Using Metadata to Simplify Digital Photography

James R. Milch and Kenneth A. Parulski
Eastman Kodak Company
Rochester, NY USA

Abstract

Digital imaging is maturing and moving into a new environment. This evolution of products and technology is introducing new requirements and new expectations. Consumers and business people would like to experience the benefits of digital imaging, but do not want to deal with the intricacies of the technology. Instead, they expect the new technology to blend in with and enhance the products they use today. For many people, digital photography is still too difficult. Most digital pictures are merely collections of pixel values. They do not leverage the capability of automated information systems to capture, retain, and deliver information just when it will do the most good. Image metadata can improve the user's experience of digital photography by providing continuity along the imaging chain and across time. It may describe the scene, the taking conditions, customer preferences, or printing information. Metadata is particularly important as the number of elements and service providers for digital photography increases. Automated transfer of control information is necessary for ease of use. The success of image metadata in the marketplace will depend on wide adoption of standard architectures, formats, and interpretations for the data.

Introduction

Digital imaging is maturing and moving into a new environment. This evolution of products and technology is introducing new requirements and new expectations. The new customers are not imaging professionals, who are in the business of managing and editing images. Instead, they are consumers and business people who would like to experience the benefits of digital imaging, but do not want to deal with the intricacies of the technology. They will not spend a lot of time learning to use tools or understanding the interactions between software applications. Instead, they expect the new technology to blend in with and enhance the products they use today. They don't want to be forced to change their habits of taking and using pictures.

The personal computer is an excellent platform for digital imaging, as practiced by professionals. They need the flexibility of the computer and are willing to learn its peculiarities. Consumers will also manage, edit, and print pictures with their personal computers, but they will expect their digital pictures to do more. Consumers are finding digital imaging services offered at retail kiosks, over the Internet, and through single-purpose imaging appliances. They can get a CD-ROM that contains high-resolution copies of the photographs they took with film. They can have these same files delivered over the Internet. They can order a variety of printed output from a CD-ROM, a compact flash card, directly from a digital camera, or from another print. These services, which together make up digital photography, are easily tailored for personal pictures enjoyed at home or business pictures used to enhance communication at work.

Most of the digital images people work with today are simple collections of pixel values. They come with no context or provenance. That is acceptable for a few images, used briefly, by one person. As soon as someone accumulates a collection of digital images, two important problems arise. First, pixel values generated by different image sources have different meanings. Second, file names and folders provide too little information to effectively find a desired picture out of the collection. These are not unique problems to the new home and office environment, but consumers demand much simpler solutions than those the professional might accept.

Image Metadata

The need for interoperability of personal computers, retail kiosks, Internet services, and imaging appliances imposes new technical requirements on the entire system of digital photography. It will not be enough to recognize the image files and display the pixel values. Digital photography must provide:

- high image quality, through proper interpretation and processing of the pixel values in a picture
- a flexible means to associate non-image information describing the context and content of a picture with the pixel values
- accurate execution of the customer's intentions for processing, printing, and delivering prints made from a picture.

We will expand on each of these requirements below. They arise, generally, from the customers' expectations that their photographic activities will produce excellent pictures with little skill or effort, and that the system will rarely ask them unnecessary or obscure questions. Digital

photography, at home or retail or over the Internet, with all its advanced capabilities, should be as easy as traditional photography: take your snapshots, drop off the removable media, and great pictures come back.

A common thread in these requirements is a means to link a collection of information, called image metadata, with each picture [1]. Metadata informs the later stages of image utilization with facts known earlier: that these pixels came from a particular digital camera; that the scene was recorded outdoors at the Grand Canyon; that picture number 23 is so wonderful that six prints should be made from it. Metadata is equally applicable to digital image capture and silver halide photography. For example, the Advanced Photo System records information about the image capture event on a magnetic layer integral with the film, and uses that information to improve print quality.

Consumers should have little explicit awareness of the metadata. They will contribute some of it, and that information will "magically" be at hand just when it is needed. Otherwise, things just work right.

