John R. Smith
IBM

Metadata Practices for Consumer Photos

Jelena Tešić
IBM T.J. Watson Research Center

The low cost of digital cameras and digital storage devices today, combined with the rapid expansion of broadband Internet connectivity, has given consumers more control over their digital photo collections than ever before. There's virtually no limit to the amount of digital images we can keep and store on our PCs. As a result, the emerging challenge is searching through these digital image collections. With no information about the image content, the only effective search method is browsing through numerous folders to find the right photo.

Digital image metadata plays a crucial role in managing digital image repositories. (See the "Defining Metadata" sidebar for a full definition.) It lets us catalog and maintain large image collections as well as search for and find relevant information. Moreover, describing a digital image with defined metadata schemes lets multiple systems with different platforms and interfaces access and process image metadata.

Metadata's wide use in commercial, academic, and educational domains as well as on the Web has propelled the development of new standards for digital image data schemes. The Japan Electronics and Information Technology Industries Association has proposed the Exchangeable Image File Format (EXIF) as a standard for storing administrative metadata in digital image files during acquisition. The International Press Telecommunications Council (IPTC) has developed a standard for storing descriptive metadata information within digital images. These metadata schemas, as well as other emerging standards, provide a standard format for creating, processing, and exchanging digital image metadata and enable image management, analysis, indexing, and search applications.

Editor's Note

The consumer photo market has seen a tremendous transformation from its traditional camera-based film process to digitally enhanced image files. In the long-promised and highly anticipated convergence of PCs and consumer electronics, consumer digital photos are clearly leading the way. We now routinely snap photos with our camera-equipped cell phones or their more serious 8- or 13-megapixel cousins, upload them to PCs or hosted Web sites, send them to friends and family over the Internet, or beam them around our homes to TVs and other devices.

The next big challenge, however, is to be able to effectively search, organize, and route these photos. That's where metadata comes in, and numerous metadata practices and standards are emerging for consumer photos. In fact, the role of metadata will become so vital as our personal digital photo repositories grow that metadata's value greatly will surpass that of the content. Jelena Tešić's article discusses these efforts and provides insight into the role of metadata and trends toward semantically indexing digital photo contents.

—John R. Smith

EXIF

EXIF is the digital still-camera (DSC) image file format standard that uses controlled vocabulary to represent administrative metadata. It's widely used as an international standard for DSCs as well as in many other fields. This standard supports recent technology advancements by adding new tags to record useful information for new DSC capabilities, printer output processing, and additional GPS information. The goal is for the wide range of DSCs to provide as much capture information as possible to improve information retrieval and image management.

In reality, a DSC can only measure a limited amount of capture information. Hence, if that information has strict values, we might not always obtain suitable information. Therefore, this standard uses somewhat general definitions so that various DSC models can record its capture conditions. Figure 1 (see p. 88) shows some of the EXIF metadata we extracted from a digital image.

1070-986X/05/$20.00 © 2005 IEEE

Published by the IEEE Computer Society

Defining Metadata

The National Information Standards Organization defines *metadata* as structured information that describes, explains, locates, or otherwise makes it easier to retrieve, use, or manage an information resource. The organization emphasizes that metadata is "a key to ensuring that resources will survive and continue to be accessible into the future."[1]

Metadata schemes are sets of metadata elements designed for a specific purpose, such as describing a particular type of information resource. The definition or meaning of the elements themselves is known as the scheme's semantics. The values given to metadata elements are the content. Metadata schemes generally specify names of elements and their semantics and might specify content rules for how to formulate content, representation rules for content, and allowable content values—for example, terms must be used from a specified controlled vocabulary.

Figure A illustrates the significance of image metadata in the creation, access, and distribution of a consumer's photo.

Reference

1. *Understanding Metadata*, Nat'l Information Standards Organization, http://www.niso.org/standards/resources/UnderstandingMetadata.pdf.

