Universidad Nacional de Colombia
Facultad de Medicina

Predicción de Patrones de Navegación en
Mega-imágenes Histopatológicas

David Edmundo Romo Bucheli
Maestría en Ingeniería Biomédica

Director:
Fabio Augusto González Osorio, PhD.
Profesor Asociado, Departamento de Ingeniería de
Sistemas e Industrial
Facultad de Ingeniería

Bogotá D.C., 18 de mayo de 2011

Nota de aceptación

_____

_____

_____

_____

_____

_____

_____

Firma del jurado

_____

Firma del jurado

_____

Firma del jurado

Bogotá D.C., febrero de 2011

"Todo cabe en lo breve. Pequeño es el niño y encierra al hombre; estrecho es el cerebro y cobija el pensamiento; no es el ojo más que un punto y abarca leguas."

Alejandro Dumas

Agradecimientos

Este trabajo ha recibido el aporte de muchas personas que de forma directa, o indirecta han contribuido en su realización. Primero quiero expresar mis agradecimientos al director de este trabajo, el profesor Fabio González Osorio, por su ejemplo de perseverancia, acertada orientación y paciencia. Al profesor Eduardo Romero, director del grupo de investigación Bioingenium, por su confianza y enorme contribución en la redacción de los artículos producto de la labor investigativa desarrollada, así como por sus esfuerzos por crear un ambiente de trabajo favorable en el grupo.

También quiero expresar mis agradecimientos a Francisco Gómez, pues sus anotaciones, sugerencias y discusiones acerca de diversos temas investigativos influyeron en el planteamiento que se tomó para desarrollar este trabajo. A mis compañeros en la maestría: Fabio, Gerardo, Fabián, Diana, Lorenza, Luz Helena por su camaraderia y ayuda durante este tiempo que hemos compartido. Hago extensivos mis agradecimientos a todas las personas en el grupo de investigación Bioingenium, pues el permanente ejercicio investigativo que se realiza inspira la generación de ideas y la perseverancia en su implementación.

Finalmente quiero agradecer a mis padres por su incondicional apoyo y paciencia, sin sus constantes mensajes de aliento y ejemplo de trabajo no hubiera sido posible la culminación de muchos proyectos en mi vida profesional y personal.

Contenido

1. Introducción **1**

1.1. Identificación del Problema 2

1.1.1. Objetivo General 4

1.1.2. Objetivos Específicos 4

1.2. Contribuciones Realizadas 5

1.3. Estructura del Documento 5

2. Sistemas de Microscopía Virtual en Patología **7**

2.1. El Proceso de Exámen Diagnóstico en Patología 8

2.2. Sistemas de Microscopía Virtual 10

2.3. Estrategias para Mejorar la Eficiencia en la Navegación de Mega-imágenes 14

3. Identificación de Regiones de Interés en Tareas de Búsqueda Visual **18**

3.1. Organización del Sistema Visual Humano 19

3.1.1. Anatomía del Ojo 19

3.1.2. Organización en la Retina 20

3.1.3. Corteza Visual y Áreas Superiores de Procesamiento Visual 22

3.2. Teorías Psicofísicas de la Atención Visual 24

3.3. Metodos Computacionales para la Identificación de Regiones de interés y Atención visual 26

3.3.1. Modelos Computacionales de Atención Visual: Modelo de Itti. 27

3.4. Modelamiento de Movimientos Oculares 33

3.5. Reconocimiento y Atención Visual 34

4. Método para Predecir Patrones de Navegación en Mega-imágenes Histopatológicas **37**
4.1. Estrategia de Detección de Relevancia 38
4.1.1. Extracción de Características y Métricas 39
4.1.2. Integración de los Mapas de Distancia 42
4.2. Predicción en una Estrategia de Prefetching 44
4.2.1. Hipótesis de Navegación - Baja Reactividad 44
4.2.2. Hipótesis de Navegación - Alta Reactividad 45
4.3. Experimentación y Resultados 46
4.3.1. Evaluación de la Estrategia de Detección de Relevancia 46
4.3.2. Evaluación de la Predicción en una Estrategia de Prefetching 50

5. Conclusiones y Discusión **56**
5.1. Conclusiones del Método Propuesto 56
5.2. Trabajo Futuro 57

Referencias **59**

A. Productos **65**

Lista de Figuras

1.1. Trayectoria de observación. 3

2.1. Visualización de una lámina virtual. 11
2.2. Procedimientos en microscopía virtual 13

3.1. Organización celular en la retina. 21
3.2. Jerarquía del sistema visual humano. 23
3.3. Modelo de atención de Itti. 28
3.4. Inhibición de retorno 32