High image quality

Three examples of metadata that can improve image quality are:

1. The manufacturer and model number of the digital camera that produced this picture. The variation in camera outputs has been reduced somewhat by the growing use of sRGB [2] as an output space. However, each camera has certain characteristics (e.g., camera resolution) and limitations (e.g., noise level), and applies distinct processing steps to its images. If the camera model is known and has been characterized, a printing system can provide optimal processing for that data.
2. Flags indicating whether the flash fired when the picture was taken and whether a flash return was received from the subject. Wide dynamic-range scenes are difficult to print correctly. Two common scene types with wide dynamic range are brightly backlit subjects and flash on a close subject. These metadata flags help the printing subsystem to apply the correct processing in each case.
3. An ICC color profile [3] describing the precise meaning of the pixel values. Some image capture devices store the image data in a color space unique to that product. This is a simple means of recording the scene information with the least degradation or modification. In these cases, an ICC profile can be associated with the file as metadata and used to interpret the pixel data.

Content and context

Content and context metadata is generally used to aid in the selection or retrieval of a picture. This information could be captured automatically by the camera, or recorded "manually" by the photographer, and associated with the picture at a later point, as the following examples make clear:

1. The time and date of image capture. This piece of metadata is very useful and can easily be provided by most image capture devices. The camera system should be designed to ensure that the recorded date is correct, with as little user involvement as possible.
2. The location at which the picture was taken. The means may be as sophisticated as a GPS receiver integral with the camera or as simple as the photographer's notebook, transferred manually to the picture file at a later time.
3. A short sound recording for annotation. This is one of the most flexible means for annotating images.

Order information

Whenever the artistic and emotive steps of photography are separated from the more routine printing and processing steps, some means must be provided to convey the customer's needs from one step to the next. This may be a quite intricate. The customer may have long-standing preferences for her prints, particular needs for this set of pictures, and specific requests for each picture. Furthermore, some of the instructions sent with the order are created implicitly by the order system, with no direct customer input. Some examples of this metadata are:

1. Preferred print aspect ratio. The Advanced Photo System provides three print aspect ratios. The photographer selects the aspect ratio by setting a switch on the camera. The switch setting is recorded with the image on the film and used to choose the correct aspect ratio at the printer.
2. The customer's e-mail address. E-mail addresses do not have the redundancy and robustness of regular mail addresses. Customers who regularly receive their pictures over the Internet should be able to enter their e-mail addresses once and have it associated with their picture files automatically thereafter.
3. Number of prints requested. Digital and APS cameras can provide means to specify the number of prints desired from each picture taken. One common use for this feature is setting it to zero, to eliminate that picture from the order. Another common use is requesting a copy of a group shot for everyone in the scene.

Interoperability

Many components of the imaging chain must work together to deliver the benefits of image metadata to the customer. Information is gathered at various places in the picture's life cycle; it is invariably used at a different place in that cycle. This places two important interoperability constraints on the components of the photographic system. First, the source of the image metadata and the user of the metadata must agree on its meaning. Second, components between the source and the point of use must faithfully pass the data through, even if they do not take advantage of it, or even understand its meaning.

In both traditional photography and digital photography, many manufacturers contribute components to the overall system. Agreement on the meaning of image metadata and its persistence along the imaging chain depends on cross-company (and for digital photography, cross-industry) standards.

This is not a new situation for photography. Consider the 35-mm photographic system. The film and physical cartridge from any film manufacturer fits in all 35-mm cameras. All color-negative films can be developed in one common chemical process. The design of these films and the color papers on which they are printed is specified, so that the scene is properly rendered in the final print. Both the camera and the processing lab use information recorded on the film and the film cartridge for process control. Even the handling of information on the photofinishing envelope has been subject to regional standards. Note that the users of 35-mm photography take these standards and information transfers for granted or are simply unaware of them. The system simply works well. This is as it should be [4].

Layers of interoperability

The systems that handle image metadata are complex and still evolving. As with all information technology, it is valuable to establish an architecture for metadata that puts different aspects of interoperability into different layers. One layer comprises the physical media and device interconnects that carry the information. A second layer covers the file formats and protocols that encapsulate and organize the image metadata. This layer includes important variations in how the metadata is bound to the pictures. The third layer defines the logical content of the metadata.

