Distributed Visual Information Management in Astronomy

Resolution scale is central to large-image visualization, offering one way to address astronomers' need to access and retrieve data. In addition, multiple-resolution information and entropy are closely related to compression rate, all three of which are related to the relevance and importance of information.

The quantity of astronomical data is rapidly increasing. This is partly owing to large digitized sky surveys in the optical and near infrared ranges. These surveys, in turn, are due to the development of digital imaging arrays such as charge-coupled devices (CCDs). The size of digital arrays is also increasing, pushed by astronomical research's demands for more data in less time.

Currently, projects such as the European DE-NIS (Deep Near Infrared Survey of the Southern Sky) and American 2MASS (Micron All Sky Survey) infrared sky surveys, or the Franco-Canadian MegaCam Survey and the American Sloan Digital Sky Survey, will each produce on the order of 10 Tbytes of image data. The American Large-Aperture Synoptic Survey Telescope, to be commissioned in 2007 and 2008, will produce approximately five Pbytes of data per year. In addition, the advent of automatic plate-scanning machines (including SuperCOSMOS in Edinburgh and several others) has made possible the routine and massive digitization of photographic plates. These machines let us digitize the enormous amount of useful astronomical data represented in a photograph of the sky, and they have opened up the full potential of large-area photographic sky surveys. However, transferring such amounts of data over computer networks becomes cumbersome and, in some cases, practically impossible. For example, transmitting a high-resolution Schmidt plate image over the Internet would take hours.

As astronomers face this enormous increase in pixels and realize that the catalogs they produce by extracting information from these pixels can be locally wrong or incomplete, their needs follow two different paths. First, they need fast access to informative pixel maps, which are more intuitively understandable than the derived catalogs. Second, they must be able to accurately refine astrometry (for example, positional data) and photometry (for example, accumulated flux data) or effectively detect missed objects.

Having briefly described the field's scientific needs, we can now look at how astronomers are explicitly using resolution and scale to assist data (image, tabular, and other) handling. These new

Fionn Murtagh
Queen's University, Belfast
Jean-Luc Starck
French Atomic Energy Commission
Mireille Louys
Université Louis Pasteur

vantage points help astronomers address the field's scientific needs. We first look at how resolution and scale are incorporated into scientific image compression. Compression is tied to information delivery, thus leading us to discuss visualization environments, partial decompression, and image-information summarization. We then exemplify how we can mathematically express information's relevance in practical applications, using entropy, and we consider storage issues and transmission channels, all in the overall context of data access and retrieval.

Compression strategies

When astronomers transfer and analyze high-resolution images, they can use different strategies to compress the data:[1,2]

- *Lossy compression*: In this case, the compression ratio is relatively low (less than 5 to 1).
- *Compression without visual loss*: This means you cannot see the difference between the original image and the decompressed one. Generally, you can obtain compression ratios between 10 and 20 to 1.
- *Good-quality compression*: The decompressed image contains no artifacts from the process, but it does lose some information. In this case, you can obtain compression ratios up to 40 to 1.
- *Fixed compression ratio*: For some technical reason or another, you might decide to compress all images with a compression ratio higher than a given value, whatever the effect on the decompressed image quality.
- *Signal–noise separation*: If noise is present in the data, noise modeling can allow for very high compression ratios just by including filtering in wavelet space during the compression.

The optimal compression method might vary according to the image type and selected strategy. A major reason for using a multiresolution framework is to obtain, in a natural way, progressive information transfer.

Signal–noise separation is particularly relevant when supporting a region of interest in an image. The JPEG 2000 standard, for example, supports a region of interest defined by a user- or automatically defined mask.[3] Noise analysis provides a natural, automated way to define the mask, and we can carry out noise analysis at each resolution scale. In the mask region, we use encoding that guarantees valid scientific interpretation, which is based on acceptable pixel-value

precision on decompression. Outside the mask region, wavelet coefficient filtering can go as far as zeroing the coefficients—for example, applying infinite quantization.

Using this principle of a mask region to define interesting and relevant signals versus less relevant regions, we can obtain compression ratios of close to 300 to 1, with guaranteed fidelity to the image's scientifically relevant properties (astrometry, photometry, and faint features). JPEG files, in contrast, rarely do better than approximately 40 to 1.

In the case of JPEGs, various studies have confirmed that beyond a compression ratio of 40 to 1, this compression method generates blocky artifacts for 12 bit-per-pixel images.[1] For the pyramidal median transform, the reconstruction artifacts appear at higher compression ratios—beyond a ratio of 260 to 1 in our images. (The pyramidal median transform is a pyramidal multiresolution algorithm based on the median transform and implemented in an analogous way to a wavelet transform.[1,4]) Figure 1 compares the visual quality of a JPEG image and a pyramidal-median-transform image.

