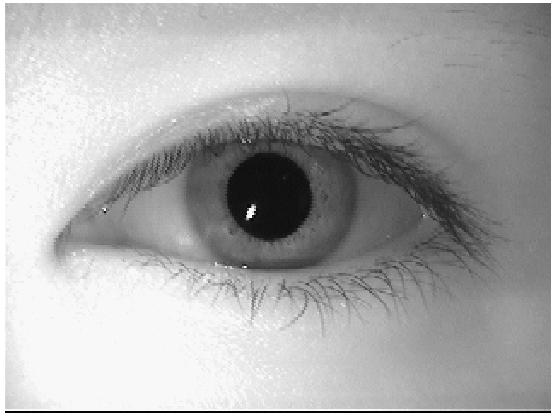
For all our experiments we used the images in 8-bit grayscale information per pixel in .bmp format since all applied software can handle these formats (except for SPIHT which requires a RAW format with removed .pgm headers). Color images have been converted to the YUV format using the Y channel as grayscale image.

- CASIA V1: 756 images, resolution 320×280
- CASIA V3 Interval: 2639 images,resolution 320×280
- MMU1: 457 images,resolution 320×240
- MMU2: 996 images, resolution 320×238
- UBIRIS: 1876 images, resolution 200×150
- BATH: 1000 images, resolution 1280×960
- ND Iris: 1801 images, resolution 640×480