4.1. Método propuesto 39
4.2. Modelo de navegación en un sistema de baja reactividad . . . 45
4.3. Modelo de navegación en un sistema de alta reactividad 46
4.4. Aplicación estrategia de relevancia. 47
4.5. Aplicación estrategia de relevancia. 48
4.6. Configuración cliente/servidor 51
4.7. Resultados con hipotesis de alta reactividad 53
4.8. Resultados con hipótesis de baja reactividad 55
4.9. Resultados con hipótesis de baja reactividad 55

Lista de Tablas

4.1. Precisión y Recall 50

Resumen
Predicción de Patrones de Navegación en Mega-imágenes Histopatológicas

La microscopía virtual puede mejorar el trabajo rutinario de los laboratorios patológicos modernos. Este objetivo ha sido severamente limitado por la gran cantidad de información contenida en las láminas histopatológicas virtuales. La adopción de técnicas para mejorar la eficiencia durante la navegación de mega-imágenes ha mostrado ser útil para reducir los tiempos de respuesta en sistemas de microscopía virtual. Este trabajo presenta un enfoque novedoso para predecir patrones de navegación en láminas histopatológicas virtuales durante tareas de evaluación diagnósticas realizadas por patólogos. A partir de la selección de imágenes de ejemplos positivos (objetivo) y negativos (distractor) realizada por el patólogo, el método construye un mapa asignando relevancia a cada una de las regiones de la mega-imagen. Durante la evaluación de la identificación de relevancia, se encontró que el método desarrollado presentó medidas promedio de precisión (55 %) y de promedio de recall (38 %) en el conjunto de datos utilizado, superando otras técnicas para detectar regiones de interés basadas en modelos computacionales de atención visual (Modelo Itti). La información contenida en el mapa de relevancia mostró una capacidad predictiva útil para la formulación de estrategias óptimas de navegación, superando estrategias tradicionales en algunas de las situaciones analizadas en el presente trabajo.

Abstract
Prediction of Navigation Patterns in Histopathological Mega-images

Virtual microscopy can improve the workflow of modern pathology laboratories, a goal limited by the large size of the virtual slides (VS). Lately some strategies to accelerate the navigation performance in large images has reduced the time. This work presents a novel method for predicting navigation patterns in VS during diagnostic tasks performed by pathologists. By selecting positive and negative image examples, the method constructs a map that assigns relevance to each image region. The evaluation of the regions of interest through Precision-recall measurements, calculated at each step of any actual navigation, obtained an average precision of 55 % and a recall of about 38 % when using the available set of navigations, outperforming other techniques based on computational models of visual attention that identify regions of interest. The predictive capability of the relevancy map was useful in the formulation of strategies to improve navigation.

Capítulo 1

Introducción

La introducción progresiva de tecnologías digitales para almacenamiento, transmisión y visualización de grandes imágenes (mega-imágenes) ha facilitado el estudio de diversas disciplinas tales como: sistemas de geoinformación, astronomía, ambientes de realidad virtual [1–3]. El campo médico no ha sido ajeno a este desarrollo tecnológico, pues la generación de mega-imágenes es usualmente requerida en numerosos procesos de evaluación diagnóstica [4]. Particularmente, por la cantidad de información obtenida al digitalizar muestras histopatológicas, en un campo que se ha denominado microscopía virtual (MV).

El análisis de especímenes histopatológicos es una herramienta fundamental dentro de la actividad médica, no solo por su valor en el proceso diagnóstico, sino por su importancia en las decisiones de tratamiento clínico [5]. Dada la relevancia de la evaluación diagnóstica de los especímenes histopatológicos, es de principal interés facilitar el acceso de las herramientas y aplicaciones que provee la MV no solo para aplicaciones educativas y de entrenamiento médico, donde ha tenido una amplia aceptación [6–10], sino también en el trabajo rutinario de patología.

La integración de la MV en el trabajo rutinario de evaluación diagnóstica se ha limitado por la lentitud del proceso de adquisición digital de la lámina, y por los tiempos de latencia en la respuesta de los sistemas computacionales de visualización, todo esto resultado de la necesidad de manejar la enorme cantidad de información contenida en estas mega-imágenes. Estos tiempos de latencia entorpecen una navegación fluida en la tarea de evaluación diagnóstica realizada por el patólogo. Por esta razón, se han utilizado mecanismos y estrategias para disminuir estos tiempos de latencia y asegurar un uso óptimo

de los recursos computacionales [11, 12].

