Multi-resolution graph-based representation for analysis of large histo-pathological images

Maria Cristina Lasso Lopera

Universidad Nacional de Colombia
Facultad de Ingeniería, Departamento de Ingeniería de Sistemas e Industrial
Bogotá D.C., Colombia
2012

Multi-resolution graph-based representation for analysis of large histo-pathological images

Maria Cristina Lasso Lopera

Tesis presentada como requisito parcial para optar al título de:
Magister en Ingeniería de Sistemas y Computación

Director:
Ms.C., Ph.D., Eduardo Romero Castro

Línea de Investigación:
Procesamiento de imágnes médicas
Grupo de Investigación:
Computer Imaging and Medical Applications Laboratory - CIM & LAB

Universidad Nacional de Colombia
Facultad de Medicina
Bogotá D.C., Colombia
2012

A mis padres y mi esposo, su apoyo fue mi fuerza para realizar este proyecto.

Agradecimientos

Quiero expresar mi profundo agradecimiento a mi director, el profesor Eduardo Romero Castro, por su apoyo y enseñanzas durante todo el desarrollo de este trabajo. Gracias profesor por tener confianza en mi y por darme la oportunidad de trabajar a su lado.
También quiero dar un especial agradecimiento al profesor Fabio González Osorio por su valioso aporte en este trabajo, sus ideas, sugerencias y apreciaciones fueron de gran ayuda.
Agradezco a los patólogos Lucia Roa Peña, Orlando Ricaurte, Viviana Arias y Yovanny Sánchez, profesores del Departamento de Patología de la Universidad Nacional de Colombia, sede Bogotá, quienes con su conocimiento contribuyeron de manera importante a este proyecto. En especial a la Dra. Lucia, por sus enseñanzas e interés en el mismo.
Quiero también agradecer a Gloria Díaz, Jose Luis Morales, Luis Daniel Martínez y Tatiana Castro por su colaboración en la recolección de los datos para el desarrollo de este trabajo.
A todos los integrantes del Grupo de Investigación BioIngenium, gracias por su acompañamiento, de cada uno aprendí muchas cosas y directa o indirectamente fueron un apoyo para la realización de este trabajo. En especial a Ricardo Gutiérrez y Andrea Rueda.
También quiero agradecer a mis maravillosos padres, ejemplo de trabajo y perseverancia y a mis hermanos, su incondicional apoyo fue invaluable para la culmiación de este trabajo.
Por último pero no menos importante, quiero agradecer a mi esposo, Fabio Martínez Carrillo, quien con su extrema paciencia e inagotable apoyo me dio la fuerza para culminar este proyecto.

Resumen

La exploración de muestras histopatológicas es potencialmente una importante fuente de información para el diagnóstico de enfermedades, educación y entrenamiento médico. Convencionalmente, este tipo de muestras se observan a través de un microscopio óptico, sin embargo, recientes avances tecnológicos en el área de imaginología han favorecido el desarrollo de la microscopía virtual, un conjunto de herramientas que permiten la navegación de versiones digitales de placas histopatológicas, también llamadas Placas Virtuales (PV). Ya que las PV son un conjunto de imágenes de campos visuales microscópicos, éstas son imágenes de alta resolución. El gran tamaño de las PV representa un cuello de botella para cualquier estrategia de navegación, introduciendo considerables retardos en la presentación de la información al usuario y produciendo en consecuencia, navegaciones poco fluidas. Por esto, el desarrollo de métodos para acelerar la interacción en este tipo de aplicaciones impulsará su uso en áreas como la telepatología. Una de las aproximaciones más aceptadas respecto a este problema son las estrategias de predicción, en las cuales solo las regiones de interés son enviadas al usuario. Por tanto el reto en este tipo de aproximaciones es identificar las regiones relevantes para el patólogo, es decir aquellas que contienen significado semantico para el diagnóstico. Los modelos de atención visual intentan emular la percepción visual humana, permitiendo identificar las pocas regiones de una escena que contienen la mayor cantidad de información. Sin embargo, la complejidad de las imágenes histopatológicas hace que este tipo de modelos sea insuficientes para identificar estructuras relevantes dentro de la muestra y se hace necesario utilizar nuevas fuentes de información. En este trabajo se presenta una novedosa aproximación estadística que permite construir un mapa de información relevante de la PV con el fin de permitir navegaciones fluidas en un microscopio virtual. Este trabajo utiliza como una primera aproximación a las regiones relevantes un mapa de saliencia obtenido a partir de un modelo de atención visual "bottom-up", el cual es modificado por conocimiento "top-down.ºbtenido a partir de navegaciones de patólogos expertos sobre PVs. Las regiones de la PV son estructuradas en un grafo tipo arbol, de acuerdo a su nivel de relevancia, en esta estructura los nodos representan regiones espaciales de la PV y los arcos representan relaciones de inclusión entre ellas. El método presentado alcanza en promedio, una capacidad de predicción del 67 % de regiones de interés que fueron visitadas en una nueva navegación, asumiendo un caché de solo el 20 % del tamaño total de la placa virtual. Adicionalmente, con respecto a un método más simple, i.e. sin la representación del grafo, el método propuesto mejora en un 3 % la capacidad de identificación de regiones relevantes, lo cual es significativo en este contexto, ya que un porcentaje como este puede representar regiones de megapixeles en una PV.

Palabras clave: Exploración de mega-imágenes, Representación de imágnes médicas, Microscopía Virtual.

Abstract

Exploration of complete histological samples is a very important source of information for diagnosis, learning and medical training. These samples have been conventionally observed through an optical microscope, but recent advances in imaging technology have enabled the development of virtual microscopy, a set of tools that allows navigation of digitized versions of tissue samples (virtual slides or VS). Since VS are the set of several microscopic visual fields they are high resolution images with huge storage spaces. The size of these images represents a bottle-neck for any navigation strategy, introducing considerable delays in displaying information to the user, which results in non-fluid navigations. Development of methodologies for accelerate interaction with such volume of data, represents a reinforcement of its use in real applications as telepathology. One of the most accepted methodologies regarding this problem are based on prediction policies, in which only the regions of current interest are provided to the user. Then, the challenge in these approaches is to identify relevant regions, which for VS exploration enclose semantic meaning for diagnosis. Classical visual attention models emulates the human visual perception allowing to identify relevant regions with highest quantity of information. However, complexity of histo-pathological images make insufficient these models to identify all relevant structures in the sample, furthermore new information sources that lies the identification of salient features is required. In this work is presented a novel statistical hybrid approach which enable to build up relevant information VS maps to achieve fluent virtual navigation. This approach, uses as first relevance approximation, a saliency map obtained from a bottom-up visual attention model, which is modified by a top-down relevance information obtained from actual VS explorations of expert pathologists. Image regions are organized in a graph structure according it's levels of relevance, in this structure nodes are related to spatial regions and edges represent belonging relationships between them. The method herein introduced achieve in average, a prediction capacity of 67 % of RoIs that were visited in a new navigation, assuming a cache of only 20 % of total size of VS. Additionally, with respect to a simpler method, i.e. without graph representation, graph-based method outperforms RoIs identification capacity in about 3 % which is significant in this context since such a percentage could represent regions of megapixels of a VS.

Keywords: Mega-images exploration, Representation of medical images, Virtual Microscopy)

Contenido

1 Introduction

Histopatological analysis is an essential source of information for diagnosis and detection of disease based on the examination of structural and/or chemical alterations of tissues. For this, thin tissue samples are placed over glass slides, and stained for further exploration using typically an optical microscope. Nowadays, these tissue sample slides can be digitized and stored as Virtual Slides (VS) [39] avoiding the natural damage of real samples, and also introducing new informatics methodologies, which allows their use in collaborative scenarios such as professional education and training, remote inter-consulting and telepathology.

Virtual microscopy is a domain which gathers together a set of devices and techniques that allow acquisition, storage and navigation of these VS [40]. Typically a histopatological specimen is digitized at the higher possible magnification to provide the pathologist with a maximum amount of information. Due to this, the result are high resolution images (mega-images), which generally require hard disk spaces of the order of the Gigabytes [53, 40], e.g. a typical digitization of a 1 cm^2 glass slide using a $\times 20$ objective magnification, results in a reconstructed image of 45000×32000 pixels and 4.3 Gbytes [22]. Thereby representing an actual challenge in terms of storage, transferring and navigation and the main reason of the it's limited use in clinical applications. For instance, in telepathology applications the remote and share access to VS is doing via typical client-server architectures, but the considerable amount of information makes the band channel represents a bottle-neck for the transfer process [15]. This fact generate delays in the display of the images causing the exploration of VS not as fluent as in the optical microscope. Interacting with such quantity of data, under the restriction of the technological and medical limitations, is still an open problem.

Improvement of the real-time interaction with VS has been addressed from different points of view, namely storage, transmission or management of client resources. Regarding the storage perspective, the classical approach concerns the data compression, taking advantage of the visual redundancy to reduce the needed data to reconstruct. Although there exist methods that reach high data compression with very little quality loss, these compression methods are not completely accepted within the medical community, event though the resource management is far from being optimal [25]. From a transmission standpoint, some protocols have focused on progressive methodologies that use multi-resolution structures to gradually increase the level of detail, improving the navigation quality [13, 37]. However, these structures are usually applied to the whole image and VS are images whose resolution exceeds by far the available displaying technologies, introducing thereby the necessity of mechanisms for transmission of specific locations. The most popular method has been the image tiling, a

simple technique that somehow divides the image and therefore transmits specific portions of the image, providing the user with the desired information. Nevertheless, most of these transmission protocols (temporal and spatial) require extra storage resources which limits its actual use in clinical scenarios [5, 36]. It is worthy to notice that the mentioned approaches are not mutually exclusive, so there exist hybrids methods which allow both compression and transmission of different resolutions and different locations of the mega-image and then permitting faster interactions [8, 22]. Although these strategies are able to provide the user with specific information, they are totally reactive, i.e. information is sent to the user only when it is specified on a query. Then, attempting to a smartly management of resources, approaches based on prediction of user's behavior as cache and pre-fetching strategies have been proposed [30, 8, 22, 16].

This work introduces a novel approach for VS representation based on the detection of the most salient regions, i.e, the informative regions for diagnosis. An exploration of a VS can be though of as a discrete observational path composed of jumps between different RoIs. The problem of determining such regions is related to the available information, that is to say, an expert pathologist can manually draw these RoIs [6, 38], but obviously this is not a realistic scenario in actual virtual microscopy applications, implying the necessity of automatic methods. Classical methodologies for the determination of RoIs consist on the detection and quantification of the edges [24, 50]. Although histopathological images are crowded of highly textured regions, not all of them enclose relevant information for diagnosis. A natural way to characterize the most relevant region in this kind of images using low-level information are Visual Attention Models (VAM). These models are based on characteristics of visual human perception to identify regions of the image which contain the highest quantity of information [33, 47].

The proposed approach stars by computing a bottom-up saliency map obtained from a visual attention model. This prior information is modulated by top-down information, constructed from actual virtual slide explorations of expert pathologists. The resulting posterior probability function is a relevance map of the regions in the image. This work evidence the importance of use hybrid information to achieve a better RoI estimations from VS.

1.1. Thesis Outline

As introduction of this work, in the chapter 2 is presented the research problem and a review of the different approaches to accelerate the navigation of high resolution images and its applicability to virtual microscope systems.

In the chapter 3, we introduce how the information contained in the image given by the dye and the texture in the structures, define diagnostic regions of interest. Then, we make a comparison of four bottom-up visual attention models according to its behaviour to identify regions of diagnostic interest in histopathological images. The results given by the algorithms were compared with fields of vision taken from experts pathologists during navigations.

Finally, we discuss about the comparative analysis done.
In the chapter 4 we introduce a flexible part-based representation of a virtual slide which allows to determine a relevance map. The presented strategy uses a first relevance approximation, generated by a prior visual attentional model, which is modified by a posterior probability function, constructed from actual virtual slide explorations of expert pathologists. The posterior probability is obtained from the observational path drawn from actual navigations. The semantics of this path is captured in graph structure whose nodes are related to spatial regions and represent the relevance weight of each of these regions. The proposed strategy was evaluated by comparing the percentage of regions defined as relevant by the two methods.
Finally, in the chapter 5 present some conclusions about this work an potential future works.

2 Navigation of large-scale medical images: a survey

Exploration of histopathology slides is an important source of information for medical diagnosis and training in anatomic pathology. These slides are very thin fixed tissue samples, placed upon a glass slide and stained with specific dyes. Conventionally, they are explored under an optical microscope with the aim of analysing the tissue architecture and microstructure. Virtual microscopy is a new domain which gathers together a set of devices and techniques that allow navigation of digitized versions of tissue samples. These so-called Virtual Slides (VS), besides solving specific problems of the actual glass slides such as physical deterioration or handling, are very useful in collaborative scenarios such as professional education and training, remote inter-consulting and telepathology [11, 40, 39]. Typically, a VS is digitized at the higher possible magnification to provide the pathologist with a maximum amount of information. The result is a very high-resolution image (mega-image), usually of the order of the Gigabytes [53, 40], thereby representing an actual challenge in terms of storage, transferring and navigation. Remote and shared access to these virtual slides is usually reached under client-server architectures and depending on the required information, the channel bandwidth may represent a bottleneck, generating annoying delays when displaying the visual information. These delays result in seamless navigations, becoming the main reason for the limited use of virtual microscopy systems in the clinical area.

Improvement of the real-time interaction with VS has been addressed from different points of view, namely storage, transmission or management of client resources. Regarding the storage perspective, the classical approach concerns the data compression, that take advantage of visual redundancy to reduce the needed data. Compression could be lossy or lossless: In lossy approach, there exist methods that reach high data compression with very little quality loss, nevertheless these compression methods are not completely accepted within the medical community [2, 52]. In lossless approach, although there is no lack of information the resource management is far from being optimal [37, 25]. From a transmission standpoint, some protocols have focused on progressive methodologies that use multi-resolution structures to gradually increase the level of detail, improving the navigation quality [13, 37]. However, these structures are usually applied to the whole image and VS are images whose resolution exceeds by far the available displaying technologies, introducing thereby the necessity of mechanisms for transmission of specific locations. The most popular method has been the image tiling, a simple technique that somehow divides the image and therefore transmits

specific portions of the image, providing the user with the desired information. Nevertheless, most of these transmission protocols (temporal and spatial) require extra storage resources which limits its actual use in clinical scenarios [5, 36]. It is worthy to notice that the mentioned approaches are not mutually exclusive, so there exist hybrids methods which allow both compression and transmission of different resolutions and different locations of the VS [8, 22]. Although these strategies are able to provide the user with specific information, they are limited regarding the large demands of resources required in actual navigations. A smart management of resources has been approached by attempting to anticipate the user needs, modeling the expert observation path to predict and pre-fetch the required portions of the image [30, 8, 22, 16].
This chapter presents a survey of the different approaches to speed up the VS navigation and its applicability to virtual microscopy systems. The chapter is organized as follows: Section 2.1 presents a set of strategies for accelerating exploration of VS, divided in three main categories according to its role within a server-client architecture, namely storage, transmission or management of client resources. Finally, section 2.2 presents some discussion and conclusions.