Consider using a rigorously lossless wavelet-based compressor, above and beyond the issues of economy, storage space, and transfer time. Wim Sweldens' lifting scheme provides a convenient algorithmic framework for many wavelet transforms.[5] Predictor and update operators replace the low-pass and band-pass operations at each resolution level when constructing the wavelet transform. When the input data consist of integer values, the wavelet transform no longer consists of integer values, so we redefine the wavelet transform algorithm to face this problem. The predictor and update operators use a floor-truncation function, and their lifting scheme formulas let us carry this out without losing information.

The Haar wavelet transform's[4,6] lifting-scheme implementation creates lower-resolution versions of an image that are mathematically exact averaged and differenced versions of the next higher resolution level.[7] So, for aperture photometry and other tasks, lower-level resolution can provide a partial analysis. We can use a low-resolution-level image scientifically because its big pixels contain the integrated average of flux covered by the higher (or finer) resolution pixels. We can thus use efficiently delivered low-resolution images for certain scientific objectives, opening up the possibility for an innovative way to analyze distributed image holdings.
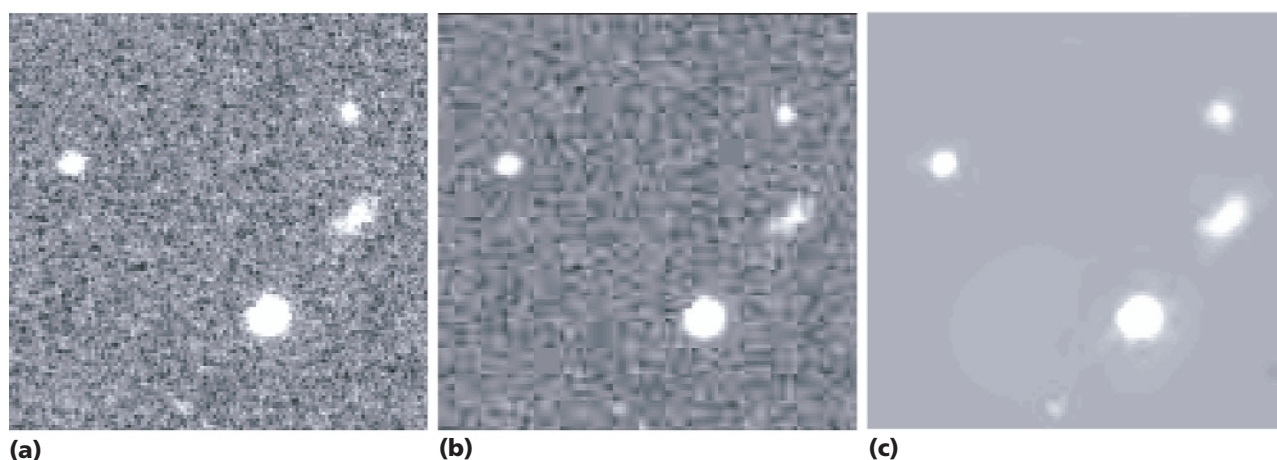
Figure 1. (a) An uncompressed image, which is a subimage extracted from a 1,024 × 1,024-pixel patch, in turn extracted from a European Southern Observatory Schmidt photographic plate (number 7992v); (b) a JPEG compressed image at a 40:1 compression ratio; and (c) a pyramidal-median-transform image at a 260:1 compression ratio.

Image visualization based on compression

With new technology developments, detectors are furnishing larger images. For example, current astronomical projects are beginning to deal with images larger than 8,000 × 8,000 pixels (ESO's Very Large Telescope 8,000 × 8,000 pixels, the MegaCam detector and the UK's Vista telescope, 16,000 × 16,000 pixels). For comparison with medical imaging, a digitized mammogram film might lead to images of approximately 5,000 × 5,000 pixels. In addition to data compression and progressive decompression, we must consider a third concept, the region of interest. Images are becoming so large that displaying them in a normal window (typically 512 × 512 pixels) is impossible, and we must be able to focus on a given area of the image at a given resolution. Moving from one area to another or increasing a particular area's resolution is an active element of decompression.

The principle of our Large Image Visualization Environment (LIVE) toolset, based on multiresolution data structure technology, is to support image navigation and full-image display at low resolution. Image navigation lets the user increase resolution (that is, improve the quality of an area of the image) or decrease it (return to the previous image), implying a fourfold increase or decrease in the size of what is viewed.

Figure 2 illustrates this concept, showing a large image (approximately 4,000 × 4,000 pixels) compressed into 500 × 500-pixel blocks (each block forming part of an 8 × 8 grid), represented at five resolution levels. The visualization window (256 × 256 pixels in our example) covers the whole image at the lowest resolution level (250 × 250 pixels) but only one block at the full resolution (or between one and four blocks, depending on the image's position). The LIVE concept consists of moving the visualization window into this pyramidal structure without loading the large image into memory. LIVE first visualizes the image at low resolution, and the user can indicate (using the mouse) which part of the visualized subimage he or she wants to enhance. At each step, the tool decompresses only wavelet coefficients of the corresponding blocks and of the new resolution level.

Decompression by scale and region

Supporting the transfer of very large images in a networked (client-server) setting requires compression and prior noise separation. Noise separation greatly aids in compression, because noise is axiomatically not compressible.

