DAPCOM
Data services

Jordi Portell (CTO – Chief Technology Officer)
jordi.portell@dapcom.es

www.dapcom.es

What is DAPCOM?

DAPCOM Data Services S.L. is a **technological spin-off company**

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH

UNIVERSITAT DE BARCELONA

Founded in February 2013

Strong background on **Space research**

8 co-founders, all of them from **Gaia**@Barcelona

Main field: **Software and Data Science**

Space data handling

Massive data processing

Efficient data compression

Main customers (so far)

IEEC (Gaia/DPAC):

IDU (Intermediate Data Updating)

Massive re-processing of accumulated raw data (>1 trillion images, >100TB raw data)

ESA (Gaia/DPAC):

XM (Cross-Matching)

Determination of **source catalogue** (>1.5 billion) from >100 billion on-board detections

High-quality and specialized software engineering

Main customers (so far)

Mind the Byte (SaaS for drug discovery)

Consultancy on efficient, massive data transfers for genomic data processing on the Cloud

Meteowhizz (for IAC)

Tailored software platform (backend + web frontend) for improved meteorological predictions in astronomical observatories

Spire Global Inc.

Adaptation of our data compression technology to their **nanosatellites constellation** (radio occultation data)

Data compression for Space systems

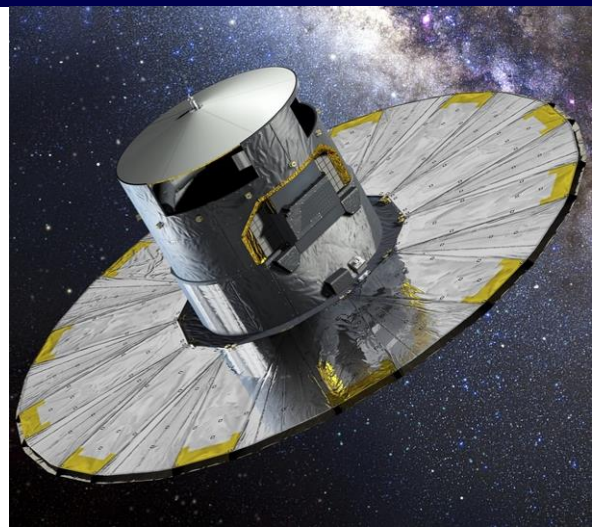
Consultative Committee for Space Data Systems (CCSDS):

Several recommendations for onboard data compression

- 121.0: Lossless data compression
  One may define and implement any data predictor
- 122.0: Lossless + lossy greyscale image data compression
  Discrete Wavelet Transform (~JPEG2000)
- 122.1: Lossless + lossy multi-band image data compression
  Extension of 122.0 for colour images
- 123.0: Lossless hyperspectral data compression
  Aimed at low complexity
- 123.1 (in prep.): Near-lossless hyperspectral compression

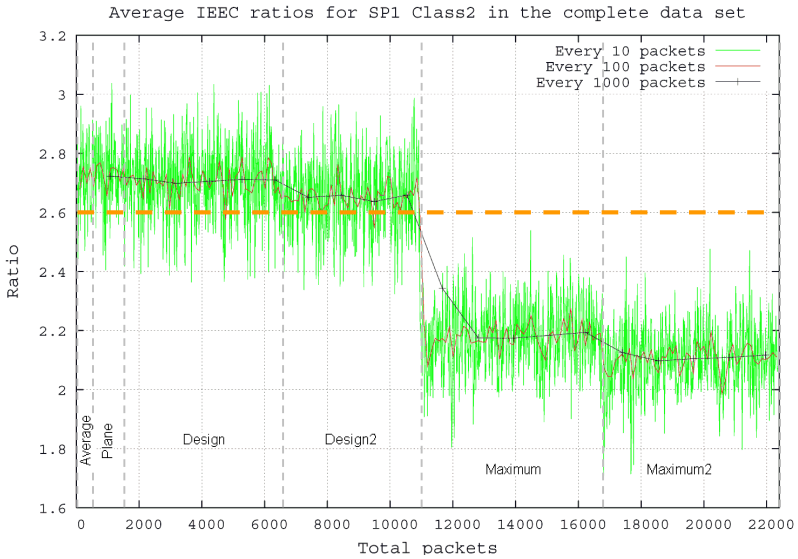
See www.ccsds.org for more details

Data compression for Gaia



- Gaia: ESA cornerstone mission
  - 3D catalogue of >1 billion (10^9) stars
  - Downlink: 60GB/day (from 1.5M km)
  - High-performance on-board data compression (tailored to Gaia, software on PowerPC, C language)

- GOCA: Gaia Optimum Compression Algorithm
  - ESA TRP Study: IEEC + GTD
  - Initial (contractor) baseline: Rice entropy coder. Many problems with *outliers* (prompt particle events)
  - IEEC solution: PEC (Prediction Error Coder)



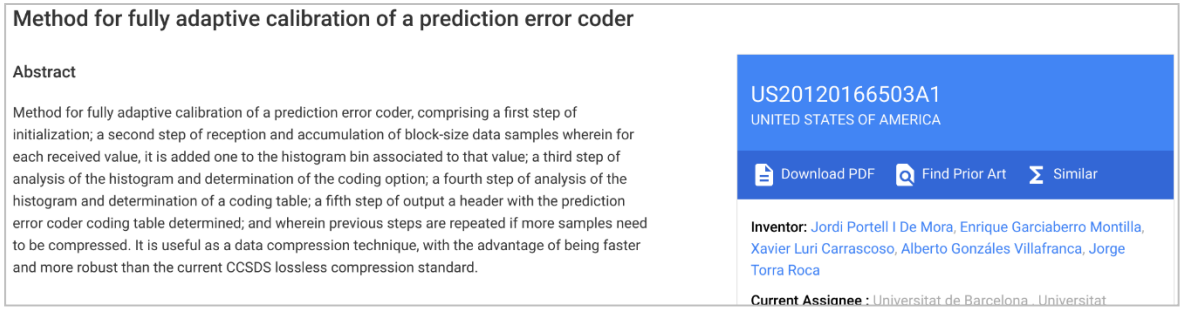
FAPEC: Our own data compression technology