2.1. Accelerating the navigation of mega-images

The recent accelerated development of image acquisition systems has introduced in many different fields a variety of high resolution images, for instance aerial, satellite and medical images [31, 29, 22]. This in due turn has led to the actual necessity of smart methodologies to interact with such quantity of visual data.
Under a client-server architecture, there exist a large diversity of approaches which address the problem from different perspectives: (1) To minimize the storage sizes with compression algorithms, (2) To provide the user just with the required information, either by resolution or spatial preferences, via optimal transfer protocols or (3) To efficiently manage client resources, for instance attempting to predict the userÂ´s navigation paths. Figure **2-1** shows this classification. Keeping these ideas on mind, in this section we will present a review of the state-of-the-art techniques to accelerate the interaction with large scale images and their applicability to virtual microscopy systems in clinical and diagnostic activities.

2.1.1. Compression

One of the most popular strategies to accelerate interaction with large scale images has been the client-server exchange of compressed information. This method minimizes the data needed to reconstruct the image, taking advantage of the visual information redundancy. There are two types of compression from a reconstruction standpoint: lossless, which results in perfect image reconstruction, and lossy, with a variable level of losses in the final reconstructed image and usually much higher compression ratios. Although there exist lossy compression

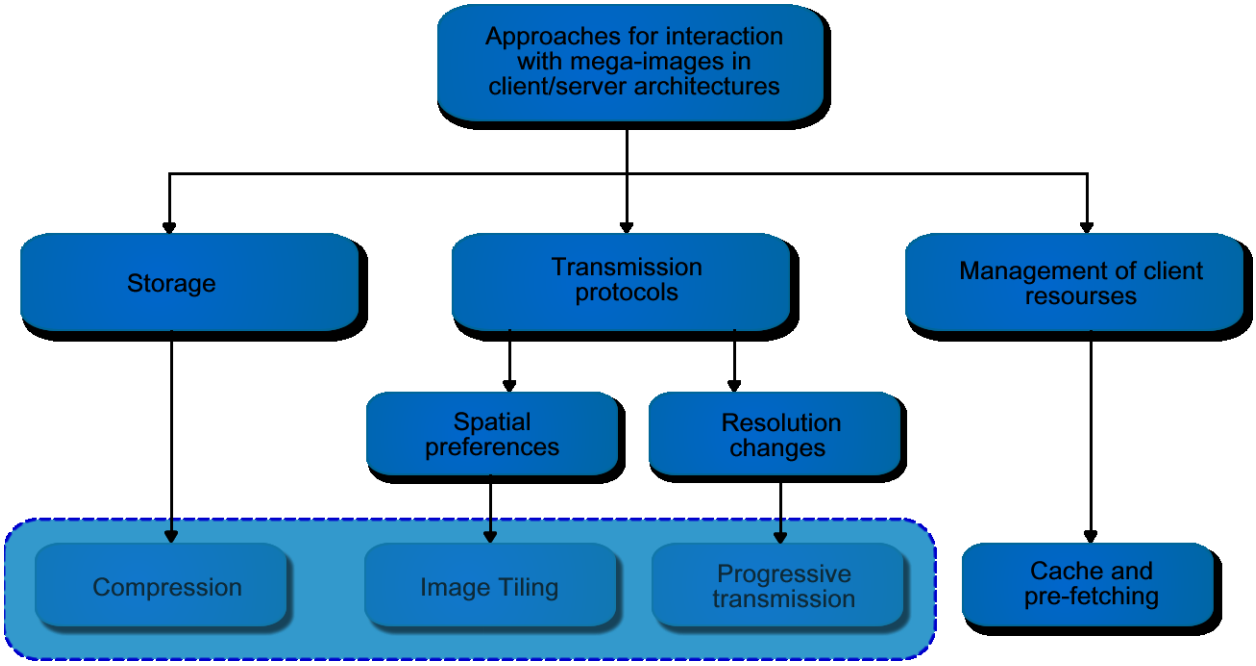


Figura 2-1: Dashed rectangle relate the set of techniques which could be incorporated in a single hybrid method.

strategies that reach high data compression with very little loss of quality, Medical environments are traditionally very reluctant to the use of these methods and Medical Doctors do not accept any kind of loss of information, because of the harsh regulations, and the possibility to mislead actual diagnoses [2, 52]. In consequence, lossless strategies have been broadly used in medical image compression applications [25, 56].

Widely used encoding formats such as GIF and PNG allow lossless compression, but the offered compression ratios are quite small, a clear drawback in medical applications in which hundreds or thousands of studies are daily acquired [37]. Different approaches have developed selective compression strategies, that is to say to apply different compression thresholds to different image regions with several relevance levels. Regions of Interest (RoIs) may be identified by different methods: texture analysis [24, 25] or comparison with images previously annotated by specialists [56], among others. The compressed images with such strategies must be totally transmitted and decoded, and for the explored cases of use, i.e., Computational tomography (CT) or Magnetic Resonance (MR) images (256×256 pixels), the time required for displaying them is still realistic in clinical scenarios. However, the same approach is unthinkable in virtual microscopy since the virtual slides are usually arrays of the order of Gigapixels [53, 40].

2.1.2. Progressive Transmission

Delays between a particular image request and the image display are annoying [37] and affect negatively the work flow in diagnosis tasks. Minimization of these delays has been approached via transmission protocols able to give the user information of the image at the very requesting moment. This so-called progressive transmission is based on multi-resolution approaches that initially send a rough version of the image and progressively increase the level of detail.

The classical data structure to progressively transmit has been the pyramid. The first step of this progressive strategy is the off-line construction of a pyramid, consisting of 0 to N levels of resolution from the original image, being 0 the coarser and N the finest level of detail. When the image is requested by the user, the 0 level is first sent to the client, followed by other higher levels, until the N level. The construction of this multi-resolution pyramid is a reducing process for which the image is firstly tiled in non-overlapping windows of $n \times m$ pixels (usually called child pixels), and every window at the k-th level is represented by one pixel at the $(k-1)$-th level (father pixel), as shown in figure **2-2**. For achieving so, many different methodologies have been introduced, like the mean pyramid, case in which every father pixel is the average of a $n \times m$ block of child pixels. The difference pyramid is built as the difference between the father pixel and each of its children, decomposing the original space into two complementary sub-spaces. Likewise, the Gaussian-Laplacian pyramid consists in building a Gaussian pyramid and the Laplacian is obtained as the difference between successive Gaussian levels. Other methods use transform matrices, for instance the forward Hadamard transform, which being a generalization of a discrete fourier transform, execute linear operations between child pixels using an orthogonal and symmetric matrix Goldberg1991. Again, the purpose of above techniques is to decompose the original information into two complementary spaces: low and high frequencies. For reconstruction of original image, only is required the highest level of the pyramid and the successive application of the inverse transform over the coefficients Goldberg1991. Although pyramidal progression methods accelerate the interaction process, this structure increases in a 33 % the storage space with respect to the original image [42]. This drawback practically limits the pyramid method to applications with image small size, obviously much smaller than those used in virtual microscopy.

The additional space needed by the classical pyramidal approaches has led to the development of other methods that built the pyramid [13, 26]. This so-called *reduced pyramids* or *virtual pyramids* achieve this structure by replacing values in the original mesh by some estimations (see Figure **2-2**), for instance the reduced-sum pyramid: in this case the sum of the values of $n \times m$ block is assigned to one of the child pixels and hence the value of the child pixel at the i-th level may be replaced by the value of father pixel and so on. Information is then transmitted in the other way around, i.e., the method uses the inverse operations [13]. In medical images, Kim et. al propose a reduced-difference pyramid which

allows lossless compression and includes some type of granularity. By adapting the information in sub-pyramids the method transmits with priority sub-pyramids that encode RoIs, defined by zoomed areas [26]. Although this reduced structures have the same size as the original image, pyramidal methods send successive versions of the image, making it redundant for transmission and representing a low efficient use of the channel.

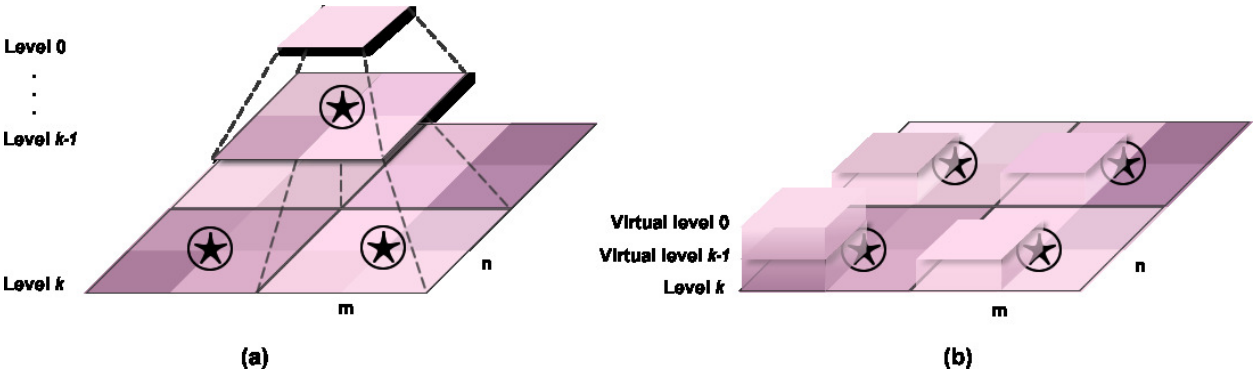


Figura 2-2: Two ways for building a pyramid structure. (a) Classical method, each level is a new version of the original image. ($\star$) represent different mathematical operations as mean, sum or difference between the overlapped pixels. (b) Virtual pyramid, the result of the corresponding operation is assigned to nodes in the same mesh.

Redundancy of pyramidal methods was successfully covered by wavelet encoding. This approach allows compression and progressive transmission, encoding the image into a base coefficient and a large number of detail coefficients. For doing so, a discrete wavelet transform (DWT) is applied to the original image, which in practice consists in applying a set of low-pass and high-pass filters to split the image in sub-bands while the image is downsampled to obtain different resolution levels. Low-frequency sub-bands contain most of the information of the original image and high-frequency sub-bands contain the details [48]. For transmission, coefficients of each sub-band are progressively sent to the client who performs the inverse transform and reconstruction. Each time, the incoming coefficients are the complementary information of what has been previously sent [37, 48], allowing a optimal use of bandwidth. Because of the great advantages in compression and transmission, wavelet strategies have been widely implemented for image storage and transmission, as explained in section 2.1.3.

Progressive transmission definitively has accelerated interaction, nonetheless most of these structures are intended to encode and transmit the whole image. It is true that image acquisition systems have improved their performance, but VS far exceed display technologies, making necessary the visualization of such images by regions, according to the required level of detail. Provided that relevance is not homogeneously distributed along the whole, it results crucial to the improve the VM interaction, the development of methodologies which

allow independent access to different image regions.

2.1.3. Image Tiling

It has been demonstrated, from a human visual perception standpoint, that there are image regions which are more informative than others [20, 45, 27]. Accordingly, more efficient transmission strategies require a flexible interaction, allowing the user the possibility of accessing just the interesting image locations [5]. In this sense, the proposed strategies generally include methods for: (1) splitting the image in independent regions and (2) having random access to these regions. Some of these strategies allow simultaneous implementation of progressive transmission algorithms, already presented in section 2.1.2.

Image tiling is an approach for fast interaction with large-scale images, broadly used in digital geographic systems where visual information is very extensive, since maps must contain topographic representation of the whole earth. In an off-line stage, maps of a particular geographic area are drawn a certain number of times to obtain different resolution levels of the same area. Each map is then tiled, indexed and stored. Afterward, in the on-line stage, the tiles sent to the user depend on the particular required location and position, which are displayed in a conventional web interface [36, 35]. More realistic digital geographic systems have been implemented using satellite and aerial imagery, instead of the drawn maps.

More related with the medical domain, Mikula et. al proposed a tile-serving algorithm applied specifically to histopathological exploration. Discrete magnification levels are obtained by a sub-sampling process of a more VS detailed version. Each level of the resulting pyramid is tiled, indexed and sent to the user as for the digital mapping system explained before [32]. Based on the same methodology, Triola et. al developed an open-source virtual microscope, based on a map engines and additionally equipped with educational features [46]. Although these approaches have shown very fast tile delivering rates (around 85 ms [32]), it is worthy to remark on the already mentioned issue of extra storage space. As every resolution level is a new image, this represents a challenge in a clinical scenario, because of (again) the huge size of the digital histopathological image.

This is why the wavelet representation has been included in modern strategies since this kind of encoding avoids the redundancy, thereby leading to a better use of resources. For instance, Chang et. al. compress the image using a (Haar) wavelet transform and splits the transformed image into blocks, each containing certain region of the image [5]. This strategy allows to send to the client side only those blocks contained in the RoI and to progressively reconstruct the whole block or partial versions with less coefficients, if the user decides that the current detail level is enough. Likwise, Wang et. al propose a web-based virtual microscope which compress high resolution pathology images, using the Daubechies wavelet transform and Huffman coding. After compression, images are split in blocks to differentiate regions. When the user performs a query, the server searches the necessary data and reconstructs the sub-image, which is coded in JPEG format and sent to the client. Wang

reports a transfer time of about 1 second, but using lossy compression ratio of 5 : 1 [49]. Advantages of the wavelet representation have also led to its introduction in new standards such as JPEG2000. This standard was developed by the Joint Photographic Expert Group and is mainly based on: (1) the Discrete Wavelet Transform (DWT), introducing a multi-resolution encoding for progressive transmission either with lossy or lossless compression and (2) the Embedded Block Coding with Optimised Truncation (EBCOT), allowing progressive quality recovery and random access to data [43, 22], figure **2-3** shows the encoding process of JPEG2000 format. The multi-dimensional granularity (resolution levels, regions of interest and quality levels) of this standard has been widely used not only for exploration of large-scale images[29, 7, 22] but also for adaptive compression and efficient storage [12, 17].

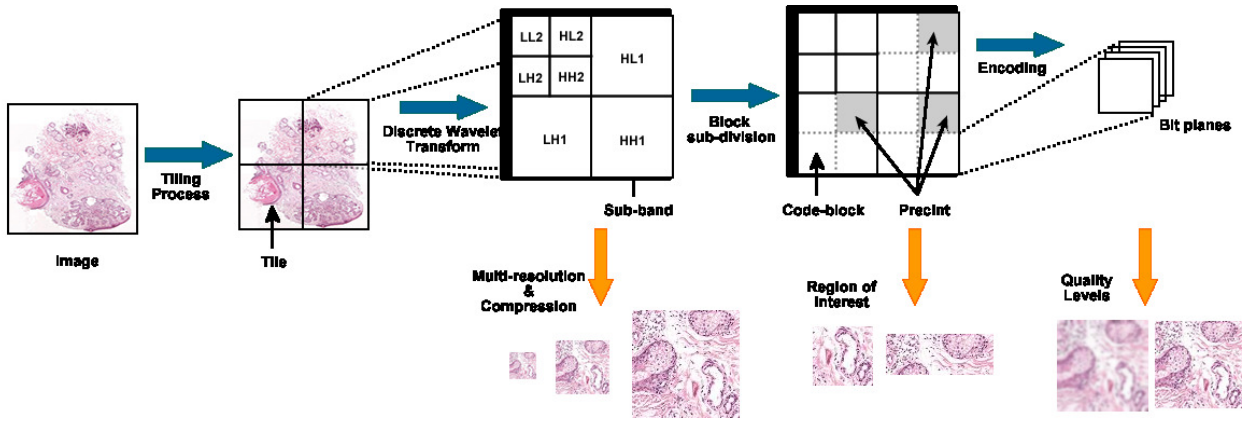


Figura 2-3: Tiling process is an image subdivision which allows independent processing of different parts of the image. DWT makes a frequency decomposition of each tile in sub-bands, allowing a multi-resolution representation and compression. Block sub-division allows configuration of coefficients in precincts, which correspond to the compressed data of different resolution levels associated to a particular image spatial zone. Encoding data in bit planes provide different quality levels

All these advantages have converted the JEPG2000 format in a more and more popular tool for optimal interaction with large-scale images, either in medical and natural scenarios [8, 16].