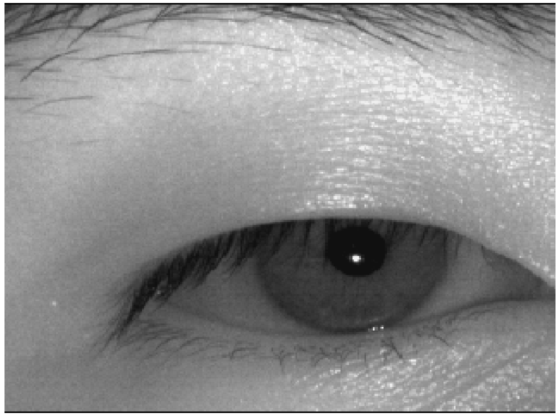
Image Examples



(a) CasiaV3

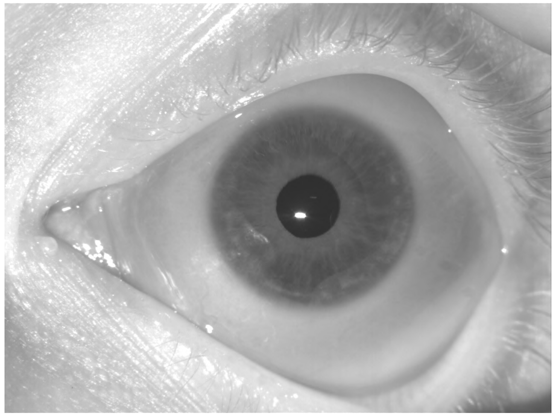


(b) MMU1

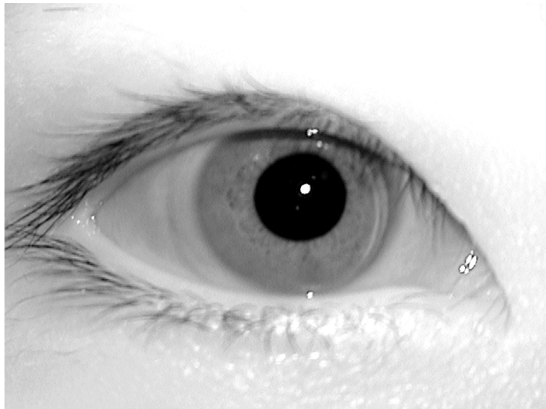


(c) MMU2

(d) UBIRIS



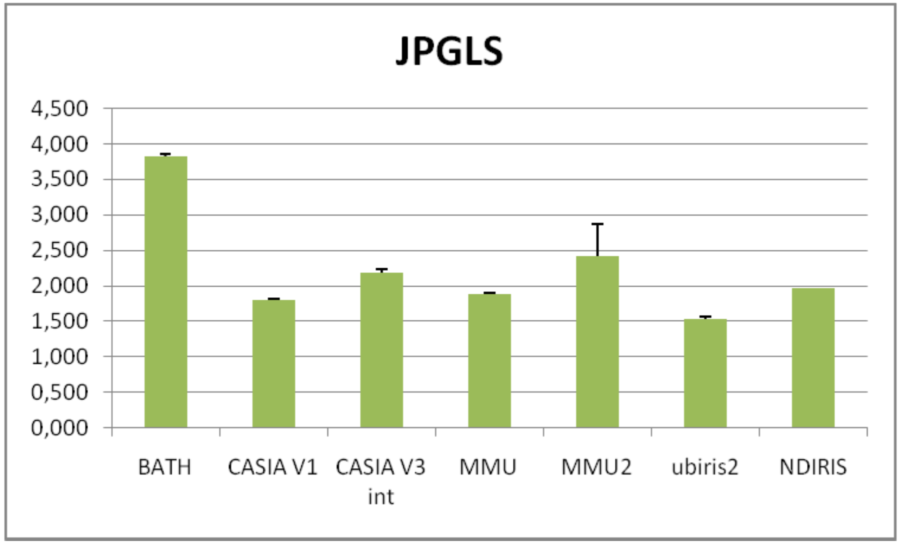
(e) BATH



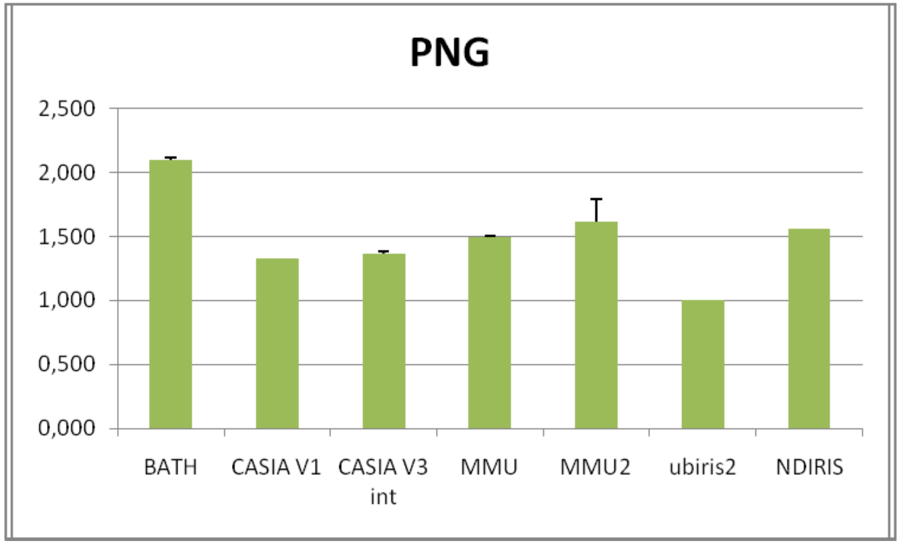
(f) ND Iris

Figure 1: Example iris images from the databases.