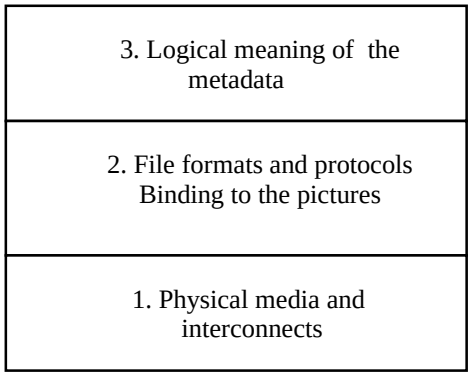

Fig. 1. Image metadata architecture

All three layers are important to successful use of image metadata in digital photography. Each presents challenges for definition and standardization. Among the challenges common to all three of them are the need to deal with legacy systems, the need to provide flexibility and extensibility, and the value of simplicity. The lower layers must operate within the larger infrastructure of the digital industry. They must be considered, therefore, in the context of computer, networking, and television standards.

An important consideration for the second layer is flexibility. We must be able to store metadata in the magnetic tracks on film, on a compact flash card, on a CD, on a hard drive in a computer, or as part of a large database of images. The physical limitations of the medium, requirements imposed by the size of the image collection, and the expectations for adding new functionality to the subsystem determine the optimal design for metadata storage. Two examples will help to clarify this point.

1. If storage space is limited and the device will never be reprogrammed, a rigid and preset binary format may be used. On the other hand, if the metadata is used in an open, rapidly changing context, a more verbose, but self-defining format such as XML [5] is a better choice.
2. With single images or small collections, it is convenient to store the metadata in the same file as the image data, so that a single item on the desktop contains all the information. In larger collections, it is better to store all the metadata together in a standard database structure, with an image thumbnail and a pointer to the full image, so that searching the collection for a specific type of image can be done efficiently.

From one point of view, the third layer is the most critical for the photographic industry to manage. The lower layers are largely inherited from the computer industry. It is relatively easy to write format conversion software if the items being converted are logically identical. For example, translation of one representation of date and time to another representation of date and time is straightforward. However, if one input source stores the date on which the image was captured and another input source stores, in the same metadata item, the date on which the prints were made, many uses of the date item are severely compromised.

Example systems

In order to illustrate how metadata can be generated and encoded, we will describe two examples of digital photography systems. The first uses direct digital capture with an electronic still camera, and the second uses an APS film camera followed by film scanning to create a Kodak picture CD.

Digital capture

The Kodak Digital Science DC260 zoomcamera can be used with a commercially available GPS unit, as shown in Figure 2. The camera includes an autofocus 3X zoom lens with a focus range of 12" to infinity. The camera body has a microphone and speaker, in order to record audio to accompany the photos. The DC260 camera includes an orientation sensor to automatically rotate images taken in the portrait orientation so that they will be "right side up" when displayed on the camera's LCD or the user's computer screen or TV. Like most digital cameras, it also includes a real-time clock. The camera supports Digita scripts that enable the camera to support many additional features, such as GPS and print ordering.

The Kodak DC260 camera stores 1536 x 1024 pixel JPEG compressed digital images on a CompactFlash format picture card. The camera supports both the Exif 2.1 [6] and the Flashpix [7] image formats. Figure 3 shows the organization of the Exif format. The Exif format complies with the JPEG standard [8], so that it is backward compatible with most imaging applications.

