Data compression on the sphere

with spherical Haar wavelets

Jason McEwen
http://www.mrao.cam.ac.uk/~jdm57/

Astrophysics Group, Cavendish Laboratory
University of Cambridge

Astronomical Data Analysis V, Crete, 2008

Motivation

- Large data-sets measured or defined inherently on the sphere arise in many applications (*e.g.* computer graphics, planetary science, geophysics, quantum chemistry, astrophysics).
- Current and forthcoming observations of the CMB of considerable size.
  WMAP: 3 mega-pixel maps; Planck: 50 mega-pixel maps
- Efficient and accurate compression of data on the sphere becoming increasingly important for both dissemination and storage of data.
- Compression on the sphere considered previously
  (notably by Schroder & Sweldens 1995 [4] for an icosahedron pixelisation of the sphere).
- We are motivated by the requirement for a compression algorithm defined on a constant latitude pixelisation and a publicly available implementation.

Outline

Haar wavelets on the sphere

- Wavelets on the sphere
  - Continuous wavelets
    e.g. Antoine & Vandergheynst 1998 [1], Wiaux *et al.* 2005 [6]
  - Discrete/discretised wavelets
    e.g. Schroder & Sweldens 1995 [4], Barreio *et al.* 2000 [2], McEwen & Eyers 2008 [3], Starck *et al.* 2006 [5], Wiaux *et al.* 2007 [7]
- Define approximation spaces on the sphere $V_j \subset L^2(S^2)$
- Construct the nested hierarchy of approximation spaces

$$V_1 \subset V_2 \subset \cdots \subset V_J \subset L^2(S^2) \, ,$$

  where coarser (finer) approximation spaces correspond to a lower (higher) resolution level j.
- For each space V_j we define a basis with basis elements given by the *scaling functions* $\varphi_{j,k} \in V_j$, where the k index corresponds to a translation on the sphere.
- Define detail space W_j to be the orthogonal complement of V_j in V_{j+1}, *i.e.* $V_{j+1} = V_j \oplus W_j$.
- For each space W_j we define a basis with basis elements given by the *wavelets* $\psi_{j,k} \in W_j$.
- Expanding the hierarchy of approximation spaces:

$$V_J = V_1 \oplus \bigoplus_{j=1}^{J-1} W_j \, .$$

Haar wavelets on the sphere

- Wavelets on the sphere
  - Continuous wavelets
    e.g. Antoine & Vandergheynst 1998 [1], Wiaux *et al.* 2005 [6]
  - Discrete/discretised wavelets
    e.g. Schroder & Sweldens 1995 [4], Barreio *et al.* 2000 [2], McEwen & Eyers 2008 [3], Starck *et al.* 2006 [5], Wiaux *et al.* 2007 [7]
- Define approximation spaces on the sphere $V_j \subset \mathrm{L}^2(\mathrm{S}^2)$
- Construct the nested hierarchy of approximation spaces

$$V_1 \subset V_2 \subset \cdots \subset V_J \subset \mathrm{L}^2(\mathrm{S}^2) \, ,$$

  where coarser (finer) approximation spaces correspond to a lower (higher) resolution level j.
- For each space V_j we define a basis with basis elements given by the *scaling functions* $\varphi_{j,k} \in V_j$, where the k index corresponds to a translation on the sphere.
- Define detail space W_j to be the orthogonal complement of V_j in V_{j+1}, *i.e.* $V_{j+1} = V_j \oplus W_j$.
- For each space W_j we define a basis with basis elements given by the *wavelets* $\psi_{j,k} \in W_j$.
- Expanding the hierarchy of approximation spaces:

$$V_J = V_1 \oplus \bigoplus_{j=1}^{J-1} W_j \, .$$

Haar wavelets on the sphere

- Wavelets on the sphere
  - Continuous wavelets
    e.g. Antoine & Vandergheynst 1998 [1], Wiaux *et al.* 2005 [6]
  - Discrete/discretised wavelets
    e.g. Schroder & Sweldens 1995 [4], Barreio *et al.* 2000 [2], McEwen & Eyers 2008 [3], Starck *et al.* 2006 [5], Wiaux *et al.* 2007 [7]
- Define approximation spaces on the sphere $V_j \subset \mathrm{L}^2(\mathrm{S}^2)$
- Construct the nested hierarchy of approximation spaces