ESA TRP study: Gaia Optimum Compression Algorithm (GOCA)

→ PEC (Prediction Error Coder), later Fully Adaptive PEC (**FAPEC**)

US patent 9,002,913 B2

License transferred to DAPCOM



Method for fully adaptive calibration of a prediction error coder

Abstract

Method for fully adaptive calibration of a prediction error coder, comprising a first step of initialization; a second step of reception and accumulation of block-size data samples wherein for each received value, it is added one to the histogram bin associated to that value; a third step of analysis of the histogram and determination of the coding option; a fourth step of analysis of the histogram and determination of a coding table; a fifth step of output a header with the prediction error coder coding table determined; and wherein previous steps are repeated if more samples need to be compressed. It is useful as a data compression technique, with the advantage of being faster and more robust than the current CCSDS lossless compression standard.

US20120166503A1
UNITED STATES OF AMERICA

Download PDF Find Prior Art Similar

Inventor: Jordi Portell I De Mora, Enrique Garciaberro Montilla, Xavier Luri Carrascoso, Alberto Gonzáles Villafranca, Jorge Torra Roca

Current Assignee : Universitat de Barcelona , Universitat

Main FAPEC benefits:

Better compression resiliency on **statistical outliers**

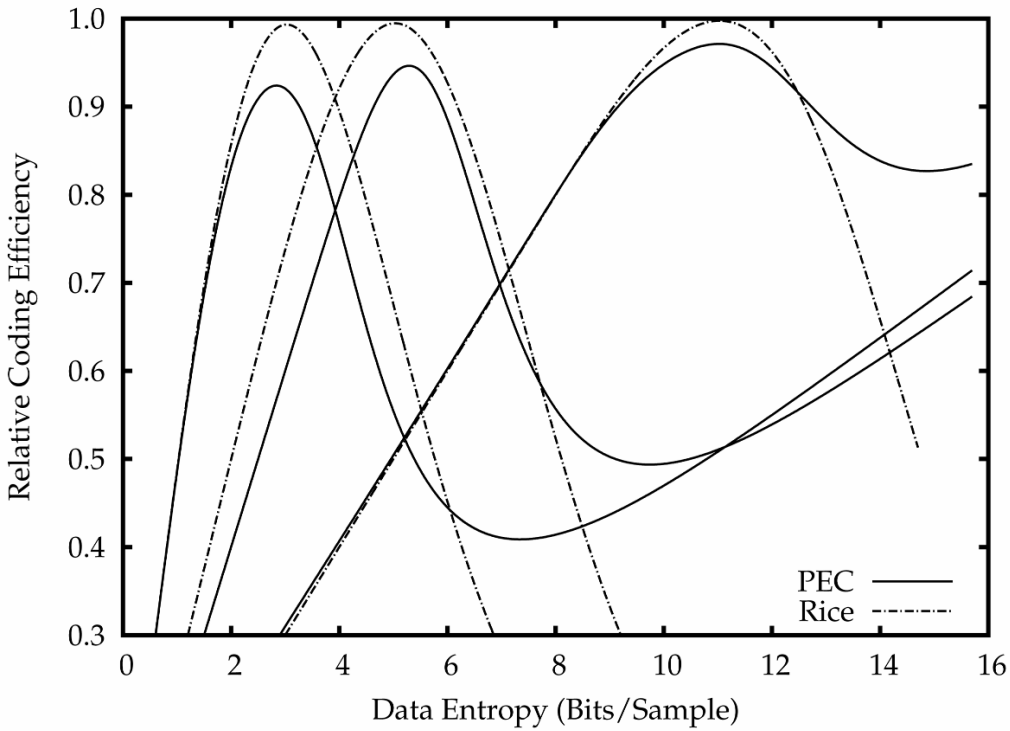
Fast and **self-calibrated** operation

ESA Business Incubation Centre (BIC) in Barcelona

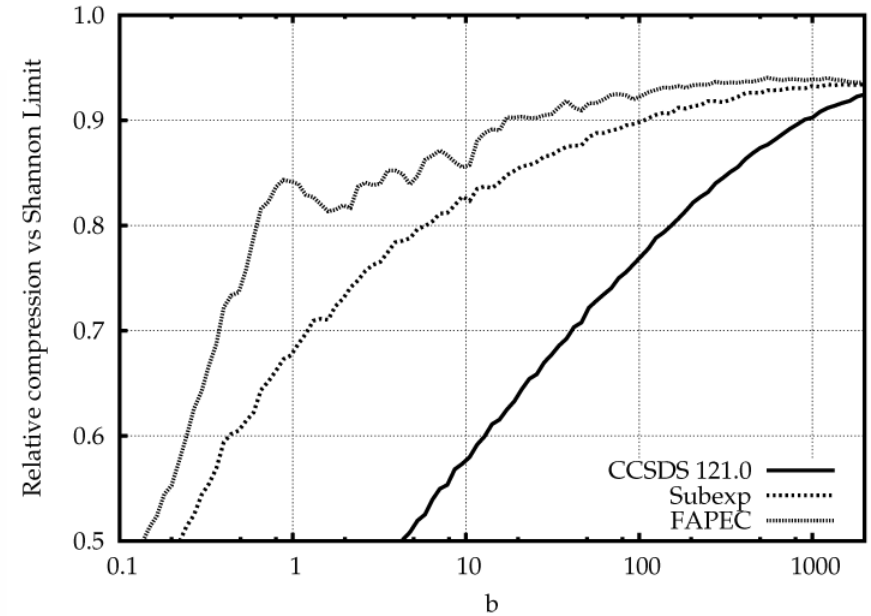
Transfer this technology to the non-Space market