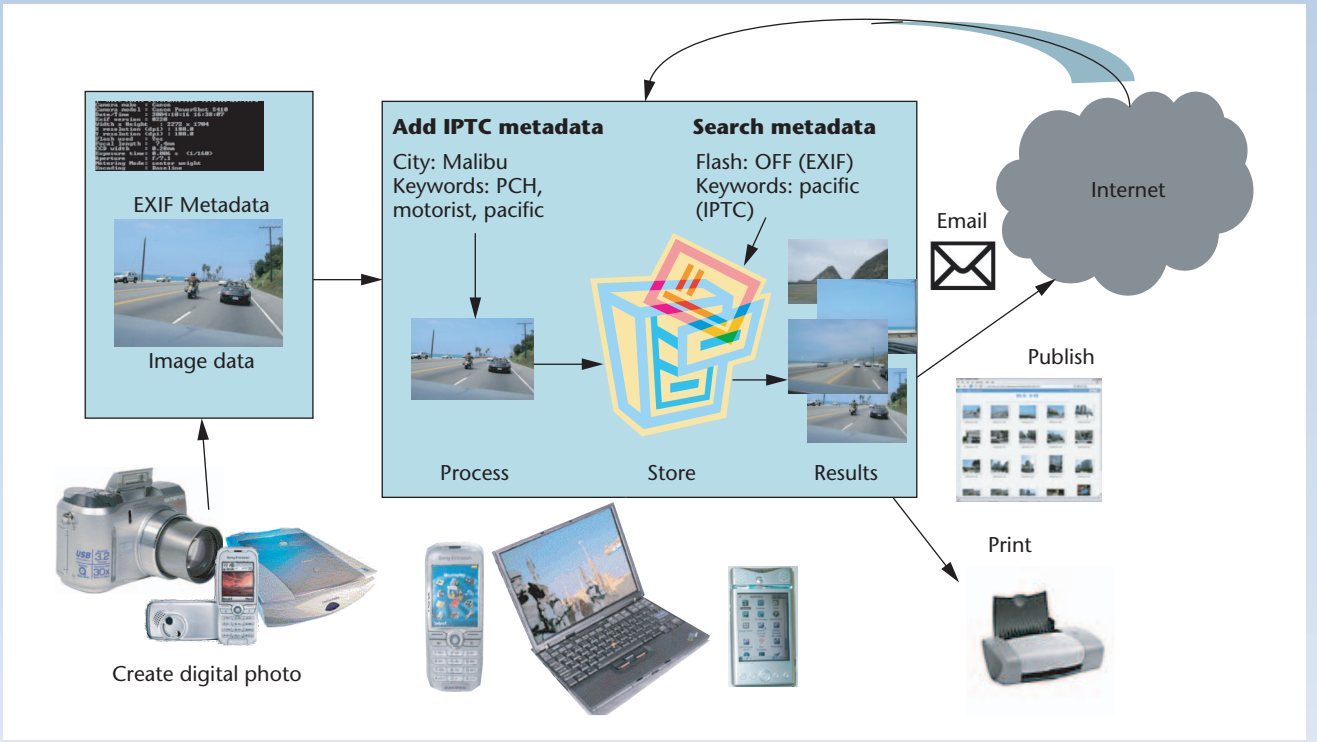

Figure A. Life cycle of digital image—importance of the metadata in creating, storing, indexing, and distributing digital images today.

EXIF administrative metadata

Administrative metadata provides information supplied by the information creator, such as when and how the image was created, digital camera specifications, and other technical information (such as location and lighting conditions). The EXIF file specification defines the method of recording these data in digital image files.[1] This standard allows image data interoperability among imaging devices (cameras, printers, cell phones, and PDAs), commercial software for image processing and cataloging, and specialized image-processing software running on a consumer's computer.

The EXIF file format conforms to existing image file specifications, allowing image file compatibility. Uncompressed files are recorded in TIFF revision 6.0 format,[2] and compressed files are recorded in JPEG (ISO/IEC 10918-1) format. Therefore, commercial image-processing applications can read, view, and process images with embedded EXIF metadata.