Results: JPEG-LS vs. PNG



(a) JPEG-LS

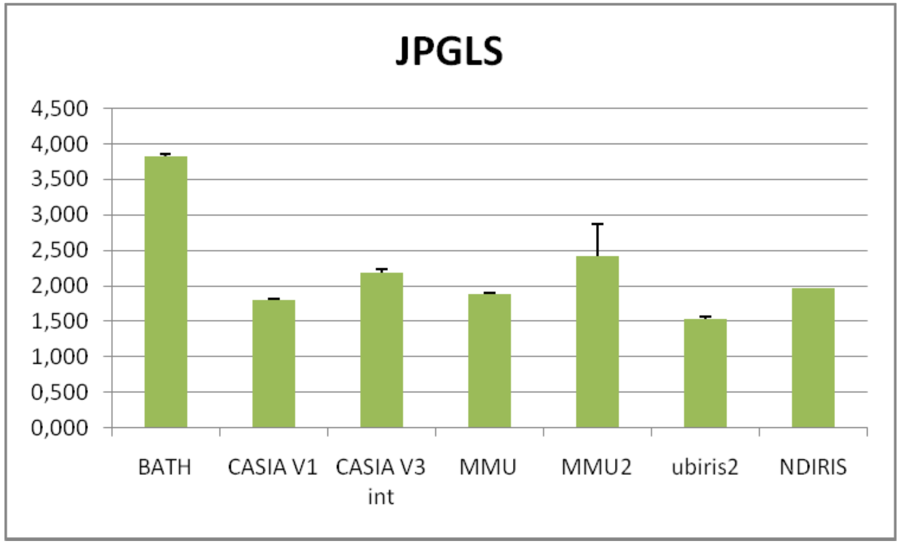


(b) PNG

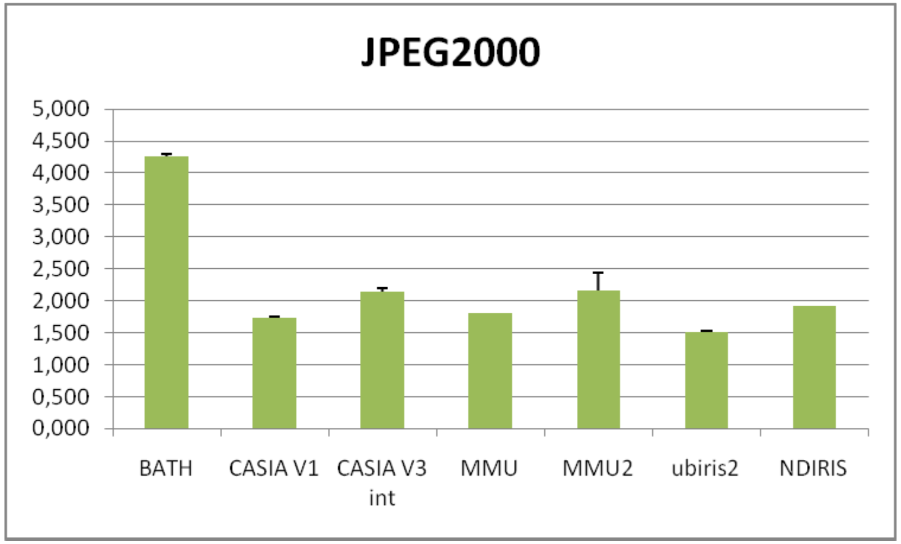
⟶ Dataset Dependence: Highest resolution dataset (BATH) gives highest compression rate, lowest resolution dataset (UBIRIS) gives lowest compression rates.

⟶ Relative Rates: PNG is clearly inferior to JPEG-LS for all datasets. Decision for standardisation is neither based on compression performance nor on computational demand (JPEG-LS is very fast).