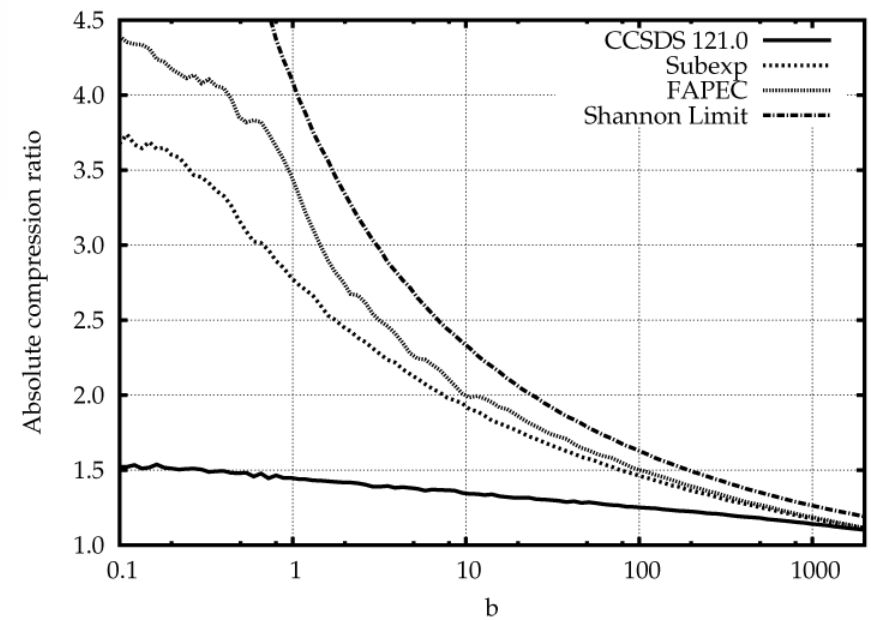
esa business incubation centre Barcelona

Entropy coding performance: Rice vs. PEC/FAPEC



$$H = \sum_i P(i) \log_2 \frac{1}{P(i)}$$

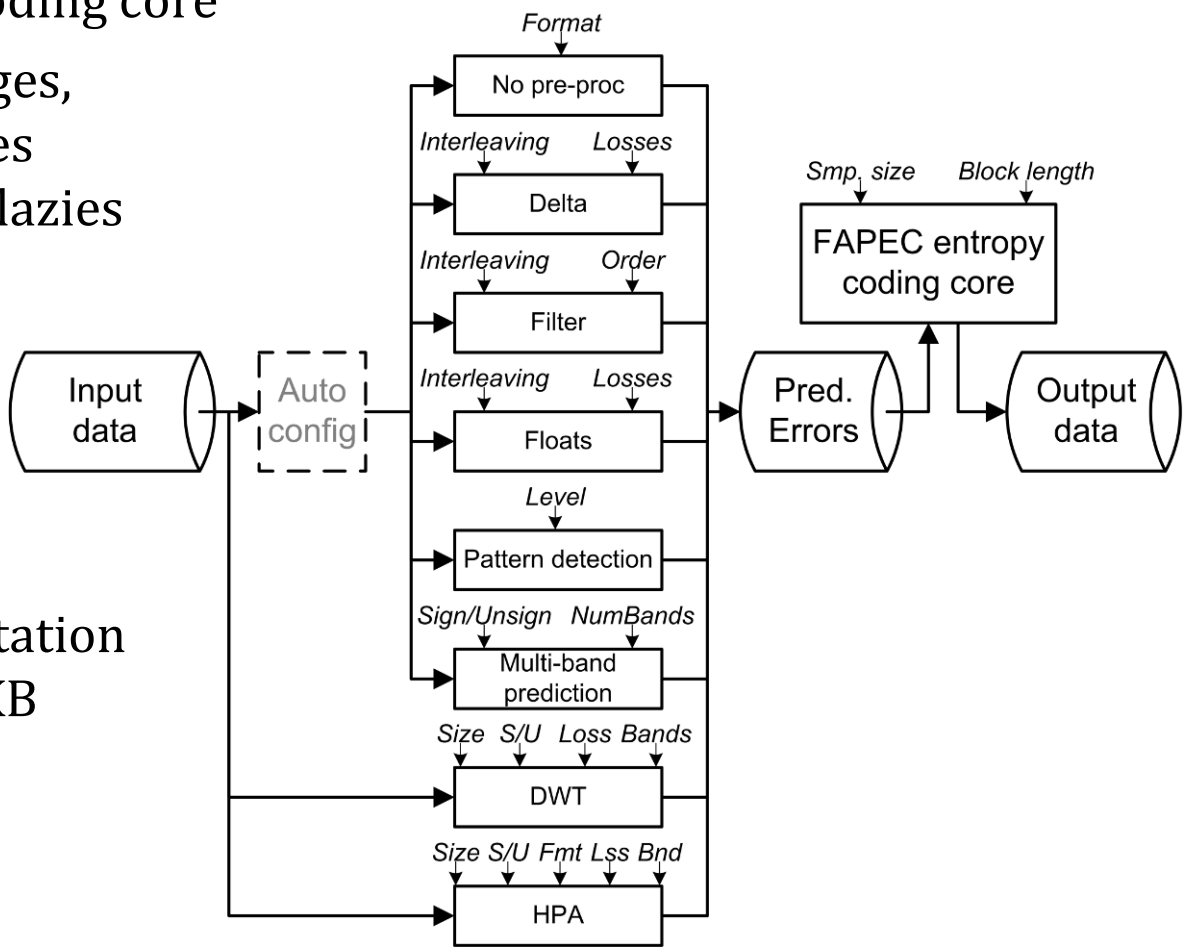




The FAPEC data compressor

Versatile data compression solution (onboard + onground applications)

- Based on FAPEC entropy coding core
- Suite of pre-processing stages, including multi-band images and auto-configuration for lazies
- 8 to 24 bits per sample (and beyond: interleaving)
- Lossless and lossy
- Typically fast operation
- ANSI C software implementation Code footprint can be <50KB
- Error-resilient file format
- Encryption, multi-thread
- Easy to add new tailored stages