2.1.4. Cache and Pre-fetching

Either with pyramids or encoding formats as JPEG2000, it is possible to accelerate the interaction by sending to the user only those regions of the image with certain level of interest. Nevertheless, the so far reviewed methods start to send information once they receive a particular request. On the contrary, a variety of approaches, based on the prediction of the expert observation paths, have been proposed and applied to cache and pre-fetching methodologies.

Cache is a memory space in which current information is stored since it is likely this information will be requested in the future of the same navigation. Provided that this space is limited, the data availability depends on the replacement policies. Classical caching algorithms are Last Recently Used (LRU), for which the last stored element is used and the first is ruled out, and Most Recently Used (MRU), which performs the opposite of LRU strategy and an element already used will not be used in the future, and the Last Frequently Used (FLU), for which the replaced element is the one that has been more frequently used [30, 15]. Soft-Caching strategies are based on exploitation of both spatial and temporal progressive transmission, allowing to store only portions of the image to be visited in the future [51]. More task related approaches that attempt to optimize the cache space have been proposed. Gómez et. al present a soft-caching algorithm for virtual slide navigation with a replacement policy that is a probabilistic model constructed from actual pathologist navigation patterns. This method exchanges the smaller unit of information encoded by JPEG2000, the packets, which in the cache are replaced according to the panning or zoom preferences during the navigation [15].

Soft-cache have been successfully complemented by pre-fetching strategies, i.e. the anticipated uploading of information which probably will be requested [8]. Such strategies are prediction policies, either general or task associated. General approaches usually take into account the last user actions, for instance a zoom in/out action may lead to a subsequent zoom action in the same direction [31]. Other method was presented by Descampe et. al, a system of 6 degrees of freedom (zoom in, zoom out, up, down, right and left) is modeled as a probability a for the event that two consecutive actions are the same and a probability $(1 - a)/5$ if the sequence of events are different [7, 8].

Gómez et. al proposed a predictive method based on the velocities profiles of the pathologist's observation path. The underlying hypothesis is that velocity is determined by agonist (motion) and antagonist (inhibit motion) neuro-muscular subsystems. Each sub-system can be represented as a log-normal function, allowing a representation of the velocity profile as a probability distribution function that involves inverse variation of these functions [14]. This method is devised for the scanning navigation phase, which corresponds to the exploration of the digital slide, a period in which examination requires the pathologist smoothly moves around the VS.

More specialized algorithms predict the RoIs during the navigation process have implemented techniques as computational visual attention models (CVAM). This kind of models attempt to emulate the selectivity of human visual system, attempting to find a semantic meaning of a scene by identifying a little set of interesting points [45, 27]. CVAM can identify RoIs from low level image features (bottom-up models), using prior knowledge of the task (top-down models), or both. Regarding pre-fetching strategies, bottom-up models have been introduced. Zang et. al proposed to use a measure of multi-scale contrast to off-line find possible interesting regions in the image, which can then be pre-loaded. A low resolution image version is presented to the user who draws the RoIs. The algorithm compares the off-line selected

regions with the preference of the user, if they match the information corresponding to that RoI is displayed, otherwise a local saliency map is calculated around that point [54]. Other approach using attentional bottom-up model is presented by Ip and Varshney who use a modified Ittiâ^{TM}s model to identify RoIs in landscape images [21]. These new approaches may extend the prediction policies in medical applications such as navigation of histopathological slides, a task which is highly guided by pathologist's knowledge and experience.

2.2. Conclusions and perspectives

A main concern when promoting the extensive use of virtual microscopy systems in diagnostic activities, is how to provide the pathologist with a system which allows fluid navigation, as usually carried out in a classical light microscope. For achieving so, it is fundamental to ensure efficient and flexible transmission of VS, avoiding delays when displaying information, typically generated by the huge size of transmitted VS. In this review we have presented a set of approaches to accelerate interaction with VS under client-server architectures and we have analyzed its applicability in virtual microscope systems for clinical applications.
Classically, compression has been the method for minimizing the quantity of transmitted data: on the one hand lossy compression algorithms reach high compression ratios with little but uncertain loss of information, likely unacceptable in medical applications, but on the other hand most of the lossless image encoding formats offer a very small compression rate. Hybrid approaches have shown better results for rapid exploration of VS. That is the case of tiling images, along with progressive transmission methods, which have been successfully used in cartography applications. Unfortunately this kind of approaches requires a lot of storage space. Other hybrid approach has consisted in taking advantage of the random access to data and progressive transmission of JPEG2000 for optimizing the information management with pre-fetching algorithms. This approach, setting aside the advantage of enabling compression without information losses, allows to anticipate transfer and load of specific VS regions. The effectiveness of this kind of methodologies lies in the capability of the prediction policies to store in memory the information to be used in the future. In that sense, it is required development of new strategies able to capture the patterns of the particular task. This is of course a very open problem since navigations store the high level pathologistâ^{TM} knowledge, interacting with a wide variety of tissues and pathologies.

3 Visual attention models for histopathological regions of interest

In exploration of histopathological images, expert pathologists are capable of rapidly find regions with larger semantic meaning, i.e. Regions of Interest (RoIs) [14, 38]. For virtual microscopy applications, the appropiate segmentation of these RoIs could be used to accelerate the navigation of Virtual Slides (VS). The challenge of determining RoIs in histopathological images is related to the available information, that is to say, an expert pathologist can manually draw these RoIs [6, 38], but obviously this is not a realistic scenario in actual virtual microscopy applications. Not only for being a time consuming activity but for the inter-observer variability [10, 3], this implies the necessity of automatic approaches.

The most classical methodologies to find RoIs in natural images using only low level characterization consist on the detection and quantification of the edges [24, 50]. Although histopathological images are crowded of highly textured regions not all of them enclose relevant information for diagnosis. A natural way to characterize the most relevant region in this kind of images using only the pixel information are Visual Attention Models (VAM). This models are based on characteristics of visual human perception to identify regions of the image which contain the highest quantity of visual information. In general, there are two types of visual attention models, "bottom-up" models in which it is assumed that the attention of the observer is guided by the information in the image and "top-down" models, in which attention is directed by previous knowledge of the task [55].

The main goal in this chapter is make a comparison of four bottom-up visual attention models according to its behaviour to identify regions of diagnostic interest in histopathological images. We want to stablish how the information contained in the image given by the dye and the texture in the structures define diagnostic regions of interest. The results given by the algorithms were compared with fields of vision taken from experts pathologists during navigations. This chapter will be presented as follows: In section 3.1 we explain the selection criteria of compared visual attention models and a brief explanation of them. In section 3.3 we present the obtained results and discussion. Finally, in section 3.4 we present the discussion of this comparative analysis.

3.1. Bottom-up Visual Attention Models

With the aim of make structures differentiation in tissue samples, they are stained with different staining protocols depending on the anatomical concept that we want to highlight. This characteristic of histopathological slides affects the way in which pathologists navigate the sample. In general, she or he looks the sample in a low resolution in order to identify structures of interest which are lately look with more detail in higher resolutions. This search visual patterns allow to intuit that it is possible to identify automatically regions or structures which are relevant for a pathologist on a first stage of exploration of a histopathological sample [15].

Some psychophysical studies referring to human selective visual attention suggest that when we observe complex scenes there are two phases: The first one, is the pre-attentional phase, in which the scene is observed in a fast and parallel way, a map that represent the conspicuousness of the regions is obtained, this is called the saliency map. After that, there is a attentional phase, that is a serial process in which specific regions of the scene are visited, these regions are suggested by the saliency map obtained in the previous phase [27, 4]. With the aim of emulate what happens in pre-attentional stage, different models to built saliency maps have been presented, from biological-based approaches [23] to computational ones [1] and hybrids [19, 18].

To have an idea of how much low level features in histological images guide the identification of regions important for the diagnosis, assess the effectiveness of three "bottom-up" visual attention algorithms. These algorithms represent recognized models in the literature for natural images, that in general, analyse the features in images such as color, given by the staining of the structures, the intensity variations between the structures and texture components. The algorithms are:

Achanta et. al model [1] is a computer model that calculates the salience map using analysis of spatial frequencies. For this, the image is transformed to color space $[L, a, b]$ in which information takes advantage of the double opposition of color and brightness of the image. In each component of this space, the method calculates the average value of the pixels and makes the difference between this value and each pixel value of a blurred version of corresponding component. Blurred version is obtained by applying a band-pass Gaussian filter, see figure **3-1**. The authors meet the requirements for the method be able to emphasize and highlight large protruding objects in a uniform and well defined edges of these objects. Although protruding objects in the images of histopathology are the largest structures in all these, the second requirement suggests that the method is suitable for identifying well-defined structures.

Itti et. al model [23] is one of the most referenced models in the literature, this model is inspired by the visual biological model of Ulman and Koch [27]. It is based in the extraction of color, orientation and intensity features of the image, this features take into account the information that dye and texture of structures can give in histopathological images. Itti et.

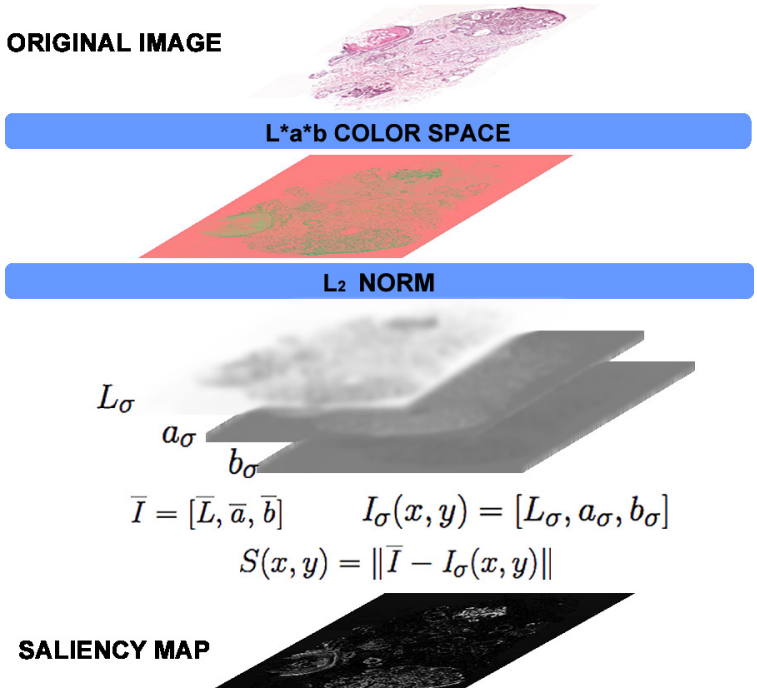


Figura 3-1: Algorithm of frequency-tuned salient region detection model proposed by Achanta et. al

al model is based in a theory which stablish that each feature in the scene compete in a independent way for drawing the attention of the observer. First, the algorithm built nine spatial scales from each feature (color, orientation and intensity) using Gaussian filters, then they do a multi-scale analysis in order to determine the differences between a pixel and its neighbourhood, this process give as result a so called "Feature maps". This maps are normalized in a specific scale giving as result the "Conspicuity maps" for each feature. The linear combination of this maps gives the saliency map. See figure **3-2**.

Harel et. al model [19] obtain feature maps according to Itti et. al model explained above, from this maps they built that they call "Activation maps". For doing this, each feature map is represented by a strong connected graph in which every pixel in the map is a node and is connected with all the other nodes (pixels), the weight of the edges is defined by a dissimilarity measure between the node values and their spatial distance. A normalization of the weights in the outgoing edges of each node is performed, this allows to interpret nodes as states and edge weights as transition probabilities. Using a Markovian approach, the method define the probability of conspicuity of a node according to how unusual is the node in its neighbourhood. A second Markovian chain is applied to make normalization of activation maps. Finally, average of activation maps within each feature and the sum across features give as a result the saliency map. See figure **3-3**.

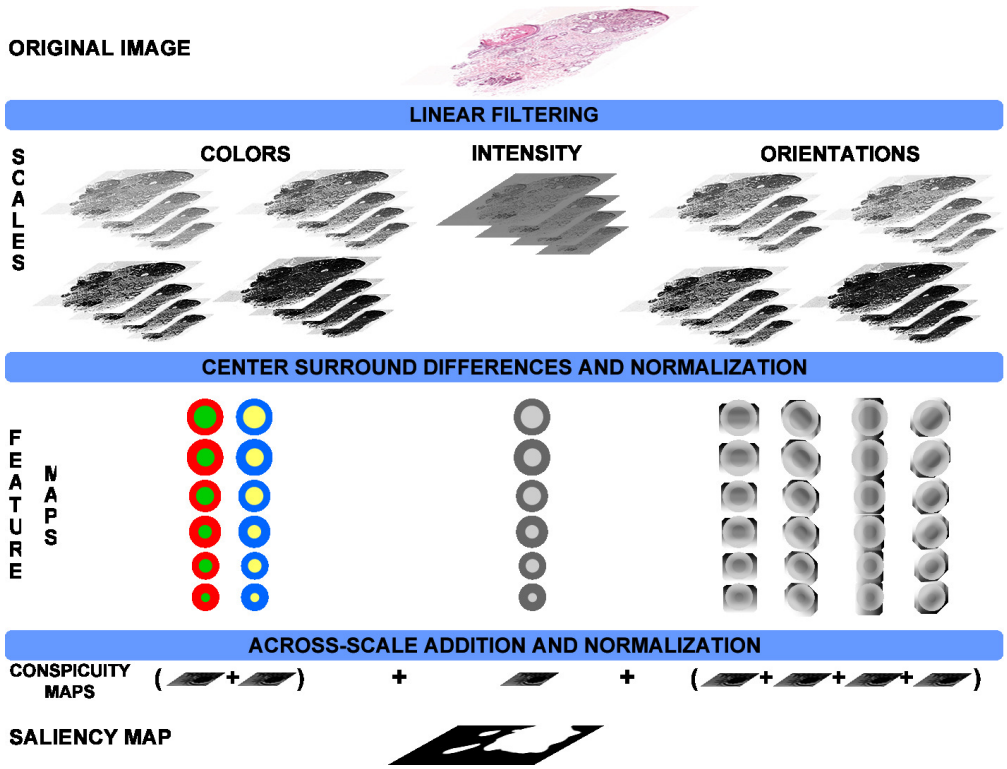


Figura 3-2: Algorithm of Saliency-based visual attention model for Rapid Scene Analysis proposed by Itti et. al

3.2. Experimental setup

To evaluate the performance of the algorithms in terms of their ability to identify regions of diagnostic interest in VS, regions determined by the VAM as relevant were compared with regions visited by the pathologist during navigation. We suppose that pathologists navigate visiting the diagnostically relevant regions, so a match between a mask of the navigation and a mask of salient regions obtained with each algorithm was performed.