Fig. 2. Photo of DC260 camera with GPS receiver attached

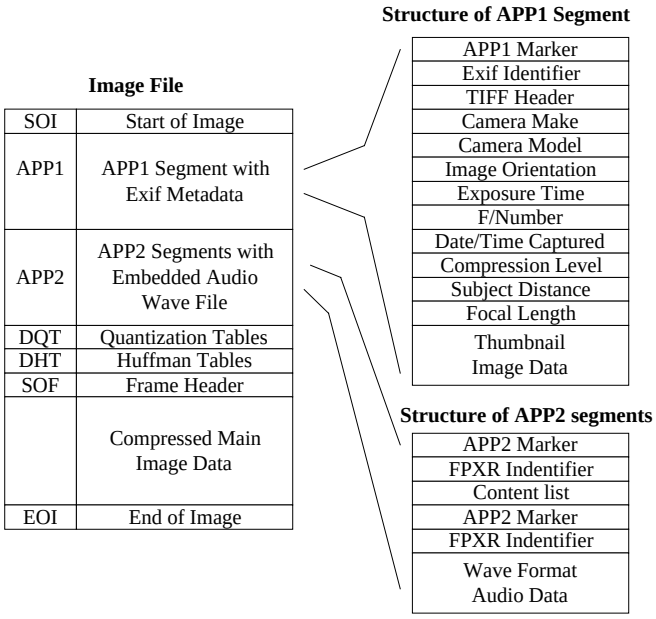

Fig. 3. DC260 camera image format organization

The file begins with the normal JPEG start-of-image marker. This is followed by an APP1 application segment as defined in the Exif specification. The APP1 segment includes an Exif ID string followed by a TIFF file, beginning with the standard TIFF header. This is followed by TIFF tag values that store extensive metadata. For example, the metadata includes make and model tags that define the image source as a Kodak DC260 camera. They also include TIFF tags that define the orientation of the image, provided by the camera's orientation sensor, and the exposure time and f/number, provided by the camera's exposure control subsytem. The real-time clock provides the date and time, and the compression quality switch setting defines the compression level. The lens control subsystem provides the focal length and subject (focus) distance values. The largest part of the APP1 segment is a TIFF IFD (Image file directory) that stores the thumbnail image data.

The APP1 segment is followed by an APP2 segment, which is used to store an embedded wave file using the Flashpix audio extension [9]. Because the JPEG standard requires that an application segment not exceed 64Kbytes, several APP2 segments may be used to store a long audio recording. Embedding the audio within the image file ensures that it will be transferred along with the image when the image is downloaded or copied. However, the audio (and other metadata) will be lost if the image is edited and rewritten by an application that is not aware of the APP1 and APP2 information.

The DC260 camera includes Digita script capability. Scripts have been developed to record GPS information within the DC260 camera image file, and to create a DPOF (Digital Print Order Format) file. DPOF is a recently introduced "digital order envelope" used to specify print making decisions. It is a simple ASCII file created by a digital camera as the images are reviewed. It is normally stored on the removable media along with the image files. It describes the type of prints (standard or index) and the number of copies, and allows for optional date/time or title overlays and image cropping.

APS film scanned to Picture CD

An alternate way of obtaining digital images is to use a film camera, such as the KODAK ADVANTIX 5800MRX Zoom Camera Outfit. This camera includes a 5X autofocus lens, a flip-up auto-fill flash, mid-roll change capability, and a wireless remote control for self-portraits. Like other APS cameras, it includes the ability to obtain prints using 3 different print aspect ratios (PAR): 4" x 6" classic prints, 4" x 7" HDTV aspect ratio prints, and 4" x 11" panoramic prints. In addition, the 5800MRX camera allows the user to select custom or pre-programmed titles that are printed on the back of the prints.

When the APS film is scanned and written to a Kodak picture CD, the images are stored using Kodak's JPEG file format shown in Figure 4. The picture CD also includes application software that allows the images to viewed, printed, e-mailed, and converted to the Flashpix format.

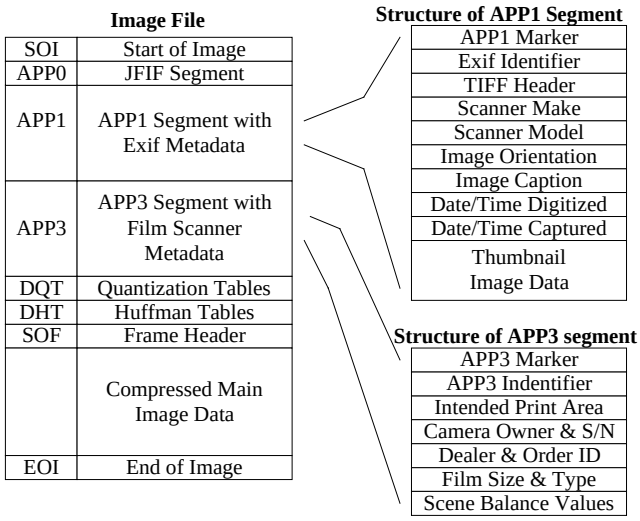

Fig. 4. Picture CD image format organization.