The EXIF digital image standard specifies

- the basic structure of digital image data files,
- tags and JPEG marker segments the standard uses, and


```
Camera make   : Canon
Camera model  : Canon PowerShot S410
Date/Time     : 2004:10:16 16:38:07
Exif version  : 0220
Width x Height    : 2272 x 1704
X resolution (dpi) : 180.0
Y resolution (dpi) : 180.0
Flash used    : Yes
Focal length :  7.4mm
CCD width     : 0.28mm
Exposure time: 0.006 s   (1/160)
Aperture      : f/7.1
Metering Mode: center weight
Encoding      : Baseline
```

(b)

Figure 1. (a) Sky above the Corn Field image and (b) partial EXIF data extracted from this digital image.

- how to define and manage format versions.

EXIF file structure

To show what EXIF is capable of recording, we can illustrate the structure of EXIF compressed digital image data files. A compressed EXIF digital image is recorded as a JPEG file, with application marker segments (APP1 and APP2). The metadata consists of JPEG application segments and are determined by application markers, ranging in binary value from 0xFFE0 to 0xFFEF. These application segments are stored before the start of stream (SOS) segment (0xFFED) that contains compressed image data. Figure 2 shows the structure of a compressed EXIF digital image file that conforms to the JPEG standard.

Every JPEG file starts with a start of image (SOI) marker (0xFFD8) with the binary value 0xFFD8 and ends with an end of image (EOI) marker with the binary value xFFD9. The EXIF standard uses APP1 (0xFFE1) and APP2 (0xFFE2) segment markers for JPEG segments to store image metadata. If EXIF metadata are available, they're stored in the APP1 segment. This segment follows immediately after the SOI marker. If multiple APP2s are present, they start immediately after APP1.

Figure 2 also shows an APP1 segment's structure. The EXIF identification 4-byte code (ASCII character EXIF) indicates that the APP1 segment interoperability is in EXIF format. It's followed by 0x00 recorded in 2 bytes, and then EXIF data follows. EXIF uses the TIFF format to store data. (See related literature[2] for more details on the TIFF format and header.)

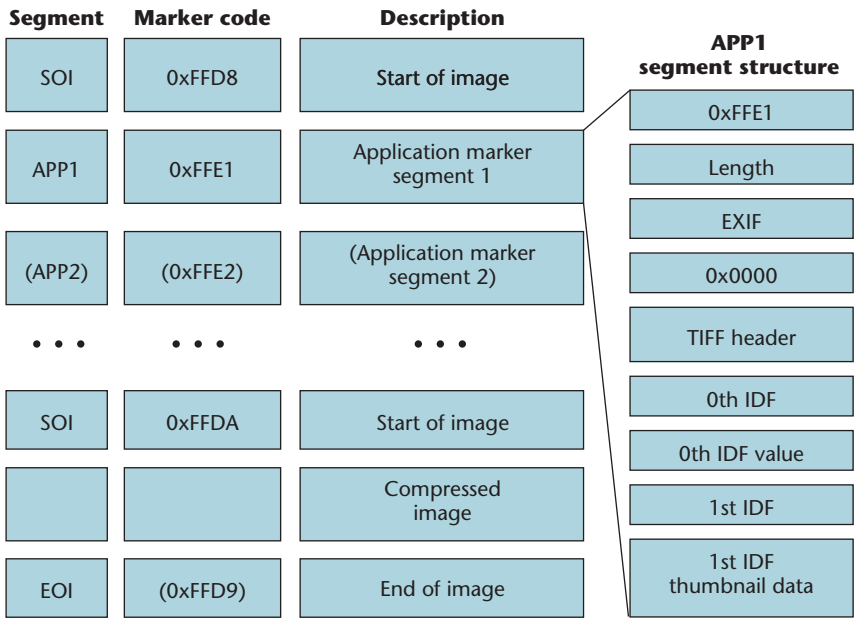

Figure 2. Structure of the compressed EXIF image data format. (EXIF = exchangeable image file format; IFD = image file direction; SOS = start of stream; and EOI = end of image.)

