

## Chapter 2

# Hardware and Software

**Abstract** Lest those interested in exploring the field not understand the nuts and bolts of the system, no book on digital technology is complete without some background on the available hardware and software. The field is changing rapidly, and specific examples may be already obsolete at the time this book goes to press. At the same time, we have found that certain principles have remained constant for a relatively long time now, and we believe that providing readers with some general technical background will help on the path to implementing successful digital pathology solutions.

**Keywords** Digital pathology • Slide scanner • Tile scanner • Line scanner • File format • Medical imaging • WSI • DP • MRXS • NDPI • SVS • BIF

This chapter talks about the technology that is used to arrive at digital pathology. As the pathologist is dependent upon his microscope, so is the digital pathologist dependent upon a digital camera or slide scanner for the creation of a single, high-magnification digital image of an entire microscopic slide or whole slide image (WSI).

This chapter is split into two parts. In Sect. 2.1, we elaborate on the various hardware components necessary to acquire virtual slides. In Sect. 2.2, we survey the various approaches to data storage and file format organization that different vendors have developed.

### 2.1 How Are Digital Pathology Images “Captured”?

Basically, WSI hardware consists of a robotic/automated microscope with specialized acquisition software. Some instruments are more specialized and purpose-specific in their design and construction than others. The simplest

---

An erratum to this chapter can be found under DOI [10.1007/978-3-319-08780-1\\_7](https://doi.org/10.1007/978-3-319-08780-1_7)



**Fig. 2.1** **a** A Zeiss AxioVision setup with a mounted camera and robotic stage. **b** A 3DHistech high-volume slide scanner (courtesy of HistoGeneX)

setups consist of add-on cameras on top of conventional microscopes. This is a great start if all you want to do is capture regions of interest (ROIs) and share them with colleagues or embed them in your publications. However, they are not necessarily suited for whole slide imaging. In order to do WSI, as well as

systematically digitize your entire workflow, you need at least a robotic staging table as well. The robot then cooperates with the software component to move the slide, capturing individual ROIs that then are stitched back together to generate a WSI. Special viewing software is usually provided so that it appears that a seamless image was obtained of the entire slide. As an alternative, there also are devices like microscopes with mounted cameras but automated stages. The advantage is that viewed ROIs can be stored even when you switch between different slides (Fig. 2.1).

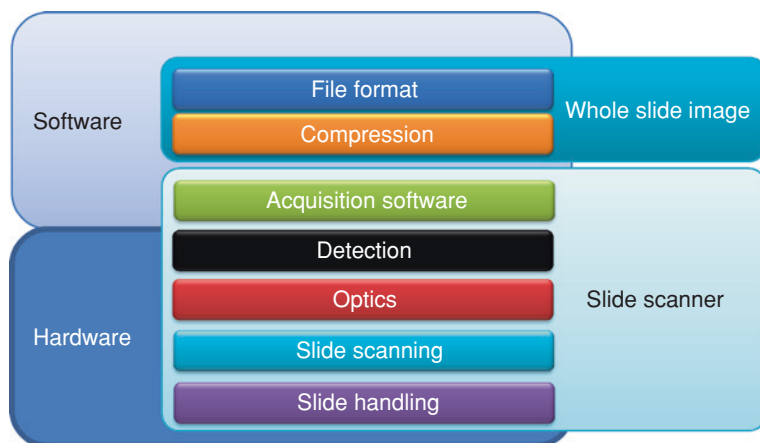
Technology has by now become sufficiently specialized so that some companies only sell complete integrated systems (scanners). However, others sell individual components as well. Examples of the latter are Hamamatsu, which sells its own nanozoomer slide scanner as an integrated system, but also sells its components to TissueGnostics for their automated solutions.

## 2.2 How Do Slide Scanners Work?

Slide scanners are the highest level of abstraction for digital microscopy. They have both a hardware and software components. We distinguish five levels, from lowest to highest: slide handling, slide scanning, optics, detection, and, finally, acquisition software. These five levels are depicted in Fig. 2.2.