$$V_1 \subset V_2 \subset \cdots \subset V_J \subset \mathrm{L}^2(\mathrm{S}^2) \, ,$$

  where coarser (finer) approximation spaces correspond to a lower (higher) resolution level j.
- For each space V_j we define a basis with basis elements given by the *scaling functions* $\varphi_{j,k} \in V_j$, where the k index corresponds to a translation on the sphere.
- Define detail space W_j to be the orthogonal complement of V_j in V_{j+1}, *i.e.* $V_{j+1} = V_j \oplus W_j$.
- For each space W_j we define a basis with basis elements given by the *wavelets* $\psi_{j,k} \in W_j$.
- Expanding the hierarchy of approximation spaces:

$$V_J = V_1 \oplus \bigoplus_{j=1}^{J-1} W_j \, .$$

Haar wavelets on the sphere

- Relate generic multiresolution decomposition to HEALPix pixelisation.
- Let V_j correspond to a HEALPix pixelised sphere with resolution parameter $N_{\text{side}} = 2^{j-1}$.
- Define the scaling function $\varphi_{j,k}$ at level j to be constant for pixel k and zero elsewhere:

$$\varphi_{j,k}(\omega) \equiv \begin{cases} 1/\sqrt{A_j} & \omega \in P_{j,k} \\ 0 & \text{elsewhere .} \end{cases}$$

- Orthonormal basis for the wavelet space W_j given by the following wavelets:

$$\psi^0_{j,k}(\omega) \equiv \left[\varphi_{j+1,k_0}(\omega) - \varphi_{j+1,k_1}(\omega) + \varphi_{j+1,k_2}(\omega) - \varphi_{j+1,k_3}(\omega)\right]/2\ ;$$
$$\psi^1_{j,k}(\omega) \equiv \left[\varphi_{j+1,k_0}(\omega) + \varphi_{j+1,k_1}(\omega) - \varphi_{j+1,k_2}(\omega) - \varphi_{j+1,k_3}(\omega)\right]/2\ ;$$
$$\psi^2_{j,k}(\omega) \equiv \left[\varphi_{j+1,k_0}(\omega) - \varphi_{j+1,k_1}(\omega) - \varphi_{j+1,k_2}(\omega) + \varphi_{j+1,k_3}(\omega)\right]/2\ .$$

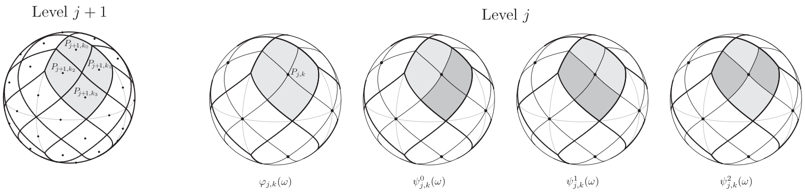


Figure: Haar scaling function $\varphi_{j,k}(\omega)$ and wavelets $\psi^m_{j,k}(\omega)$

Haar wavelets on the sphere

- Multiresolution decomposition of a function defined on a HEALPix data-sphere at resolution J, *i.e.* $f_J \in V_J$ proceeds as follows.
- Approximation coefficients at the coarser level j are given by the projection of f_{j+1} onto the scaling functions $\varphi_{j,k}$:

$$\lambda_{j,k} = \int_{S^2} f_{j+1}(\omega)\ \varphi_{j,k}(\omega)\ d\Omega\ .$$

- Detail coefficients at level j are given by the projection of f_{j+1} onto the wavelets $\psi^m_{j,k}$:

$$\gamma^m_{j,k} = \int_{S^2} f_{j+1}(\omega)\ \psi^m_{j,k}(\omega)\ d\Omega\ .$$

