Journal of Biomedical Informatics 71 (2017) 190–197

Contents lists available at ScienceDirect

Journal of Biomedical Informatics

journal homepage: www.elsevier.com/locate/yjbin

An efficient architecture to support digital pathology in standard medical imaging repositories

Tiago Marques Godinho [a,*], Rui Lebre [a], Luís Bastião Silva [b], Carlos Costa [a]

[a] University of Aveiro, Portugal
[b] BMD Software, Portugal

ARTICLE INFO

Article history:
Received 24 February 2017
Revised 2 June 2017
Accepted 5 June 2017
Available online 7 June 2017

Keywords:
Whole-slide imaging
DICOM
PACS
Pathology
Web
Viewer

ABSTRACT

In the past decade, digital pathology and whole-slide imaging (WSI) have been gaining momentum with the proliferation of digital scanners from different manufacturers. The literature reports significant advantages associated with the adoption of digital images in pathology, namely, improvements in diagnostic accuracy and better support for telepathology. Moreover, it also offers new clinical and research applications. However, numerous barriers have been slowing the adoption of WSI, among which the most important are performance issues associated with storage and distribution of huge volumes of data, and lack of interoperability with other hospital information systems, most notably Picture Archive and Communications Systems (PACS) based on the DICOM standard.

This article proposes an architecture of a Web Pathology PACS fully compliant with DICOM standard communications and data formats. The solution includes a PACS Archive responsible for storing whole-slide imaging data in DICOM WSI format and offers a communication interface based on the most recent DICOM Web services. The second component is a zero-footprint viewer that runs in any web-browser. It consumes data using the PACS archive standard web services. Moreover, it features a tiling engine especially suited to deal with the WSI image pyramids. These components were designed with special focus on efficiency and usability. The performance of our system was assessed through a comparative analysis of the state-of-the-art solutions. The results demonstrate that it is possible to have a very competitive solution based on standard workflows.

1. Introduction

Pathology focuses mainly on the identification structural anomalies, through the naked eye or a microscope, and on the detection of possible relationships with functional disorders of tissues, therefore, identifying diseases. The aim of pathology has remained unchanged over time; focused on the analysis and comparison of tissue specimens on specific glass slides. For this, the use of optical microscopes has been fundamental since it was the only available instrumentation for centuries [1]. Despite using very methodical analysis workflows, it is possible for the same professional to draw different conclusions about the same specimen at different times. Moreover, asking for second opinions is common practice, and specific cases could be part of conferences or external quality assurance programs [2]. Consequently, there is the requirement for glass slides storage and delivery infrastructures. However, the specimen storage process is expensive, requiring accessibility, cleaning, and protection, which entails greater care by specialized staff. In contrast, digital storage and distribution cuts these costs and increases the throughput of pathology laboratories.

In this context, the branch of digital pathology and whole-slide imaging arises. These recent concepts refer to the digital capture of an image from a classic glass slide as well as to the field of information systems for managing the associated data. The arrival of Digital Slide Scanners (DSS) has introduced the concept of whole-slide images (WSI). As the name suggests, these images capture the slide as a whole, rather than specific artifacts found by pathologists. As such, they can be captured unattendedly and screened later. These large images have characteristics, forms of handling and operation similar to images produced by optical microscopes [3]. It is worth mentioning that a typical slide scanned at ×40 (approx. 1600 megapixels) produces a file with several gigabytes [4,5] and requires a viewer application with special functionality to fulfill the pathologist's needs [6].

Digital pathology and whole-slide imaging have been gaining momentum with the proliferation of digital scanners from differ-

* Corresponding author.
E-mail addresses: tmgodinho@ua.pt (T. Marques Godinho), ruilebre@ua.pt (R. Lebre), bastiao@bmd-software.com (L.B. Silva), carlos.costa@ua.pt (C. Costa).

http://dx.doi.org/10.1016/j.jbi.2017.06.009
1532-0464/

ent manufacturers [5,7]. The literature reports significant advantages associated with the adoption of digital images in pathology [8–11], namely, improvements in diagnostic accuracy, promotion of distributed work processes (e.g. telepathology), integration of images with hospital information systems and economic efficiency gains. Moreover, WSI offers new clinical and research applications to the pathology community [5,12]. However, numerous barriers have been slowing this process, including performance issues, workflow efficiency, infrastructure, and integration with other software [8,9].

Despite being a medical imaging modality, pathology studies have been kept away from standard Picture Archive and Communications Systems (PACS), in proprietary formats and information systems [5,7]. In part because the DICOM standard did not address all the necessary requirements, but also due to the technical challenges raised by whole-slide imaging. As a result, major DDS vendors preferred to use proprietary archives and image formats (e.g. NDPI, SVS, and TIFF) [5,13]. However, in the last few years, whole-slide imaging has been incorporated in the DICOM standard, and vendors are slowly starting to support DICOM in their scanners. This opens the door to the use of vendor-neutral DICOM-compliant archives and services for storage processes and for exchanging data with viewer applications [6]. Moreover, such a system would support third-party applications to access the WSI images directly. Nonetheless, WSI in general purpose PACS is still a mirage, because traditional solutions are not ready to handle the remarkable image resolution and volumes of data generated from each study, requiring new architectures for storage, efficient distribution, and visualization across heterogeneous systems [3].

