Informatics Challenges—Lossy Compression in Medical Imaging

Don Dennison · Kinson Ho

Published online: 3 May 2014
© Society for Imaging Informatics in Medicine 2014

There has been much discussion on the topic of lossy image compression within the medical imaging community, with much of the focus on the potential clinical impact and business benefits. The implications of applying irreversible lossy compression to images as it relates to informatics and management is often not considered thoroughly.

Why Use Lossy Compression?

Common motivations include as follows:

- Less storage required—Smaller files mean reduced capital and operational costs.
- More data available on-line—If a typical cache and archive storage model is used, smaller files mean more studies on-line and fewer retrievals from the archive.
- Faster retrieval—Less data mean faster retrieval from storage (cache and archive).
- Less network bandwidth consumed—Smaller files take up less of your “pipe” and get to the destination faster.
- Faster display—Smaller files typically consume fewer resources in the viewing application, resulting in faster access and rendering compared to an equivalent lossless compressed image.

Also, while lossy compression was traditionally considered for use with prior exams, more and more organizations are considering its use for new exams used in primary diagnosis.

What Do the Standards Say?

First, some Digital Imaging and Communications in Medicine (DICOM) basics. In the DICOM standard,

- A SOP Instance UID is used to uniquely identify an object such as an image
- A SOP Instance UID is a globally unique identifier

The DICOM standard states that if an object (image) is lossy compressed, it must be assigned a new SOP Instance UID and a number of lossy compression-associated attributes (DICOM 2011 Part 3 Section C.7.6.1.1.5). This seemingly small detail has many implications on data and systems that manage this data.

DICOM Query/Retrieve

If a Picture Archiving and Communication System (PACS) or Vendor Neutral Archive (VNA) stores images only as lossy compressed and discards the original, when requesting systems that do not support lossy compressed objects for the particular SOP Class, the PACS is required to decompress them to Implicit VR Little Endian (the DIOM default transfer syntax). This results in the following:

- Extra processing on the sending PACS side
- A larger, uncompressed image with any artifacts introduced by lossy compression; meaning that the benefits of lossy compression are not realized.

If the PACS stores images in both original and lossy format, when the requesting system wants the lossy compressed object, it has to explicitly negotiate the lossy compressed transfer syntax and not include the original/lossless

D. Dennison (✉) · K. Ho
205 Fern Cres, Waterloo, ON N2V 2P9, Canada
e-mail: don@dondennison.com

transfer syntax in order to retrieve the lossy compressed version. Otherwise, it is up to the discretion of the PACS server to decide what version of the study will be returned.

However, this is still problematic because there will be two distinct set of objects, one in the original format and another set in lossy format. Therefore, any study level retrieve will return both sets, regardless of what is negotiated in the association. Furthermore, there is no information about the transfer syntax available in the query result. Therefore, for a requesting system to determine which subset of objects in the study are lossy compressed, some additional consideration is necessary within the PACS with regard to how the lossy compressed objects are being organized. One method is to generate the lossy compressed objects in a separate series with its distinct series description indicating that this series contains the equivalent set of objects in lossy compressed format. Then, if the requesting system issues a series level query, it can identify which series contains the lossy compressed objects and only retrieve this series rather than the full study.

Given that different modalities can organize their acquired images using various different methods and post-acquisition activities (e.g., Computer Aided Detection (CAD), user creation of markup captured in Grayscale Softcopy Presentation State (GSPS) objects) can result in new objects being created, the effectiveness of using the series level retrieve method mentioned above highly depends on the sophistication of the requesting system.

Managing Both Lossless and Lossy Version of Images Within a PACS

PACS manages images internally in different ways, depending on the design. For those that manage the original DICOM Part 10 objects, versus a proprietary format, it is typical that the PACS will manage each object as a unique record, defined by the SOP Instance UID. This means that if two copies of the same image with the same Study UID and Series UID, but different SOP Instance UIDs, the PACS will treat them as two images within the same study and series (Fig. 1).

According to DICOM, all objects within the same series should come from the same device. This is because there are several attributes directly associated to the manufacturer, device, and software version that generates the objects. So, if lossy versions of objects are created afterwards, for example by a PACS or similar system, these should be stored in a separate series, defined by a new unique Series Instance UID.

In addition to the informatics requirements mentioned above, DICOM also specifies a number of optional DICOM attributes that can be used to provide additional contextual information that may be valuable to consuming devices when accessing lossy versions of images.

