White paper

A brief history of digital pathology and its applications

Yves Sucaet[1], Agelos Pappas[2], Wim Waelput[1,*]

[1]Pathomation bvba, Pr J Charlottelaan 10, 2600 Berchem, Belgium.

[2]SmartCode, Smyrnis 14, 47100 Arta, Greece

INTRODUCTION

One of the newest fields in medicine is Digital Pathology (DP). The concept and functionality of DP is growing rapidly, and currently includes for the visualization of specimens in digital form, electronic management of two and three-dimensional specimens, their real-time evaluation, comparison, two to three-dimensional reconstruction, archiving, and dissemination for widespread viewing and consultation. Data can be compiled with other patient data, data mined, and used for education, clinical diagnosis and patient management, research. Obstacles to implementation exist, but are surmountable and very much worth overcoming. The Pathomation family of focused software-products can help you accomplish your DP-goals.

1   HISTORY

The first pathology slides were digitized as black and white digital images back in 1968, and subsequently transferred from Boston's Logan Airport to Massachusetts General Hospital way. It was a huge step towards recognizing the advantages of such technology: the ability to convert pathology slides into digital images that could then be transmitted across distances. This was feasible because a rudimentary laboratory information system, called the Massachusetts General Hospital Utility Multi-Programming System (MUMPS), had been created as a collaboration between Massachusetts General Hospital and a company called Bolt Beranek Newman. Inadequate programming tools and computer technology held back DP for most of the 1970s and even 1980s. It was creation of the World Wide Web in the 1990s that led to the widespread creation and use of data transmission systems and now comprises one of the cornerstones of DP.

In the 21st century, DP has the potential to fundamentally change the way in which pathological specimens are viewed. Instead of viewing glass slides or other specimens through a microscope, they can now be examined through a digital monitor. The most essential element of this process is a device to digitally capture images (image digitization). These come in two forms: digital cameras and digital scanners. Digital cameras are typically available as add-on hardware for your already purchased microscopes, and provide a low-risk way to ease into DP and familiarize yourself and your staff with digital representations of physical slides. Beyond cameras are slide scanners, which can drastically reduce the need for and occupation rates of microscopes at your facilities.

Figure 1: different configuration of cameras and slide scanners

2   GOING DIGITAL

One of the many advances in digital pathology that have occurred over time relate to the creation of the pathologist's new work station, which has been called a digital cockpit or digital dashboard, so that it enhances each pathologist's ability to access, visualize, interpret and share digital pathology images and thereby utilize digital

---

*To whom correspondence should be addressed: wim@pathomation.com .

informatics systems effectively and efficiently. One major problem with dashboards, and in fact with most components of telepathology and digital pathology systems, has been their lack of standardization.

Specialized slide scanners are both high-resolution and high-speed, and can digitize images across large arrays and combine these images, through a process called stitching, into even larger arrays. Entire slides can be visualized in a single pass this way. They can be scanned at multiple levels of magnification and in all three planes (x, y and z)(14). Another advantage of stitching numerous images together (rather than trying to capture entire slides at once) relates to the former allowing for images to be brought into focus and then captured at higher levels of magnification, rather than having to capture the entire slide at lower magnification and later enlarge it, thereby tampering with focus and losing resolution. However, stitching creates its own complications, like ensuring that the lighting, focus and colouring of each of the partial images match, especially since the topology of the specimen may vary from one part of the specimen to another. This has led to numerous process refinements, like shading and lighting normalization, auto-focusing and independent dual-sensor scanning that allows for image acquisition and focusing to be performed sequentially, rather than in the same step. Various vendors have come up with different solutions to the problems, which has unfortunately resulted in even more incompatible file formats. This is where our company wants to make a difference, by offering a vendor-agnostic and format-independent image broker for DP data.

Figure 2: An H&E-stained slide visualized in the HistoWebViewer browser-based application of Pathomation.

3 TECHNICAL CHALLENGES FOR IMAGE CAPTURING AND FILE FORMATS

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 tenths to hundredths of thousands pixels per dimension (width, height). Various techniques are being employed to make it easier and quicker to process such images using computer software.

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

Storing precomputed zoom levels increases 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.