The mask of salient regions was obtained as follows: having on mind that in saliency maps regions with high intensity values represent more conspicuous regions, we select a set of salient regions through binarization of the saliency maps using the Otsu method. With this method it is assumed that pixels in the image belong to two classes and the idea is to find a threshold that minimize the inter-class variance and maximize the inter-class variance [34]. The mask of navigation was obtained by setting value 1 for visited pixels and value 0 for non-visited pixels. Overlapping of both masks results in a match image, in which pixels were classified as $TruePositive(VP)$, $TrueNegatives(TN)$, $FalsePositives(FP)$ and $FalseNegatives(FN)$.

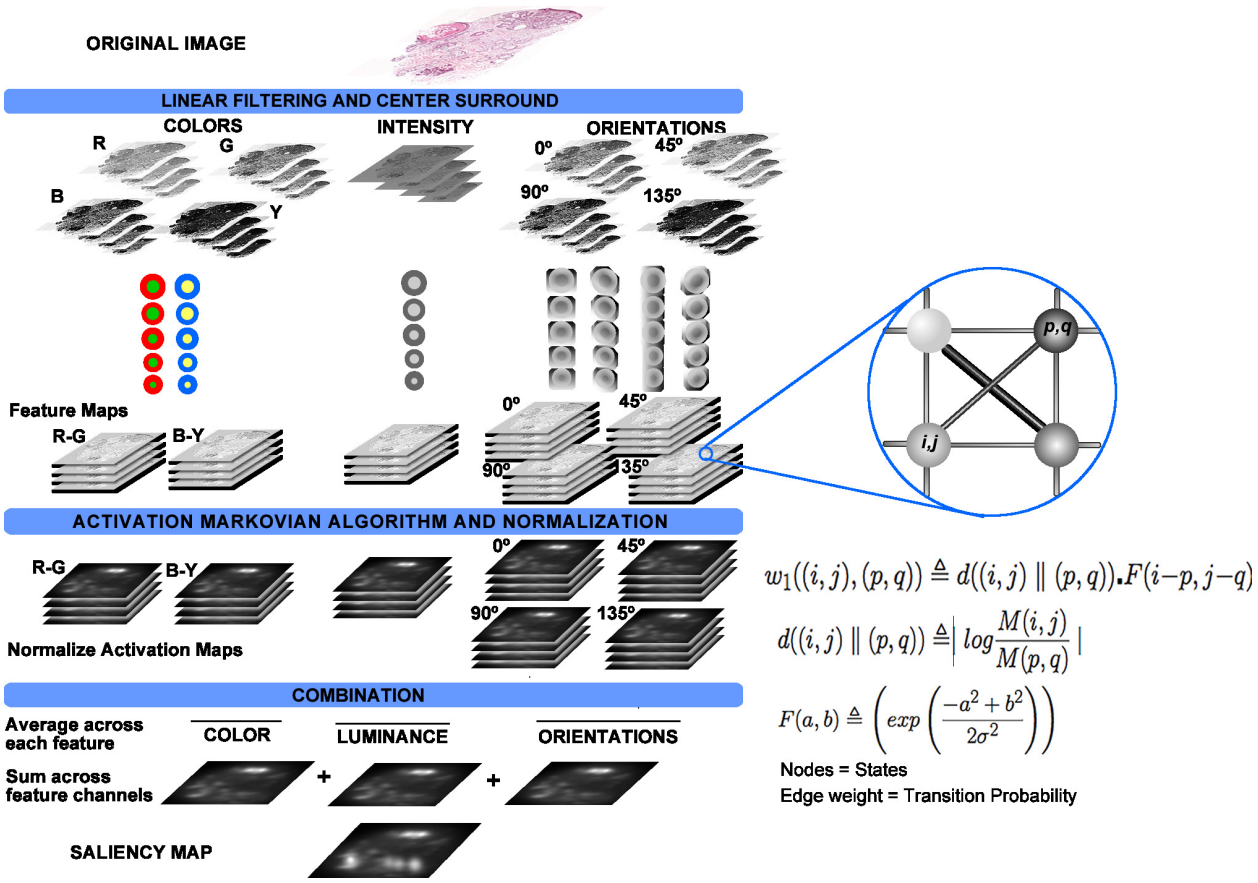


Figura 3-3: Algorithm of graph-based visual saliency model proposed by Harel et. al

From this, different measures was calculated: Precision (P) which specifies the fraction of pixels that the algorithm identify as relevant and they really are. Recall (R) which specifies the fraction of pixels that are relevant and were successfully identified by the algorithm. Accuracy (A) which indicates the capacity of the algorithm to identify if a regions is relevant or not when it really is. Specificity (S) which specifies the quantity pixels that are not relevant and were successfully identified by the algorithm.

Original VS were digitalized using a microscope consisting of an optic microscope LEICA DM LS with a NIKON Coolpix 950 camera attached. Four of the images are stained with Peroxidase technique and six of them are stained with Hematoxilin-Eosin technique. The navigations were recorded as a set of visual fields visited during the exploration of the slide as shown in Figure 3.4(b). Each field of view is a portion of the slide that is observed through a virtual microscope at a given instant and with a particular magnification. Four expert pathologists navigated 10 virtual slides.

Algorithms were tested with low resolution images in agree with the first stage of slide exploration of the pathologists. Images of 1/32 of original resolution of the mega-image was

used. Configuration parameters of the implemented algorithms are showed in table **3-1**.

Algorithm	Parameters	Value
Achanta et. al	—	—
Itti et. al	Features	$\{C, I, O\}$
Harel et. al	σ	0.02
	Features	$\{C, I, O\}$
	Feature maps weight	{1,1,1}

Tabla 3-1: Configuration parameters of implemented algorithms. Parameters C, I, O refer to Color, Intensity y Orientation respectively.

3.3. Results

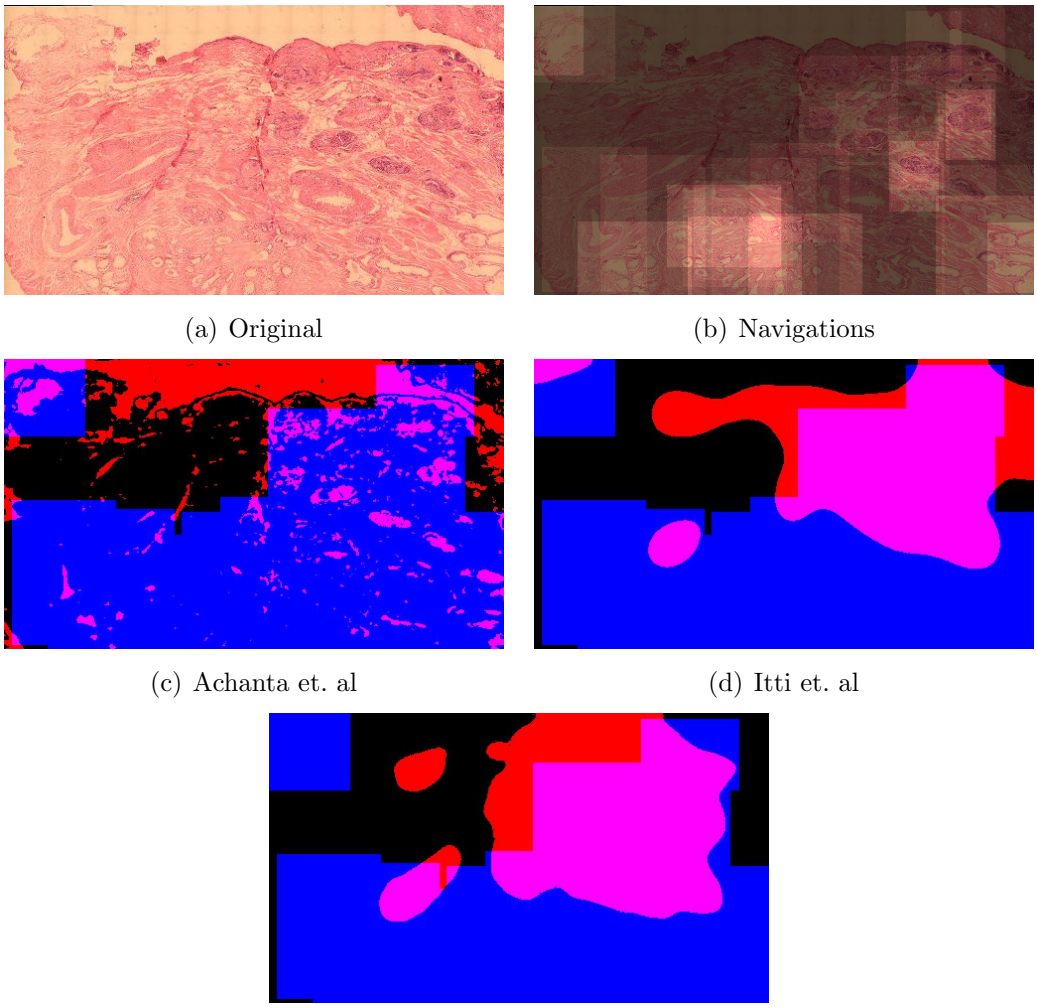
Fig. **3-4** shows an example of result match images obtained by overlapping the navigation mask from 3.4(b) and salient regions masks from the evaluated methods. Violet pixels are those that are relevant according to the navigation and were classified as relevant by the method, likewise, black pixels are those that are not relevant and were well classified by the method. Blue pixels are those that being relevant in the navigation were classified as non-relevant by the method, and red pixels are those that are not-relevant for the navigation but were classified as relevant by the method.
Numeric results of mean algorithms behaviour are specified in **3-2**.

Algorithm	P [%]	R [%]	A [%]	S [%]
Achanta et. al	42.97 ± 9.79	6.12 ± 0.62	36.75 ± 2.26	88.99 ± 1.80
Itti et. al	66.75 ± 5.19	26.17 ± 1.69	45.38 ± 1.11	79.78 ± 1.52
Harel et. al	75.04 ± 4.98	27.19 ± 1.75	47.70 ± 1.91	83.06 ± 2.27

Tabla 3-2: Results of implemented algorithms according to precision, recall, accuracy and specificity.

3.4. Discussion

Achanta et. al algorithm received the lowest scores compared to others, generally identified as relevant background and assigned a lower value of salience to the histological structures. This is because for this model are relevant objects at low frequencies and because of the Gaussian filters used, only preserves the high frequencies that define edges, eliminating the contribution



(a) Original

(b) Navigations

(c) Achanta et. al

(d) Itti et. al

(e) Harel et. al

Figura 3-4: Results of different visual attention algorithms applied to a histopathological image. (a) Image of histopathological sample stained with Hematoxilin-Eosin technique.(b) Lighter regions represent fields of view which were more visited or for much time by pathologists. In figures (c), (d), (e) black color represents TN pixels, blue color represents FN pixels, red color represents FP pixels and violet color represents TP pixels.

of texture histopathological structures. The method does not include information other than the color that represents a constraint to their application to specific histopathology images. On the other hand, Itti et. al algorithm, higher values of salience assigned to regions with varying intensity. Because these images are poor in color information, since the staining of the samples studied the spectrum limited to two colors (blue-brown for peroxidase and purple-pink for hematoxylin-eosin) this feature is low compared to the intensity.
Harel et. al implementation, obtained the best results, especially in *Precision*, exhibits behaviour similar to the algorithm of Itti et. with respect to the characteristics of color, orientation and intensity, however identifies regions of interest in a few fields of vision visited by pathologists, because the factor that governs the variance measure of closeness of the nodes in the graph. The outstanding results can be derived from what the authors call "central bias" that explains how the nodes are generally closer to certain central nodes of the image to the periphery. In most of the images tested fields of vision are concentrated in the center of the image.
According to metric used, Harel et. al algorithm shows a better performance to identify regions which are relevant for diagnosis in histopathology images. Nevertheless, its relative good performance may be due to the emphasis of the method in central regions, avoiding the peripheral regions which usually are part of the glass slide without tissue. In general, all tested methods present low performance to identify all the fields of view visited by the pathologist.
These results suggest that Bottom-up VAM are insufficient to identify regions in a VS that are relevant for the diagnostic task. Due to the diagnostic path is not guided only for the low level information in the image but in an important way by the pathologist's knowledge and training. This top-down information should to be included somehow in order to have accurate automatic identification of RoIs.

4 Estimation of Histopathological Image Relevance: A Hybrid Bottom-Up/Top-Down Approach

Approaches to improving real-time VS interaction require clever management of resources, prediction of the user's examination patterns using for instance strategies such as cache and pre-fetching [30, 8, 22, 16]. Nevertheless, exploration of histopathological slides during a diagnostic task is not an evident process, since visual information is not uniformly distributed, i.e., there exist specific areas in the tissue sample which are entailed with larger semantic meaning. The challenge in such predictive approaches is to determine the regions of the image with a high level of information in terms of diagnosis. The diagnostic task can then be thought of as an efficient process of interaction with a virtual slide, by means of which a pathologist defines an examination or observational path using more informative regions in the sample [28, 14, 38]. The problem of determining such regions is related to the high level of knowledge involved in such a process. Of course an expert pathologist can manually draw these RoIs [38], but obviously this is not a realistic scenario in actual virtual microscopy applications, in which the only possible information comes from the navigations. The question addressed in this chapter is then how to use the observational path defined by a particular navigation to determine which are the relevant RoIs, i.e., the relevance map of a VS.

The way in which attention is focused on specific regions in a scene, has been widely studied in psychophysical studies of the human visual perception. Different experiments [45, 44, 9] have shown that when the visual process is associated to a task, identification of Regions of Interest (RoI) involves a complex interaction between two complementary sources of information, on the one hand the bottom-up source, provided by integration of the information coming up from basic image features such as color, intensity or texture and, on the other hand, a top-down source whose main role is to feedback the primary bottom-up relevance map with semantic knowledge about the current task. Specifically in diagnostic exploration of histopathology slides, the bottom-up contribution is given by the low level image information produced by the tissue staining which allows to highlight particular structures. The top-down contribution is given by the expert, who is able to establish relationships between semantic knowledge and the tissue architecture and appearance [15].

The main contribution of this chapter is a flexible part-based representation of a virtual

slide which allows to determine a relevance map. The presented strategy uses a first relevance approximation, generated by a prior visual attentional model, which is modified by an evidence function, constructed from actual virtual slide explorations of expert pathologists. The posterior probability is obtained from the observational path drawn from actual navigations. The semantics of this path is captured in a graph structure whose nodes are related to spatial regions and whose edges represent the relevance path through the regions. The proposed strategy was evaluated by comparing the percentage of regions defined as relevant by the method with visited regions in new navigations.
This chapter is organized as follows. Section 4.1 introduces the part based representation under a probabilistic framework. Section 4.2 expose the experimental setup, section 4.3 shows the obtained results. Theses are then discussed in chapter **??** in which also conclusions are presented.