A strategy that involves maintaining both lossy and lossless compressed objects in different systems needs to consider that if both objects are stored to one system (by any means, including a routing, retrieval, import from portable media), it is likely that the study will simply have twice as many objects. The PACS server or viewer may have logic to prevent or manage this, but it would need to be confirmed.

If a PACS does have the capability to manage lossy and lossless compressed versions of the same image, this may fail to achieve the desired benefit or reducing storage costs. It also requires that PACS have some logic to return the desired version (lossy or lossless compressed, or both) to external requesting systems—normally, the requesting system will simply ask for the study. Returning both may cause issues in the requesting systems and will certainly consume extra network bandwidth.

Storing both versions (lossy and lossless compressed) of the study objects will also consume extra database resources as each object (header data, storage location data, etc.) will be registered in the database. This will result in storage for the database tables being consumed more rapidly and may have an impact of the performance of the database (if not managed effectively).

New Objects, Same Study: Exception Workflows

Having a new object (defined by a different SOP Instance UID) for a study that is already in PACS may trigger some exception workflow. For example, the PACS may interpret the new object as a new image that has not been reviewed. If the study has already been read, the PACS may trigger some workflow to have the radiologist review the new images to either confirm their prior diagnosis or create an addendum. In this case, they would see what likely appears to be a second set of duplicate images, likely resulting in user confusion and frustration.

Evidence Documents

One of the key implications of changing the SOP Instance UID, as required by DICOM, is that the SOP Instance UID of an object is often referenced in other objects; for example, DICOM Key Object Selection (KOS), GSPS and Structured Report (SR), etc. This means that if the original object's SOP Instance UID is changed, any objects referencing these objects also need to be changed to reference the new SOP Instance UID.

Enhanced Multi-frame SOP Class Objects

As the new Enhanced SOP Class objects (for CT, MR, XA) allow for large studies to be organized into multiple multi-

Fig. 1 Objects with the same Study Instance UID but different SOP Instance UID will be organized into the same study within the PACS

frame objects, called concatenations, care must be taken to ensure that lossy and lossless versions of the image frames are not mixed within a study.

IHE XDS-I

In enterprises that are sharing images using the Integrating the Healthcare Enterprise (IHE) cross-enterprise document sharing for imaging (XDS-I) integration profile, there are some specific implications that impact the XDS systems.

As the IHE XDS-I manifest references individual objects, having a separate set of lossy compressed objects, requires additional consideration. If it is desirable to maintain only one manifest for any study, then the Imaging Document Source will need to publish a new replacement manifest (deprecating the original manifest) if objects are lossy compressed and SOP Instance UIDs are changed after

the study is already registered in the XDS infrastructure. To do this properly, the XDS-I Imaging Document Source needs to support the Replace Document option in order to make sure the new manifest with references to the original instances are properly deprecated. Special consideration is necessary with regard to the time of registration to XDS. If lossy compression only applies to studies after they have been reported but the registration of studies to XDS-I happened once the study is successfully received, this implies that every manifest will have to be deprecated at least once. This incurs more work done and resources consumed by the Imaging Document Source, Document Repository, and Document Registry.

On the other hand, if it is desirable to have the original images and lossy compressed versions independently discoverable, then the Imaging Document Source will need to publish a separate manifest for the lossy compressed objects. Ideally, this new manifest is explicitly associated to the

Fig. 2 Progressive wavelet involves sending a portion of the image file, which can then be used to display a smaller resolution version of the image

Table 1 Comparison of methods used to handle lossy compressed images

Method	Storage cost savings	Network cost savings	Reprocess study instance object UIDs and references in PACS	Client display	Lossless images available for primary diagnosis
Create lossy images at modality	Yes	Yes	No	Can view images in browser with WADO-URI renditions	No
Convert lossy images from lossless images in PACS	Yes	Yes	Yes	Can view images in browser with WADO-URI renditions	No
Use JPEG2000 images	Minimal	When lossy images requested	No	JPIP client required; cannot view images in browser	Yes
Create lossy images on demand in PACS	No	Yes	No	Can view images in browser with WADO-URI renditions	Yes

original manifest as a transformation. This transformation relationship helps the Imaging Document Consumer to identify the purpose of the new manifest.

Post-processing

While research has been done on the clinical impact of lossy compression on primary diagnosis, there is little research published on the impact on clinical applications or post-processing software (e.g., 3D reconstruction, CAD, etc.). It may be that lossy compressed images may not provide the same signal-to-noise ratio due to the artifacts on the images introduced by the lossy compression.