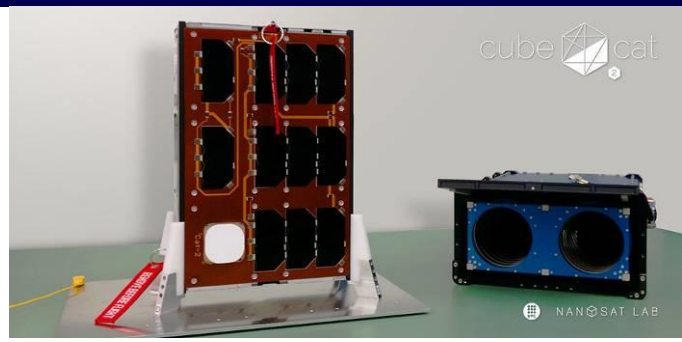


Implementation complexity

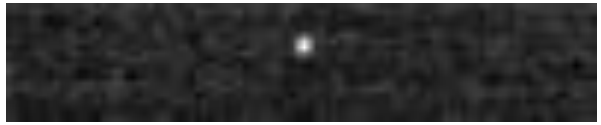
- Tests with *Valgrind* tool
- FAPEC compression is typically in the 50-150 IPS range
- Decompression as well (note dictionary decompression IPS!)
- Compressible data leads to fewer instructions

Pre-processing stage	Data used	SS	Ratio	bps	IPS
None (signed)	Solar image	8	1.03	7.76	68
None (unsigned)	Solar image	8	1.05	7.62	72
Delta	Planetary image	8	1.07	7.48	65
Delta (auto-configured, 327-kB file)	Planetary image	8	1.07	7.48	92
Delta	Telecommand file	8	1.15	6.96	61
Delta	Meteorological image	8	1.44	5.55	55
Delta	Stellar image	16	2.98	5.37	54
Delta *(decompression)*	Stellar image	16	2.98	5.37	96
Delta interleaving 4	Stellar image	16	2.76	5.80	69
Delta interleaving 224	Hyperspectral image	16	2.01	7.96	74
Delta lossy 4 LSB	Stellar image	16	6.94	2.31	58
Delta lossy + interleaving	Stellar image	16	6.19	2.58	58
3rd order linear filter	Stellar image	16	2.92	5.48	80
Filter + interleaving	Stellar image	16	2.59	6.18	87
Floats	Radioastronomy data	32	1.15	27.8	160
Floats + interleaving	Radioastronomy data	32	1.15	27.8	168
Dictionary level 0	Telecommand file	8	2.30	3.48	59
Dictionary level 2	Telecommand file	8	2.62	3.05	104
Dictionary level 2 *(decompression)*	Telecommand file	8	2.62	3.05	32
3 bands prediction	Meteorological image	8	1.97	4.06	71
3 bands prediction (auto-conf, 5.1MB file)	Meteorological image	8	1.97	4.06	73
224 bands prediction	Hyperspectral image	16	2.34	6.84	98
DWT 1 band	Solar image	8	1.30	6.15	131
HPA 1 band	Solar image	8	1.26	6.35	118
DWT 3 bands	Meteorological image	8	1.82	4.40	110
HPA 3 bands	Meteorological image	8	2.48	3.23	99
DWT 224 bands	Hyperspectral image	16	2.41	6.64	157
HPA 224 bands	Hyperspectral image	16	2.47	6.48	115
DWT 1 band lossy	Solar image	8	3.79	2.11	134
HPA 1 band lossy	Solar image	8	3.50	2.29	144
DWT 3 bands lossy	Meteorological image	8	9.50	0.84	113
HPA 3 bands lossy	Meteorological image	8	6.06	1.32	131
DWT 224 bands lossy	Hyperspectral image	16	31.0	0.52	152
HPA 224 bands lossy	Hyperspectral image	16	20.8	0.77	134

(Almost) operations: ^{3}Cat-2



- First satellite using FAPEC
  - Launched 15-Aug-2016
- GNSS-R (reflectometry)
  - Tiny (100x20x16bpp) and noisy images



- Original baseline for on-board compression: PNG (Zip-like)
- FAPEC was accepted as the new baseline
  - Better ratios, faster operation
  - Lossy option chosen (noisy data)
  - Also chosen for experimental star tracker images!

Simulated data:	**Worst ratio**	**Best ratio**
PNG (original baseline)	0.94	0.99
Lossless delta FAPEC	1.08	1.20
4 LSB lossy delta FAPEC	1.48	1.72

Acknowledgements:
R. Olivé and A. Camps
UPC NanoSat Lab
www.tsc.upc.edu/nanosatlab