El acceso remoto de mega-imágenes de histopatología plantea desafíos al uso óptimo de los recursos computacionales. En primer lugar, es deseable un almacenamiento eficiente porque al reducir la cantidad de bytes que representan la mega-imagen, no solo se optimiza la utilización de espacio en el servidor sino que disminuye la demanda en la tasa de transmisión del canal de comunicación. Para ello, se debe tener en cuenta que existe un compromiso entre la compresión y la calidad, puesto que la utilización de la mega-imagen histopatológica en análisis diagnóstico exige alta fidelidad en la información.

Al transmitir la información comprimida, surge la necesidad de reconstruir la mega-imagen para su visualización. Los algoritmos de descompresión imponen una intensa carga computacional, que junto a la latencia propia del canal de comunicación, generan tiempos de retardo en la visualización de la mega-imagen durante el proceso de navegación. Estos retardos, pueden llegar a interferir en el proceso de evaluación diagnóstico del patólogo. Por esta razón, es deseable estimar y anticipar las trayectorias de navegación que el patólogo pueda llegar a realizar.

Un método para la detección eficiente y precisa de las trayectorias de navegación permitirá el desarrollo e implementación de técnicas para la anticipación en la transmisión y reconstrucción de regiones de la mega-imagen histopatológica, mejorando la calidad de la navegación y disminuyendo la interferencia de los retardos en la evaluación diagnóstica realizada por el patólogo.

La tesis desarrollada en este documento, aborda la tarea de predecir qué regiones de una mega-imágen serán requeridas en alta magnificación por los patólogos durante una tarea de evaluación diagnóstica. Estas regiones usualmente están asociadas a conceptos histopatológicos determinantes para el diagnóstico. El enfoque utilizado se basa en el conocimiento del experto (patólogo) expresado usando imágenes de ejemplo. Una vez identificadas las regiones de interés, se analizan las condiciones bajo las cuales, esta capacidad predictiva puede resultar útil como política en una estrategia de prefetching/caching.

1.1. Identificación del Problema

Durante el exámen diagnóstico, el patólogo realiza operaciones sobre el microscopio para examinar con detalle diversas regiones de la lámina. En

los sistemas de microscopía virtual también es necesario realizar operaciones similares sobre la imagen pues la resolución de los dispostivos de visualización es menor que la resolución total de la mega-imágen. Estas operaciones se ven reflejadas en peticiones que se realizan al software para ajustar la ventana de visualización en ciertas regiones de la mega-imagen, con unas determinadas especificaciones de magnificación, calidad, tamaño de la ventana.

El conjunto de regiones visitadas junto con las respectivas características de magnificación, calidad, posición, y área de la ventana (en algunos casos se incluye información anotada por el experto [13]) se denomina trayectoria diagnóstica o de observación (Diagnostic Path). Un ejemplo de una trayectoria diagnóstica se muestra en la figura 1.1.

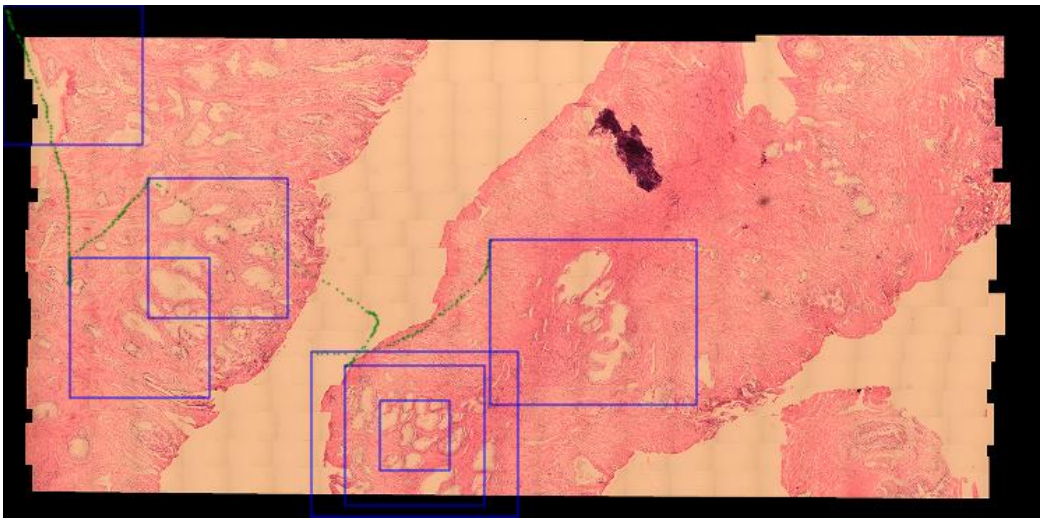

Figura 1.1: Ejemplo de trayectoria diagnóstica o de observación.

