# Portable Image Archiving: Annotation, Search and Data Retrieval

*Vladimir Misic, James M. Kang, Ankur M. Teredesai, and Jordan Sissel*
*CS Department, Rochester Institute of Technology*
*Rochester, New York, USA*

## Abstract

JPEG 2000 is an emerging image compression standard that offers not only a competitive image compression algorithm, but the complex weave of its file format opens the possibilities for direct data content management - which were previously unthinkable with the use of other common image file formats like JPEG, GIF, PNG, etc.

In this paper, we discuss the comparative advantages of JPEG 2000 over other commonly used image formats, in metadata-oriented tasks such as image annotation, metadata search, and image and data archival. Also, we present the original set of software tools necessary to create and maintain a JPEG 2000 – based image database.

## Introduction

To answer the ever-increasing requirements --for higher performance and new features-- of multimedia technologies, a new standard for image compression has been developed: JPEG 2000. It is not only designed to provide rate-distortion and subjective image quality performance superior to existing standards, but also to provide the features and functionality unattainable by current standards.

Before proceeding to the technical aspects of the JPEG 2000 file format, one should understand that this compression standard was designed with certain functionality in mind, and with clearly defined set of features that it should posses. The most prominent of these required features are described as:

**Superior Low Bit-Rate Performance**

JPEG 2000 should offer significantly better performance (image quality) than the current standards, especially at the low-bit rates (high compression). Some of the applications that rely on this feature are network and wireless image transmission, and remote sensing.

**Lossless and Lossy Compression in the Same Codestream**

The possibility of the lossless coding is of vital interest for some applications such as image archiving and medical imaging. JPEG 2000 provides both lossy and lossless versions of the same image. They are both easily accessible in the same codestream. This eliminates the need for multiple downloads and simplifies data management.

**Progressive Data Transmission**

Progressive transmission allows images to be reconstructed with increasing accuracy, as the transmission progresses. This feature allows the reconstruction of images with different quality (defined by resolution or accuracy of the pixel intensity), and depending on the target device (printer, camera, any device attached to the network, etc.).

**Region-of-Interest (ROI) Coding**

The ROI is part of the image that is more important than the rest of the image. This feature enables the more important part of the image to be transmitted before the rest of the image. Result is the better image quality and image rendering of the ROI, for the same overall bit-rate.

**Random Codestream Access**

Due to this feature, the ROI can be dynamically accessed and that part of the image can be decompressed with less distortion than the rest of the image.

**Content-Based Image Description**

Metadata information can be embedded in the compressed image codestream, so it helps image search, indexing and archiving.

**Continuous-Tone and Bi-Level Compression**

Most of the existing compression schemes cannot be used at all, or cannot be used with the same efficiency, for the compression of continuous-tone and bi-level images. JPEG 2000 is designed to perform reasonably well on any type of images.

**Robustness to Transmission Errors**

The compressed codestream is designed to be highly resilient to the bit-errors in transmission.

The JPEG 2000 file format has been engineered around these features, with the idea to maximize support for the image compression modalities the standard offers. In this paper we investigate the potential of this file format for digital image archiving tasks.

## JPEG 2000 File Format

Unlike the JPEG, JPEG 2000 standard precisely defines not only the compression algorithm, but also a variety of file formats depending on intended use. The JPEG 2000 file format is a foundation for storing application specific data in association with a JPEG 2000 codestream (compressed data).

The basic JPEG 2000 file format (JP2) defines the organization and content of a minimal set of information that accompanies compressed data. The JP2 file format provides a method by which applications can interchange image files in such a way that all conforming readers can properly interpret and display the image.

Conceptually, JPEG 2000 conforming file encapsulates the codestream and additional information about that codestream. The building block of the JPEG 2000 file format is called a *box* (Fig. 1), and all information contained within the file is encapsulated in boxes. The box structure is:

- LBox: Box Length. This field specifies the length of the box, stored as a 4-byte big endian unsigned integer.
- TBox: Box Type. This field specifies the type of information found in the DBox field. The value of this field is encoded as a 4-byte big endian unsigned integer.
- XLBox: Box Extended Length. This field specifies the actual length of the box if the value of the LBox field is one. This field is stored as an 8-byte big endian unsigned integer.
- DBox: Box Contents. This field contains the actual information contained within this box. The codestream (i.e. compressed image) is in the DBox field of appropriate box (Contiguous Codestream Box).

The standard defines several types of boxes; the definition of each specific box type defines the kinds of information that may be found within a box of that type. A JP2 file represents a collection of boxes. Some of those boxes are independent, and some of those boxes may contain other boxes. The binary structure of a file is a contiguous sequence of boxes. This structure enables a JPEG 2000 vendor to add arbitrary content to the file without jeopardizing the validity of the file.