The first slide scanner was designed by James Bacus in 1994, during a period of rapid Internet expansion worldwide. The corresponding BLISS system, which is now recognized as having been the first virtual microscopy system ever developed, was designed to generate virtual microscope slides. Meanwhile, a WebSlide Server, Browser, and ActiveX Viewer were developed to allow for viewing virtual microscope slides over the Internet. Over the next several years, the Bacus group developed and patented the methods and apparatus to perform automated assays of biological specimens, immunoploidy analysis, measurements of tissue thickness, and tests of neoplasm progression, as well as devices to allow for the remote control of microscopes, the creation of virtual microscopy slides, the magnification of specimen images, and the Internet, intranet, and local viewing of such slides [1–19]. Moreover, Bacus Laboratories not only created the first virtual slide system, they also created the first market for it. They did this by framing their system as an educational tool. Their plan was to ultimately replace standard microscopes with virtual microscopy in medical education. They achieved this by combining their virtual microscopy system with a collection of educational “slides,” for which institutions could lease access licenses annually. Because of its successful business model, Bacus Laboratories was purchased by Olympus America Inc. in 2006.

In terms of features, the capacities of slide scanners today vary widely. For example, some can do bright-field images only, some can do fluorescence images only, and some can do both. The price of a model generally correlates with its slide-loading capacity, which can range from one to 400 slides per batch. Slides



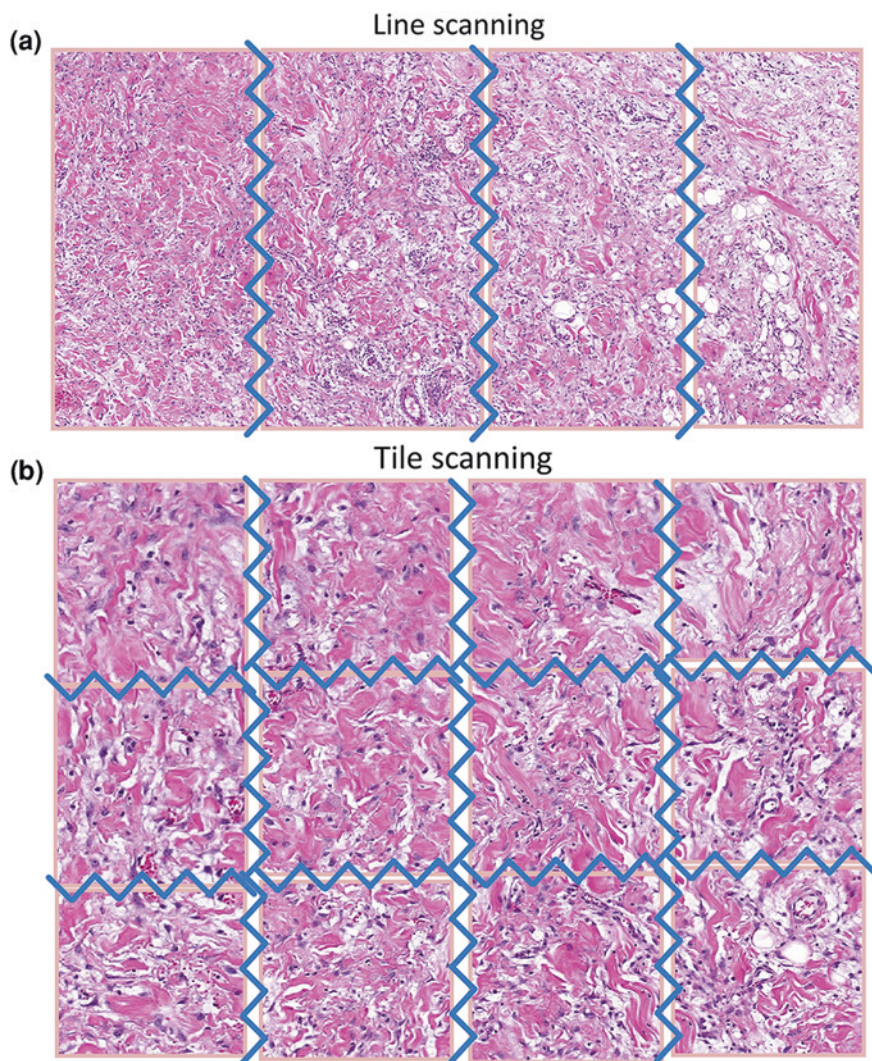
**Fig. 2.2** Different layers of processing in digital pathology

can be handled as a single slide/stage, as standalone autoloaders (“hotels”), as slide trays, or as slide magazines. The type of slide can vary as well: While most scanners today still handle basic 1” × 3” slides only, others—like Aperio and Huron Technologies—also support 2” × 3” and even larger (whole mount) slides. The more variability that is allowed for physical slide media, the harder it is to batch process large numbers of slides.