De forma concisa, la pregunta de investigación que se aborda en este trabajo consiste en: *¿Cómo anticipar esta trayectoria diagnóstica*, o de forma más explícita: *¿Cómo anticipar el requerimiento de regiones de la mega-imágen histopatológica (en alta magnificación) por parte del patólogo, durante la tarea de evaluación diagnóstica?*.

En la sección 2.1 se describe brevemente algunas características de la navegación de mega-imágenes histopatológicas, y se menciona que debido a la enorme cantidad de información en las láminas histopatológicas, el exámen diagnóstico se lleva a cabo mediante un examen secuencial de regiones de interés. Este comportamiento es similar a los procesos de búsqueda visual

descritos en el capítulo 3. Lo anterior es más claro si tenemos en cuenta que un componente importante en el análisis de especímenes histopatológicos es el proceso de búsqueda de regiones de interés. Podemos abordar de forma similar nuestro problema de investigación al problema de la atención visual selectiva, y simplificar nuestra pregunta de investigación mediante la formulación de las siguientes dos preguntas:

- ¿Cómo identificar regiones de interés en mega-imágenes histopatológicas?
- Una vez determinadas las regiones de interés, ¿Cómo determinar la secuencia de petición en alta magnificación de las regiones de interés?

Teniendo en cuenta estas dos preguntas, los objetivos del trabajo de investigación desarrollado se definen a continuación:

1.1.1. Objetivo General

- Proponer e implementar un método para la predicción de patrones de navegación en mega-imágenes histopatológicas.

1.1.2. Objetivos Específicos

- Definir e implementar una estrategia, usando conocimiento a priori del patólogo, para la asignación de relevancia, por regiones, en la mega-imagen.
- Proponer un modelo de navegación entre regiones de la mega-imagen que estime la probabilidad de que una región específica de la imagen sea visitada.
- Plantear un método de predicción de la trayectoria de navegación basado en la información entregada por el modelo de navegación.
- Implementar y evaluar la precisión del método propuesto en mega-imágenes histopatológicas.

1.2. Contribuciones Realizadas

La principal contribución del presente trabajo es la habilidad de introducir información específica del campo histopatológico y utilizar esta información para la formulación de estrategias caché/prefetching que puedan ser útiles para mejorar la reactividad de sistemas de microscopía virtual. Es cierto que en la literatura se encuentran sugerencias que discuten el uso de información específica de la imagen en estrategias de caché/prefetching, pero no se conoce la existencia de un método general que permita introducir este tipo de conocimiento.

En el desarrollo del trabajo, se analizó un enfoque novedoso para identificar regiones de interés en mega-imágenes histopatológicas, basado en la caracterización de ejemplos seleccionados por patólogos para la definición de regiones relevantes sobre la mega-imagen. Esta caracterización esta influenciada por la estructura de modelos de atención visual, entre cuyas características más prominentes se puede enunciar: El uso de múltiples características a bajo nivel relacionadas con orientación, color, intensidad y textura; la introducción de métodos para integrar la información multimodal de las diferentes características; y el uso de elementos objeto y distractores para aumentar el poder discriminativo del método.

También se estudiaron las condiciones del sistema del sistema de microscopía virtual bajo las cuales la información del campo específico puede resultar beneficiosa, determinandose que es más apropiada para sistemas con una reactividad baja. En sistemas con reactividad alta, fuentes de información relativas a la posición y velocidad del dispositivo de captura (mouse) pueden resultar más sencillos y confiables en el momento de realizar una predicción.