We developed one prototype in the MR/1 software package with a Java client[8] and another[9] using the Smithsonian Astrophysical Observatory's DS9 software, SAO DS9, to visualize large images (see http://hea-www.harvard.edu/RD/ds9). In developing these prototypes, we examined compression performance on numerous astronomical images. Consider, for example, a 12,451 × 8,268-pixel image from the CFH12K detector

at the Canada-France-Hawaii Telescope (CFHT), Hawaii. A single image is 412 Mbytes. Given a typical exposure time—a few minutes or less—we can quickly calculate the approximate amount of data expected in a typical observing night.

Some typical computation time requirements follow. Using denoising compression, we compressed the CFH12K image to 4.1 Mbytes—that is, to less than 1 percent of its original size. Compression took 13 minutes and 9 seconds on an UltraSparc 10. Decompression to the fifth resolution scale (that is, dimensions divided by 2^5) took 0.43 seconds. For rigorously lossless compression, compression to 97.8 Mbytes (23.75 percent of the original size) took 3 minutes and 44 seconds, and decompression to full resolution took 3 minutes and 34 seconds. Decompression to full resolution by block was near real time.

We developed a user interface[9] as a plug-in for the SAO-DS9 image viewer for images that the software package MR/1 compressed.[8] This interface lets the user load a compressed file and choose not only the image's scale but also its size and the portion to be displayed, resulting in reduced memory and processing requirements. Astrometry and SAO-DS9 functionality are still simultaneously available. Available functionality includes

- *Compression*: MR/1 includes compression and decompression tools. It implements wavelet, pyramidal-median, and lifting schemes, with lossy or lossless options. It stores the final file in a customized format.
- *An image viewer*: There are many astronomical image viewers. We looked at JSky (because it is written in Java) and SAOImage-DS9; we selected the latter because it is well maintained and easier for programmers to use. DS9 is a Tcl/Tk application that uses the SAOTk widget set. It also incorporates the new X Public Access (XPA) mechanism to let external processes access and control its data and graphical user interface functions.
- *An interface*: DS9 supports external file formats using an ASCII description file. It works with the MR/1 compressed format but can load only one scale of the image. The solution we selected was a Tcl/Tk script file, which interacts with XPA. The SAO team recommends Tcl/Tk, which is free and portable. This interface lets the user select a file, select the displayed window's maximum size, zoom in on a selected region (inside the displayed window), and unzoom.
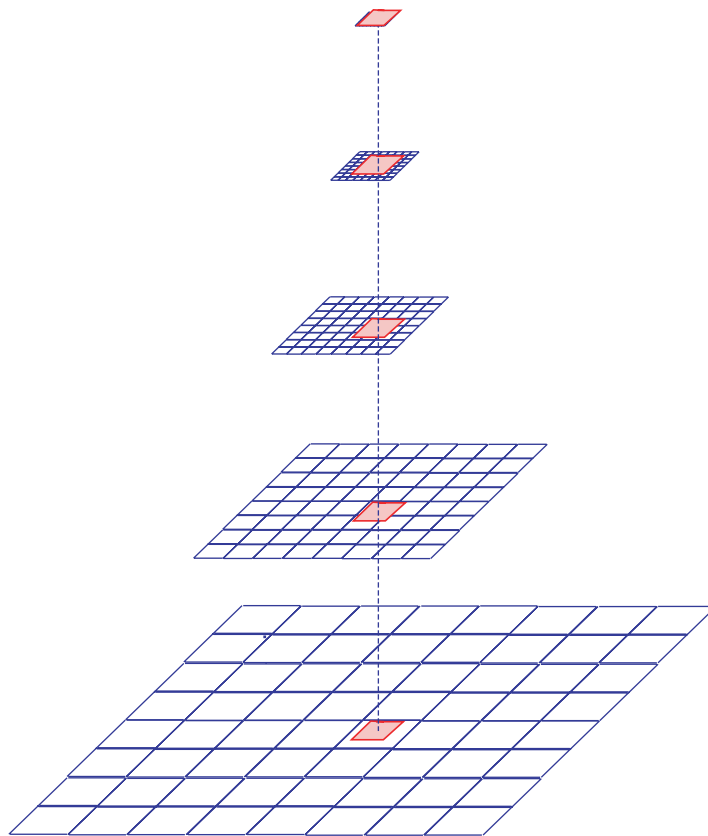

Figure 2. A large image compressed by blocks, represented at five resolution levels. At each level, the visualization window is superimposed at a given position. At low resolution, the window covers the whole image; at full resolution level, it covers only one block.

Astronomers have used the Tcl/Tk script file with DS9 and the decompressed module on Solaris (Sun Microsystems Sparc platform), Linux (Intel PC platform), and Windows NT and 2000 (with some tuning). It can also work on HP-UX and ALPHA-OSF1. On a three-year-old PC, the latency is approximately one second.