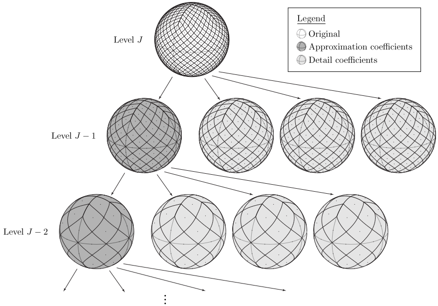


Figure: Haar multiresolution decomposition

- The function $f_J \in V_J$ may then be synthesised from its approximation and detail coefficients:

$$f_J(\omega) = \sum_{k=0}^{N_{J_0}-1} \lambda_{J_0 k}\varphi_{J_0 k}(\omega) + \sum_{j=J_0}^{J-1}\sum_{k=0}^{N_j-1}\sum_{m=0}^{2} \gamma^m_{j,k}\psi^m_{j,k}(\omega)\ .$$

Compression algorithms

- Haar wavelet transform to compress energy content.
- Lossless compression

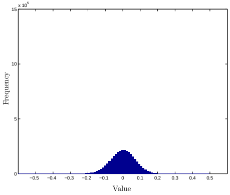
Lossless compression algorithm

1. Haar wavelet transform on sphere
2. Quantise detail coefficients to numerical precision (precision parameter p)
3. Huffman encoding

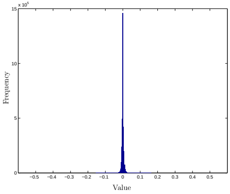
- Lossy compression
  - Controlled degradation to quality of original data allows higher compression ratios.
  - Discard detail coefficients close to zero.

Lossy compression algorithm

1. Haar wavelet transform on sphere
2. Thresholding
3. Quantise detail coefficients to numerical precision
4. Run-length encoding (RLE)
5. Huffman encoding



(a) Original data



(b) Wavelet coefficients

Figure: Histograms

Compression algorithms

- Haar wavelet transform to compress energy content.
- Lossless compression

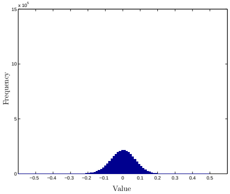
Lossless compression algorithm

1. Haar wavelet transform on sphere
2. Quantise detail coefficients to numerical precision (precision parameter p)
3. Huffman encoding

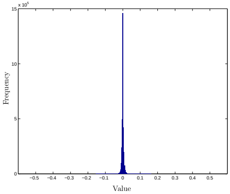
- Lossy compression
  - Controlled degradation to quality of original data allows higher compression ratios.
  - Discard detail coefficients close to zero.

Lossy compression algorithm

1. Haar wavelet transform on sphere
2. Thresholding
3. Quantise detail coefficients to numerical precision
4. Run-length encoding (RLE)
5. Huffman encoding



(a) Original data



(b) Wavelet coefficients

Figure: Histograms

Compression algorithms

- Haar wavelet transform to compress energy content.
- Lossless compression

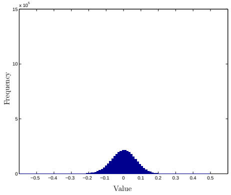
Lossless compression algorithm

1. Haar wavelet transform on sphere
2. Quantise detail coefficients to numerical precision (precision parameter p)
3. Huffman encoding

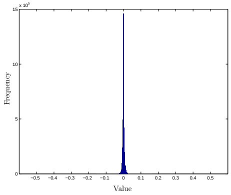
- Lossy compression
  - Controlled degradation to quality of original data allows higher compression ratios.
  - Discard detail coefficients close to zero.

Lossy compression algorithm

1. Haar wavelet transform on sphere
2. Thresholding
3. Quantise detail coefficients to numerical precision
4. Run-length encoding (RLE)
5. Huffman encoding



(a) Original data

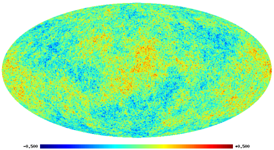


(b) Wavelet coefficients

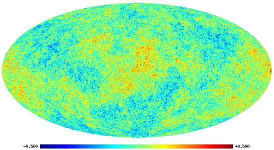
Figure: Histograms

Compression of CMB data: compression performance

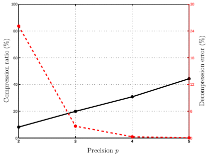
- Lossless to a user specified numerical precision only.