Two approaches exist to scanning a slide: tile scanning (Bacus patents) [20] and line scanning (Aperio patents) [21]. In both cases, the resulting images (tiles or strips) are fitted together into a single large image (i.e., the WSI). With a tile scanner, the slide is scanned as a series of rectangular tiles. For each tile, the highest physical magnification desired is used (e.g., 40× or 20×). The tiles are then stacked into a WSI, like bricks forming a wall. This is done either concurrently with or after the scanning process, via the acquisition software. Conversely, with line scanning, after magnification, strips are combined side by side into a single image. Proponents of the latter approach claim that it generates fewer seams and, hence, fewer optical aberrations (Fig. 2.3).

One particular problem related to scanning is focusing. A pathologist looking through a microscope automatically adjusts the focus depending upon the area of the slide he or she is looking at, the thickness of the specimen, the type of glass slide used, etc. With a scanner, this process must be automated. With both tile and line scanners, it is possible to autofocus each field after moving the stage, but this can be very time-consuming, especially with tile scanners. A better approach is to focus on every *n*th field being scanned. This is both faster and simpler; but the placement of focus points lacks context, and it is still possible to waste time on larger areas that, by chance, do not require refocusing. A focus map is another solution. With this approach, focus points are distributed over the tissue forming a surface. Focus is only recalculated for intervening tissue. The number of selected





**Fig. 2.3** a/b Line versus tile scanning: note the huge decrease in the number of seams with line scanning

focus points can be controlled via the scanner software. A trade-off is usually made between more focus points (less speed) and greater accuracy. The settings can be tissue dependent, and a technician can maintain a preset list of “profiles” that can be referred to, depending upon the type of specimen that needs to be scanned.

Z-stacking is becoming increasingly commonplace, but this poses its own unique challenges to file format organization (see later in this chapter). The new frontier is now spectral imaging.

## 2.3 Virtual Slide Formats

### 2.3.1 *How Are WSI Data Organized?*

After the acquisition software in the scanner obtains a digital image representation of a slide, it needs to store this information somewhere. This again can be seen as a two-step process, whereby first data compression takes place, and subsequently data are stored, usually in a vendor-specific file format.

Digital slide image formats typically consist of one or more files that contain high-resolution scanned areas as well as image information in the form of meta-data. The resolution of such images varies, but usually ranges from ten to hundreds of thousands of pixels per dimension (width, height). Various techniques are currently employed to make it easier and quicker to process such images using computer software.

