

Securing and Authenticating Images with On-Image Metadata

Charles Bradley and James Mancuso

Advanced Microscopy Techniques, 3 Electronics Ave., Danvers, MA 01923

Though worth "a thousand words," a photomicrograph still requires the additional context of "who, what, when, where, and how" to fulfill its scientific purpose. Additionally, image integrity has legal and ethical implications, when using images as evidence for business, funding, or medical decisions [1,2]. An advantage of a digital format is that images can be made self-explanatory and secure by adding this information as metadata into image files.

Adobe's Tiff 6.0 [3], which allows 65535 "tags" for metadata, is the *de facto* standard for saving digital micrographs. Unfortunately, only the 11 "required baseline" tags specified by Adobe for general photographic applications have universal support. This limits the utility of any non-baseline tags to a single application. Even worse, non-baseline metadata can be silently stripped from the file when the image is re-saved on other platforms that do not support these tags, as shown in Fig 1. Thus, all but the baseline TIFF tags lack portability to widely available, low cost applications. Portability requires a comprehensive standard for the content and format of metadata as well as a guaranty that metadata are preserved.

Our goal is to develop a rigorous and robust protocol for saving images that addresses the following:

- 1) Original picture data must be preserved.
- 2) Experimental conditions plus sample and authorship information must be preserved.
- 3) It must be possible to authenticate both picture and metadata.
- 4) It must have the ability to record image processing steps.

We note that that burning in captions and scale bars is a well-accepted practice in emulsion-based recording of micrographs. Moreover, this information scales with and survives darkroom procedures. In fact these essentially comprise a small set of well-preserved metadata in an analog format. Many digital systems emulate this and display a scale bar and caption with the image, although this information is not always burned-in or saved with the TIFF file. Currently digital emulation is unnecessarily limited to a handful of parameters, even though a digital system can save far more information than conveniently shown in captions. We propose the following general features:

- 1) A single TIFF file contains both digital image and standard metadata.
- 2) Metadata reside in a fixed caption that remains in the "picture" section of the TIFF file.
This metadata representation in the caption is therefore visible.
- 3) The metadata should survive conversion to loss-less image formats.
- 4) The metadata set contains security and self-authenticating features to ensure image integrity.

Figure 2 shows a prototype version of this protocol using a "Secure Data Strip," which is 40 lines wide. Metadata are encoded into grayscale values in a reserved region of caption portion of a single image. Each pixel is given an 8 bit ASCII character value (in an 8 bit image). In this 2k wide image the strip can incorporate 2000 characters per line, so this data strip has a capacity of 80kB. This is

more than enough to record typical experimental conditions, sample information, facility, and authorship.

There are several advantages to this approach. Most importantly, existing TIFF reading programs can be used "as is," because the metadata field is simply another part of the picture. Second, there is no need to carry and keep track of companion files containing metadata. Third, metadata removal is immediately apparent, because the strip is visible. Fourth, as long as one avoids file formats that use "lossy" compression, any format can be used to re-save the image. Finally, use of this approach does not preclude other protocols of incorporating metadata into a file.

A simple stand-alone application can read the encoded information, but practically, plug-ins or functions need to be included for each image acquisition or image manipulation package. Thus this method should be non-proprietary, so that all suppliers of digital imaging equipment can freely implement the reading and writing of metadata.

References

- [1] J.M. MacKenzie et al., *Microscopy Today* 14 (1) (2006) 40.
- [2] U.S. Food and Drug Administration, (CFR 21 Part 11) http://www.fda.gov/ora/compliance_ref/Part11/
- [3] Adobe TIFF Developer, <http://partners.adobe.com/public/developer/tiff/index.html#spec>

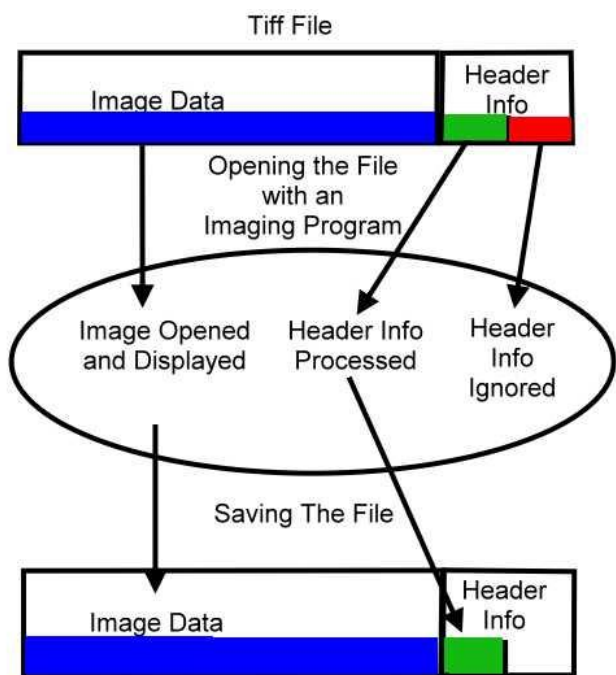


FIG. 1. Metadata stored in a tag in a tiff header can be silently stripped from the file.

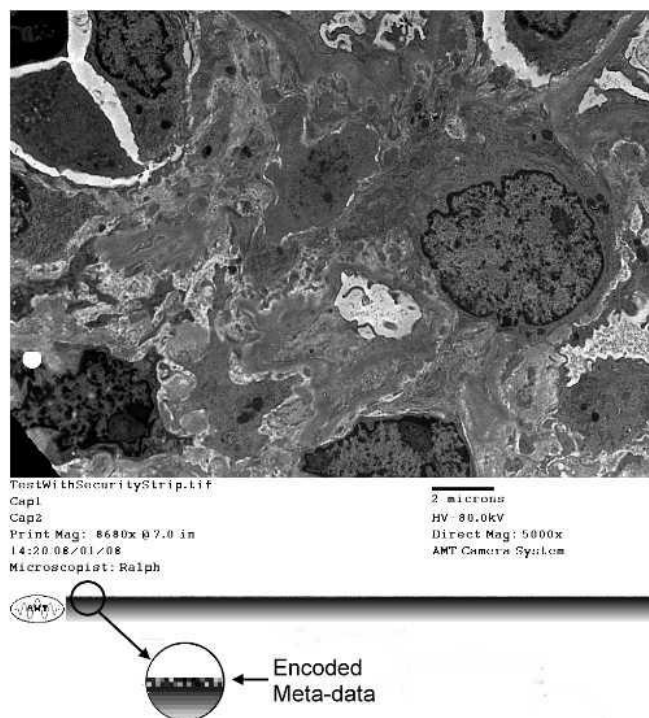


FIG. 2. Encoded metadata incorporated as part of an image.