A Few Thoughts About Image File Storage

Michael Bode
Soft Imaging System Corp.
mb@soft-imaging.com

A recent thread on the MSA list server about problems with image formats (in this case TIF or TIFF) showed, that there is a bit of confusion in the microscopy community about the best file format for digital images. I will try to shed a bit of light onto this problem.

Digital images are at the core a large array of numbers. One number per pixel for b/w images, 3 numbers per pixel for color images. The simplest file format consists of 2 numbers that define the width and height of the image, and then just a listing of numbers for each pixel. By using the right conventions, the image can be recreated from a data file like this. In essence, this is the format of a bitmap image (BMP), and other formats derived from it. It is immediately obvious, that this is a reasonable format for vacation pictures, but it is not sufficient for scientific images. The minimum information we need to keep with the file is a calibration value, that tells us what the real world size of each pixel is, which is not the case for a simple file like this.

As most file formats were developed without a microscopist in mind, they simply do not have an option to store this information within the file itself. Some manufacturers have found a workaround by providing an additional text file where all this information is kept. Others are using more complex file types, which do provide methods for keeping the information. One of those file formats is TIF (or TIFF). This acronym stands for Tagged Image File Format and has grown to be the *de-facto* standard for imaging, mostly due to its flexibility.

The TIF file format defines a number of standard "tags"in the so-called header part of the TIF file, such as width, height, and other properties of the image. It also provides a method for anybody to incorporate their own, "private" tags. However, these private tags are usually not accessible or even known to the general public, so they only provide a consistent way of storing information for files from those manufacturers who share that information. TIF specifies, that unknown tags be ignored by the software that can't identify them. So, if company A stored magnification in a private tag, Software from company B will not read or correctly interpret this unless the information about it is shared between the companies.

Other uses of the TIF format lead to misinterpretation. For example, tag number 282 is the "Xresolution". While this seems to be a natural tag to store the image resolution or calibration, it leads to unsuspected results. Software, such as WORD or other programs use this value, together with the number of pixels in *x* and *y* to calculate the actual size of the image on paper. If a TEM image is acquired with, for example a 1 nm pixel resolution and an image width of 1000 pixels, and the information "1 nm" is stored in tag 282, it will be printed on paper at a size of 1nm x 1000 = 1 micron, definitely too small for any publication. Another issue can be the way 12 or 14 bit b/w images are interpreted. As there is no 12 or 14 bit file format, the images need to be stored as 16-bit files. This can then lead to black images (for example in Photoshop), because the application displays the full 16-bit. The 12 bits of information only fill the darker 6% (1/16th) of this range, and the image appears black. This can usually be fixed easily with an adjustment of the display levels. There is a new standard under consideration (EXIF), which combines features found in JPEG with TIFF features, and it remains to be seen if this new standard addresses these issues. The good news is, that all TIFF files should be compatible, as long as the writer and reader software adhere to the TIFF standards.

To avoid cumbersome re-calibration of images when they are transferred from one application to another, I suggest, that the Microscopy Society of America, as a representative organization of microscopists, develop a minimum set of parameters that need to be stored with an image. A subset of "MSA tags" in TIF files could then be defined to share information across applications. These tags should also be submitted to the body that defines the standard TIF tags and other standard setting authorities for incorporation into general standards.

Compression is also an issue that is often misunderstood. In short, there are two types of compression: lossy and loss-less. It is important to understand the differences. A "loss-less" compression, such as the one employed in the popular "ZIP" utilities, can restore a file to the exact form it was before. The compression ratios depend on the contents of the file, but are typically 2:1. A "lossy" compression, such as JPG, can result in much higher compression ratios (10:1 and higher), but it will change the content of the files. The choice is therefore to either preserve all the information in an image, or create a small file size. Both have their places: For publications or displays of images one can probably accept minor changes to the content, therefore JPG would be a reasonable choice. If the image needs to be processed further, any changes to the file could potentially introduce artifacts, which make the choice of loss-less compression obvious. In fact, MSA recently addressed the problem with a statement regarding ethical digital imaging (*Microscopy Today*, Vol 11, Number 6, page 61): "Ethical imaging requires that the original uncompressed image file be stored on archival media ...".

To demonstrate the losses, Fig.1 shows a test image in its original form. The uncompressed file size of this image is about 863 KB. Using loss-less compression, the file size can be reduced to about 140 KB. Using JPG compression and a "quality factor" of 50%, the file size can be reduced further to about 23 KB, which is almost a factor 40 smaller than the original image. However, the compression leads to artifacts and losses in the image. Fig. 2 shows the compressed image, and Fig. 3 shows the difference between the original and the compressed image. The contrast of this difference image has been increased to demonstrate the effect. The image shows the information that is lost by the compression, and it is immediately clear, that some of the information lost may prevent further image analysis.