4.1. Materials and Methods

A successful extraction of RoIs in VS leads to a significant improvement of the interaction quality in diagnosis tasks in virtual microscopy systems [14]. The main objective of this work is to find an optimal VS representation to extract the most relevant RoIs in histopathological images. For doing so, this work introduces a Bayesian framework that integrates two VS relevant information: a bottom-up visual attention model (VAM) and a top-down evidence pathologist's exploration in diagnostic tasks. The Bayesian fusion strategy is analysed through two approaches with different complexity level: a patch-based model and a Graph-based approach.

4.1.1. Bottom-up flow

Bottom-up information comes from the low level captured features. Since the VAMs are robust representations devised to emulate the human visual system, they result appropriate as a first descriptor of the relevant regions. The representation herein used integrates information coming up from color, intensity and orientation features [19] to calculate a saliency map, a 2D function whose intensity values stand for the levels of relevance. In the present investigation it has been used a saliency map computed from low resolution versions of VS, since pathologists usually identify interesting structures at this magnification and use larger ones to further explore [38].

4.1.2. Top-down flow

Other important source of information that contributes to the identification of relevant regions in VS comes from the pathologistâ$^{\mathrm{TM}}$s expertise. From this point of view, regions which are more frequently visited or regions in which the expert spends a larger time should have

a higher probability of being relevant than regions that are not. In this work, a navigation map is computed iteratively, where each new navigation update the current probability map.

4.1.3. Bayesian strategy

Integration of top-down and bottom-up informations is achieved by means of a Bayesian rule, reads as:

$$P(\Theta|\mathcal{X}) = \frac{P(\mathcal{X}|\Theta)P(\Theta)}{P(\mathcal{X})}$$

Were $P(\Theta|\mathcal{X})$ is the posterior probability of the particular parameters given the navigation sample $\mathcal{X}$. $P(\mathcal{X}|\Theta)$ is the likelihood, defined as the probability of having a particular navigation $\mathcal{X}$, given a distribution with parameters Θ. $P(\Theta)$ is the a priori probability of the parameters, and $P(\mathcal{X})$is the evidence, i.e. the recorded navigation.

Patch-based approach

This model estimates a relevance map in which the probability is assigned to each region according to the number of times the region was visited in previous navigations. The idea behind this framework is that relevant regions for diagnosis are those that are more frequently visited by pathologist during exploration. The first step consist in the maximum likelihood estimation. Then, being $P(x_i|\Theta)$ the probability of visiting region i given a distribution with parameters Θ and modelling this as a multinomial probability, the probability of the region is

$$P(x_1^t, x_2^t, \cdots, x_k^t|\Theta) = Multi(\mathbf{x}|\Theta) = \prod_{i=1}^{k} p_i^{x_i}$$

$x_1, x_2, ..., x_k$ are indicator variables, where x_i is 1 if the region i was visited or 0 otherwise. $\Theta = (p_1, p_2, \cdots, p_k)$ and p_i is the probability of visiting region i. Now, notice that a navigation is a set of visited regions

$$\mathcal{X} = (\mathbf{x}^1, \mathbf{x}^2, \cdots, \mathbf{x}^N)$$

Where $\mathbf{x}^t = (x_1^t, x_2^t, \cdots, x_k^t)$, x_i^t are indicator variables and $\sum_{i=1}^{k} x_i^t = 1$. According to this, the likelihood of a navigation is then given by:

$$P(\mathcal{X}|\Theta) = \prod_{i=1}^{k} p_i^{N_i}$$

Where $N_i = \sum_{t=1}^{N} x_i^t$ indicates the number of times the $i-th$ region was visited during a navigation, see figure **4-1**. This information comes from the recorded diagnostic explorations performed by the expert pathologists and is represented by the vector $\mathbf{n} = (N_1, N_2, \cdots, N_k)$.

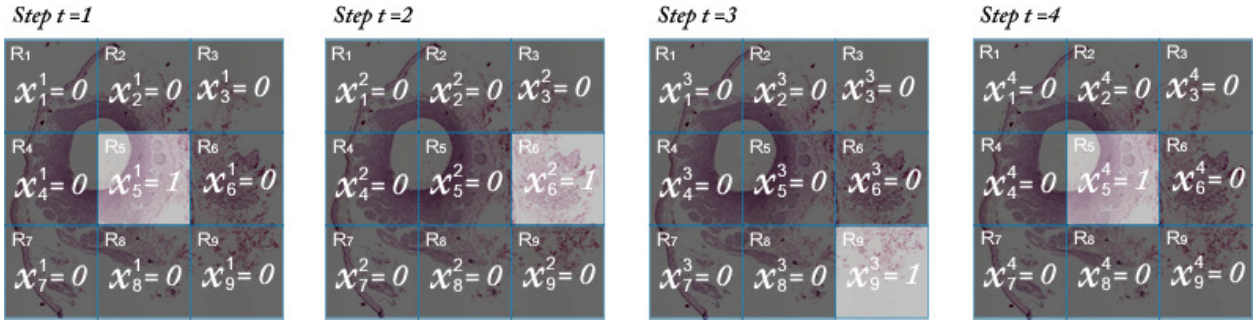


Figura 4-1: A VS divided in 9 regions is navigated in four steps. In each step only one region is visited. In this example, at the end of navigation region 5 was visited twice (step $t = 1$ and $t = 4$), then $N_5 = \sum_{t=1}^{4} x_5^t = 2$. Regions can be any distribution of sets of contiguous pixels. In the extreme case, regions corresponds to individual pixels.

Since Dirichlet is the conjugate prior of multinomial distribution, the prior distribution is

$$P(\Theta|\alpha) = Dirichlet(\Theta|\alpha) = \frac{\Gamma(\alpha_0)}{\Gamma(\alpha_1)\cdots\Gamma(\alpha_k)} \prod_{i=1}^{k} p_i^{\alpha_i - 1}$$

Where $\Theta = (p_1, p_2, \cdots, p_k)$, $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_k)$ and $\alpha_0 = \sum_{i=1}^{k} \alpha_i$. Given the prior and the likelihood, the posterior, which also has a Dirichlet distribution is given by

$$P(\Theta|\mathcal{X}, \alpha) = \frac{\Gamma(\alpha_0 + N)}{\Gamma(\alpha_1 + N_1)\ldots\Gamma(\alpha_k + N_k)} \prod_{i=1}^{k} p_i^{\alpha_i + N_i - 1}$$

Vector $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_k)$ are the initial parameters of prior distribution given by the saliency map, α_i corresponds to the visiting probability assigned by the VAM to $i - th$ region.
We estimate p_i using a Bayesian parameter estimation as follows,

$$\Theta_{Bayes} = E\left[\Theta|\mathcal{X}, \alpha\right] = \int \Theta P(\Theta|\mathcal{X}, \alpha) d\Theta$$

which solution is given by $p_i = \frac{\alpha_i + N_i}{\alpha_0 + N}$. Then, the estimated parameter for each region is given by

$$P(p_i|\mathcal{X}, \alpha) = \frac{\Gamma(\alpha_0 + N)}{\Gamma(\alpha_1 + N_1)\cdots\Gamma(\alpha_k + N_k)} \left(\frac{\lambda\alpha_i + (1-\lambda)N_i}{\lambda\alpha_0 + (1-\lambda)N}\right)^{\alpha_i + N_i - 1}$$

A λ parameter was included to control the relative importance of the prior with respect to the evidence. The higher is the value, the more important is the prior as shown in figure **4-2**.

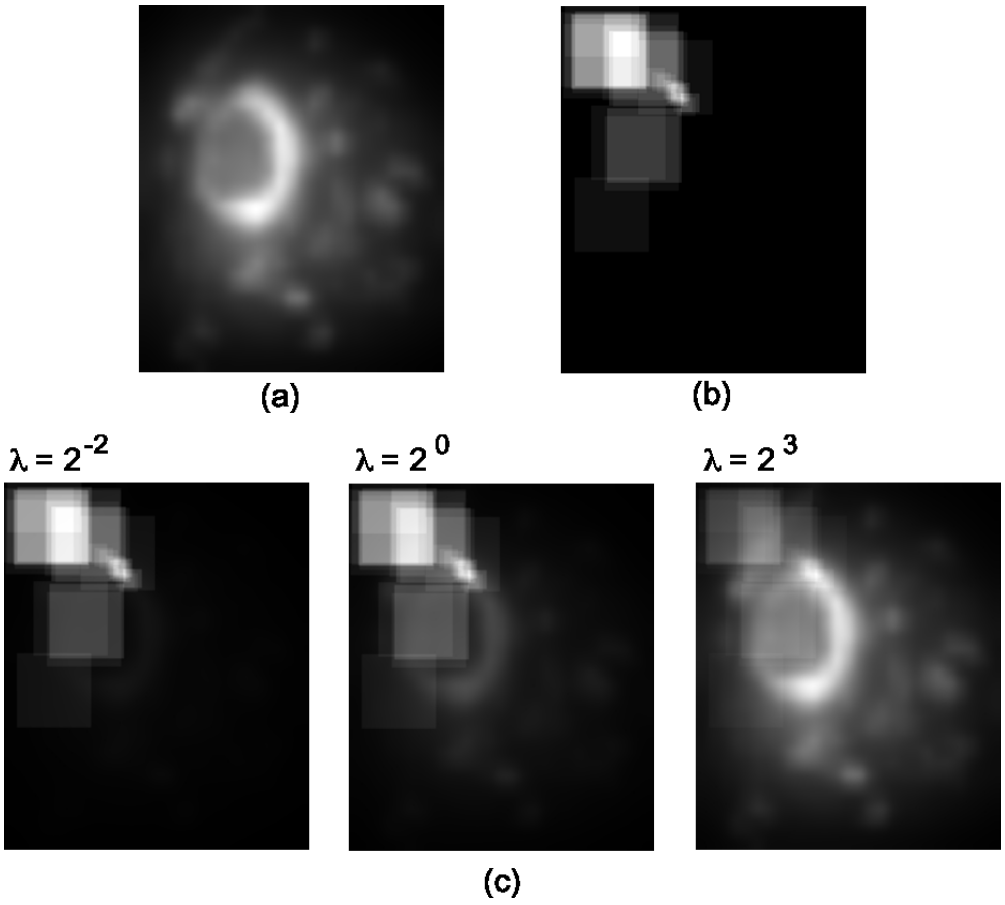


Figura 4-2: (a) Saliency map calculated by a bottom-up VAM. Intensity values correspond to the levels of relevance. (b) Recorded navigation as square fields of view, intensity values represent how many times a region was visited. (c) Posteriori probability functions for different λ values.

Every time there is a new navigation available, successive refinements of the probability can be done. With a prior α and a navigation $\mathbf{n_1}$, a estimate Θ_1 s calculated, with a new navigation $\mathbf{n_2}$ a new estimate Θ_2 can be calculated using Θ_1 as prior, and so on.
Interestingly, this approach achieve good performance in the task of extracting relevant regions, but in some cases, these are scattered small regions (patches) which do not represent coherent pre-fetching for new navigations. Then, a more complex model is herein introduced, which allows the definition of more meaningful regions for exploration with diagnostic purposes.

4.1.4. Graph-based approach

The graph-based approach derives from the Bayesian framework exposed in 4.1.3, but additionally, performs a splitting of the VS that allows a robust representation in terms of relevant regions. This model defines a graph $G = R, E$, where R is the set of regions of the VS with a relevance score associated and E is the set edges representing hierarchy relationships of relevance between nodes.
The first step is to built a discrete saliency map and structure it in a graph, this is done as follows: using a continuous saliency map obtained from a bottom-up VAM, a set of local minima are determined by a two-dimensional first-order derivative. Each local minimum determines a saliency level which define an image region with similar relevance value (iso-relevance). Every iso-relevant region is represented by a node of the graph. They are connected according to region inclusion relationships through relevance levels. The resulting graph corresponds to a tree that encodes the hierarchical relationships between image regions in terms of relevance. The pipeline of this first step is shown in Figure **4-3**.
The idea is to enrich the tree structure described above, with information from pathologist's navigations as follows: the path followed by expert pathologists during a diagnostic exploration is recorded, each area visited by the pathologist during the navigation is compared with the nodes of the tree (image regions). So, by a similar treatment as shown in patch-based model, the aim is to estimate the posterior probability $P(\mathbf{v}|\mathcal{X})$ given the navigation sample $\mathcal{X}$, where $\mathbf{v} = v_1, v_2, \cdots, v_k$ and v_i is the probability of visiting node i. Modelled as a Dirichlet probability, posterior probability for node i is stated as:

$$P(v_i|\mathcal{X}, \alpha) = \frac{\Gamma(\alpha_0 + N)}{\Gamma(\alpha_1 + N_1) \cdots \Gamma(\alpha_k + N_k)} \left(\frac{\lambda\alpha_i + (1-\lambda)N_i}{\lambda\alpha_0 + (1-\lambda)N} \right)^{\alpha_i + N_i - 1}$$

Where $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_k)$ are the initial parameters of the prior distribution given by the discrete saliency map. Notice that, a navigation path may overlap the image area corresponding to several nodes or just small portion of the area associated to one node, see figure **4-4**(d). A more accurate estimation of the probability distribution is obtained by weighting the visited area by the number of times this area was visited (in the same way as explained in section 4.1.3) as well as the intersected area, that is to say, non overlapped areas take their

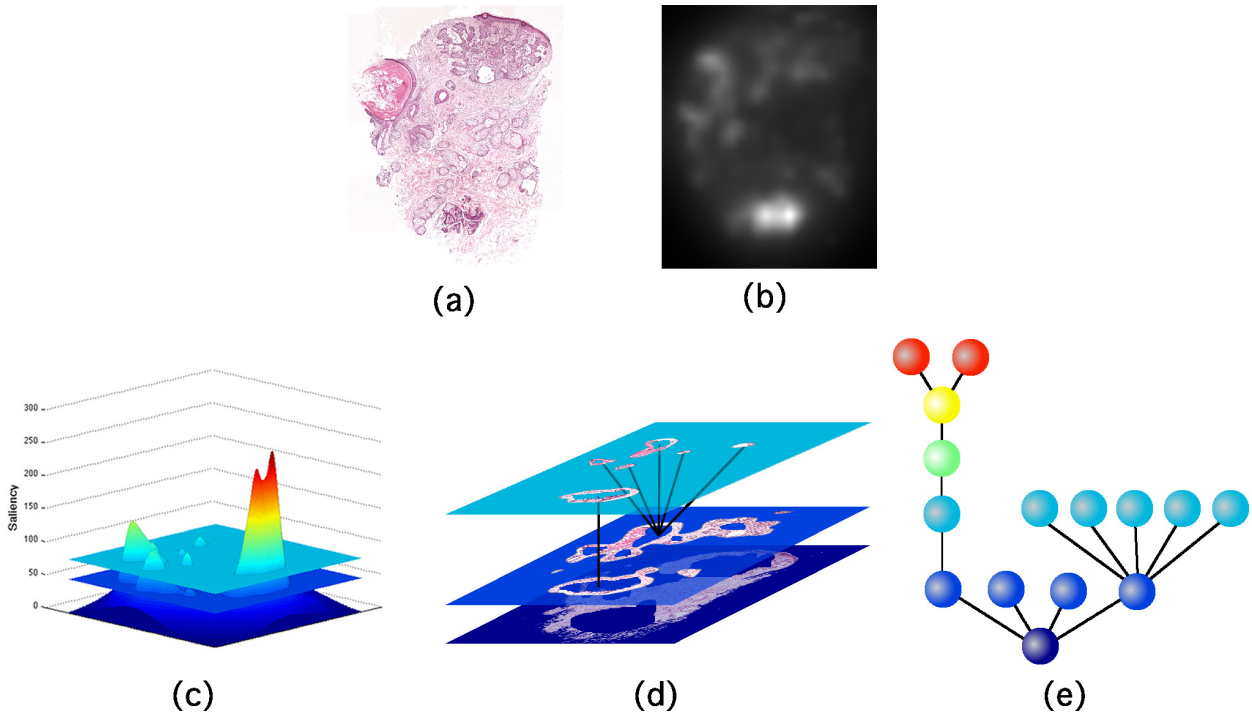


Figura 4-3: (a) input image, (b) saliency map calculated by a visual attentional model, (c) saliency map with different $z - levels$ identified, (d) different regions defined by iso-saliency contours at different saliency levels, lines indicate inclusion relationships between regions, (e) tree representation of regions and inclusion relationships.