### 2.3.2 *The Pyramidal Format*

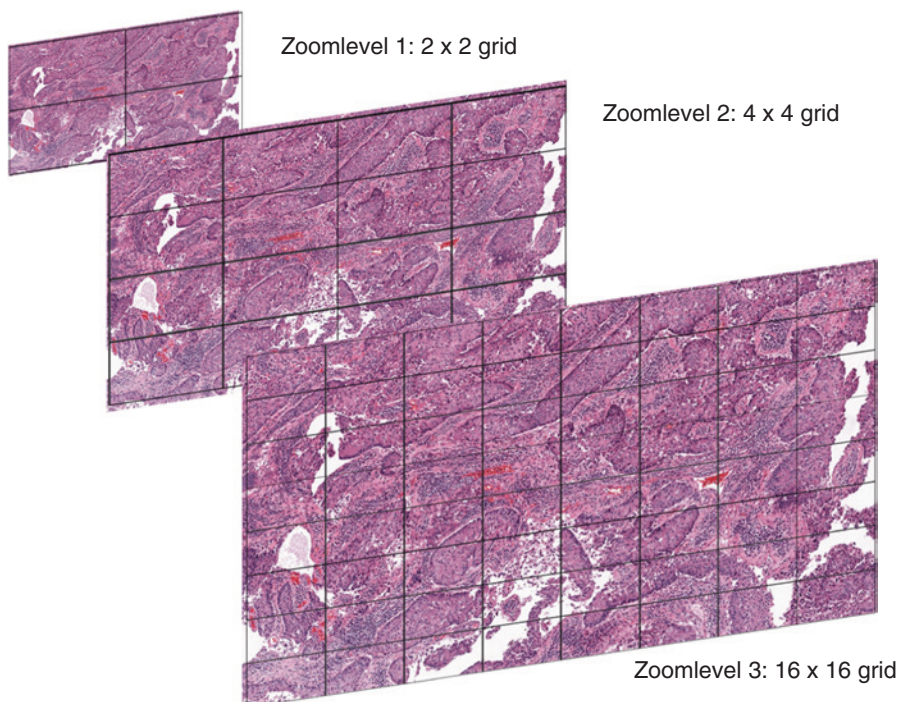
Scaled versions of the original image (called “zoom levels”) are often created and stored in a single-“container” format. This is usually called a “pyramidal format,” since every scaled-down image is smaller than its previous level, just like a pyramid gets smaller and smaller the higher up you go. By storing pre-computed scaled-down versions of the high-resolution image, a computer program can quickly render a smaller version of the image by reading pixel data from the zoom level closest to the scale currently being displayed (Fig. 2.4).

The pyramidal format increases display performance at the cost of storage efficiency. For this reason, many vendors try to minimize the actual scan area that is being stored. This is done by spotting the significant areas while scanning the slide and only storing these in high resolution. This leads to a digital slide image with many sparse high-resolution areas, which may follow the pyramidal format independently. For different tissue types, the tissue detection parameters (called “profiles” by some vendors) often must be fine-tuned.

### 2.3.3 *Tiles*

To further optimize random access and minimize disk read operations (input/output or I/O), digital slides split the image into smaller rectangular areas (tiles). Every zoom level is therefore a grid of tiles of the same size. When a computer program needs to display a small part of a high-resolution image, it is able to reduce the data being read by selecting only those tiles that intersect with the current viewport.

Slide scanning is performed in steps. The scanner’s camera moves along the slide and takes pictures which are then stitched together by the scanner’s software. Some



**Fig. 2.4** A pyramidal stack represented symbolically on a sample whole slide image

vendors decide to store overlapping images of the slide and let the viewing software do the stitching. This is done because selecting stitching offsets that depend on the visible parts of the image every time may reduce stitching artifacts. This, in turn, would have otherwise been introduced if stitching had happened during scanning. In this case, stitching hints are stored as metadata along with the image.

Regardless of when the stitching process takes place (during scanning or while the image is viewed), images acquired by the scanner require adjustments. Overlapping regions might have differences in brightness and contrast, known as *shading*, due to the different positions of the camera each time a photograph is taken. Various techniques are employed to address this issue, such as blending and histogram equalization.

### 2.3.4 Color Spaces

The most common illumination techniques found in digital slides are bright-field and fluorescence. Bright-field microscopy images typically store pixels in the red, green, and blue (RGB) model (color space) or YC<sub>B</sub>C<sub>R</sub> (another family of color

spaces) for JPEG images. Grayscale (points of equal RGB values; essentially a subspace of the RGB model) is especially used in the case of fluorescence microscopy slides to store the intensity of the reflected emission. This is then multiplied by a constant factor in order to be colorized for display purposes.

### **2.3.5 Compression Schemes**

Digital pathology images employ various compression and image representation schemes, which may or may not lead to color information loss. Some of the compression schemes that are used are Raw, JPEG, JPEG2000, PNG, LZW, and DEFLATE.

## **2.4 Vendor-Specific File Format Implementations**

The images that are generated by digital slide scanners are very different from the JPEG images typically obtained using your cell phone or digital camera. To start with, they are vastly larger and more complex. This is aptly demonstrated in Fig. 2.3, in which a typical high-resolution digital camera image is compared to a typical digital WSI. Note that the WSI has almost 1,500 times as many pixels ( $200,000 \times 100,000$  vs.  $4,600 \times 3,000$ ) as the camera shot (Fig. 2.5).