So, what is the best file format to store images? Although there is no general answer to this question, TIF seems to be the best all-around format at this point, even with the shortcomings explained above. ■

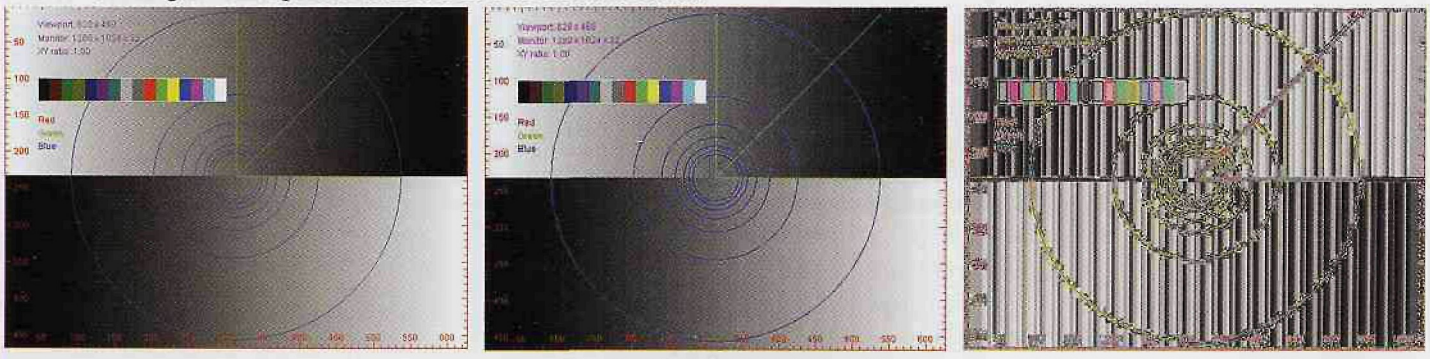


Fig.1 (Left) Test image . Fig.2 (Center) JPEG compressed image, quality factor 50%. Fig.3 (Right) Losses to Image in Fig 1 due to compression (contrast equalized to show details)

INCAFeature

Visionary microanalysis technology See it today

Available right now with:

Unique automatic identification of all elements
Simple image assisted set-up
Fast feature detection and data collection algorithm

INCAFeature

The automated feature microanalysis system for *your* electron microscope.

And it's the only feature microanalysis system which assumes no prior knowledge of the elements within your samples.

Analysis and classification

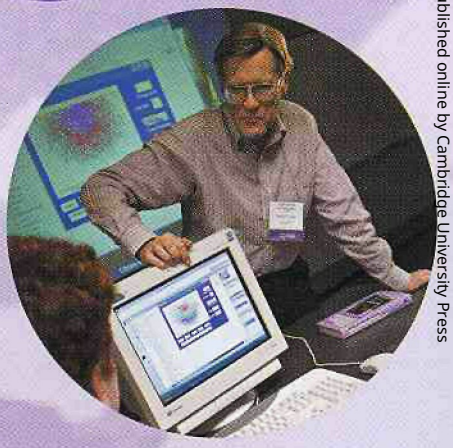
- Particles
- Inclusions
- Gunshot residue
- Filters
- Wear metals

Call us today for more details about the world leading INCAFeature microanalysis system.

Oxford Instruments Analytical
130A Baker Avenue Ext. Concord, MA 01742 USA
Tel: +1 978 368 9933 Fax: +1 978 369 8287 Toll free: +1 800 447 4717
email: info@ma.oxinst.com
www.oxford-instruments.com

OXFORD INSTRUMENTS

LEHIGH Microscopy SCHOOL



June 6-18, 2004

Nine Courses in SEM, X-ray Analysis, AEM, AFM for Materials Engineers, Geologists, Biologists, Polymer Scientists

The World's Best SEM, AEM, SPM Courses

- *http://www.lehigh.edu/microscopy*
- *Taken by over 4000 engineers, scientists, and technicians*
- *Participants have come from 48 states and 29 countries*
- *Free to every registrant: the Lehigh CD ROM containing software, images, and data*
- *Textbook authored by the lecturers is included*

For a detailed brochure contact:

Ms. Sharon L. Coe
Lehigh Microscopy School
Whitaker Lab
Lehigh University
5 East Packer Avenue
Bethlehem, PA 18015-3195

P: 610.758.5133
F: 610.758.4244
E: sharon.coe@lehigh.edu

"Overall, an excellent course. Residential facility, food and social hours are 'A' plus."