Results: JPEG-LS vs. JPEG2000



(c) JPEG-LS

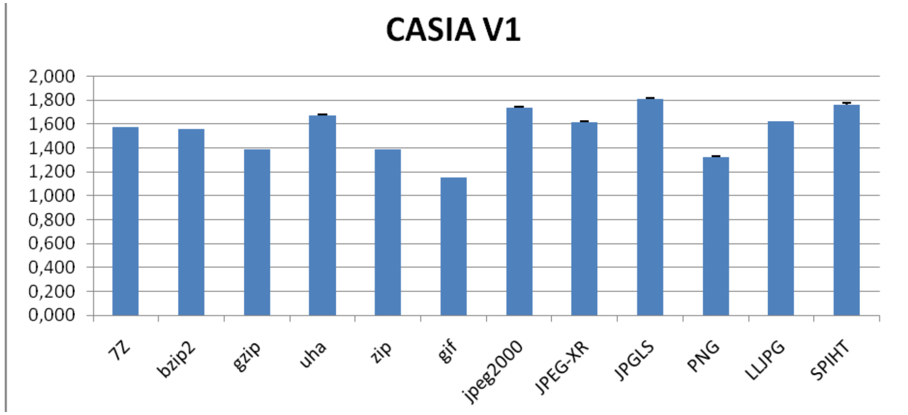


(d) JPEG2000

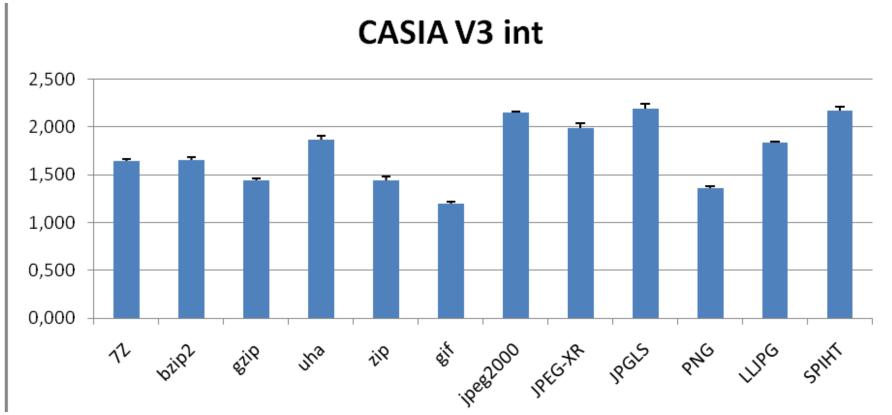
⟶ Dataset Dependence: Is almost identical for JPEG-LS and JPEG2000.

⟶ Relative Rates: JPEG-LS and JPEG2000 exhibit almost identical performance for different datasets, JPEG-LS sligthly better except for BATH images.

Results: CASIA Datasets



(e) CasiaV1



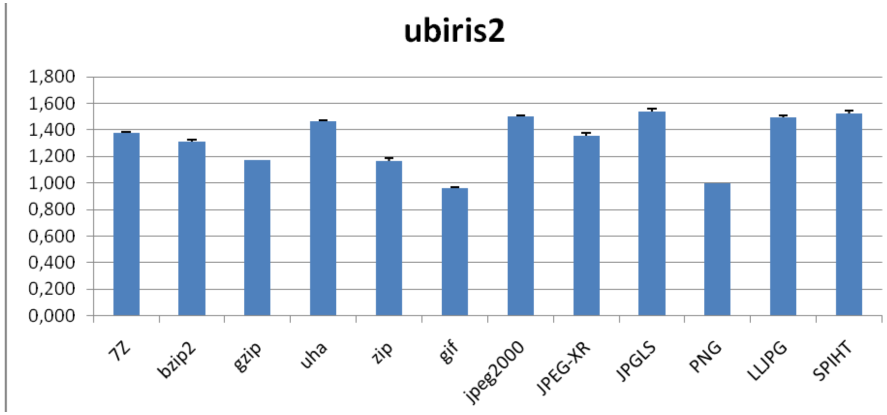
(f) CasiaV3

⟶ JPEG-LS is best closely followed by JPEG2000 and SPIHT.