*Figure 1. Structure of the JPE 2000 box*

Because all information is encapsulated in boxes, and all boxes have types, the format provides a simple mechanism for a reader application to extract relevant information, while ignoring any box that contains information that is not understood by that particular reader. In this way, new boxes containing metadata can be created. In addition, any new box added to a JP2 file should not change the visual appearance of the image.

Of particular interest to us are XML and UUID boxes. The XML box contains vendor specific information in XML form, and the UUID box contains any other information that cannot be described and is not contained within any other box. Thus, any metadata associated with the image should be contained in these two boxes. However, since there is no clear "standard" format for XML metadata, and since our project is just "proof of the concept", we decided to organize associated metadata in a ndor (RIT) specific box.

## JPEG 2000 Image Annotation

Unlike other common image compression formats (TIFF, JPEG, GIF, etc.) JPEG 2000 enables "native" association between data and image. JPEG 2000 enables definition of the ROI that is downloaded first / with better quality than the rest of the image. We are further developing this concept and attaching metadata to the ROI.

In our data management system implementation we have identified few principle types of annotation:

- Expert annotation: annotation requires human expert to input relevant information into system.
- Automatic annotation: based on the image features: colors, textures, edges, etc.
- Automatic annotation: based on the semantic properties of the image.

We limited ourselves to expert-based annotation, considering automatic annotation to be a straightforward extension of the existing model.

Also, in the first phase of the project we limited ourselves to textual annotation only, with the idea to later expand annotation to the voice form, as well as to enable voice-to-text and text-to-voice conversions.

## Annotation, Search and Data Retrieval

Our data management system can be modeled as in Fig. 2 and Fig. 3. In the annotation phase (Fig. 2), the user is responsible for image annotation and image upload to the server. The server application accepts the image and extracts metadata into a database.

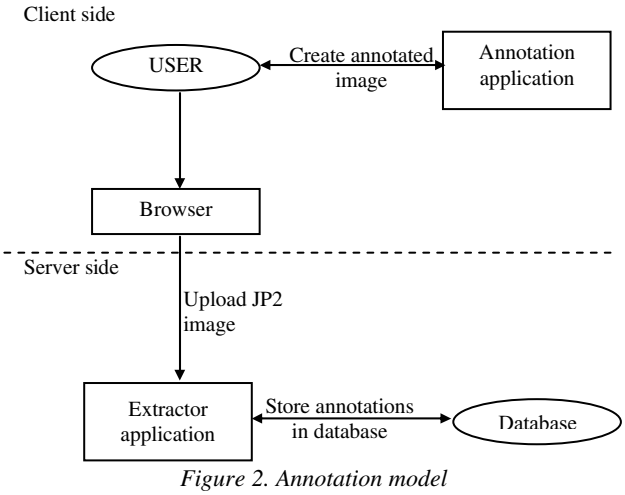


*Figure 2. Annotation model*

In the search and data retrieval phase (Fig. 3), user queries the database that uses extracted metadata to match not only image but also the specific region of the image to the user's query. Then, a plug-in is used to accept the small portion of the JPEG 2000 image codestream and to create preview of the search result. If the user wants so, he or she can decide to proceed with download and improve the quality of the resulting image.

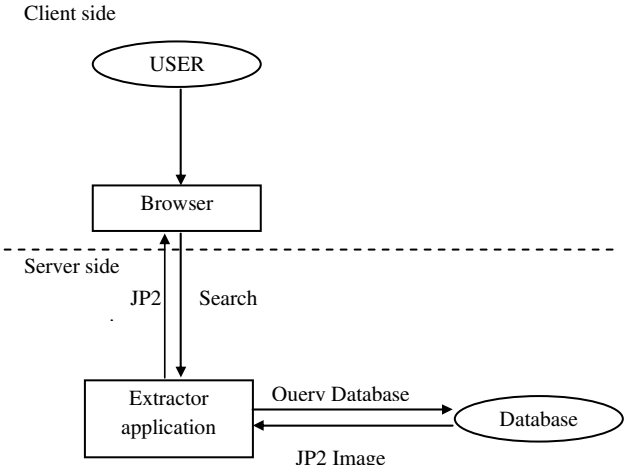


*Figure 3. Data retrieval model*

**Annotation**

In the annotation phase we used GUI based on a modified Kakadu® implementation of JPEG 2000 coding standard. The interface is simple and intuitive. A user selects a rectangular region of the image he or she wants to annotate and a dialog box opens, as demonstrated in Fig 4.