-From 2003 participant

LOOK FOR REGISTRATION FORM AND COURSE OUTLINES AT WWW.LEHIGH.EDU/MICROSCOPY

MAIN COURSE

Scanning Electron Microscopy and X-ray Microanalysis
June 7-11, 2004

Designed for those who use scanning electron microscopy and x-ray microanalysis in academic, governmental, or industrial laboratories: engineers, technicians, physical and biological scientists, clinicians, and technical managers. This course will provide a working knowledge of the principles of these two related techniques. Special optional sessions will be conducted for students interested specifically in specialized imaging techniques, x-ray analysis of difficult specimens, and organic materials.

Introduction to SEM and EDS for the New SEM Operator
June 6, 2004

One-day course with lectures and labs related to the basic operation of the SEM. Enrollment will be limited to 60 participants who are also attending the main SEM course beginning on Monday.

ADVANCED COURSES

Problem Solving with SEM and X-Ray Microanalysis
June 14-18, 2004

For scientists, engineers, technicians, and technical managers who seek an intermediate/advanced course in the techniques of SEM and x-ray microanalysis applied to problem solving. High-resolution SEM, low-voltage SEM, variable-pressure SEM (environmental), electron backscatter diffraction (EBSD), and x-ray spectrometry by EDS, WDS, silicon drift detectors, and microcalorimeters will be described in lectures and laboratories. Emphasis will be given to specific problem solving. Students are encouraged to bring specimens and/or data to the course for discussion. Items could include SEM images, x-ray spectra (export as MSA/MAS format), EBSPs, etc.

Quantitative X-ray Microanalysis of Bulk Specimens and Particles
June 14-18, 2004

This course is for SEM and microprobe operators and supervisors who wish to advance their knowledge of x-ray microanalysis. Emphasis will be placed on the physical concepts and practical details required for successful quantitative analysis. Topics include ZAF and $\phi(\rho z)$ corrections, WDS and EDS, particle analysis, and analysis of thin films on substrates. Computer-aided imaging and microanalysis will also be discussed.

Analytical Electron Microscopy Analysis for TEM Specimens
June 14-17, 2004

This course gives an integrated, in-depth description of the capabilities of the latest analytical transmission electron microscopes (TEMs). These instruments offer high spatial resolution elemental and chemical microanalysis through x-ray and electron spectrometry. Convergent-beam electron diffraction allows point and space-group analysis of micro-crystalline regions. The student should have a basic understanding of the fundamentals of TEM and/or SEM.

Characterization of Nanostructures
June 14-17, 2004

In this course, an introduction is given to the different classes of nanostructured materials (e.g., nanoparticles, nanowires, nanotubes) with special emphasis on explaining why nanoscale materials have different physical properties from their bulk counterparts. *'Top-down'* fabrication processes such as photolithography, electron beam lithography, and focused ion beam lithography are compared with *'bottom-up'* fabrication techniques such as self-assembly, nanolithography, and nanomanipulation. The wide variety of tools available to characterize the structure, properties, and chemistry of nanomaterials are then reviewed. Crystallographic techniques (XRD, CBED, neutron diffraction), imaging techniques (STM, AFM, SEM, HREM, STEM), and spectroscopic methods (XEDS, EELS, atom probe, Auger and XPS) will all be described in lectures, and demonstrated, wherever possible, during laboratory sessions.

Focused Ion Beam Instrumentation and Applications
June 14-17, 2004

This course is designed for those using or who are interested in using the FIB or FIB/SEM platforms in academic, government, or industrial laboratories in either the physical or biological sciences. Basic ion-solid interaction theory will be introduced and used in describing methods for specimen preparation (SEM, TEM, AFM, Auger, SIMS, atom probe, etc.), FIB/SEM analytical characterization, and milling/deposition techniques for nanotechnology.

Particle Characterization, Preparation, Microscopy, and Analysis
June 14-17, 2004

This course provides an integrated overview of the various methods used for the imaging and microanalysis of particles. Emphasis will be placed on preparation techniques and the principles behind the analytical methodologies for the analysis of both individual particles and particle populations. Topics include the analysis of individual particles by electron microscopy techniques, including electron probe microanalysis and analytical scanning electron microscopy, electron backscatter diffraction, and nanoparticle analysis by transmission analytical electron microscopy. This course will cover polarized light microscopy as it is used to support specimen preparation, electron beam analysis, and the analysis of particle populations by automated electron probe microanalysis.

Atomic Force Microscopy and Other Scanned Probe Microscopies
June 14-17, 2004

This course provides an understanding of the concepts, instrumentation, and applications of the rapidly expanding field of scanned probe microscopes (SPM). AFM and STM will be covered extensively, while other SPM techniques will be covered in varying depths depending on the interest of students. The theory of operation for both imaging and spectroscopy will be addressed, with attention paid to instrumental artifacts and methods to avoid them.