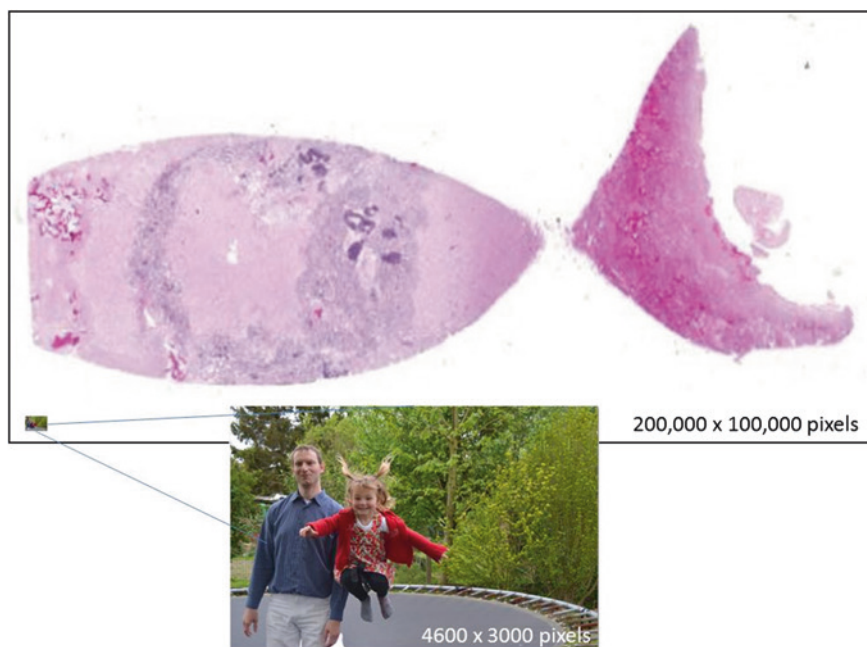
There also are a large number of different WSI file formats, which are necessary because of the multiple applications for which these images are used, beyond simple viewing. For example, Zeiss format (czi and zvi) images (see further in this chapter) can encompass as many as six or seven dimensions, versus the 2-dimensional JPEG you generate with your home camera, to accommodate their use for microbiology, time lapse, fluorescence, and other applications. Below, we briefly describe just some of these various formats, highlighting their basic characteristics, as well as their advantages, disadvantages, and differences.

### **2.4.1 TIFF-Based Formats**

#### **2.4.1.1 TIFF**

TIFF images are used by scanner vendors to store digital slides. The TIFF format natively supports storing images in grids of tiles and is generally well suited for random access. It allows for multiple images (directories) to be stored within a single file and for various compression schemes to be used. Since a slide's size may overcome the maximum 4 GB threshold, the BigTIFF format is also common. It essentially uses 64-bit pointers to store offsets within the file.





**Fig. 2.5** Comparing a traditional high-resolution digital camera photograph (of coauthor YS and his bouncing daughter) against a typical digital whole slide image

A tiled TIFF directory specifies four important tags:

- Tile width (TIFF specifications tag number 322)
- Tile length (also known as height) (323)
- Tile offsets (324)
- Tile byte counts (325)

Typically, one TIFF directory is the high-resolution image, while several others may follow that are down-scaled versions of the original. One downside of the plain TIFF format is that there is no definitive way to specify which directory is for the high-resolution image and which are for the down-scaled images, because the specifications do not anticipate relationships between the directories. The display software attempts to overcome this by making assumptions; for example, the largest directory (in width or height) may be considered the original image and every other directory a smaller zoom level.

#### 2.4.1.2 Open Microscopy Environment TIFF (Extensions .tif, .tiff)

The OME-TIFF format was created to maximize the respective strengths of OME-XML and TIFF. It takes advantage of the rich metadata defined in OME-XML,

while retaining the pixels in multi-page TIFF format for compatibility with many more applications.

An OME-TIFF dataset has the following characteristics:

- Image planes are stored within one multi-page TIFF file, or across multiple TIFF files. Any image organization is feasible.
- A complete OME-XML metadata block describing the dataset is embedded within each TIFF file's header. Thus, even if some of the TIFF files in a dataset are misplaced, the metadata remain intact.
- The OME-XML metadata block may contain anything allowed in a standard OME-XML file.
- OME-TIFF uses the standard TIFF mechanism for storing one or more image planes in each of the constituent files, instead of encoding pixels as base 64 chunks within the XML. Since TIFF is an image format, it makes sense to only use OME-TIFF as opposed to OME-XML, when there is at least one image plane.