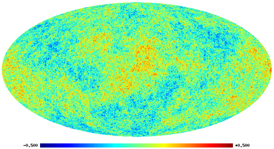
(a) Original for $N_{\rm side} = 512$ (13MB)



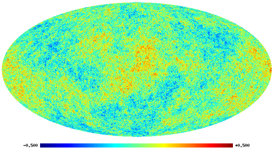
(b) Compressed for $N_{\rm side} = 512$ (2.5MB)



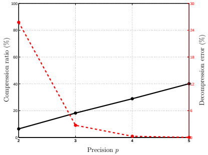
(c) Compression performance for $N_{\rm side} = 512$



(d) Original for $N_{\rm side} = 1024$ (50MB)



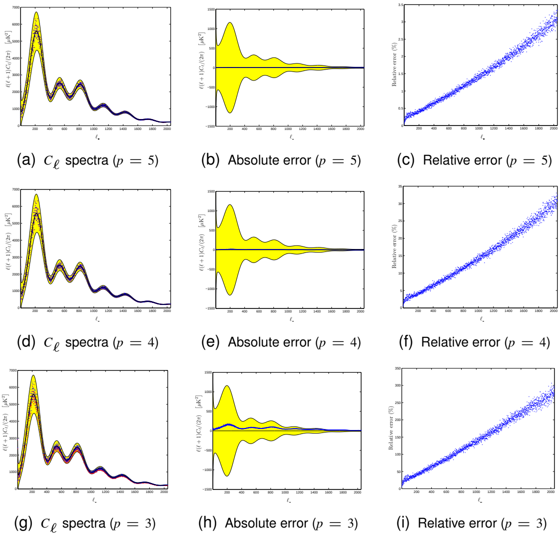
(e) Compressed for $N_{\rm side} = 1024$ (9.1MB)



(f) Compression performance for $N_{\rm side} = 1024$

Figure: Lossless compression of simulated Gaussian CMB data

Compression of CMB data: cosmological information content



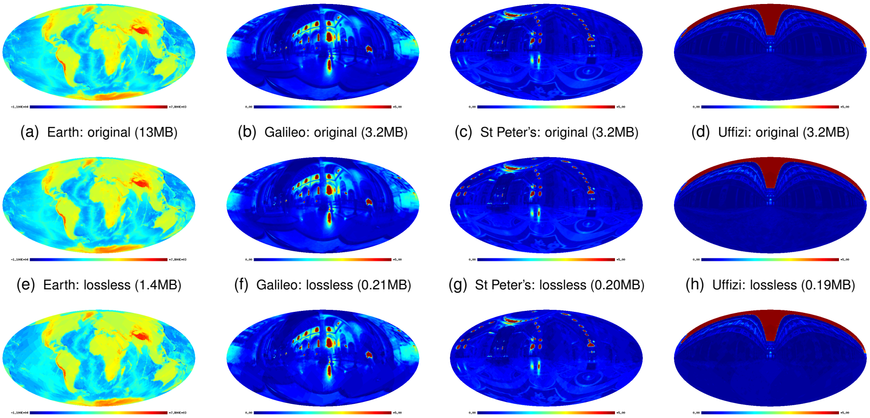
(a) C_ℓ spectra ($p = 5$) (b) Absolute error ($p = 5$) (c) Relative error ($p = 5$)

(d) C_ℓ spectra ($p = 4$) (e) Absolute error ($p = 4$) (f) Relative error ($p = 4$)

(g) C_ℓ spectra ($p = 3$) (h) Absolute error ($p = 3$) (i) Relative error ($p = 3$)