Attribute information is recorded in two image file directories (IFDs) in metadata fields specified in the TIFF Revision 6.0. IFDs consist of a 2-byte count of metadata fields, metadata fields (each a 12-byte field), and a 4-byte offset to the next IFD, in conformance with TIFF Revision 6.0 (see Figure 3). Each of the 12-byte fields consists of four elements:[1] tag (bytes 0-1), type (bytes 2-3), count (bytes 4-7), and value offset (bytes 8-11).

The initial IFD in a file is the 0th IFD, and it contains an image's general metadata information. The next IFD, the 1st IFD, contains a thumbnail image's attribute information.

Specialized metadata informa-

tion is saved in private IFDs within the 0th IFD data value field. Attribute information specific to Exif is recorded in the metadata fields in the EXIF IFD. The 0th IFD contains metadata fields that point to private EXIF and GPS IFDs (see Figure 3). Table 1 gives a partial list of defined EXIF tags and metadata fields. (See the full standard document for a complete list.[1]) As an example, Table 2 gives the meanings of the LightSource metadata field tag 37384 (9208.H). This tag is short with a count 1.

The GPS IFD contains a set of metadata fields used to record GPS information. It's pointed to by the offset, which a GPS IFD private tag value in 0th IFD and (sometimes) in the EXIF IFD indicate. Because of the APP1 segment's limited size, GPS IFD can appear in the APP2 segment. GPS tags include the image's latitude, longitude, altitude, GPS time, and name of the GPS area.

Implementation example

We can demonstrate the EXIF data's value for the consumer with the following example. Eva goes to California every year for vacation and takes thousands of photos with her digital camera, which embeds EXIF information into each image when she takes a photo. Eva dumps all the images on her computer without processing them.

After some time, Eva wants to make a surprise digital collage of photos of Santa Barbara for a friend who's moving there. She remembers taking many pictures in Santa Barbara, and she knows that they are somewhere on her hard drive. Instead of browsing through all the accumulated images (with inane names such as DSC_526.jpg) to find the right ones, Eva can use any of the available image-processing and cataloging applications (both commercial and free) to filter her collection in seconds based on EXIF data.

For example, using embedded GPS information, she can filter the Santa Barbara photos based on EXIF GPS information. Because she wants to make a photo collage only from the day photos, Eva can use the brightness value and Flash EXIF tag information to further filter out the night images from the collection. If she wants to include only photos taken in the last two years, DateTimeOriginal information can help her filter out the old photos. In the end, using just part of the available EXIF information, Eva can retrieve the best photos in her collection in a matter of seconds.

IPTC header

The IPTC and the Newspaper Association of

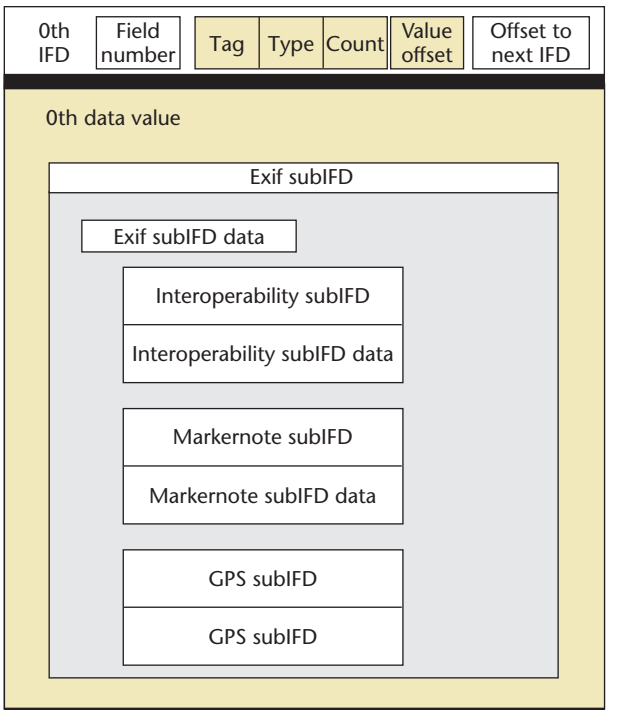

Figure 3. Structure of a IFD segment.