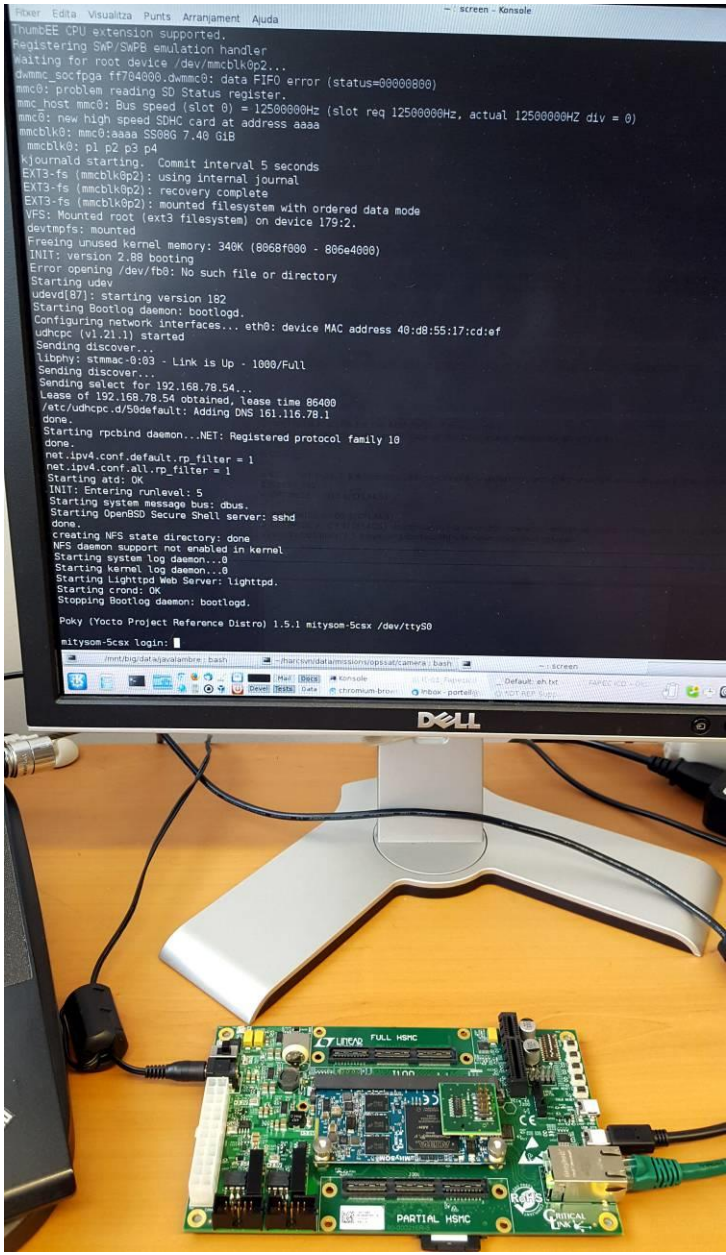
NANOSAT LAB

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH
Departament de Teoria del Senyal i Comunicacions
RSLab

Tests: OPS-SAT

- ESA technology demonstrator
- Powerful payload:
  - Dual-core ARM9 800MHz
  - 1 GB RAM + 8 GB SD
  - 2048x1944px camera (Bayer)
  - Fine ADCS, GPS, SDR...
- Several experiments
  - FAPEC included

	Lab 1 image		Lab 2 image	
	Ratio	**Speed**	**Ratio**	**Speed**
Lossless delta	2.05	15.4 MB/s	2.74	17.6 MB/s
4 LSB lossy delta	4.24	18.6 MB/s	7.6	21.2 MB/s
Lossless HPA	2.01	9.3 MB/s	2.70	9.9 MB/s
Level 4 lossy HPA	2.73	7.8 MB/s	4.18	8.0 MB/s
Level 7 lossy HPA	5.36	8.0 MB/s	12.2	9.0 MB/s