#### 2.4.1.3 Leica SCN (Extension .scn)

Leica slides are stored in single-file BigTIFF format. The image description tag for the first TIFF directory contains an XML document that defines the structure of the slide. The slides are structured as a collection of images, each of which has multiple dimensions (pyramid levels). The collection has a size, and images have both a size and position. Each dimension has a size in pixels, an optional focal plane number, and a TIFF directory that contains the image data. Fluorescence images have different dimensions (and thus different TIFF directories) for each channel.

Bright-field slides have at least two images: a low-resolution macro-image and one or more main images corresponding to regions of the macro-image. The macro-image has a position of (0, 0) and a size that matches the size of the collection. Fluorescence slides can have two macro-images: one bright-field and one fluorescence.

#### 2.4.1.4 Aperio SVS (Extension .svs)

Aperio slides are stored in single-file TIFF format. The first directory in an SVS file is always the full-resolution image. This image is always tiled, usually with a tile size of  $240 \times 240$  pixels. The second image is always a thumbnail, typically with dimensions of about  $1,024 \times 768$  pixels. Unlike the other slide images, the thumbnail image is always stored in strips, instead of tiles. Following the thumbnail, there may be one or more intermediate "pyramid" images. These are always compressed with the same type of compression as the baseline image and have a tiled organization of identical tile size.

Optionally at the end of an SVS file, there may be a slide label image, which is a low-resolution picture taken of the slide's label, and/or a macro-camera image, which is a low-resolution picture taken of the entire slide. The label and macro-images are always stored in strips.

### 2.4.1.5 Ventana BIF (Extension .bif)

Ventana slides are stored in single-file BigTIFF format. The first directory contains a label image, usually in tiled format. This is a thumbnail which includes the actual physical label for the glass slide. The directory specifies the XMP tag (700) and stores valid XML metadata about the slide. Next comes a thumbnail, and the high-resolution image follows.

BIF images contain overlapping tiles, and an appropriate algorithm is required to correctly render them. The directory containing the high-resolution images also specifies the XMP tag which contains tile stitching hints between neighboring tiles. The rendering algorithm calculates global coordinates for every tile, based on these hints. This may result in stitching artifacts in parts of the image. Subsequent image directories do not have such information, and tile positions are calculated via reduction to the base level.

### 2.4.1.6 Hamamatsu NDPI (Extension .ndpi)

Hamamatsu NDPI slides are stored in single-file TIFF format. Unlike other TIFF-based formats, NDPI does not use TIFF's tile format. Instead, every directory contains a single JPEG image, declared as stripped. These images are not necessarily valid. The JPEG specification does not allow images larger than  $65,536 \times 65,536$  pixels. Therefore, when a directory exceeds these limits, NDPI sets the width and height of the image to zero. This means that at least some JPEG libraries are unable to read the image data. It is still possible to parse these with out-of-the-box solutions like a LIBJPEG. When doing so, the programmer has to invoke a callback mechanism to inform the JPEG decoder of the actual size whenever 0, 0 (a 'reset marker') is found in the stream.

Restart markers are placed inside the JPEG stream to achieve support for random-access reading. Each interval contains one tile, effectively implementing a tiled format within the JPEG image. The offsets of the markers are stored in a custom TIFF tag (65426), so that the whole stream does not need be parsed for them to be located.

## 2.4.2 Other Format Types

### 2.4.2.1 3DHISTECH Mirax (Extension .mrxs)

3DHISTECH Mirax slides are stored in a multi-file JPEG format with proprietary metadata and indexes. JPEG images are packed into a small number of data files, while the index file provides offsets into the data files for each required piece of data.

The camera on MIRAX scanners takes overlapping photographs and records the position of each one. Each photograph is then split into multiple JPEG images which do not overlap. Overlaps only occur between images that come from different photographs.

To generate level  $n + 1$ , each JPEG image from level  $n$  is down-sampled by two and then concatenated into a new JPEG image, with four old images per new JPEG image ( $2 \times 2$ ). This process is repeated for each level, irrespective of image overlaps. Therefore, at sufficiently high levels, a single image can contain one or more embedded overlaps of non-integral width.