1.3. Estructura del Documento

El documento esta organizado de la siguiente forma: El capítulo 2 describe a grandes rasgos los sistemas de microscopía virtual, sus aplicaciones, limitaciones, nuevas tendencias y estrategias de caché/prefetching para facilitar la navegación de mega-imágenes histopatológicas. También se analiza el proceso de examén diagnóstico en patología, mediante la descripción de patrones de observación de los patólogos. La existencia de estos patrones de exploración permite abordar el problema de predicción de trayectorias diagnósticas. En el capitulo 3 se analizan estrategias de predicción/estimación usadas en el

modelamiento computacional de la atención visual. En el capítulo 4 se describe detalladamente el método utilizado para identificar regiones de interés en mega-imágenes histopatológicas y el uso de esta información para anticiparse durante la navegación interactiva en mega-imágenes histopatológicas, durante una evaluación diagnóstica. Finalmente en el capítulo 5 se enuncian algunas conclusiones y perspectivas para el trabajo futuro en este tema.

Capítulo 2

Sistemas de Microscopía Virtual en Patología

En el contexto de la patología médica, la microscopía óptica es una herramienta utilizada para estudiar las manifestaciones de las enfermedades que son visibles a nivel tisular y celular. Aún hoy en día, con la disponibilidad de diversas técnicas y procedimientos que permiten el exámen diagnóstico no invasivo, el análisis del especimen histopatológico es usualmente el último paso que prueba/desaprueba de forma categórica la existencia de una patología. La utilidad de este análisis no solo tiene valor diagnóstico, pues una valoración apropiada de los especímenes puede determinar la selección de una terapia efectiva. Aún más, el análisis estadístico de estas muestras histopatológicas en grandes poblaciones provee importante información acerca de los patrones y calidad de la salud en un país [5].

El exámen diagnóstico del material histopatológico no sería posible sin un proceso de preparación de los tejidos que termina en la construcción de láminas histopatológicas. Durante este proceso, el especímen es cortado en finas capas (cortes entre $1 - 10\mu m$) y sometido a la acción de tinciones con el objetivo de destacar las estructuras y arquitectura de los tejidos (que normalmente son transparentes) [14]. Al finalizar este proceso, se obtiene una placa o lámina de vidrio de algunos milímetros de espesor. Después de este proceso, el patólogo busca en la lámina, mediante el uso del microscopio óptico, características visuales que le permitan establecer un diagnóstico [15].

Si bien el ejercicio de la patología y el inseparable uso del microscopio óptico ha sido un sello distintivo en el cuidado médico de calidad por más de un siglo, el advenimiento de la era digital y las tecnologías de informa-

ción prometen cambiar este paradigma. Durante las últimas 2 décadas se han realizado crecientes esfuerzos por suministrar herramientas para hacer de la telepatología, definida como la práctica de la patología a distancia usando sistemas de captura de video y telecomunicaciones [16], una realidad. Particularmente, destacan las tecnologías relativas a la digitalización completa de las láminas histopatológicas y su utilización en un campo denominado microscopía virtual, cuyo principal objetivo consiste en emular la funcionalidad que provee el microscopio óptico tradicional mediante la manipulación de mega-imagenes digitales (láminas virtuales) en un computador convencional [17].

Este capítulo inicia con un análisis del proceso de evaluación diagnóstico realizado por patólogos, enfocandose particularmente en la descripción de patrones o mecanismos de navegación que utilizan los patólogos para llegar a un diagnóstico. Posteriormente, se realizará un breve exámen de los requerimientos y los procedimientos utilizados para que los sistemas de microscopía virtual sean implementados. Finalmente, se cerrará este capítulo revisando múltiples estrategias utilizadas para acelerar el proceso de navegación en mega-imágenes, incluyendo la identificación automática de regiones de interés diagnóstica.

2.1. El Proceso de Exámen Diagnóstico en Patología

Al abordar el problema de predicción de navegación en mega-imágenes histopatológicas, se debe considerar si el patólogo utiliza algún tipo de mecanismo estructurado para el proceso de análisis diagnóstico, susceptible de ser anticipado. Los estudios realizados sobre el tema muestran que aunque en principio el movimiento es caótico e intrincado, existe una diferencia clara entre los patrones de observación de un patólogo experto y un novato. El entrenamiento recibido y la experiencia adquirida modulan sensiblemente la tarea de navegación en el proceso de análisis diagnóstico histopatológico [18, 19].