relevance value in function of their Euclidean distance to the overlapped areas. The more distant non-overlapped areas the lower relevance value. This generates a smoother saliency region function, allowing to redistribute the probability in an inter and intra-node fashion. Letâ$^{\text{TM}}$s say:

$$\gamma = \frac{\Gamma(\alpha_0 + N)}{\Gamma(\alpha_1 + N_1) \cdots \Gamma(\alpha_k + N_k)}$$

The posterior probability for each node is given by:

$$P(v_i(x,y))|\mathcal{X}, \alpha) = \gamma \left(\frac{\lambda\alpha_i(x,y) + (1-\lambda)N_i(x,y)}{\lambda\alpha_0 + (1-\lambda)N} \right)^{\alpha_i + N_i - 1} \times \frac{1}{\mathbf{D}((x,y)(p,q))}$$

Where $\mathbf{D} = ((x,y)(p,q))$ is the euclidean distance between the current pixel and its nearest pixel overlapped by the navigation path. After this, a normalization is performed.
In this way hierarchy of the graph is updated. The nodes corresponding to regions more visited will have higher relevance score. A new saliency map is built by the union of every region, See figure **4-4**. As long as a new navigation is available, probability is redistribute

either inter and intra-node as an iterative sum of navigations. As in the patch-based method, the λ parameter allows to weight the probability maps to define a criterion of importance.

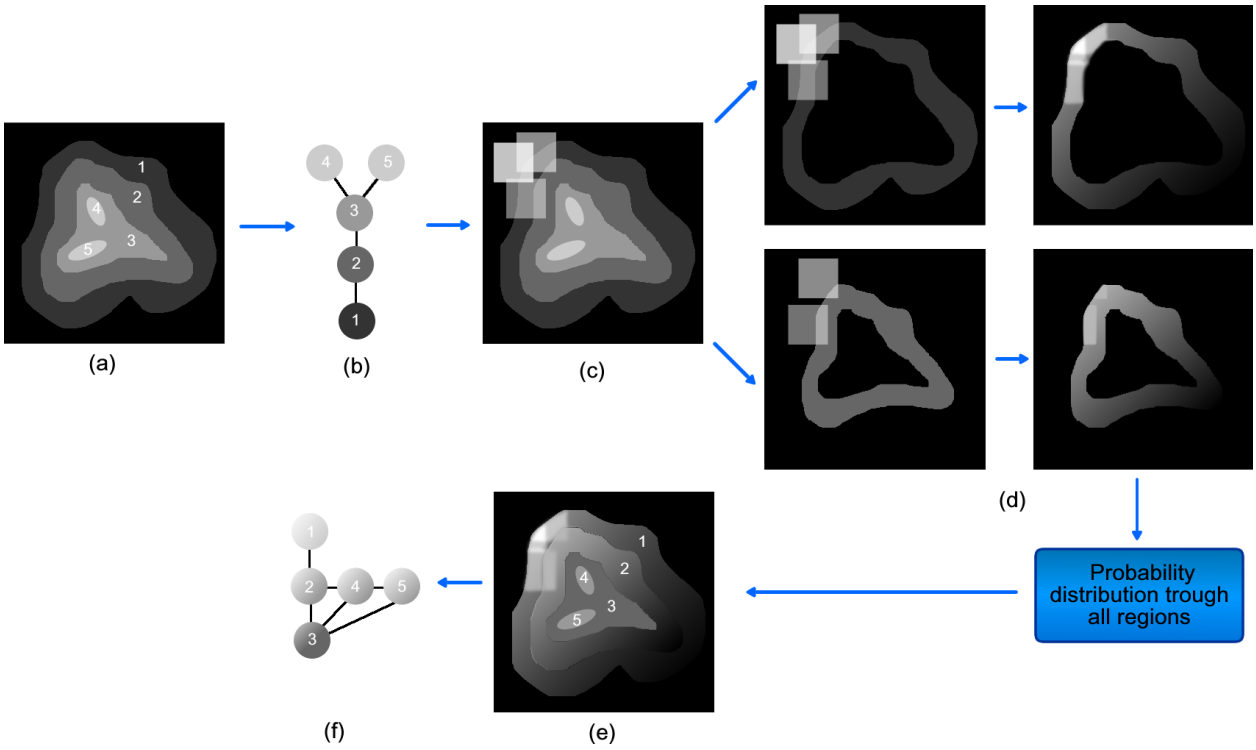


Figura 4-4: (a) and (b) Saliency function obtained by steps in figure **4-3** and its corresponding graph, respectively. (c) A recorded navigation (gray-level squares) over the discrete saliency prior. (d) Intra-region probability distribution by multinomial analysis and distance function. (e) and (f) final relevance map and its corresponding graph representation.

4.2. Data set

The data used consisted in a total of twenty skin biopsies of patients diagnosed with different types of basal cell carcinoma, embedded in paraffin and stained with Hematoxylin-Eosin. The set of histological samples was provided by the pathology department of Universidad Nacional de Colombia and is representative of what pathologists usually observe in their clinical routine. Each specimen was completely digitized at 20× magnification using a standard microscope (LEICA DM LM), coupled to a colour digital camera (NIKON Coolpix 950). Virtual slides were obtained by assembling secuencial microscopical fields of view. Resulting images have between 82 and 458 megapixels with storage sizes between 73 MB and 342 MB, depending on the sample size. Each VS were stored in J2K format [22] for later access

and navigation, in this experimentation, VS were formatted with lossless compression, four wavelet levels, tiles of 512×512 pixels, codeblock size of 32×32 pixels and single layer of quality.
Four expert pathologists from the same pathology school, with similar training level, navigated these virtual slides using a custom virtual microscope, presented below.

4.2.1. Virtual microscope

The virtual microscope used to capture pathologistâ$^{\mathrm{TM}}$s navigations has a design that exploits the importance of low magnifications for exploration and scanning, and high resolutions for diagnosis [22]. The Graphical User Interface (GUI) is composed of three panels: thumbnail, exploration and control. Thumbnail panel, at right side of the interface, shows a low resolution version of the VS and a square boundary that delimits the current Window of Interest (WoI) which is shown in high resolution on exploration panel, at left side of the interface, see **4-5**. The square WoI allows spatial location of the current exploration panel respect to the whole VS. Displacements of the WoI in the thumbnail are done by clicking in the desired point, drag and drop actions are only allowed within the exploration panel. Control panel allows, among other actions, select the magnification level, as thumbnail and exploration panels are fixed size frames, only the WoI has a variable size, the higher is the magnification level, the smaller is the square WoI.
Each pathologist was previously trained on virtual microscopy using two test images. They did not have any information about the organ or the specific pathology they were observing. The twenty VS were randomly displayed. Each pathologist was asked to run over the megaimage up to a diagnosis was set. During examinations, every pathologist action was recorded for later analysis, namely, the WoI location (upper-left corner window coordinates), the WoI size related to the original image, the time elapsed in each WoI and the level of magnification displayed on exploration panel.

4.3. Evaluation and Results

The performance of the proposed methods was evaluated as their ability to predict new pathologists' navigation paths. For doing so, we implemented a simulated navigation model using real navigations. The model defines a cache space in which those regions, marked as the most relevant, are stored. Navigation requirements were compared with what was available in the cache space. The performance measure corresponds to the percentage of cache hits, i.e., the percentage of navigated image pixels that were available in the cache space. Results were evaluated for each of the twenty available images.
Analysis of the method potential to integrating new evidence in this Bayesian framework, it was assessed the prediction capability of the simplest method (patch-based) in consecutive training steps. For doing so, it was measured the percentage of cache hits in a fixed cache size.

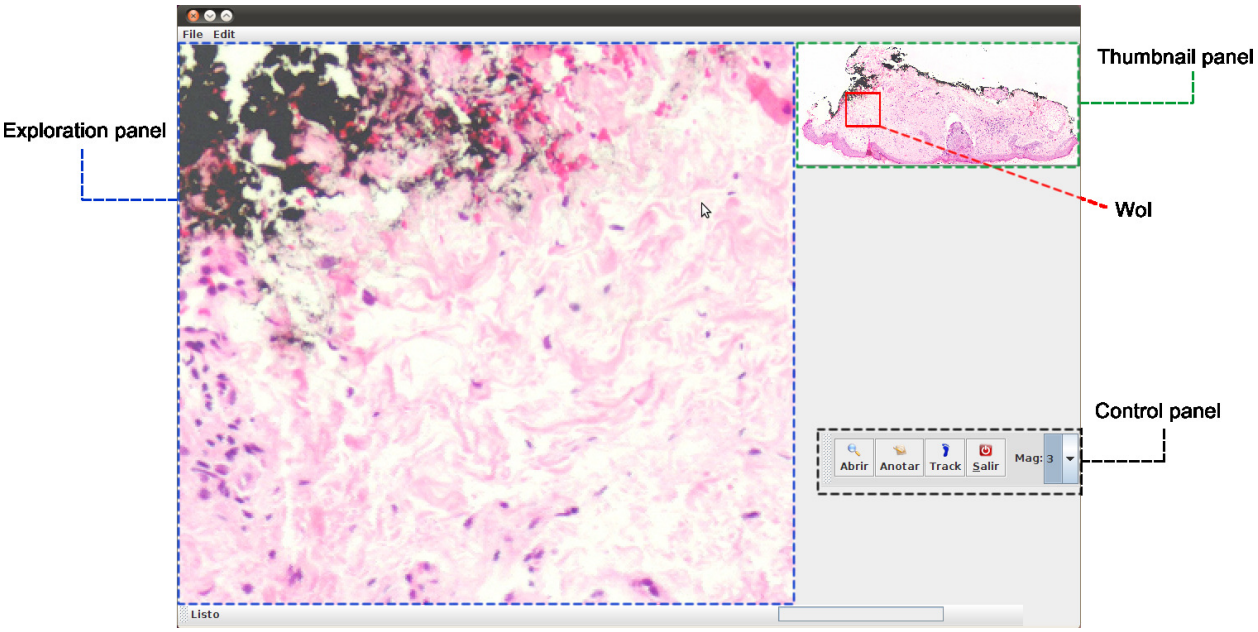


Figura 4-5: Exploration panel shows in the specified magnification, the square region enclosed by the WoI (red square) located in the thumbnail panel. This panel allows spatial localization of current WoI within the entire VS. Control panel presents to the user functions to open, close, annotate and record navigations of VS, also allows to change the magnification level of the current view.

First, an initial experiment using only the saliency map given by the bottom-up VAM was carried out. Then, a first a posteriori probability $P(\Theta_1|\mathcal{X}_1)$ is calculated using one navigation and tested with a second navigation. Likewise, a second posterior probability $P(\Theta_2|\mathcal{X}_2)$ was calculated using these two previous navigations and evaluated in a third one, and so on. The bottom-up VAM used in this experimentation is the graph-based visual saliency model proposed by Harel et. al [19].

Each experiment was performed using 9 different values of λ to evaluate the response of the contribution of each source of information. Results shown in figure **4-6**, demonstrate that when new knowledge is integrated, the predictive capability of the model increases, especially with $\lambda = 0{,}7$. This indicates that preservation of prior information, namely saliency map for the first iteration or navigations for further iterations, is important to improve the prediction of relevant regions that probably will be later visited.

A second evaluation compares the two proposed Bayesian models, with a fixed λ value. A cross-validation leave-one-out scheme was implemented as follows: for each fold, three navigations were used to build and train a relevance map (with both, patch-based and graph-based method) and one remaining navigation was used to measure the percentage of cache hits. The average across the cross-validation folds and images is reported.

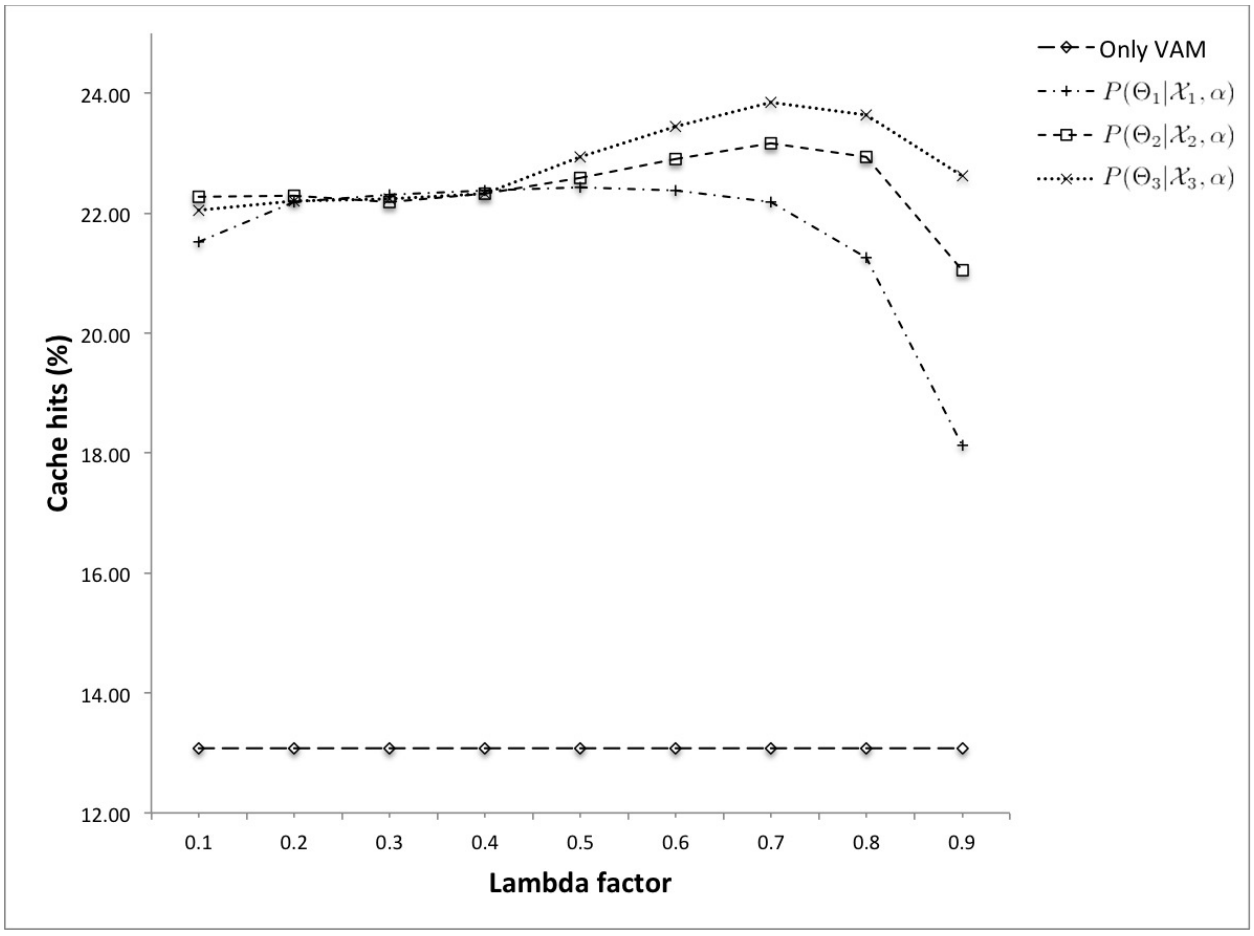


Figura 4-6: Results using only the prior information of a saliency map (Only VAM) and after one ($P(\Theta_1|\mathcal{X}_1)$), two ($P(\Theta_2|\mathcal{X}_2)$) and three ($P(\Theta_3|\mathcal{X}_3)$) learning iterations. Cache size was fixed to 5 % of image size.