Table 1. Selected EXIF tags.

Tag Field Name	Tag Identification (Hexadecimal)	Type	Count
Make	0x10F	ASCII	Any
Model	0x110	ASCII	Any
Software	0x131	ASCII	Any
PixelXDimension	0xA002	Short or long	1
PixelYDimension	0xA003	Short or long	1
DateTimeOriginal	0x9003	ASCII	20
ExposureTime	0x829A	Rational	1
FNumber	0x829D	Rational	1
BrightnessValue	0x9203	Rational	1
LightSource	0x9208	Short	1
Flash	0x9209	Short	1
FocalLength	0x920A	Rational	1

Table 2. Meaning of the LightSource EXIF metadata fields.

Metadata Field Values	Meaning
1	Daylight
2	Fluorescent
3	Tungsten (incandescent light)
4	Flash
9	Fine weather
10	Cloudy weather
11	Shade

Table 3. Selected IPTC tags.

Tag Field Name	Tag Identification (Hexadecimal)	Number of Characters
Category	0xf	3
Keywords	0x19	64
Byline	0x50	64
City	0x5A	32
Province/State	0x5f	32
Country	0x65	64
Headline	0x69	256
Credit	0x6e	32
Source	0x73	32
Copyright	0x74	128
Caption	0x78	2,000

America (NAA) proposed the Information Interchange Model (IIM) in 1991. IIM[3] is a metadata structure that defines a descriptive metadata scheme for digital image metadata that describes digital image content. The IIM model provides interoperability that embraces all types of data and uses existing formats during transition. The IIM information about the digital image consists of *datasets*, each with its own identifier. Metadata elements of IIM are well known as IPTC headers for digital image files.

Adobe Systems invented their own mechanism for inserting metadata structures into Photoshop, JPEG, and TIFF files but adopted IIM's data structure and several of its metadata elements such as author, caption, keywords, categories, comment, and some free-format fields. Table 3 shows sample IPTC fields.

If EXIF is a DSC image file format standard, the IPTC header standard is a standard for storing and accessing metadata information in digital images because many commercial applications have adopted the Adobe mechanism of inserting and reading the IPTC metadata headers. The IPTC standard lets consumers embed captions, keywords, and text descriptions into their digital images.

In principle, the IPTC header is saved within a digital image the same way as the EXIF metadata. We can analyze how Photoshop (version 6 and higher) saves the IPTC metadata into a JPEG image;[4] IPTC metadata is saved within an APP13 JPEG application segment, as Figure 4 illustrates. The segment starts with an APP13 marker (0xFFED), followed by the Photoshop 3.0 0x00 identifier and the list of Photoshop segments. These segments have a fixed structure (see Figure 4).

Photoshop always uses its signature segment marker, a 4-byte-long 8BIM string , followed by a 2-byte unique metadata identifier—for example, 0x0404 for an IPTC data block. Zero padding is inserted to make the block size even. The 4-byte IPTC data segment then follows. An IPTC data segment is a sequence of field units. Each field consists of a unique tag and a data field. Table 2 shows sample IPTC fields. A standard tag header is used when the data field size is less than 32,768 bytes; otherwise, an extended tag header is used. The datasets don't need to show up in numerical order according to their tag.