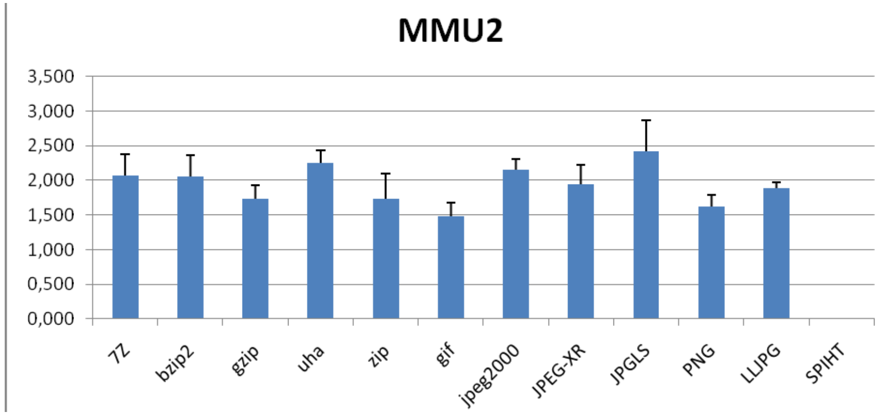
⟶ File compression algorithms are close to JPEG XR and lossless JPEG.

⟶ PNG and GIF are clearly the worst algorithms considered.

Results: UBIRIS and MMU2



(g) UBIRIS



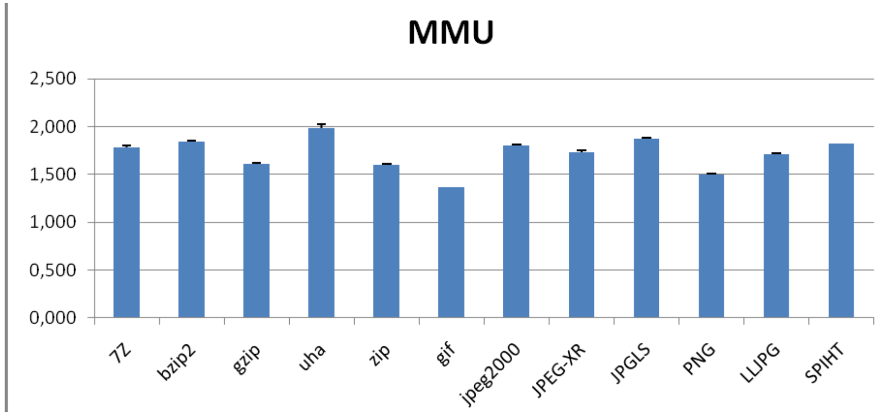
(h) MMU2

⟶ JPEG-LS is best closely followed by JPEG2000 and SPIHT, for UBIRIS2, lossless JPEG does a very good job.

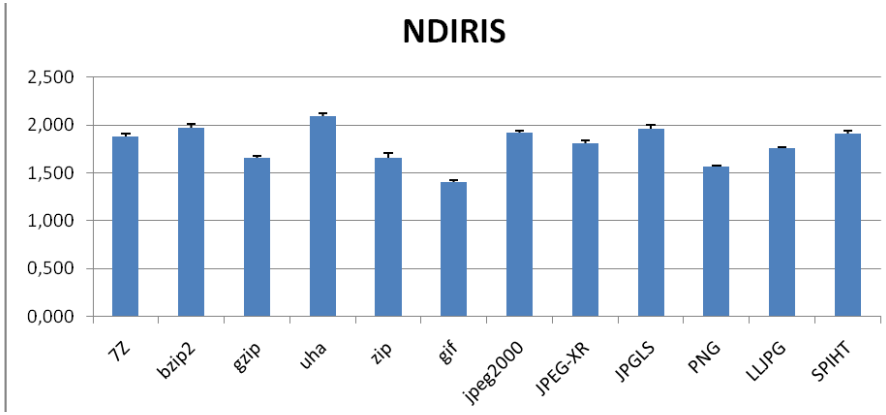
⟶ UHA (file compression algorithm !) is second best for MMU2 !

⟶ Again, PNG and GIF are clearly the worst algorithms considered.