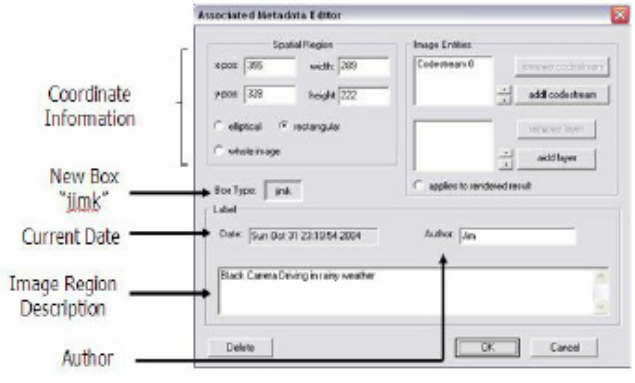


*Figure 4. Annotation dialogue*

A user can annotate as many regions as he or she wants. The information a user can record are: region coordinates (automatic, based on region selection), date (current date by default), information about annotation author, and region description in textual form.

Once in the system, annotation data are recorded in the XML form (Fig. 5.).

```
<keywords>Blue Jaguar, XJ8</keywords>
<date>Sun Oct 3123:19:30</date.
<author>Vladimir</author>
<xcoord>100</xcoord>
<ycoord>123</ycoord>
<width>200</width>
<height>150</height>
```

*Figure 5. Example of XML annotation supported by our system.*

An alternative way of annotation is a use of our on-line image conversion service, where a user can upload any image (in any format) and convert it to JPEG 2000 on the server side, with simultaneous (optional) annotation definitions, associated with the specific region. Drawback of service-based annotation and image conversion is that a user cannot interactively define the ROI (using the GUI), but has to figure out its size and coordinates in a separate process (typically, using some image-handling application). However, a user can specify not only ROI size and ROI position, but also can come up with the name for the box containing the embedded information in the XML form.

**Data Extraction and Image Storage**

Once the image is annotated, it can be saved locally and uploaded to the database using any browser and provided portal, or it can be uploaded to the server directly from the annotation application. In both cases, image is stored in the server's file system and not in the database itself. However, the PHP based portal and Java application are used to parse the incoming image, extract XML data, and store it in the database.

The extractor application recognizes generic XML boxes as well as proprietary "jimk", "mks" and "rit" boxes as the metadata carriers, and extracts XML information content into the database.

Any changes in the supported XML scheme (tags) can be easily mirrored on the database scheme, without jeopardizing current database content.

**Search and Image Retrieval**

Search application is written in the form of web portal and implemented in PHP programming language. The user performs textual search based on keyword phrase, which is matched against the extracted image metadata. Search results are then presented in the preview form (thumbnail), and accompanied with the associated information (annotation). Currently, we offer two previews for each hit: complete image preview, and the ROI preview (Fig. 5) extracted from the original JPX (extended basic JPEG 2000 file format) images stored at the server. However, on the server side, images are accessed using standard web-based server, instead of JPEG 2000 specialized JPIP server - due to the lack of the appropriate server module. There is no

requirement for separate storing of thumbnail images such as there would be with any other image file format. If the user "likes" what he or she sees in the thumbnail (successful search), he or she can proceed with the image download, without the need to restart downloading from the beginning. Also, if the user is not sure about the search result, he or she can incrementally improve the previewed image quality.

This feature alone is often enough to provide huge comparative advantage for many archiving systems. The elimination of the multiple versions of the same image simplifies the metadata storage scheme (database scheme) and information management. This comes as direct consequence of having only one copy of annotated data, embedded (and not even related to) in the image it describes, and one copy of the image on the server.

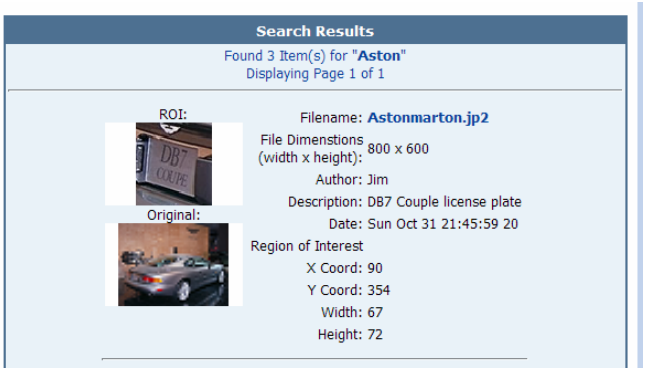


*Figure 5. Example of search result: ROI and image previews*

## JPEG 2000 vs. Other Image Format In Archiving Tasks

Most of the existing digital archives and libraries rely on the established practices that use well-known image formats such as TIFF and JPEG to store and search archive content. TIFF supports accurate source reproduction (a.k.a. lossless compression), extended bit-depth (beyond 24bpp) and capability of embedding metadata into image. JPEG enables better image compression (a.k.a. lossy compression) in the standardized form that is supported by all operating systems, and virtually all image viewing and image manipulating applications.