Millions of professional and personal digital photos in Photoshop, JPEG, and TIFF image formats over a range of distribution mechanisms use IPTC headers. The IPTC standard lets users insert the metadata into digital photos based on image

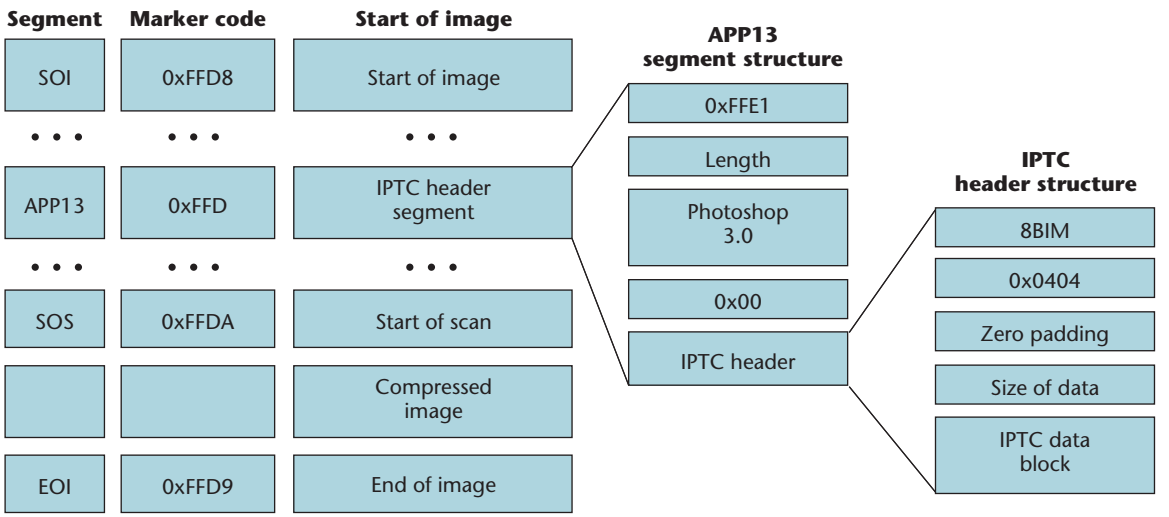

Figure 4. An IPTC segment of JPEG compressed data. (IPTC = International Press and Telecommunications Council.)

Related URLs

The following URLs provide additional information from metadata resources:

- Site dedicated to EXIF and related resources: http://www.exif.org
- History behind the IPTC header: http://www.iptc.org/IIM/
- Updates on the IPTC Core Schema for XMP standard: http://www.iptc.org/IPTC4XMP/
- Overview of MPEG-7 standard: http://www.chiariglione.org/mpeg/standards/mpeg-7/mpeg-7.htm
- The Dublin Core Metadata Initiative (DCMI) site: http://dublincore.org/

content, thus enabling a more meaningful content-based image search.

Figure 4 also shows the value of using IPTC metadata in consumer photo applications. Consumers can enter semantically meaningful photo descriptions and keywords, such as place (Malibu), region description (Pacific), or personal descriptions (motorist). IPTC fields can also specify high-level semantics that include names and an event description, such as "a surprise birthday party for Nikola;" people's relationships in a photograph, such as "lab mates from UCSB at Shawn's wedding;" or a full photo description, such as "Mom, Dad, Uncle Peng, and his girlfriend at the Empire State Building, summer 2005." Searching these files is then reduced to the capability of software applications to parse IPTC schema, extract fields and their meanings, and use them to conduct a search over image datasets.

Efforts in image metadata standardization

Besides EXIF and IPTC, other existing and emerging standards have the potential to further extend personal digital image cataloging. (See the "Related URLs" sidebar for more details.) The Dublin Core Metadata Initiative (DCMI; http://dublincore.org) provides a forum for defining semantics, both for a general description core and subject-specific extensions. DCMI has led the development of structured metadata to support resource discovery. In general, the DCMI vocabulary defines a hierarchy of terms that describe a document's purpose, context, and origin rather than describing the document itself. Several community interest groups (including educational organizations, libraries, government institutions, and scientific researchers) have widely used it as the basis for descriptive systems.