Wait, Doesn't Wavelet Image Compression, Like JPEG200, Save the Day?

Wavelet compression formats like JPEG2000, which is included in the DICOM standard as a valid Transfer Syntax, attempts to solve the “dual personality” challenge by supporting both lossless and lossy compressed data in a single

format. Although it is great in theory, there are challenges with supporting JPEG2000 compressed objects in practice:

- Many modalities do not support the creation of JPEG2000 compressed objects, resulting in a necessary conversion upon arrival in PACS.
- Most PACS do not support JPEG2000 compressed objects, resulting in a conversion from JPEG2000 into another transfer syntax (lossless LPEG, uncompressed, etc.) upon transmission
- The math involved in the compression and decompression of wavelet-based images, such as JPEG2000, is much more computational intensive, meaning the decompression of a JPEG2000 image of the same file size as JPEG takes longer (generally 2 to 4 times longer).

Progressive Wavelet Display: How it Works

In order to take advantage of a progressive wavelet format like JPEG2000, the client needs to request “just enough” of the file to meet its needs (Fig. 2). This requires a “smart” client using a special protocol called JPIP (JPEG2000 Interactive Protocol).

Fig. 3 Reduced resolution images compressed in lossy JPEG or even a lossless format, like PNG, can be done in near real-time on retrieval

For example, an image viewer on a standard resolution (1 megapixel) monitor displaying four mammography images in a 2×2 layout only needs a fraction of the available pixels to display each image in the layout—until the user zooms in or changes the layout so that the image is larger on the screen, downloading additional pixels will not gain any benefit for the user. The monitor resolution is the limiting factor.

The image viewer has to have information on the resolution and size of the available pixels within the layout in order to make an intelligent request for enough data to render the image within that space. It expresses that request through parameters using a protocol called JPIP. The server has to provide a JPIP service that enables it to return the requested data.

Images compressed in JPEG2000 without using JPIP to retrieve them offer limited benefits. Some image types, particularly larger-resolution ones like mammography, compress a little more efficiently in JPEG2000 than in JPEG, but the storage efficiency gained is minimal overall.

Alternative Approaches

So, what if—after reading the above and looking at your systems and workflows—you deem that lossy compression is not feasible in your environment, but you still want some of the benefits.

To achieve *most* of the benefits of lossy compression, without the informatics implications described above, an alternative approach is to invest in a fast archive access interface with the ability to generate lossy compressed images at the time of retrieval.

A modern access interface, with advanced archiving and retrieval algorithms, can return studies efficiently, regardless of location. In these configurations, the notion of cache vs. archive is obsolete—data is returned just as fast from both (given that a similar storage technology used).

Applying lossy compression as images are retrieved for display provides the desired benefit of using lossy compressed objects, specifically a reduced bandwidth usage (especially good for remote, wireless, and mobile access).

Considering typical PACS implementations, it may be simpler to provide access to lossy images with conversion

on-the-fly through protocols like WADO-URI than more traditional protocols like DICOM C-MOVE. Table 1 provides a comparison of various methods of compression.

Why is On-the-Fly Lossy Compression Not More Common?

The technology to apply lossy compression on-the-fly is not new or particularly difficult, yet it is not commonly available in systems today. So, what changed? The answer is processor power.

Today's Central Processing Units (CPUs) can provide the processing power to transform images from one format to another in just a few milliseconds. Combined with increasingly efficient image processing techniques, it is now feasible to create a lossy version of an image in near real-time (Fig. 3).

With this capability, there is no need to convert all newly stored images into JPEG2000. And no need to maintain additional storage for lossy compressed objects in addition to existing lossless objects. But perhaps the most important aspect is there is no need to convert the millions to billions of historic images to a new format to gain many of the desired benefits.

It should be noted that JPEG lossy compression only supports up to 12 bits; 16-bit images are not supported

Summary and Legal Considerations

Some organizations and legal experts have made statements that the use of lossy compression, if desired for clinical and/or business reasons, should be applied prior to a primary diagnosis. The rationale is that the record stored should be identical to the record used in diagnosis.

It is important to note that this is a position borne out of legal perspectives, and not clinical safety or technical or business benefits. In reality, considering all the other factors that go into the presentation of the images—monitor calibration, environmental lighting, and access to clinical history at the time of diagnosis as well as experience, fatigue, and eyesight of the interpreter—it is unlikely that the exact display conditions or interpreter perceptions as originally diagnosed can ever be assured later on.