The file begins with the normal JPEG start of image marker. The format includes an 18 byte JFIF APP0 segment. This segment is included to remove an interoperability limitation of a few popular software applications, which expect to find a JFIF APP0 segment immediately following the SOI marker. The APP0 segment is followed by an APP1 application segment as defined in the Exif specification. The APP1 segment has the same structure as described earlier for the DC260 camera. However, a slightly different set of metadata is stored in the file. For example, the metadata includes the make and model of the scanner that created the file. It also includes the image orientation determined during scanning, and the image caption (Christmas, Baby's first step) that may have been selected when the picture was taken with the APS camera. The APP1 segment stores both the date and time the picture was taken, provided by the real-time clock in the APS camera, and the date and time the image was scanned. The APP1 segment also includes a TIFF IFD with the thumbnail image data.

The Exif image format was designed for digital camera use, and therefore does not include certain metadata items that are critical when providing image files from scanned APS film. Therefore, an APP3 segment was defined by Kodak for use with picture CD. Like APP1, the APP3 segment begins with an ID string, followed by a TIFF file starting with the standard TIFF header.

The most important APP3 metadata is the intended print area. This provides the print aspect ratio selected by the user when the APS picture was taken. When the film is scanned, the entire image area is digitized and stored as the compressed main image data. When the image is viewed, however, the default display should normally show only the intended print area of the image, to match the prints provided to the user. As a special feature, the user might be allowed to recompose the image to provide a different print aspect ratio.

The APP3 segment can also store the camera owner and serial number, the dealer and order ID number, the film size and type (color negative, reversal, etc.), and the scene balance settings used when the exposure and color balance of the image were adjusted when the image was scanned.

Metadata persistence

The metadata provided by the digital and APS film cameras can be used to provide improved ease of use, better image quality, and new system features, as we described earlier. In order to enable these simplified, improved systems, however, the metadata must remain bound with the image through the various stages of the imaging system. This is relatively simple when the system consists of only a digital camera and a home appliance printer made by the same manufacturer. However, for more open systems, metadata persistence is both more difficult and more important, since the metadata is potentially much more valuable.

Conclusion

Digital technology has provided more than just a new way to capture images. It has changed every element of the imaging chain. One new opportunity that spans the imaging chain is the ability to associate image metadata with the bare picture information. The metadata must be captured and used consistently. It must also be transported with the picture persistently. Common usage of metadata will improve the user experience and enable digital photography to deliver new benefits to its customers.

References

1. Our use of the term "metadata" differs from that in formal databases. Image metadata is information other than pixel values that is bound or linked to the picture.
2. M. Anderson et al., "Proposal for a Standard Default Color Space for the Internet-sRGB", *Proceedings of the Fourth Color Imaging Conference*, (1996).
3. ICC Profile Format Specification, International Color Consortium, Version 3.2, (1995).
4. The standards of 35-mm photography are a mixture of formal standards, informal agreements, and universal industry practice, developed over a period of time. The Advanced Photo System is another example of a standardized photographic system, developed by a small consortium of companies. The process by which standards for digital photography are created will be some mixture of these two models.
5. A clear introduction to XML may be found at www.arbortext.com.
6. JEIDA, "Exchangeable Image File Format for Digital Still Cameras: Exif , Version 2.1", (1998).
7. J. R. Milch, "A Storage Format for Resolution Independent Imaging Systems", *Proceedings of ISEP*, (1996).
8. G. Wallace, "The JPEG Still-Picture Compression Standard, *Communications of the ACM*, **34** (1991).
9. Flashpix specifications, including the Flashpix audio extension, may be found at www.digitalimaging.org.