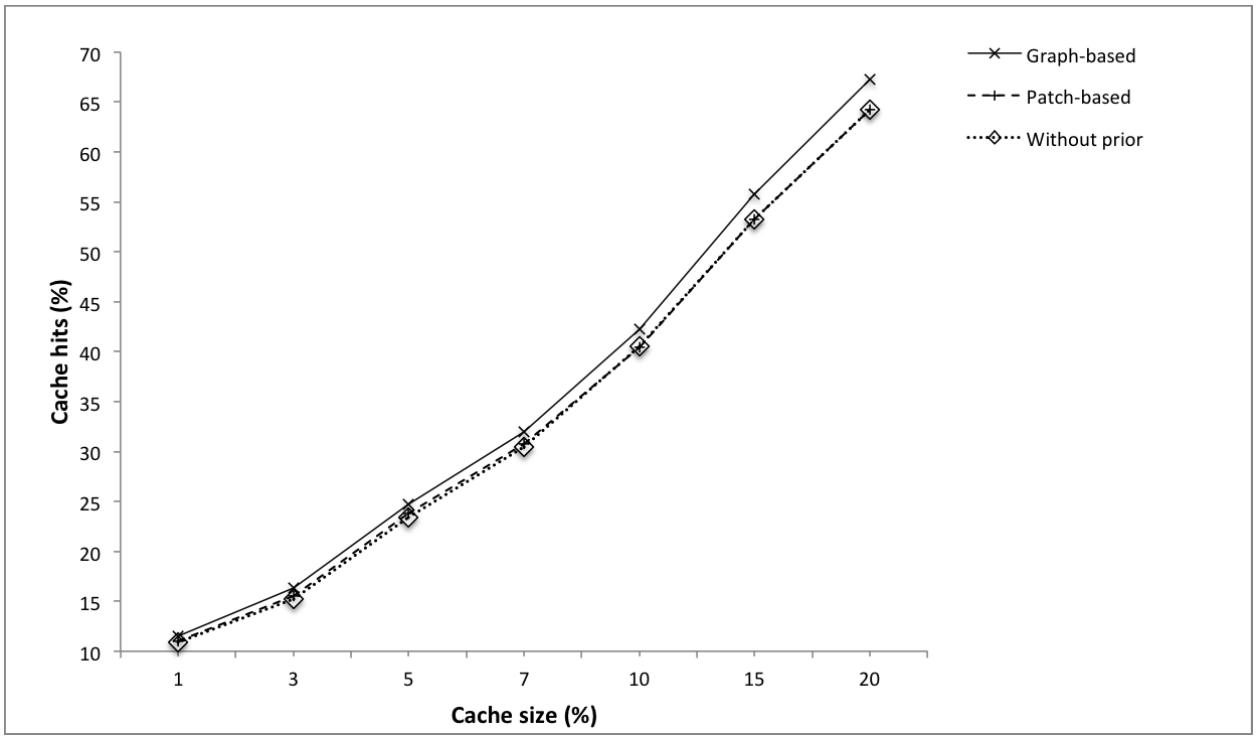


Figura 4-7: Performance is measured as the percentage of hits within a certain cache size.

The graph-based approach outperforms the patch-based strategy from about a 1 %, when the cache size was set to a 5 % of the original VS size, to a 3 % when the cache size was set to to 20 % of VS size. Figure **4-7** shows the percentage of pixels that in average were required during the two test navigations.

A navigation, a small set of jumps between regions of interest, constitutes a limited source of information which is not properly captured by the Patch-based approach. In contrast, the Graph-based approach is able to smoothly re-define the probabilistic map by optimally propagating the information learned from the number of visited regions. This is further observed since the graph-based model not only outperforms the patch model, but it also avoids scattered small regions as shown in figure **4-8**.

Finally, Figure **4-9** shows a typical result obtained with a sample image and a sample navigation. Subfigures (a), (b) and (c) show the sample image, the saliency map generated using a visual attentional model and a navigation performed by a pathologist respectively. Subfigure (d) shows a relevance map obtained by integrating the saliency map and the navigation method using the Bayesian model from Equation **??**. Subfigure (e) shows the relevance map calculated using the proposed method.

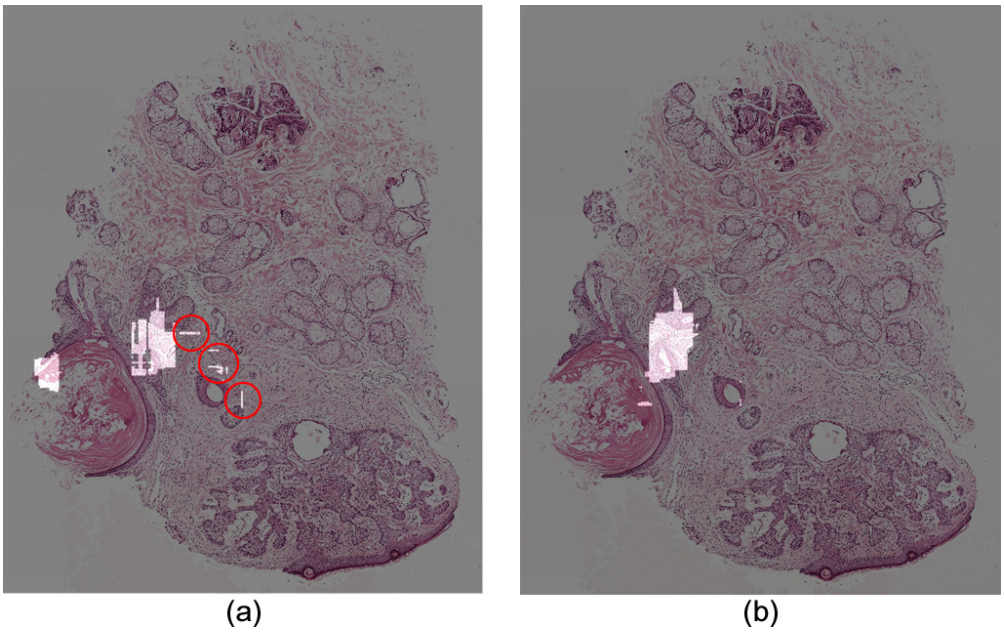


Figura 4-8: Patch-based model (a), pre-fetches to cache space some scattered small regions (patches) as surrounded by red circles. Graph-based model (b) tends to pre-fetch more compact regions.

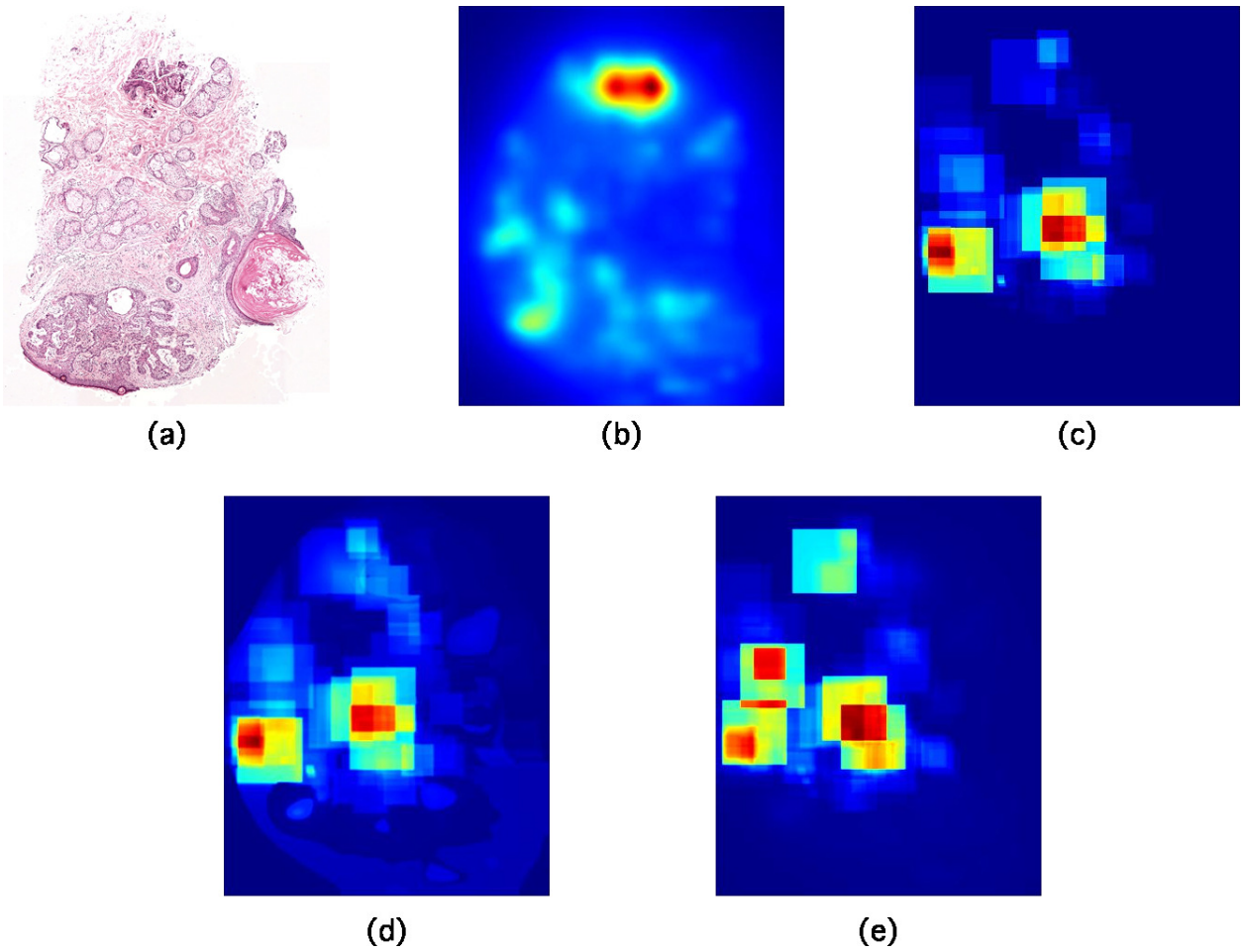


Figura 4-9: (a) original image, (b) saliency map calculated by a visual attentional model, (c) a particular pathologists navigation represented by the frequency of regions visited and the time elapsed in each region. (d) integrated relevance map generated using the graph-based method, (e) integrated relevance map calculated by patch-based method.

5 Conclusions and Perspectives

Fluid exploration of digitized histological samples (Virtual Slides) has become an important research topic from the application of virtual microscopy in clinical scenarios. The challenge is to manage large VS samples captured in different resolutions to allows an appropriate analysis from the physicians and also to allow the development of real telepathology applications. Several approaches to improving real-time high resolution images interaction have been proposed in the literature and it was described in chapter **??** as storage, transmission, management of client resources but standing out cache and pre-fetching methods which based on prediction policies, provide the user with only the regions of current interest.
Since the Visual Attentional Models are robust representations devised to emulate the human visual system, they result appropriate as a first descriptor of the VS relevant regions. A complete analysis of different state of the art methodologies in chapter **??** shown that the algorithm of Harel et. the best performance demonstrated for the identification of the relevant regions compared with explorations of expert pathologist. Nevertheless, results suggest that Bottom-up VAM are insufficient to identify regions in a VS that are relevant for the diagnostic task. Due to the important role of pathologist's knowledge and training in the diagnostic path. This top-down information should to be included somehow in order to have accurate automatic identification of RoIs. It is worth to notice that since the samples have irregular shapes there exist always blank regions in the rectangular VS. That means, although without integration of bottom-up information the method has a poor performance identifying RoIs, it is able to ignore regions that are not relevant at all.
The work presented in chapter **??** has introduced a novel learning method that improves detection of relevant regions in histopathology virtual slides. Over a Bayesian framework this work integrates two sources of information: the bottom-up computed from a visual attention model and top-down captured from expert pathologists' navigations. Integration of these information flows result in a relevance map able to highlight possible RoIs specifically for diagnostic exploration of VS. Such identification of RoIs is an important information for applications as: (1) selective compression for efficient storage of VS. (2) Strategies for acceleration of VS navigation, since prefetching of RoIs represent a minimization of the latency upon server-client cache architectures. (3) Training.
Romo et. al proposed a RoIs learning approach based in the integration of low level VS features and experts pathologist knowledge [41]. In this work, the prior it is introduced by the selection of local target and distractor regions. However, top-down knowledge is restricted to an expert image annotation, a highly time consuming task. In contrast, our method use as

high-level knowledge recorded navigations that pathologists perform in their daily practice. In this framework, a first approximation presents a patch based approach which assigns to each pixel a probability of visiting according to a multinomial probability distribution. Results show that the method is able to learn and its prediction capability improves every time new evidence is available. Nevertheless, in some cases extracted regions are scattered small regions (patches) which do not represent coherent compact section of VS. So, we proposed a second approach, which is also based on multinomial distribution, but presenting a robust VS representation. The method starts by setting a graph structure that represents the regions of interest extracted with a visual model. The vertices of the initial graph structure are weighted with the steady state probability distribution generated by a random walk, a strategy used in artificial vision to determine salient regions. The present framework mixed up this bottom-up hierarchy with a top down strategy that aimed to capture the expert knowledge. Assuming a cache where only 20 % of total size of VS can be stored, graph-based method is able to predict in average, the 67 % of regions that were visited in a new navigation. Additionally, this method outperforms RoIs identification capacity of patch-based method in about 3 % which is significant in this context since 3 % of a VS represent regions of megapixels.
Limitations of graph-based method are related with the consistency between the shape of the regions initially defined by the method and the square regions visited by the pathologist during diagnostic exploration. Future work includes:

- Analysis of new methods for splitting of the VS in order to achieve a more dynamic behaviour of the graph, not only by updating the relevance hierarchy but also updating the defined region within each node.
- Adaptation of the method for online updating the relevance scores.