Figure 3 shows an example SAO-DS9 operation. The image shows a five-minute exposure (five 60-second dithered and stacked images), R-band filter, taken with a CFH12K wide-field camera (100 million pixels) at the primary focus of the CFHT in July 2000. Shown is a rich zone of our galaxy, containing star formation regions, dark nebulae (molecular clouds and dust regions), emission nebulae, and evolved stars.

Resolution scale in data archives

Unlike in Earth observation or meteorology, astronomers do not want to delete data after
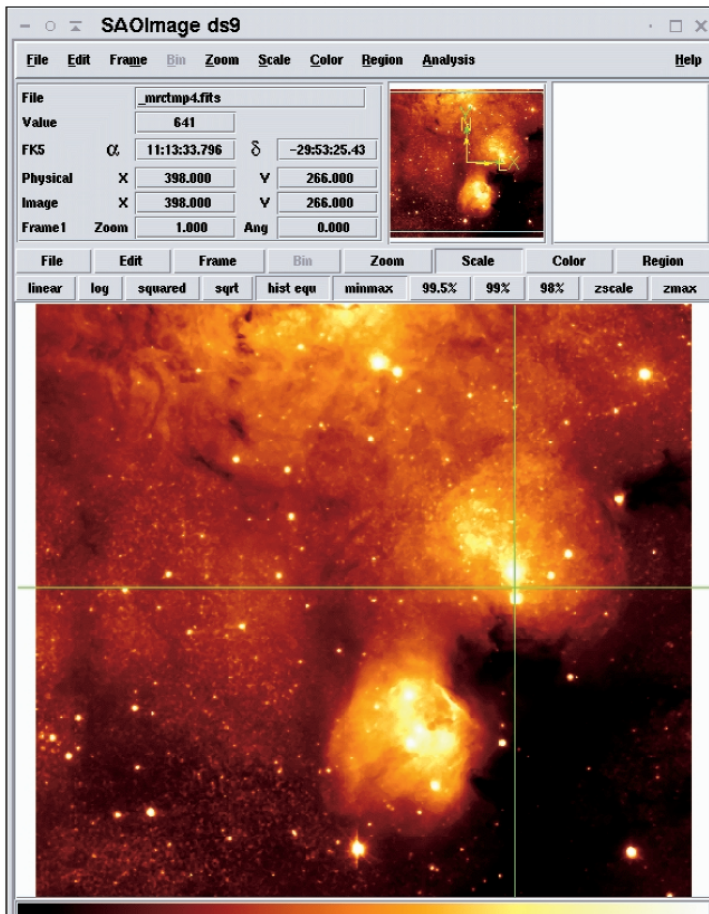
Figure 3. The Smithsonian Astrophysical Observatory's DS9 software with the XLIVE-DS9 user interface. Image courtesy of Jean-Charles Cuillandre.

they've interpreted it. Variable objects (supernovas, comets, and so forth) prove the need for astronomical data to be available indefinitely. The unavoidable problem is the overwhelming quantity of data that we now collect. The only basis for selecting what to keep long-term (and at what resolution and refinement levels) is to associate the data capture more closely with information extraction and knowledge discovery.

Research in data warehousing is now beginning to address this problem. Janne Skyt and Christian Jensen[10] discuss replacing aging, low-interest detailed data with aggregated data. Traditional databases are append-only, and deletion is a logical rather than physical operation—that is, the act of removing a link is not necessarily the freeing up of storage space. A new approach is based on a temporal vacuuming

specification, where access consists of both removal specification and keep specification. Removal is carried out in this new, storage-economizing approach in an asynchronous or lazy manner. A set of temporal relations, vacuumed according to specification, define a vacuumed temporal database.

So far, so good: we have a conceptual framework for keeping aggregated data long-term, based on an aggregation specification. One example is Web click-stream data,[10] where the aggregation is based on access hits. In astronomy imaging, we have already noted how the Haar wavelet transform, based on a lifting-scheme implementation, provides functionality for data aggregation. Aggregated flux uses "big" pixels, and local flux conservation is guaranteed.

Astronomers have yet to formally apply data aggregation to the vacuuming of scientific databases in practice.

Multiple-resolution information and entropy

Compression and resolution ought to be inherently linked to information content and, consequently, to entropy. The latter provides quality criteria (by asking, for example, if one compression result is better than another) and inherent limits to data coding. We first look at a link we developed between compression and entropy.

Elsewhere, we introduced a theory of multiscale entropy filtering, based on three stages:[11,12]

1. Model the signal or image as a realization (sample) from a random field, which has an associated joint *probability density function*, and compute entropy from this PDF, not directly from the signal or image pixel intensities themselves.
2. Use a basic vision model, which takes a signal, X, as a sum of components: $X = S + B + N$, where S is the signal proper, B is the background, and N is noise.
3. Extend this decomposition to further decompose entropy by resolution scale.

Stage 3 is based on defining the entropy in wavelet transform space. The wavelet transform's direct-current component (or continuum) provides a natural definition of signal background. A consequence of considering resolution scale is that it then accounts for signal correlation. Stage 2 rests on a sensor (or data capture) noise model.