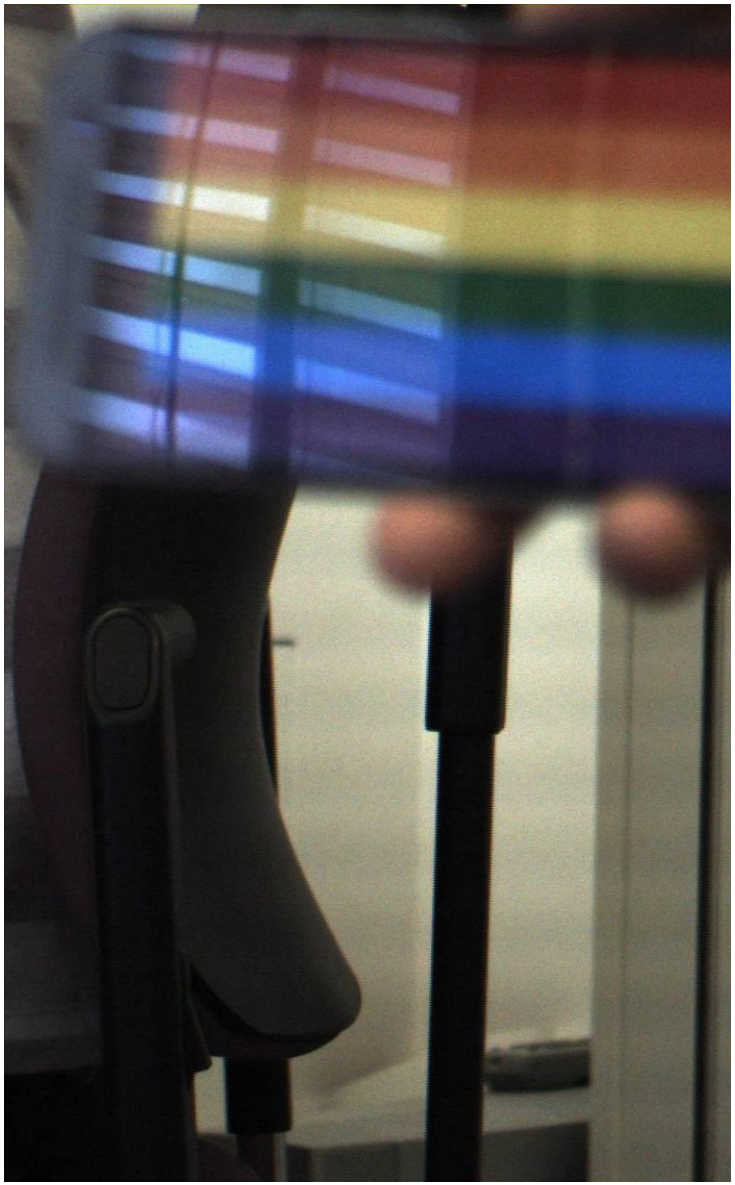


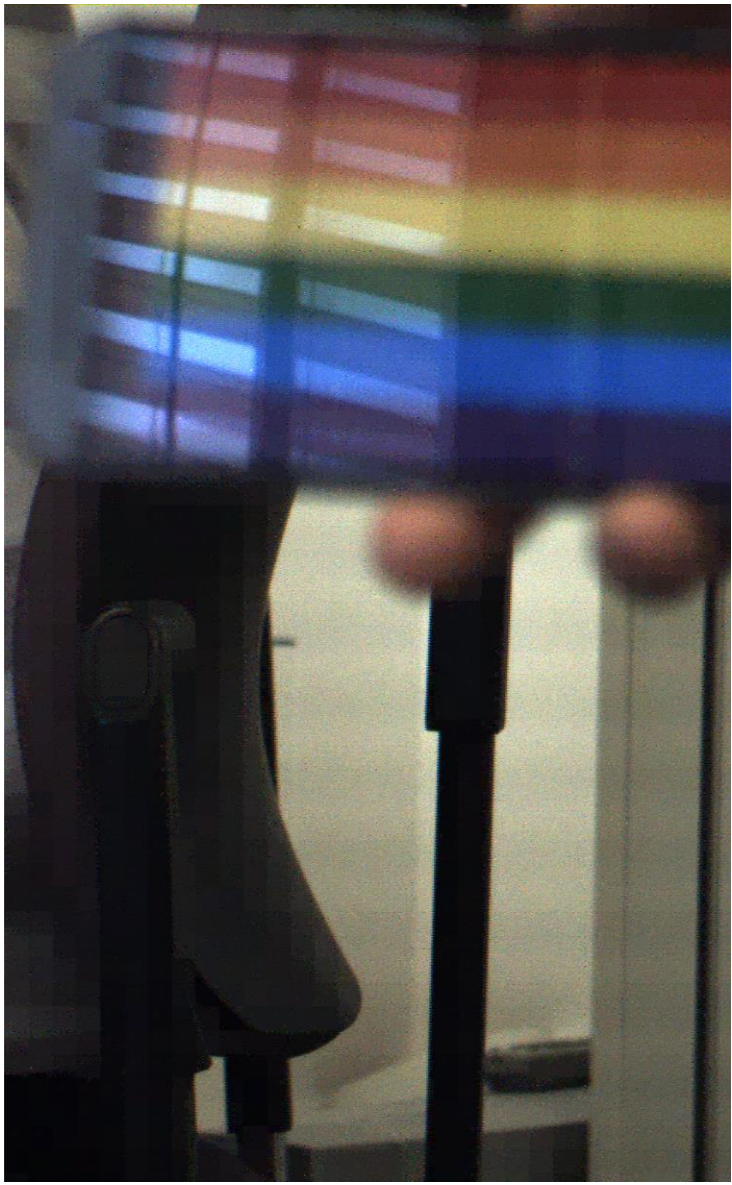
Tests: OPS-SAT

Lossy tests

Left:
original

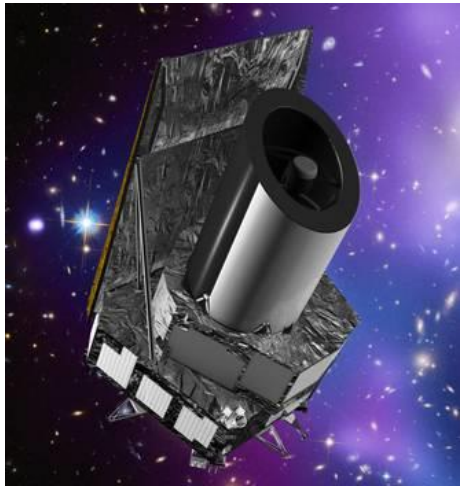
Right:
HPA ratio 12





Tests: Euclid

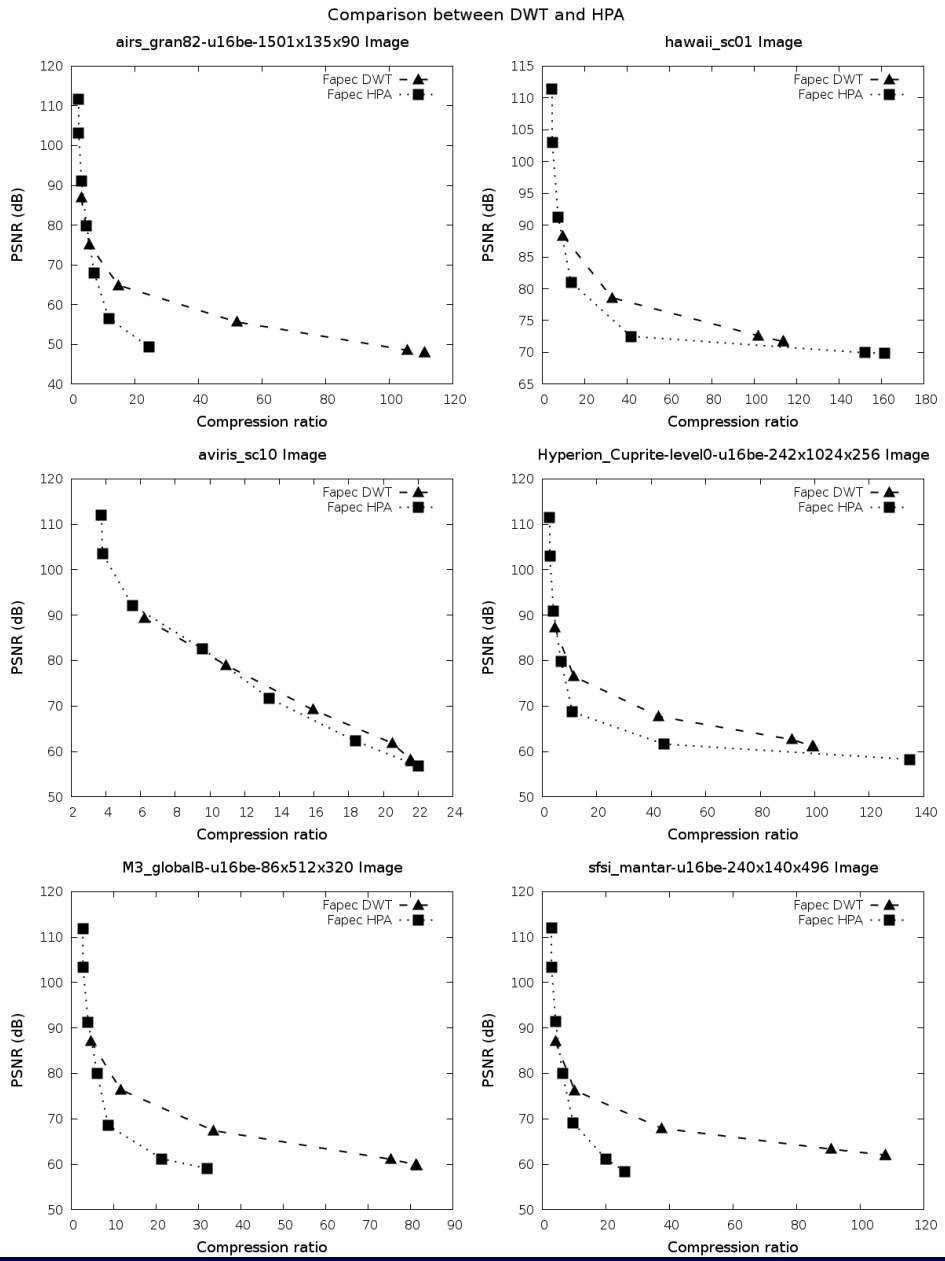
- ESA mission
- FAPEC tested on VIS + NISP simulated data
- Also integrated in the VIS engineering model board
  - Better ratios + slightly better times than Rice baseline