Results: MMU and NDIris



(i) MMU

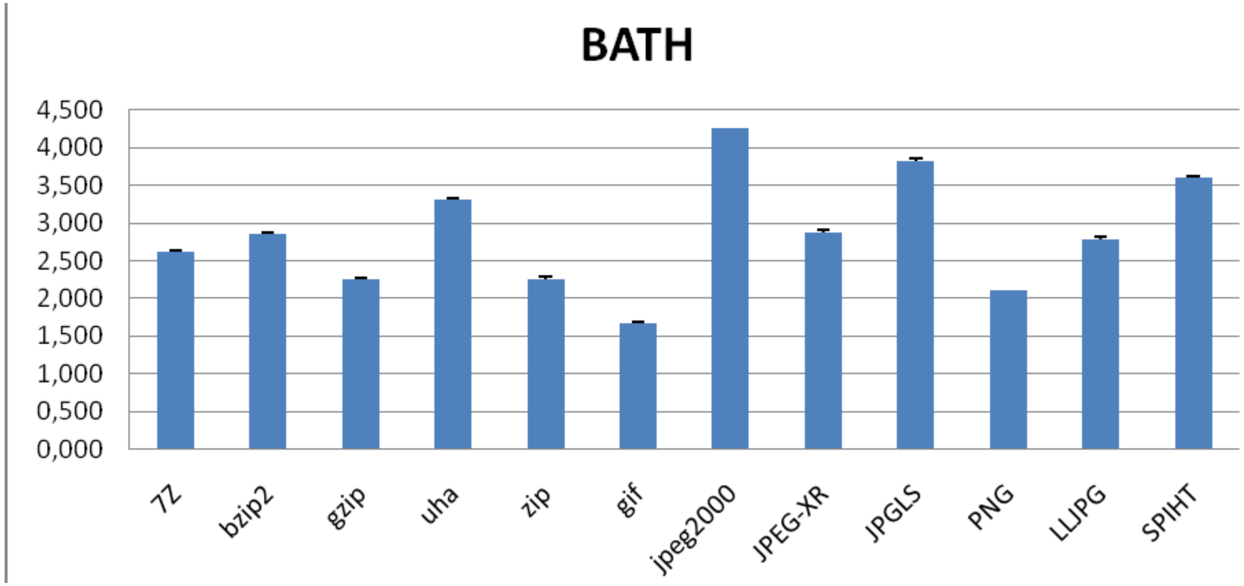


(j) NDIris

$\longrightarrow$ UHA is the best algorithm, closely followed by JPEG-LS and BZip2. JPEG2000 and SPIHT are close.

$\longrightarrow$ Again, PNG and GIF are clearly the worst algorithms considered.

Results: BATH Dataset



(k) BATH

⟶ JPEG2000 is clearly best, followed by JPEG-LS and SPIHT (note that JPEG2000 is the original file format of the BATH set, before conversion to .bmp !).

⟶ Again, PNG and GIF are clearly the worst algorithms considered.

Results: Overall

	Best	Ratio	Worst	Ratio
CASIA V1	JPEG-LS	1.81	GIF	1.15
CASIA V3 Int.	JPEG-LS	2.19	GIF	1.20
MMU	UHA	1.99	GIF	1.36
MMU2	JPEG-LS	2.42	GIF	1.47
UBIRIS	JPEG-LS	1.54	GIF	0.96
BATH	JPEG2000	4.25	GIF	1.66
ND Iris	UHA	2.09	GIF	1.40

Conclusion

- Results depend on the database considered, in most cases the best techniques are JPEG-LS or UHA.

- PNG is a poorly performing scheme for this kind of data and its standardisation should be re-considered. JPEG-LS or JPEG2000, both being international ITU and ISO standards, are much better suited for the datasets considered.

- General purpose file compression algorithms do a tremendeous job for all datasets, being even top performing for two of them.

- The ranking of the compression schemes tends to be very stable across all databases, at least considering the top and least performing groups of techniques.

- As expected, higher resolution leads to higher absolute compression ratios.

Thank you for your attention !

Questions ?