6 Products

- Estimation of Histopathological Image Relevance: A Hybrid Bottom-Up/Top-Down Approach. Cristina Lasso, Fabio Gonález and Eduardo Romero. Under review Journal of Microscopy.
- Navigation of large-scale medical images: A survey. Cristina Lasso and Eduardo Romero. Under review revista Med.
- Sistema Modular de Telepatología para la exploración Remota de Placas Histopatológicas sobre RENATA. Cristina Lasso, Carlos Vargas, Lucia Roa-PeÃ±$a, Olga Álvarez y Eduardo Romero. R

- Hybrid method for representation of relevant regions in histopathological images. Cristina Lasso, Fabio Gonález and Eduardo Romero. 7th International Seminar on Medical Information Processing and Analysis - SIPAIM 2011.
- Análisis comparativo de modelos de atención visual para identificar regiones de interés en imágenes de histopatología. Cristina Lasso, Ricardo Gutiérrez y Eduardo Romero. VI Seminario Internacional de Procesamiento y Análisis de Imágenes Médicas - SIPAIM 2010.

Bibliografía

[1] Achanta, Radhakrishna ; Hemami, Sheila ; Estrada, Francisco ; Süsstrunk, Sabine: Frequency-tuned Salient Region Detection. En: *IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, 2009

[2] Buvat, I. ; Chameroy, V. ; Aubry, F. ; Pélégrini, M. ; Fakhri, G. E. ; Huguenin, C. ; Benali, H. ; Todd-Pokropek, A. ; Paola., R. D.: The need to develop guidelines for evaluations of medical image processing procedures. En: *SPIE Medical Imaging*, 1999

[3] C., Lettieri ; G., Veerappan ; J., Parker ; T., Franks ; D., Hayden ; W., Travis ; A., Shorr: Discordance between general and pulmonary pathologists in the diagnosis of interstitial lung disease. En: *Respiratory Medicine* 99 (2005), p. 1425–1430

[4] Calvin, Nodine ; Harold, Kundel: Using eye movements to study visual search and to improve tumor detection. En: *RadloGraphics* 7 (1987), November, Nr. 6, p. 1241–1250

[5] Chang, E. C. ; Yap, C. K. ; j. Yen, T.: Realtime Visualization of Large Images over a Thinwire. En: *In IEEE Visualization '97 (Late Breaking Hot Topics*, 1997, p. 45–48

[6] Crowley, Rebecca S. ; Naus, Gregory J. ; III, Jimmie S. ; Friedman., Charles P.: Development of visual diagnostic expertise in pathology an information-processing study. En: *American Medical Informatics Association* 10(1) (2003), p. 39â“51

[7] Descampe, A. ; Ou, J. ; Chevalier, P. ; Macq, B.: Data prefetching for smooth navigation of large scale JPEG 2000 images. En: *Multimedia and Expo, 2005. ICME 2005. IEEE International Conference on*, 2005, p. 4 pp.

[8] Descampe, A. ; Vleeschouwer, C. D. ; Iregui, M. ; Macq, B. ; Marques, F.: Pre-fetching and caching strategies for remote and interactive browsing of JPEG2000 images. En: *IEEE Trans. on Image Processing* 16 (2007), May, Nr. 5, p. 1339–1354

[9] Duncan, John ; Humphreys, Glyn W.: Visual Search and Stimulus Similarity. En: *Psychological Review* 96 (1989), Nr. 3, p. 433 – 458. – ISSN 0033–295X

[10] Erasmus, J. J. ; Gladish, G. W. ; Broemeling, L.: Interobserver and intraobserver variability in measurement of non small cell carcinoma lung lesions: Implications for assessment of tumor response. En: *Journal of Clinical Oncology* 21 (2003), p. 2574â“2582

[11] Fontelo, P. ; DiNino, E. ; Johansen, K. ; Khan, A. ; Ackerman, M.: Virtual Microscopy: Potential Applications in Medical Education and Telemedicine in Countries with Developing Economies. En: *Proceedings of the 38th Hawaii International Conference on System Sciences* (2005), p. 1–7

[12] Gokturk, S.B. ; Tomasi, C. ; Girod, B. ; Beaulieu, C.: Medical image compression based on region of interest, with application to colon CT images. En: *Engineering in Medicine and Biology Society, 2001. Proceedings of the 23rd Annual International Conference of the IEEE* Vol. 3, 2001. – ISSN 1094–687X, p. 2453 – 2456 vol.3

[13] Goldberg, M. ; Wang, L.: Comparative performance of pyramid data structures for progressive image transmission. En: *Communications, IEEE Transactions on* 39 (1991), apr, Nr. 4, p. 540 –548. – ISSN 0090–6778

[14] Gómez, Francisco ; Iregui, Marcela ; Romero, Eduardo: Prediction of pathologist navigation patterns in virtual microscopy based on a soft-computing approach. En: *HCI '08: Proceedings of the Third IASTED International Conference on Human Computer Interaction*. Anaheim, CA, USA : ACTA Press, 2008. – ISBN 978–0–88986–726–0, p. 150–155

[15] Gómez, Francisco ; Marín, Diana ; Romero, Eduardo: A Soft-cache Strategy for Pathologistâ^{TM}s Navigation in Virtual Microscopy. En: *Microscopy Research and Technique* (2010)

[16] Gómez, Francisco ; Romero, Eduardo: A model for predicting pathologist's velocity profiles when navigating virtual slides. En: *Microscopy Research and Technique* 73 (2009), p. 85â“98

[17] Gómez, Francisco ; Villalón, Julio ; Gutiérrez, Ricardo ; Romero, Eduardo: Finding Regions of Interest in Pathological images: an attentional model approach. En: *SPIE Medical Images, Proceedings of SPIE* Vol. 7260, 2009, p. 7 – 12

[18] Gutiérrez, Ricardo E.: *Compresión selectiva en imágenes digitales de microscopio en patología*, Universidad Nacional de Colombia, Tesis de Grado, 2009

[19] Harel, Jonathan ; Koch, Christof ; Perona, Pietro: Graph-based visual saliency. En: *Advances in Neural Information Processing Systems 19*, MIT Press, 2007, p. 545–552

[20] Hoffman, James E.: A two-stage model of visual search. En: *Perception and Psychophysics* 25 (1979), p. 319–317

[21] Ip, Cheuk Y. ; Varshney, Amitabh: Saliency-Assisted Navigation of Very Large Landscape Images. En: *Visualization and Computer Graphics, IEEE Transactions on* 17 (2011), dec., Nr. 12, p. 1737 –1746. – ISSN 1077–2626

[22] Iregui, Marcela ; Gómez, Francisco ; Romero, Eduardo: Strategies for efficient virtual microscopy in pathological samples using JPEG2000. En: *Micron* 38 (2007), Nr. 7, p. 700 – 713. – ISSN 0968–4328

[23] Itti, L. ; Koch, C. ; Niebur, E.: A model of saliency-based visual attention for rapid scene analysis. En: *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 20 (1998), nov, Nr. 11, p. 1254 –1259. – ISSN 0162–8828

[24] Karras, D.A. ; Karkanis, S.A. ; Maroulis, D.E.: Efficient Image Compression of Medical Images Using the Wavelet Transform and Fuzzy c-means Clustering on Regions of Interest. En: *EUROMICRO Conference* 2 (2000), p. 2469. ISBN 0–7695–0780–8

[25] Kil, Se-Kee ; Lee, Jong-Shill ; Shen, Dong-Fan ; Lee, Je-Goon RyuEung-Hyuk ; Min, Hong-Ki ; Hong, Seung-Hong: Lossless Medical Image Compression using Redundancy Analysis. En: *International Journal of Computer Science and Network Security* 6 (2006), p. 50 –56

[26] Kim, Yong-Sung ; Kim, Whoi-Yul: An efficient image data format for lossless compression and its application to interactive viewing. En: *Image Processing, 1996. Proceedings., International Conference on* Vol. 1, 1996, p. 73 –76 vol.1

[27] Koch, C. ; Ullman, S. *Shifts in selective visual attention: towards the underlying neural circuitry.* 1985

[28] Krupinski, Elizabeth A. ; Tillack, Allison A. ; Richter, Lynne ; Henderson, Jeffrey T. ; Bhattacharyya, Achyut K. ; Scott, Katherine M. ; Graham, Anna R. ; Descour, Michael R. ; Davis, John R. ; Weinstein, Ronald S.: Eye-movement study and human performance using telepathology virtual slides. Implications for medical education and differences with experience. En: *Human Pathology* 37 (2006), Nr. 12, p. 1543 – 1556. – ISSN 0046–8177

[29] Li, Jin ; Sun, Hong-Hui: On interactive browsing of large images. En: *Multimedia, IEEE Transactions on* 5 (2003), dec., Nr. 4, p. 581 – 590. – ISSN 1520–9210

[30] Li, Tsai-Yen ; wei Chiang, Chih: Data Management for Visualizing Large Virtual Environments. En: *Proceedings of International Symposium on Multimedia Information Processing*, 1999

[31] Lin, Chengjiang ; Zheng, Yuan F.: Fast Browsing of Large Scale Images Using Server Prefetching and Client Caching Techniques. En: *in Applications of Digital Image Processing XXII SPIE*, 1999, p. 376–387

[32] Mikula, Shawn ; Trotts, Issac ; Stone, James M. ; Jones, Edward G.: Internet-enabled high-resolution brain mapping and virtual microscopy. En: *NeuroImage* 35 (2007), Nr. 1, p. 9 – 15. – ISSN 1053–8119

[33] Milanese, R. ; Wechsler, H. ; Gill, S. ; Bost, J. M. ; Pun, T.: Integration of bottom-up and top-down cues for visual attention using non-linear relaxation. En: *Computer Vision and Pattern Recognition, 1994. Proceedings CVPR '94., 1994 IEEE Computer Society Conference on*, 1994, p. 781 –785

[34] Otsu, N.: A Threshold Selection Method from Gray-Level Histograms. En: *Systems, Man and Cybernetics, IEEE Transactions on* 9 (1979), jan., Nr. 1, p. 62 –66. – ISSN 0018–9472

[35] Rasmussen, CA US) Rasmussen Lars Eilstrup (Fairlight AU) Ma Stephen (Kingsford A. *Generating and serving tiles in a digital mapping system.* October 2009

[36] Rasmussen, CA US) Rasmussen Lars Eilstrup (Fairlight AU) Taylor Bret Steven (Mountain View CA US) Norris James Christopher (Mountain View CA US) Ma Stephen (Kingsford AU) Kirmse Andrew Robert (Emerald Hills CA US) Gordon Noel Phillip (Hunters Hill AU) Laforge Seth Michael (Seattle WA U. *Digital mapping system.* January 2007

[37] Rempel, Allan G.: *Fast Progressive Transmission of Images Using Wavelets With Sorted Coefficients*, The University of Saskatchewan, Tesis de Grado, 1993

[38] Roa-Pena, Lucia ; Gómez, Francisco ; Romero, Eduardo: An experimental study of pathologist's navigation patterns in virtual microscopy. En: *Diagnostic Pathology* 5 (2010), Nr. 1, p. 71. – ISSN 1746–1596

[39] Rocha, Rafael ; Vassallo, José ; Soares, Fernando ; Miller, Keith ; Gobbi, Helenice: Digital slides: Present status of a tool for consultation, teaching, and quality control in pathology. En: *Pathology - Research and Practice* 205 (2009), Nr. 11, p. 735 – 741. – ISSN 0344–0338

[40] Romero, E. ; Gómez, F. ; Iregui, M.: Virtual Microscopy in Medical Images: a Survey. En: *Modern Research and Educational Topics in Microscopy* 2 (2007), p. 996–1006

[41] Romo, David ; Romero, Eduardo ; González, Fabio: Learning regions of interest from low level maps in virtual microscopy. En: *Diagnostic Pathology* 6 (2011), March, Nr. S1

[42] Sloan, K.R.Jr. ; Tanimoto, S.L.: Progressive Refinement of Raster Images. En: *Computers, IEEE Transactions on* C-28 (1979), nov., Nr. 11, p. 871 –874. – ISSN 0018–9340

[43] Taubman, D.: High performance scalable image compression with EBCOT. En: *Image Processing, 1999. ICIP 99. Proceedings. 1999 International Conference on* Vol. 3, 1999, p. 344 –348 vol.3

[44] Treisman, Anne ; Gormican, Stephen: Feature Analysis in Early Vision: Evidence From Search Asymmetries. En: *Psychological Review* 95 (1988), Nr. 1, p. 15 – 48. – ISSN 0033–295X

[45] Treisman, Anne M. ; Gelade, Garry: A feature-integration theory of attention. En: *Cognitive Psychology* 12 (1980), Nr. 1, p. 97 – 136. – ISSN 0010–0285

[46] Triola, Marc ; Holloway, William: Enhanced virtual microscopy for collaborative education. En: *BMC Medical Education* 11 (2011), Nr. 1, p. 4. – ISSN 1472–6920

[47] Tsotsos, John K. ; Culhane, Scan M. ; Wai, Winky Yan K. ; Lai, Yuzhong ; Davis, Neal ; Nuflo, Fernando: Modeling visual attention via selective tuning. En: *Artificial Intelligence* 78 (1995), Nr. 1-2, p. 507 – 545. – Special Volume on Computer Vision. – ISSN 0004–3702

[48] Kap. 12 In: Wang, James Z. ; Wiederhold, Gio ; Li, Jia: *Wavelet-Based Progressive Transmission and Security Filtering for Medical Image Distribution.* 1998, p. 303–324

[49] Wang, James Z. ; Nguyen, John ; koi Lo, Kin ; Law, Christine ; Regula, Donald: Multiresolution Browsing of Pathology Images using Wavelets. En: *Journal of the American Medical Informatics Association, Proceedings of the AMIA Annual Symposium* Vol. 1999 symposium suppl., 1999, p. 430â“434

[50] Watanabe, Hiroshi ; Ogita, Takeo: Transcoding by Automatic ROI Extraction from JPEG2000 Bitstream. En: *Pro-ceedings of Picture Coding Symposium 2003*, 2003

[51] Weidmann, C. ; Vetterli, M. ; Ortega, A. ; Carignano, F.: Soft caching: image caching in a rate-distortion framework, 1997, p. 696 –699 vol.2

[52] White, P: Legal issues in teleradiology-distant thoughts. En: *The British Journal of Radiology* 75 (2002), p. 201â“206

[53] Wildermoth, M. ; Bradley, A. ; Mills, P.: Virtual Microscopy with Extended Depth of Field. En: *Proceedings of the Digital Imaging Computing: Techniques and Applications (DICTA 2005)* (2005)

[54] Zang, Cunxun ; Liu, Qingshan ; Cheng, Jian ; Lu, Hanqing: Human-centered image navigation on mobile devices. En: *Multimedia and Expo, 2008 IEEE International Conference on*, 2008, p. 321 –324

[55] Zelinsky, G. J. ; Zhang, W. ; Yu, B. ; Chen, X. ; Samaras, D.: The Role of Top-down and Bottom-up Processes in Guiding Eye Movements during Visual Search. En: *Advances in neural information processing systems* 18 (2006), p. 1569â“ 1576

[56] Zukoski, M. J. ; Boult, T. ; Iyriboz, T.: A novel approach to medical image compression. En: *Int. J. Bioinformatics Res. Appl.* 2 (2006), March, p. 89–103. – ISSN 1744–5485