La complejidad de la tarea diagnóstica en patología radica en que el patólogo obtiene información acerca de la anatomia microscópica de un órgano, que se encuentra descrita en un espacio tridimensional, a partir de información contenida en las dos dimensiones de una lámina histopatológica [20]. Es

por esta razón que durante la tarea diagnóstica, el experto no solo realiza un procedimiento de búsqueda de patrones (pattern matching), sino que involucra ampliamente sus conocimientos acerca de la arquitectura del tejido en observación, su conocimiento visual y su razonamiento para establecer hipótesis que permitan describir las características de la imagen histopatológica.

El uso del microscopio óptico hace que la evalución de láminas histopatológicas sea una tarea inusual dentro de las actividades médicas diagnósticas, pues no es viable examinar todo el campo (lámina) en su totalidad de una sola vez, y se requiere de un proceso de búsqueda visual para examinar el tejido. Con la ayuda del microscopio, el patólogo se mueve entre áreas de la placa histológica usando objetivos de diferente magnificación, buscando características que le permitan elaborar un diagnóstico [15]. Este aspecto determina la lentitud del proceso de evaluación diagnóstica. A diferencia de muchas tareas de clasificación visual, la imagen es codificada y entendida en pequeñas piezas a través de un proceso de búsqueda serial. El patólogo se enfoca en **áreas de interés** e ignora detalles irrelevantes, identifica características importantes y les asigna significado, y eventualmente converge en un diagnóstico que se adapta al patrón visual encontrado [18].

La utilización de tecnología de seguimiento de movimientos oculares para analizar los patrones de navegación durante el exámen diagnóstico de imágenes histopatológicas, revela un proceso de selección previa de áreas de interés durante la observación de la imagen en baja magnificación [21]. Durante este proceso, la atención del patólogo se dirige hacia regiones de interés candidatas que tienen un alta probabilidad de contener información diagnóstica. A continuación, y en cuestión de segundos, el patólogo selecciona estas regiones de interés y las visualiza a un nivel de magnificación mayor para realizar un exámen detallado.

Este patrón de selección también se observa en el uso de sistemas de microscopía virtual. Los patólogos tienden a observar en detalle regiones específicas de una lámina virtual, y aún más, estas regiones representan un interés común para un grupo de patólogos [22]. También es interesante resaltar que la transición entre regiones de interés visitadas en alta magnificación se realiza de forma casi lineal [22, 23], y que la velocidad que se observa en estas transiciones presenta patrones isotónicos [24]. Esta información de velocidad y dirección durante la navegación ha sido usada exitosamente en el diseño de estrategias para mejorar los tiempos de exploración de mega-imágenes histopatológicas [24].

Como cualquier proceso de búsqueda visual, por ejemplo los movimien-

tos sacádicos de los ojos en análisis de imagenes naturales, la evaluación diagnóstica en patología esté sujeta a una inmensa variabilidad debido a que los patrones de búsqueda de cada uno de los patólogos son idiosincráticos [25]. Aún así, la existencia de regiones de interés diagnóstico en la mega-imágen, denota la posibilidad de una predicción confiable, puesto que éstas son un elemento común de búsqueda en el análisis diagnóstico de muestras histopatológicas.

2.2. Sistemas de Microscopía Virtual

La microscopía virtual se basa en el uso de sistemas de software y hardware que permiten crear y/o observar láminas histopatológicas virtuales. Estas láminas virtuales son mega-imágenes que corresponden a la digitalización de láminas histopatológicas, y que pretenden transmitir fielmente la información contenida en éstas [26]. El desarrollo de herramientas en MV provee mecanismos para el intercambio de información, el análisis múltiple de muestras histológicas y la educación en el área histopatológica [8–10, 27–29].Estas tecnologías han generado grandes expectativas en la comunidad médica dedicada a la histopatología, pues se estima que jugarán un rol importante en el diagnóstico, enseñanza, entrenamiento e investigación médica [26].

El principal objetivo durante la construcción de una lámina virtual es el de conservar todas las características importantes de la placa real, y a la vez ofrecer las ventajas que tiene la digitalización de la información. Para ello, las láminas virtuales requieren de herramientas de software especializado.