For the resolution-scale-related decomposition, we have the following definition. Denoting h as the information relative to a single wavelet coefficient, we define

$$H(X) = \sum_{j=1}^{l} \sum_{k=1}^{N_j} h(w_{j,k}) \qquad (1)$$

with $h(w_{j,k}) = -\ln p(w_{j,k})$. l is the number of scales, N_j is the number of samples in band (scale) j, and $p(w_{j,k})$ is the probability that the wavelet coefficient $w_{j,k}$ is due to noise. The smaller this probability, the more important the information relative to the wavelet coefficient. For Gaussian noise, we get

$$h(w_{j,k}) = \frac{w_{j,k}^2}{2\sigma_j^2} + \text{Const.}, \qquad (2)$$

where σ_j is the noise at scale j. (In the case of an orthogonal or bio-orthogonal wavelet transform using an L^2 normalization, we have $\sigma_j = \sigma$ for all j, where σ is the noise standard deviation in the input data.) We can introduce multiscale entropy into filtering and deconvolution, and, by implication, into feature and faint signal detection.[11]

Elsewhere, we have considered a range of examples based on simulated signals, the widely used Lena image, and case studies from astronomy.[11,12] Later, two of us extended this framework to include both a range of noise models other than Gaussian and the role of vision models.[13] In the case of astronomy,[14] we looked at multiple band data, based on the Planck orbital observatory (a European Space Agency mission, planned for 2007, to study cosmic background radiation). We then introduced a joint wavelet and Karhunen-Loève transform (the WT-KLT transform) to handle cross-band correlation when filtering such data. We also looked at background-fluctuation analysis in astronomy, where we might not be able to observe the presence of astronomical sources but we know they are there (for instance, owing to observations in other parts of the electromagnetic spectrum).[14]

Multiscale entropy as a measure of relevant information

Because multiscale entropy extracts the infor-mation only from the signal, it was a challenge to see if an image's astronomical content was related to its multiscale entropy.

We studied the astronomical content of 200 images of $1{,}024 \times 1{,}024$ pixels extracted from scans of eight different photographic plates carried out by the MAMA digitization facility (Institut d'Astrophysique, Paris) and stored at the Strasbourg Data Center (Strasbourg Observatory, France). We estimated the content of these images in three ways, counting

1. Objects in an astronomical catalog (United States Naval Observatory A2.0 catalog) in the image. The USNO catalog was originally obtained by source extraction from the same survey plates we used in our study.
2. Objects that the Sextractor[15] object detection package found in the image. As in the case of the USNO catalog, these detections were mainly point sources (that is, stars as opposed to spatially extended objects such as galaxies).
3. Structures detected at several scales using the MR/1 multiresolution-analysis package.[7]

Figure 4 shows the results of plotting these numbers for each image against the image's multiscale-signal entropy. The MR/1 package obtained the best results, followed by Sextractor and then by the number of sources extracted from USNO. The latter two basically miss the content at large scales, which MR/1 considers. Unlike MR/1, Sextractor does not attempt to separate signal from noise.

We also applied Sextractor and multiresolution methods to a set of CCD images from CFH UH8K, 2MASS, and DENIS near infrared surveys. The results we obtained were similar to the results presented in Figure 4. This lends support to the quality of the results based on MR/1, which considers noise and scale, and to multiscale entropy being a good measure of content of such a class of images.