In this paper, we propose an architecture to include digital pathology workflows in a general purpose PACS using exclusively DICOM standard data formats and communications. The PACS archive provides storage and access to the WSI data via the most recent DICOM Web services which follow the RESTful Web services principles. This flexible interface allows a specific region (i.e. tile) of the large WSI matrix to be requested. A web viewer application that includes a tiling engine specially designed to fulfill the WSI navigation requirements was also developed. It consumes the WSI data directly from the PACS via its DICOM services. This viewer application features all typical tools available in proprietary solutions. The whole solution usability, performance, and efficiency were evaluated. The performance implication of serving WSI was assessed and the results demonstrate an architecture capable of efficiently supporting pathology imaging workflows. The main goal was to demonstrate that it is possible to support a digital pathology laboratory in a common hospital PACS, opening doors to the proliferation of vendor-neutral equipment and application to digital pathology.

2. Background

2.1. PACS and DICOM

The DICOM Standard has been the principle driving force behind the adoption of Picture Archive and Communication Systems [14]. DICOM defines not only the storage format for medical images but also the communication protocol for exchanging images and related meta-data among different PACS applications. DICOM objects come coupled with a wide-range of meta-data related to the image itself, the acquisition procedure, and the different stakeholders involved, such as patients, physicians and institutions. The standard defines thoroughly which information should be included in each modality, as well as the image organization in the DICOM Files. This allows DICOM to address the needs of medical imaging workflows and provide general purpose and vendor image formats [14].

Initially, DICOM was developed as a standard for radiology modalities. Nonetheless, it has rapidly spread its coverage to other medical specialties, such as cardiology or ophthalmology [14], due to the benefit of aggregating all patient imaging history in the PACS. The first attempt to introduce microscopy into DICOM dates to 1999, with the addition of the visible light supplement 15 [15]. As stated in [15], the scope of this supplement was to support visible light images acquired by endoscopes, microscopes or photographic cameras. This extension to the standard introduced four new modalities: Endoscopy (ES), Microscopy (GM), Automated-Stage Microscopy (SM), and Photography (XC). WSI was not included in this supplement since the technology was in its early stages. Therefore, the requirements of pathologists were not fulfilled and they continued to work without standard digital imaging end systems. Nevertheless, the automated stage microscopy modality opened the door to automated whole-slide scanners in DICOM.

The rising interest in WSI motivated the constitution of DICOM Workgroup 26 with the intention of supporting whole-slide imaging in DICOM. This required alterations to the standard, not only at the image storage level, to handle images with very high resolutions, but also to the information model [16]. The traditional DICOM information model is patient-centric as opposed to the microscopy environment where the specimen is the most relevant subject [16]. The DICOM supplement 122 [17] was the first contribution of this workgroup and introduces the concept of a specimen and container. Moreover, it also includes new data elements necessary to more accurately describe the workflows associated with WSI preparation and acquisition [14].

2.2. Whole-slide imaging visualization

WSI presented major challenges for medical imaging visualization. These images cannot be fully displayed in their maximum resolution, not only because there are no displays supporting such resolutions, but also because their size is much greater than the amount of memory available in most computers [13,18]. As a result, pathologists screening these images use a pan & zoom strategy. This involves browsing and selecting sub-regions with diagnostic interest in a lower resolution image and then zooming into a higher resolution whenever more detail is needed. Nowadays, most Internet users are accustomed to this process because of the map browsers available online, such as Google Maps.

The arrangement of the image pixel data in the storage data structure has significant performance implications for the visualization processes. The simplest, and most common, way of storing two-dimensional images is in a single frame/page, where the image pixels are stored in a sequential array. Normally, pixels are oriented horizontally, by rows, although they could follow a vertical orientation just as easily. However, this organization has limitations in the WSI screening paradigm, because it does not provide direct access to 2D sub-regions. As a result, all the regions' overlapping rows must be loaded completely, which due to the gigantic size of WSI may not fit into the system capabilities [18]. An illustration of this limitation is provided on the left of Fig. 1. In order to retrieve the green region, all the pixels in red must be retrieved because they are arranged sequentially.

A more efficient way of storing high-resolution WSI is using a tiled organization, illustrated in Fig. 1(middle), where the entire image is stored in rectangular regions sequentially. Despite being a more complex organization, it enables direct access to these sub-regions. As a result, in order to load the green region, it is only necessary to retrieve a smaller image subset composed only of the overlapping tiles [3], as shown in Fig. 1. The tile size has a signifi-