Los sistemas para la visualización de estas mega-imágenes deben asegurar al menos una capacidad de manipulación similar a la del microscopio óptico. Debido a que la resolución de las mega-imágenes histopatológicas es mayor que la que se tiene en los dispositivos de visualización convencionales, no es factible presentar la lámina virtual en su totalidad. Aún si fuera posible, una visualización completa sería poco práctica, pues no explotaría el comportamiento de los patólogos durante la navegación, que consiste en un exámen en baja resolución para definir regiones de interés, seguido de un examen en alta magnificación de las regiones potencialmente relevantes para la tarea diagnóstica [15]. Para facilitar la exploración de mega-imágenes, es común el uso de herramientas que permitan interactuar con el experto/usuario [1]. Un tipo de interfaz gráfica ampliamente utilizado en este tipo de aplicaciones consiste de dos ventanas: La primera contiene una versión en baja mag-

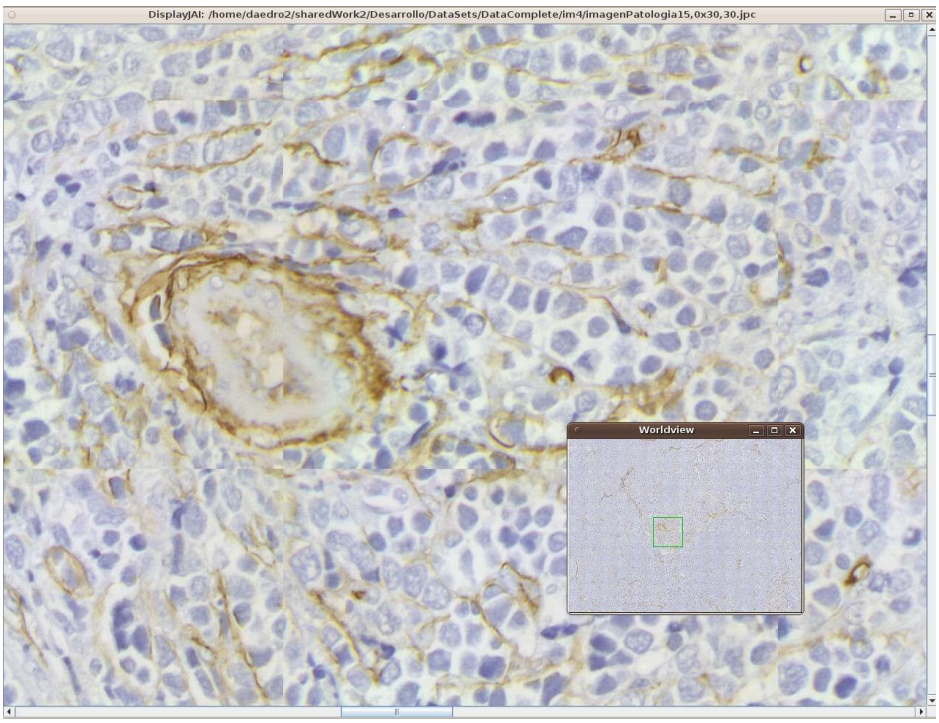

Figura 2.1: Visualización de una lámina virtual. Observese que esta interfaz provee de una visión de la imagen en baja y alta magnificación.

nificación de toda la mega-imagen (thumbnail). En esta ventana existe una herramienta en forma rectangular que permite seleccionar una porción de la mega-imagen. En la segunda ventana, la región seleccionada es observada en alta magnificación. Un ejemplo de este tipo de interfaz se observa en la figura 2.1.

De manera concreta las láminas histopatológicas virtuales y los sistemas de microscopía virtual deben poseer diversas características para asegurar su usabilidad, entre ellas [4, 26]:

- Reproducción fiel de la placa real.
- Claridad y resolución suficientemente alta.
- Acceso aleatorio a la información en la imagen.
- Representación robusta a niveles diferentes de magnificación.

- Transmisión rápida sobre la red.
- Facilidad para la preparación, observación, exámen y almacenamiento.
- Manipulación intuitiva y fácil de llevar a cabo.
- Capacidad de realizar anotaciones.
- Facilidad para realizar comparaciones entre láminas virtuales.
- Integración con información del paciente.
- Interfaz Adaptable

La disponibilidad de la tecnología de microscopía virtual está sujeta a la implementación de diferentes procedimientos, desde el proceso de digitalización hasta la visualización (ver figura 2.2) [4]. A continuación se describen a grandes rasgos estos procedimientos:

- Ensamble de la imagen:

  En este paso, varias imágenes son ensambladas en una sola mega-imagen de alta resolución. Cada una de las imágenes es obtenida a través de la captura del campo de visión del microscopio en alta magnificación sobre una región específica de la placa histopatológica. Las imágenes en su totalidad corresponden a un escáneo completo de la superficie de la placa. La coherencia en la información durante el proceso de ensamble se asegura mediante técnicas que permiten la correcta alineación o registro de las diferentes imágenes, usando para ello la información común en dos campos microscópicos adyacentes. Después de digitalizada, una lámina virtual tiene un tamaño de entre 1-20GB [26]. Este tamaño puede resultar considerable para un uso rutinario.