Subsequently, we looked for the relation between the multiscale entropy and an image's optimal compression ratio, which we can obtain using multiresolution techniques. (By optimal compression ratio, we mean a compression ratio that preserves all the sources and does not degrade the astrometry [object positions] and photometry [object intensities].) Mireille Louys and some of her colleagues have estimated this optimal compression ratio using the MR/1 package's compression program.[1]
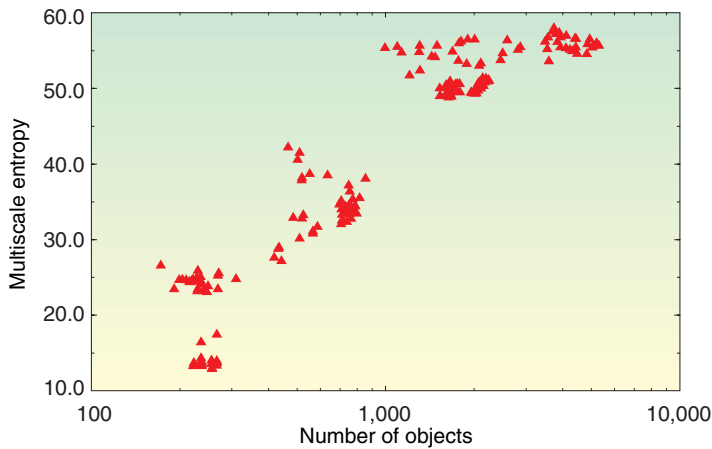
(a)
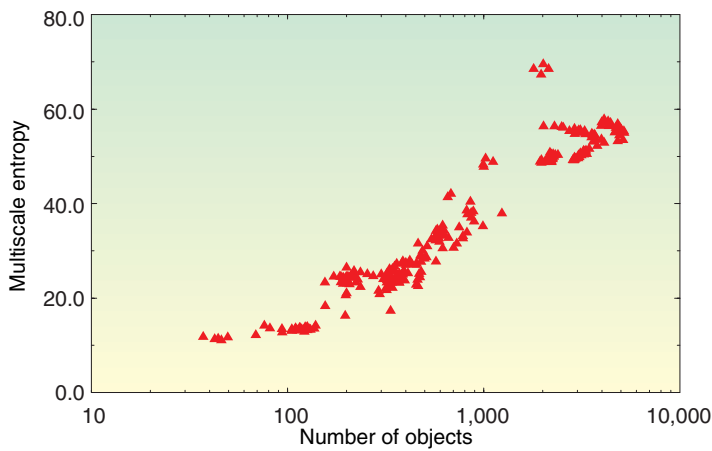
(b)
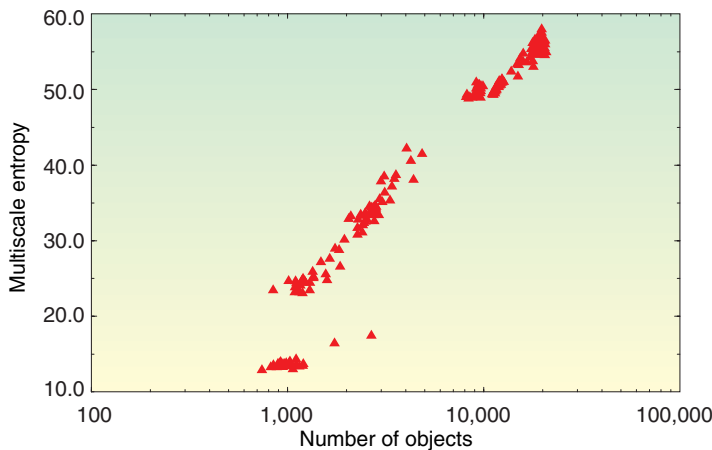
(c)

Figure 4. Multiscale entropy versus the number of objects: the number of objects obtained from (a) the United States Naval Observatory catalog, (b) the Sextractor package, and (c) the MR/1 package.

Figure 5 shows the relation between multiscale entropy and the optimal compression ratio for all images used in our previous tests, both digitized-plate and CCD images. The power law relation is obvious, letting us conclude that

- The compression ratio depends strongly on the image's astronomical content. Thus, compressibility is also an estimator of the image's content.
- The multiscale entropy confirms, and lets us predict, the image's optimal compression ratio.

Multiscale entropy for image database querying

We have seen that we must measure information from the transformed data—not from the data itself—so that we can consider a priori knowledge of the data's physical aspects. We could have used the Shannon entropy (perhaps generalized) to measure the information at a given scale and derive the histogram's bins from the noise's standard deviation. However, we thought it better to directly introduce noise probability into our information measure. This leads, for Gaussian noise, to a physically meaningful relation between the information and the wavelet coefficients (see Equation 2). First of all, information is proportional to the energy of the wavelet coefficients normalized by the noise's standard deviation. Second, we can generalize this to many other kinds of noise, including such cases as multiplicative noise, nonstationary noise, or images with few photons or events. Finally, our experiments have confirmed that this approach gives good results.

In the work presented in the preceding section, which was related to the semantics of numerous digital and digitized photographic images, we took already prepared (external) results and used two other processing pipelines to detect astronomical objects in these images. Therefore, we had three sets of interpretations of these images. We then used multiscale entropy to tell us something about these three sets of results. We found that multiscale entropy provided interesting insight into the performances of these different analysis procedures. Based on strength of correlation between multiscale entropy and the analysis result, we argue that this provided evidence of one analysis result being superior to the others.

Finally, we used multiscale entropy to measure optimal image compressibility. From our previ-

ous studies,[1,11,13] we already had a set of images with compression ratios consistent with the best recoverability of astronomical properties. These astronomical properties were based on positional and intensity information—astrometry and photometry. Therefore, we had optimal compression ratios, and for the corresponding images, we measured the multiscale entropy and found a strong correlation.

The breadth and depth of our applications lend credence to the claim that multiscale entropy is a good measure of image or signal content. The image data studied are typical not just of astronomy but of other areas of the physical and medical sciences. We have built certain aspects of the semantics of such data into our analysis procedures.

Could we go beyond this and directly use multiscale entropy in the context of content-based image retrieval? Yes, if the user's query is for data meeting certain signal-to-noise ratio requirements, or with certain evidence (which we can provide) of signal presence in noisy data. For more general content-based querying, our work opens up another avenue of research: in querying large data collections, we can allow greater recall at the expense of precision. Our semantics-related multiscale entropy measure can rank any large recall set. Therefore, we can use it in an interactive image-retrieval environment.

Total information of image and accumulated accesses

The vast quantities of visual data collected now and in the future present us with new problems and opportunities. Critical needs in our software systems include compression and progressive transmission, support for differential detail and user navigation in data spaces, and "thinwire" transmission and visualization. The technological infrastructure is just one side of the picture.