#### **2.4.2.2 Olympus VSI (Extensions .vsi, .ets)**

The VSI file format is based on the well-known TIFF image file format. Besides the standard TIFF IFDs (image file directories) necessary to access the raster pixel data, the VSI format contains a chain of custom tags used to store additional information in the image files.

VSI files may contain multiple regions of the same physical slide, each scanned at a different level of resolution. High-resolution pixel data are not stored within the .vsi file, but in files with an .ets extension, stored in subdirectories defined within the VSI.

Extensible tile server (ETS) is a proprietary file format that is used to store multi-dimensional data organized in tiles. Usually, a single region of a slide is stored in tiled pyramidal fashion inside an ETS file.

#### **2.4.2.3 Carl Zeiss CZI/ZISRAW (Extension .czi)**

This format has been designed to be as close as possible to the open microscopy environment (OME) specification (<http://www.openmicroscopy.org>). The XML schema definition (XSD) was defined to exhibit maximum compatibility with the OME tiff and XML data formats, while maintaining the essential requirements to run Carl Zeiss ZEN software optimally.

The chosen architecture is a chain of “segments.” Each segment is identified by a header with a defined identifier (SID). Segments are aligned on 32-byte boundaries. This improves the speed of the recovery process when rescanning the file, which can be useful in situations like when the system crashes. A single search step in case of a missing segment header can move to the next multiple of 32 bytes instead of advancing byte by byte.

#### **2.4.2.4 Carl Zeiss ZVI (Extension .zvi)**

The ZVI format is a standard OLE compound file containing a storage named “image” for the container image and commonly used information. The image data (pixel array) items are contained in substorages named “item( $n$ )” where ( $n$ ) is a value from 0 to Count – 1.

### 2.4.3 The Role of DICOM

As the proverbial 800-pound gorilla in the room, DICOM deserves its own paragraph. DICOM stands for *Digital Imaging and Communication in Medicine* and is a network maintained by the National Electronic Manufacturer's Association (NEMA) and supported by large-image management systems called *picture archive and communication systems* (PACS). Various PACS systems are used in hospitals and laboratories to manage images used for clinical and research purposes in medicine; this includes, among other functions, their storage and retrieval.

Since 2009, a new supplement has been added to the DICOM standard. This supplement is known as "Supp 145, Whole Slide Imaging in Pathology." The supplement was developed by Workgroup 26 and describes how an extension has been made to the DICOM standard to allow for the storage of very large images. The DICOM standard defines a family set for the image, called "instances" as per the DICOM vocabulary. All these instances follow an information object definition which is defined in PS 3.3; currently, version 2011 is the latest available. In all those IODs, DICOM instances have columns and rows defined as unsigned short values. What this means is that, in theory, all images must be 64K columns and rows. WSI frequently has images much larger than these pixel dimensions.

Rather than following what occurs during the TIFF to BigTIFF (64 bits extension) transition, the DICOM committee chose a different, very conservative approach, whereby the unsigned short value for the column and row does not change. Instead, new attributes are added to store the actual pixel dimension. In this scenario, a single WSI cannot be stored within a single "instance." Instead, a single WSI is inserted in fragments at a series level.

The proposed approach guarantees that all legacy software remains able to process any incoming WSI series, since the attributes in columns and rows are still defined as unsigned short.

One should notice, though, that this supplement pushes the DICOM standard to the edge, since uncompressed pixel data stored within a single DICOM instance are limited to a  $2^{32} - 1$  byte (4 GB minus a byte, 0xFFFFFFFF being a reserved value). Therefore, the lower level of this pyramid is unlikely to be saved in uncompressed form, since its total size will likely exceed that limit considerably. In such a case, it is assumed that another transfer syntax will be used for those larger pyramid levels (e.g., JPEG). When using an encapsulated transfer syntax (e.g., JPEG type), there is no such limit, and all individual tiles can be stored within a single DICOM instance.

## 2.5 Bits, Bytes, and Wires

This book is not intended as guidelines on how you can build your own scanner or write your own WSI-viewer software. Nevertheless, no review of digital pathology can be complete without also addressing the hardware and software involved.