The Digital Imaging Group (DIG35) Initiative of the International Imaging Industry Association (I3A) has also defined a metadata standard for digital images. The DIG35 Specification includes a standard set of metadata for digital images that promotes interoperability and extensibility between various digital imaging devices.

MPEG-7 is an ISO/IEC standard, developed by the Moving Picture Experts Group (MPEG), that provides a rich set of standardized tools to describe multimedia content. MPEG-7 describes audiovisual information regardless of storage, coding, display, transmission, medium, or technology.[5] MPEG-7 uses the XML Schema for content description. This allows MPEG-7 applications to leverage a large body of existing tools and server technology built around the World Wide Web Consortium's XML standards.

The Extensible Metadata Platform (XMP) is an Adobe metadata XML-based schema for storing image metadata. XMP metadata is encoded as XML-formatted text, using the W3C standard Resource Description Framework (RDF). In 2004, a joint effort of the IPTC and Adobe started to work on the IPTC Metadata for XMP standard (called IPTC4XMP) for a smooth and explicit transfer of metadata values from the IPTC headers to the XMP framework.

In April 2005, this effort resulted in the IPTC Core Schema for XMP specification on how to synchronize IPTC header values with XMP.[5] The Core Schema specifies a standardized set of metadata properties to be used within the scope of the XMP framework. The IPTC Core Schema for XMP will improve metadata flexibility and accessibility by expanding the scope of information metadata.

These new XML-based standards enable us to capture metadata such as searchable keywords, author and copyright information, and meaning-

ful descriptions in a format that humans, software, and Internet applications can read. Internet search engines will be able to index digital images based on its embedded metadata, not just the file name.

Also, content management systems will be able to easily integrate keyword searches over digital image collections to identify relevant photos. For example, the same desktop search engines that we use to search document files will be able to use data from XML-based metadata schemes to search through image content using similar principles. These functionalities can extend to mobile devices, including a mobile phone with an internal hard drive or portable media storage devices.

Looking forward

The existing EXIF and IPTC digital image standards provide a standardized metadata schema for describing image content Very soon, the newly developed IPTC4XML standard will allow wide use of digital image content in Internet search and Semantic Web applications. These standards provide interoperability across multiple sources and platforms. An average user will see value in these standards once the photo management programs can extract knowledge from heterogeneous data sources, which reduces the cost of annotation and labeling in an interactive environment.

Integrating calendar, email, image, video, and audio data on users' PCs will let them derive higher-level information and model semantic concepts that support more intuitive search over photo collections. For example, a "sunset in Hawaii" search will return all the images that contain the automatically detected concept "sunset" and that have a Hawaii tag or GPS information points to a location in Hawaii. Also, an "anniversary" tag could point out an exact calendar date, letting a system retrieve only the photos taken on that date if other abstract metadata information doesn't exist. **MM**

References

1. Exif Version 2.2 Digital Still Camera Image File Format Standard (Exif) Version 2.2, Japan Electronics and Information Technology Industries Assoc., http://www.exif.org/Exif2-2.pdf, 2002.
2. TIFF Revision 6.0, Aldus Developers Desk, http://www.scarse.org/docs/tiff/tiff6.pdf, 1992.
3. IPTC–NAA Information Interchange Model Version 4, Int'l Press Telecomm. Council, http://www.iptc.org/IIM/, 1997.
4. B.S. Manjunath, P. Salembier, and T. Sikora, eds., *Introduction to MPEG–7: Multimedia Content Description Interface*, John Wiley & Sons, 2002.
5. IPTC Core Schema for XMP Version 1.0, Int'l Press Telecomm. Council, http://www.iptc.org/IPTC4XMP/, 2005.

Readers may contact Jelena Tešić at the IBM T.J. Watson Research Center, 19 Skyline Dr., New York, NY 10532; jtesic@us.ibm.com.

Contact Standards editor John R. Smith at jsmith@us.ibm.com.