Another side is a human's limited ability to interpret vast quantities of data. A study by David Williams has quantified the maximum possible volume of data that researchers at CERN can conceivably interpret. This points to another, more fundamental justification for addressing the critical technical needs we've indicated. This is that the related themes of selective summarization and prioritized transmission are increasingly becoming a key factor in human understanding of the real world, as mediated through our computing and networking base. We must receive condensed, summarized data
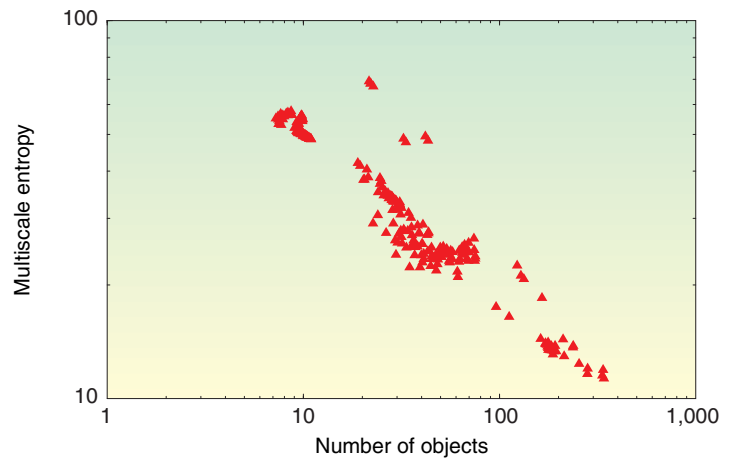

Figure 5. Multiscale entropy of astronomical images versus the optimal compression ratio. Images that contain numerous sources have a small ratio and a high multiscale entropy value. With logarithmic numbers of sources, the relation is almost linear.

first, which will then give us more detail, added progressively, to help us better understand the data. A hyperlinked and networked world makes this need for summarization more acute. We must consider resolution scale in our information and knowledge spaces. These are key aspects of progressive transmission.

Iconized and quick-look functionality imply a greater reliance on, and increased access to, low-resolution versions of images and other data. We have considerable expertise in the information content and hence compressibility of single images.[11,12] However, what is the total system's compressibility, for both storing and transferring files, when many users benefit from varying low-resolution versions of the data? We are interested in ensemble averages over large-image collections, many users, and many storage and transfer strategies. In other words, we are interested in the compressibility and information content of single-image files and the topology of search, feedback, and access spaces.

Researchers have traditionally applied coding theory to single image files. Jean Carlson at UC Santa Barbara and John Doyle at Caltech have provided an enhanced framework,[16,17] raising such questions as how do we link progressively coded images as separate files, and how do we group the resolution and scale components in single files? They point out that a Web layout allows, first and foremost, the logical cutting of 1D objects, such as a large image, into pieces for individual downloading. Such cutting embodies some progressive multiresolution coding—

that is, summary information first. Various Web design models that could be of interest in this context include simplified designs based on chain structures, tree structures, more general graph structures, and geometrical (or partition) structures.

We started by using resolution and scale in astronomy images, and it has led us to consider optimal Web site designs. Doyle and his colleagues find that this problem of visual information management is typical of complex systems that are robust and have a certain tolerance to uncertainty.[17] Access patterns show inherently bursty behavior at all levels, so we can't apply traditional Poisson models, which get smoothed out by data aggregation or by aggregation over time. Consequently, data aggregation, such as the use of the flux-preserving Haar wavelet transform (discussed earlier), will not reduce the information available. This is bad news from the viewpoint of total efficiency in our image retrieval systems, because such data aggregation will lead to evident gains in data storage but additional access and transfer overheads. The good news is that data aggregation does not go hand in hand with destroying information. There is no theoretical reason why we should not benefit from it in its proper context.

The *virtual observatory* in astronomy is premised on the fact that all usable astronomy data are digital (the term "virtual" meaning using reduced or processed online data). High-performance information cross-correlation and fusion, and long-term availability of information, are required.

A second trend with major implications is that of the Grid. The computational Grid aims to provide an algorithmic and processing infrastructure for the scientific "collaboratories" of the future. The data Grid aims to allow ready access to information from our tera- and petabyte data stores. Finally, the information Grid should actively and dynamically retrieve information, not just pointers to where information might exist.

The evolution of how we do science, driven by these themes, is inextricably linked to the problems and recently developed algorithmic solutions we surveyed in this article. **C**₍SE₎

References

1. M. Louys et al., "Astronomical Image Compression," *Astronomy and Astrophysics Supplement Series*, vol. 136, no. 3, May 1999, pp. 579–590.

2. F. Murtagh, J.L. Starck, and M. Louys, "Very High Quality Image Compression Based on Noise Modeling," *Int'l J. Imaging Systems and Technology*, vol. 9, no. 1, 1998, pp. 38–45.