We have tried to introduce you to some of the intricacies of engineering that were required to develop slide scanners in the first place. Then, we moved on to the software side of things: How are WSI data stored? This is something that we all get exposed to, if only by transferring slides to a colleague via a USB memory key.

Slide scanners have not been around all that long. Two basic modes of operation exist for scanners, and virtually all scanners on the market today can be traced back to one or two sets of patents. Data captured by the slide scanner must be stored on the hard disk of the user's computer and organized so that it can be visualized optimally. File formats have been devised by different vendors to accomplish this. Because of the pixel size of the raw data (roughly 1,500 times more than the digital camera that you use on vacation) and the different features of the scanners (bright-field, fluorescence, confocal, etc.), various solutions have been thought of. However, these differences make it hard to move from one digital pathology platform to another, and one risks vendor lock-in because of this. Some initiatives for standardization have already been undertaken and are expected to become more center stage in the future.

## References

1. Bacus, J.V., Bacus, J.W.: Method and apparatus for processing an image of a tissue sample microarray. US Patent 6,466,690 (2000)
2. Bacus, J.V., Bacus, J.W.: Method and apparatus for acquiring and reconstructing magnified specimen images from a computer-controlled microscope. US Patent 6,101,265 (2000)
3. Bacus, J.V., Bacus, J.W.: Method and apparatus for acquiring and reconstructing magnified specimen images from a computer-controlled microscope. US Patent 6,226,392 (2001)
4. Bacus, J.V., Bacus, J.W.: Method and apparatus for creating a virtual microscope slide. US Patent 6,272,235 (2001)
5. Bacus, J.V., Bacus, J.W.: Method and apparatus for acquiring and reconstructing magnified specimen images from a computer-controlled microscope. US Patent 6,404,906 (2002)
6. Bacus, J.V., Bacus, J.W.: Apparatus for remote control of a microscope. US Patent 6,674,884 (2004)
7. Bacus, J.V., Bacus, J.W.: Method and apparatus for creating a virtual microscope slide. US Patent 6,775,402 (2004)
8. Bacus, J.V., Bacus, J.W.: Apparatus for remote control of a microscope. US Patent 7,110,586 (2006)
9. Bacus, J.V., Bacus, J.W.: Method and apparatus for creating a virtual microscope slide. US Patent 7,146,372 (2006)
10. Bacus, J.V., Bacus, J.W.: Method and apparatus for internet, intranet, and local viewing of virtual microscope slides. US Patent 7,149,332 (2006)
11. Bacus, J.W., Bacus, J.V.: Method and apparatus for automated assay of biological specimens. US Patent 5,473,706 (1995)
12. Bacus, J.W.: Method and apparatus for automated analysis of biological specimens. US Patent 5,526,258 (1996)
13. Bacus, J.W., Marder, J.M.: Method and apparatus for immunoploidy analysis. US Patent 5,541,064 (1996)
14. Bacus, J.W., Bacus, J.V.: Method and apparatus for measuring tissue section thickness. US Patent 5,546,323 (1996)

15. Bacus, J.W., Oud, P.S.: Apparatus and method for analysis of biological specimens. US Patent 5,485,527 (1996)
16. Bacus, J.W., Bacus, J.V.: Method and apparatus for testing a progression of neoplasia including cancer chemoprevention testing. US Patent 6,031,930 (2000)
17. Bacus, J.W., Bacus, J.V.: Method and apparatus for internet, intranet, and local viewing of virtual microscope slides. US Patent 6,396,941 (2002)
18. Bacus, J.W., Bacus, J.V.: Method and apparatus for creating a virtual microscope slide. US Patent 6,522,774 (2003)
19. Bacus, J.W., Bacus, J.V.: Method and apparatus for internet, intranet, and local viewing of virtual microscope slides. US Patent 6,674,881 (2004)
20. James W.: Bacus patents. JamesBacus.com. (2014). Ref type: Electronic Citation
21. Aperio Technologies Inc., Patent Applications. Patentdocs. (2014). Ref type: Electronic Citation

Digital Pathology

Sucaet, Y.; Waelput, W.

2014, X, 83 p. 10 illus. in color., Softcover

ISBN: 978-3-319-08779-5