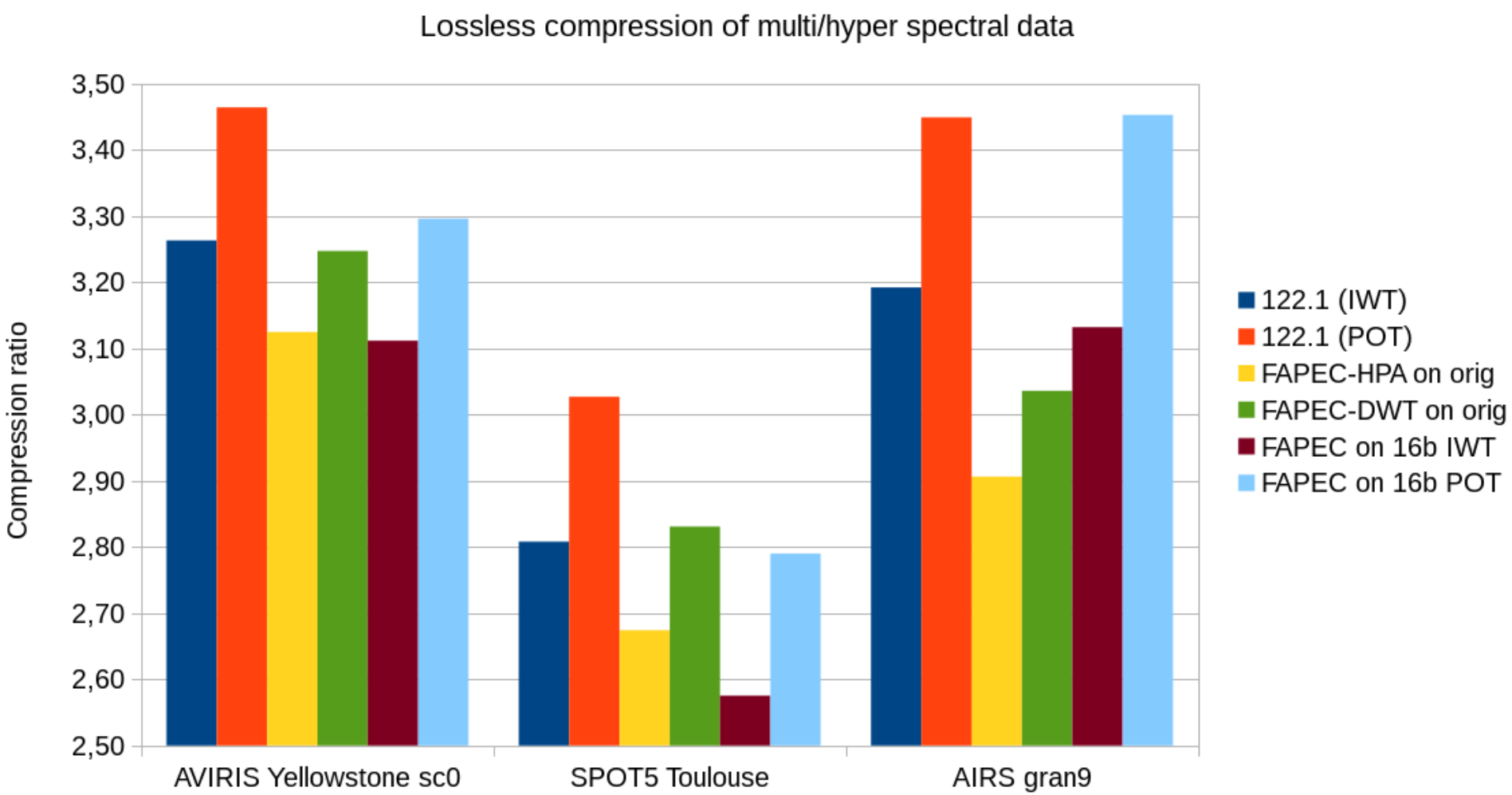
	Rice	FAPEC Delta	FAPEC 2nd order filter
Charge inj.	4.60	4.73	4.92
Charge rel.	2.72	3.07	3.10
Flat field	1.34	1.32	1.35
Star field 1	**2.55**	**2.80**	**2.84**
Star field 2	**2.75**	**3.17**	**3.15**

Tests: Lossy hyperspectral (I)

- FAPEC tests to compare its DWT and HPA stages
- DWT outperforms HPA for high ratios
- HPA provides smoother lossless-to-lossy transition
- HPA supports more coding formats
- HPA uses inter-band correlation also for few bands
- HPA typically faster



Tests: Lossless hyperspectral (II)



Tests: “onboard” compression speed

C3SatP (IEEC transversal initiative):