Figure: Reconstructed angular power spectrum of compressed CMB data

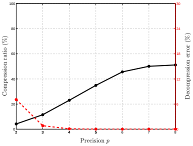
Lossy compression applications



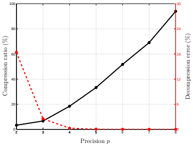
(a) Earth: original (13MB)
(b) Galileo: original (3.2MB)
(c) St Peter's: original (3.2MB)
(d) Uffizi: original (3.2MB)
(e) Earth: lossless (1.4MB)
(f) Galileo: lossless (0.21MB)
(g) St Peter's: lossless (0.20MB)
(h) Uffizi: lossless (0.19MB)
(i) Earth: lossy (0.33MB)
(j) Galileo: lossy (0.07MB)
(k) St Peter's: lossy (0.08MB)
(l) Uffizi: lossy (0.10MB)

Figure: Compressed data for lossy compression applications

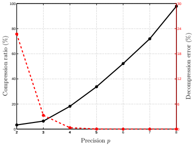
Lossy compression applications



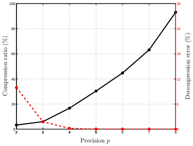
(a) Earth: lossless



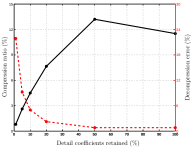
(b) Galileo: lossless



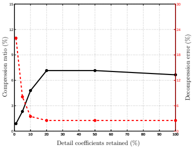
(c) St Peter's: lossless



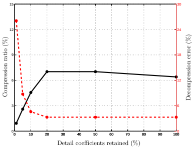
(d) Uffizi: lossless



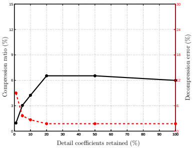
(e) Earth: lossy



(f) Galileo: lossy



(g) St Peter's: lossy



(h) Uffizi: lossy

Figure: Compression performance for lossy compression applications

Summary

- Developed algorithms to perform lossless and lossy compression of data defined on the sphere.
- Performance evaluated on various data and trade-off between compression ratio and fidelity of decompressed data examined.
- Compress CMB data to approximately 40% of its original size, while ensuring that essentially no cosmological information content is lost. Compress to below 20% if small loss of cosmological information content is tolerated.
- For lossy compression of Earth topography and environmental illumination data compression ratios of 40:1 ($\sim$2-3%) can be achieved for a relative error of $\sim$5%.
- Future improvements:
  - Other invertible wavelet transforms on the sphere (*e.g.* scale discretised wavelets of Wiaux *et al.* 2007 [7])
  - More sophisticated lossy compression algorithms
  - Optimise storage of encoding tables
- Implementation will be made available publicly very soon.

References

[1] J.-P. Antoine and P. Vandergheynst.
Wavelets on the n-sphere and related manifolds.
J. Math. Phys., 39(8):3987–4008, 1998.

[2] R. B. Barreiro, M. P. Hobson, A. N. Lasenby, A. J. Banday, K. M. Górski, and G. Hinshaw.
Testing the Gaussianity of the COBE-DMR data with spherical wavelets.
Mon. Not. Roy. Astron. Soc., 318:475–481, 2000.

[3] J. D. McEwen and D. M. Eyers.
Data compression on the sphere.
Mon. Not. Roy. Astron. Soc., submitted, 2008.

[4] P. Schröder and W. Sweldens.
Spherical wavelets: efficiently representing functions on the sphere.
In *Computer Graphics Proceedings (SIGGRAPH '95)*, pages 161–172, 1995.

[5] J.-L. Starck, Y. Moudden, P. Abrial, and M. Nguyen.
Wavelets, ridgelets and curvelets on the sphere.
Astron. & Astrophys., 446:1191–1204, February 2006.

[6] Y. Wiaux, L. Jacques, and P. Vandergheynst.
Correspondence principle between spherical and Euclidean wavelets.
Astrophys. J., 632:15–28, 2005.

[7] Y. Wiaux, J. D. McEwen, P. Vandergheynst, and O. Blanc.
Exact reconstruction with directional wavelets on the sphere.
Mon. Not. Roy. Astron. Soc., 2007.