3. C. Christopoulos, J. Askelöf, and M. Larsson, "Efficient Methods for Encoding Regions of Interest in the Upcoming JPEG 2000 Still Image Coding Standard," *IEEE Signal Processing Letters*, vol. 7, no. 9, Sept. 2000, pp. 247–249.

4. J.L. Starck, F. Murtagh, and A. Bijaoui, *Image and Data Analysis: The Multiscale Approach*, Cambridge Univ. Press, Cambridge, UK, 1998.

5. W. Sweldens, "The Lifting Scheme: A Custom-Design Construction of Biorthogonal Wavelets," *Applied and Computational Harmonic Analysis*, vol. 3, no. 2, Apr. 1996, pp. 186–200.

6. M. Louys, J.L. Starck, and F. Murtagh, "Lossless Compression of Astronomical Images," *Irish Astronomical J.*, vol. 26, no. 2, 1 July 1999, pp. 119–122.

7. MR/1, *Multiresolution Image and Data Analysis Software Package, Version 3.0*, Multi Resolutions Ltd., 2001; www.multiresolution.com.

8. J.L. Starck and F. Murtagh, *Astronomical Image and Data Analysis*, Springer-Verlag, New York, 2002.

9. R.D. Gastaud, F.S. Popoff, and J.L. Starck, "A Widget Interface for Compressed Image Based on SAO-DS9," to be published in *Astronomical Data Analysis Software and Systems Conf. XI*, Astronomical Soc. of the Pacific, San Francisco, 2001.

10. J. Skyt and C.S. Jensen, "Persistent Views: A Mechanism for Managing Aging Data," *Computer J.*, vol. 45, no. 5, 2002, pp. 481–493.

11. J.L. Starck, F. Murtagh and R. Gastaud, "A New Entropy Measure Based on the Wavelet Transform and Noise Modeling," *IEEE Trans. Circuits and Systems Part II*, vol. 45, no. 8, Aug. 1998, pp. 1118–1124.

12. J.L. Starck and F. Murtagh, "Multiscale Entropy Filtering," *Signal Processing*, vol. 76, no. 2, 1 July 1999, pp. 147–165.

13. J.L. Starck and F. Murtagh, "Astronomical Image and Signal Processing: Looking at Noise, Information, and Scale," *IEEE Signal Processing*, vol. 18, no. 2, Mar. 2001, pp. 30–40.

14. J.L. Starck et al., "Entropy and Astronomical Data Analysis: Perspectives from Multiresolution Analysis," *Astronomy and Astrophysics*, vol. 368, no. 2, Mar. 2001, pp. 730–746.

15. E. Bertin and S. Arnouts, "Sextractor: Software for Source Extraction," *Astronomy and Astrophysics Supplement Series*, vol. 117, no. 2, 1 June 1996, pp. 393–404.

16. J. Doyle and J.M. Carlson, "Power Laws, Highly Optimized Tolerance, and Generalized Source Coding," *Physical Rev. Letters*, vol. 84, no. 24, 12 June 2000, pp. 5656–5659.

17. X. Zhu, J. Yu, and J. Doyle, "Heavy Tails, Generalized Coding, and Optimal Web Layout," *Proc. 20th Ann. Joint Conf. IEEE Computer and Communications Societies* (INFOCOM 01), vol. 3, IEEE Press, Piscataway, N.J., 2001, pp. 1617–1626.

Fionn Murtagh is a professor of computer science at Queen's University, Belfast. He is also an adjunct professor at Strasbourg Astronomical Observatory, Strasbourg, France. He holds a BA and BAI in mathematics

and engineering science, and an MSc in computer science, all from Trinity College Dublin, a PhD in mathematical statistics from Université P & M Curie, and an Habilitation from Université L. Pasteur. He chairs the iAstro project (www.iAstro.org) and is the editor-in-chief of *Computer Journal*. Contact him at the School of Computer Science, Queen's Univ., Belfast, Belfast BT7 1NN, Northern Ireland, UK; f.murtagh@qub.ac.uk.

Jean-Luc Starck is a senior researcher at the French national energy agency, CEA. The projects he has worked on include ISO, XMM, Planck, and Terapix. He holds a PhD from the University of Nice at Sophia Antipolis, and an Habilitation (DSc) from the University of Paris XI. Contact him at DAPNIA/SEI-SAP, CEA-Saclay, 91191 Gif-sur-Yvette Cedex, France; jstarck@cea.fr.

Mireille Louys is an assistant professor at the École Nationale Supérieure de Physique de Strasbourg and a researcher at the Laboratoire des Sciences de l'Image, de l'Informatique et de la Télédétection in Strasbourg. She has been involved in metadata standardization work and interoperability in the framework of the International Astronomical Virtual Observatory Alliance. She received her PhD in digital image analysis and processing at the Université Louis Pasteur, Strasbourg, France. Contact her at LSIIT, École Nationale Supérieure de Physique de Strasbourg, Bd. Sebastien Brandt, 67400 Illkirch; mireille.louys@astro.u-strasbg.fr.

For more information on this or any other computing topic, please visit our Digital Library at http://computer.org/publications/dlib.