- High-performance **on-board computer**, **data handling** and SDR platform
- Aimed at **cubesats**
- On-board data handling: Xilinx Zynq UltraSCALE+ (**4 ARM cores, up to 1.5GHz**)
- Max power consumption: ~8W (TBC)

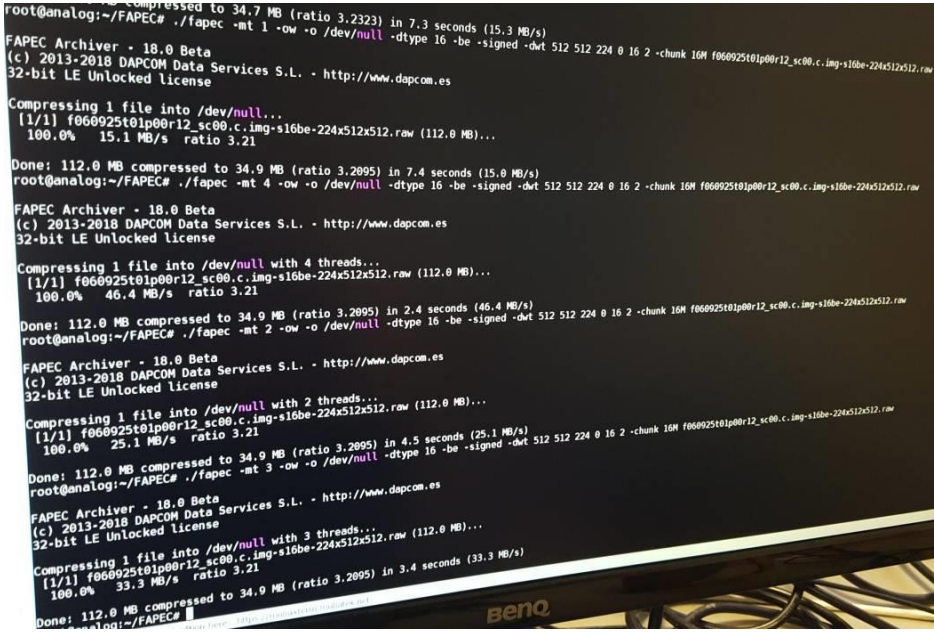
Multispectral Earth Observation is an excellent use case for this platform
→ test hyperspectral data compression here

C3SatP software development platform

Tests: “onboard” compression speed

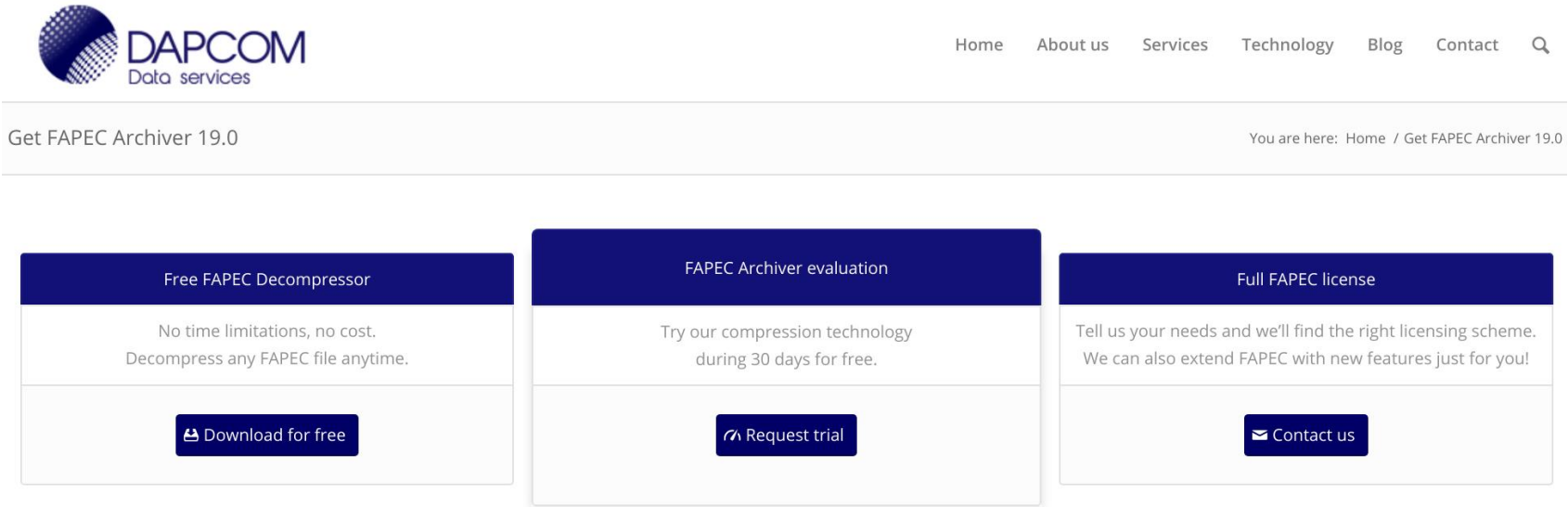
FAPEC-DWT compression test on **AVIRIS** file (**C software** as-is):

- Single-chunk (single-thread): **15.7 MB/s** (**7.8 Msamples/s**)
  → full AVIRIS 512x512x224 scene in ~7s
- Additional tests with ~16MB chunks (ratio reduced from 3.25 to 3.21):
  - 2 threads: 25 MB/s (x1.6)
  - 3 threads: 33 MB/s (x2.1)
  - 4 threads: 46 MB/s (x2.9)
    → up to 23 Msamples/s
- Reminder: ~90 MB/s (x5.7) on a laptop (Intel Core i7 2640M 2.8GHz), single-thread



- CCSDS **122.1** <u>hardware</u> reference (RTAX2000S, **VHDL**): 18.4 Msamples/s
- Another <u>hardware</u> reference (**FAPEC** on ProASIC M1A3P1000L, **VHDL**, only entropy coding core): 12 Msamples/s

Give it a try!



https://www.dapcom.es/get-fapec