JPEG 2000 also supports bit-depths beyond 24bpp for accurate color and multi-spectral reproduction. It also supports embedded metadata. However, JPEG 2000 does not have limitation that metadata should be related to the whole image, or any limitation on the type of the metadata. Users are free to embed any information they want, and that does not affect the image reproduction. JPEG 2000 offers superior compression not only to TIFF, but also to JPEG. Moreover, JPEG 2000 offers both lossy and lossless compressions in the same codestream. That means that JPEG 2000 eliminates need for multiple image versions and formats to store the same image at different quality levels, different compression levels and different resolutions.

JPEG 2000 offers random codestream access; meaning that the original (typically lossless) image can be reorganized "on the fly" and sent to the final user according to users recognized needs (e.g. preview, display or printing, etc.) and capabilities (e.g. bendwidth, or physical memory limitations). Metadata embedded into the image is not limited to textual information only, but can be extended to any type of digital recording (audio, video, or even another document – e.g. embedded pdf).

Another advantage of JPEG 2000 based archival is that image annotation can be done concurrently, using distributed computer and human resources. Annotation process is unified through the share of the metadata boxes between annotators: there is no need to move original images around the system. This system capability directly minimizes data management costs and simplifies the database management process.

In a JPEG 2000 database, the content is inherently portable, since all searchable image information is contained within the image file, and is not externally linked with the image. There are numerous advantages of this approach. Once the user downloads the image, he or she also downloads all the associated metadata. As a consequence, the whole database content can be easily ported to another location by the simple mean of copying image files. Any change in the database content is also reflected in (or is the result of) the altered image file. Changes in database scheme do not affect relation between data and image, and can be done dynamically (add/drop columns and tables). Database migration (to another DBMS system) or recovery becomes almost trivial, practically equivalent to simple file copying. Thus, database maintenance is extremely simplified, resulting in significantly reduced operational costs.

Compared to standard archival practices, management of digital rights is dramatically improved since a digital curator can easily limit user access to different spatial regions of the image, or to different quality (and resolution) parts of the image codestream. Also, associated metadata can, but does not have to, be the part of the delivered image, providing the base for the commercialization of the "added value" services (e.g. company which is not copyright owner, might have advanced system for image categorization based on specific image features). However, these issues are out of the scope of this paper and will not be discussed in detail.

## Conclusion

In this paper, we discussed potentials of JPEG 2000 file format in archiving practices. We demonstrated use of user-defined boxes to hold region-oriented metadata in the XML form, and to support description-based image search.

Future research will be oriented toward systems for automatic image annotation using AI (artificial intelligence) methods, voice-based image annotations and textual information retrieval from such annotations, as well as toward autonomous agents for distributed image annotation.

## References


1. Vladimir Misic, James M. Kang, Jordan Sissel, Ankur M. Teredesai, JPEG 2000 – New Paradigm In Image Archiving, *1st CCIS*, Rochester NY, 2005
2. Vladimir Misic, James M. Kang, Jordan Sissel, JPEG 2000: Web Services for Image Conversion and Document Database Search, JPEG 2000 Symposium, University of Connecticut, Storrs CT, (2004).
3. Vladimir Misic, MRC and JPEG2000 file format, Workshop on Teaching With Technology, *CTLTR*, Univ. of Rochester, Rochester NY, (2001).
4. ISO/IEC 15444-1:2004 | ITU-T Rec. T.800, Information technology – JPEG 2000 image coding system: Core coding system
5. ISO/IEC 15444-2:2004 | ITU-T Rec. T.800, Information technology – JPEG 2000 image coding system: Extensions
6. D. Taubman, Kakadu: A foundation for JPEG 2000 applications, The University of New South Wales, Sydney, (2001).


## Biography

**Vladimir Misic** received his BS degree in electrical engineering from the University of Novi Sad, Serbia, and his MS and Ph.D. degrees in electrical and computer engineering from the University of Rochester, in 2003. He is currently a visiting professor at the Computer Science Department, Rochester Institute of Technology, and an adjunct assistant professor the Electrical & Computer Engineering Department of the University of Rochester. His research interests are in digital content management, image processing algorithms, and digital halftoning.

He is a member of the Institute of Electrical and Electronic Engineers and the Imaging Science and Technology Society.

.