To further optimize random access and minimize 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 the 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 vendors decide to store overlapping images of the slide and let the viewing software do the stitching. This is done because by selecting stitching offsets depending on the visible parts of the image every time, may reduce stitching artifacts which would have otherwise been introduced if stitching would have happened during scanning. In this case stitching hints are stored as meta-data along with the image.

Finally, DP 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 being used are Raw, JPEG, JPEG2000, PNG, LZW and DEFLATE. The color space in which image information is encoded is usually RGB, YcbCr (for JPEG images) and gray

scale. Gray scale is especially used in the case of fluorescent microscopy slides to store the intensity of the reflected emission. This is then multiplied by a constant factor in order to be colorized for displaying purposes.

4   WHY DP DIFFERS FROM RADIOLOGY

Notice how we call it "radiology", not "digital radiology". That's because radiology already *is* digital. We take it for granted, so we omit the "digital" adjective. We estimate pathology is about ten years behind radiology in terms of digitization. It is tempting therefore to think that the solutions for radiology can be replicated or even re-used for pathology.

However, it is important that pathology is also fundamentally different from radiology, and there are reasons why digital pathology lags behind. First, the source material is different: radiology usually works with live specimens (patients), whereby pathology usually concerns specimen samples (biopsies, cytologies). Protocols are fundamentally different in pathology, whereby often additional stainings can be asked by the pathologist, based on original observations in an H&E stained slide. This constant interchange between digital observations and wet lab techniques pose new challenges for laboratory information systems, which is just one reason was RIS-software is not a good fit for today's pathology departments (and therefore cannot easily be ported either).

Digital pathology has a number of intrinsic advantages as well. The whole slide images are representations of physical (stained) microscopy slides, which are archived by most hospitals. This means that in the simple case where an image is unfocused, it suffices often to simple re-scan the original slide. This cannot be done in radiology, where a patient has to be called in again, rescheduled etc.

Finally, DP has always had the advantage of the over coupling DICOM-organization, which sets data representation standards for a wide array of imaging data. However, one main advantage here is that these data typically involves grayscale data. Chromogenic stains in DP are by definition in color. Recognizing these essential differences, DICOM has recently branched of a different workgroup to examine the needs for DP-specific applications (WG-26).

5   A BRIGHT FUTURE

The ability to scan tissues at different levels and in all three planes has led to the generation of three-dimensional image reproductions of original tissue, which is achieved by scanning multiple focal planes into images and then stacking them, a process that is invaluable for the evaluation of cytological specimens, frozen sections and other thick specimens where the pathologist needs to assess cellular architecture in multiple planes; in this way, entire tissue sections can be visualize. Thick specimens can not only be scanned throughout, but the focal plane can be rotated in any direction. In addition, stains can be both detected and quantified using recent innovations like automated histopathology pattern recognition; color enhancement and standardization techniques, as well as color content analysis that allows for the detection and quantification of histochemical stains; and multiplexed biomarker testing so that several tissue characteristics, biomarkers or stains can be sought and detected on the same slide, thereby replacing the tedious-to-make and difficult-to-maintain cell and tissue (paraffin) blocks of traditional microscopy.

6   THE PATHOMATION ADVANTAGE

Pathomation offers a range of software products that can help you manage and share your digital pathology data. Central in our product offering is a vendor-agnostic image server, which recognizes DP-specific file formats, including SCN, MRXS, NDPI, TIFF, SVS, ZVI, VMS and VMU. The data server sits as a broker between our different visualization and data transfer modules. We offer DP-specific slide viewers, that target your desktop, your intranet/internet-environment, or the public cloud. In addition to viewers, we've developed plugins for different popular image processing software, like PhotoShop or ImageJ.

About Pathomation

Pathomation was founded in 2012 by two practicing pathologists and a bioinformatician. We develop digital pathology software for healthcare and life science environments. Our product offering includes both local installations and cloud-deployments. We believe in controlled flexibility, whereby versatility is key (different organ systems require difference evaluation methods and modalities), but control and supervision is possibly even more important with increasing regulatory requirements.

Please visit out website at http://www.pathomation.com. A public portal of our capabilities is presented at http://public.pathomation.com. Our cloud-based solution is available via http://cloud.pathomation.com

---