- Almacenamiento:

  La enorme cantidad de información contenida en una mega-imagen de histopatología hace imprescindible el uso de estrategias de compresión para un óptimo uso de los recursos computacionales. Sin embargo, esta compresión en la imagen debe estar sometida a la restricción de un acceso eficiente y aleatorio a la información. Además se debe asegurar

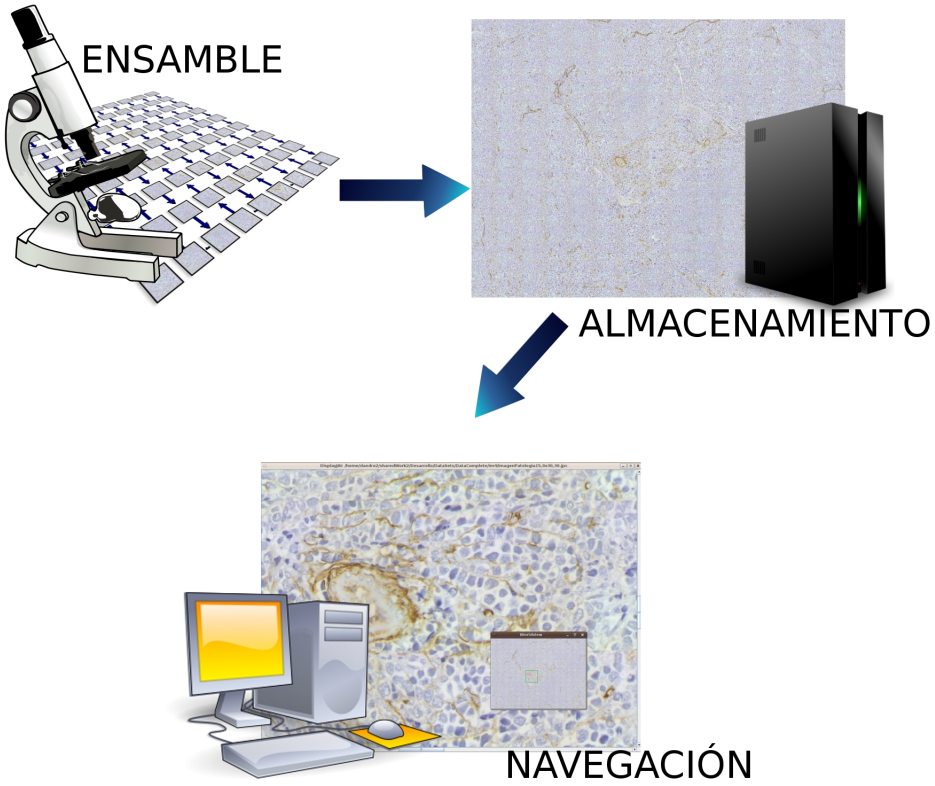

Figura 2.2: Procedimientos en microscopía virtual

la mayor fidelidad posible a la lámina histopatológica real, pues distorsiones en la lámina virtual pueden inducir errores en el proceso de diagnóstico [30].

Si bien es cierto que la capacidad de almacenamiento ha venido creciendo constantemente, el uso de la microscopía virtual de forma rutinaria requeriría una cantidad enorme de espacio, si se tiene en cuenta que un hospital especializado produce entre 10000 y 50000 láminas histopatológicas cada año [31].

- Navegación:

  Este componente, permite a un usuario llevar a cabo un exámen microscópico de la muestra de forma similar a la utilización de un microscopio convencional: Movimientos traslacionales (eje XY) o de acercamiento/alejamiento (eje Z).

La utilización de sistemas de microscopía virtual en aplicaciones acceso remoto, además de un apropiado proceso de compresión, requiere de la utilización de una configuración cliente-servidor [7, 32]. En el servidor, las láminas virtuales son almacenadas y en el cliente, una interfaz gráfica adecuada se encarga de la interpretación de la información enviada por el servidor para la navegación de la imagen y de la interacción con el usuario.

El tamaño de una lámina virtual, el proceso de transmisión, decompresión y visualización introducen retrasos en la respuesta que hacen difícil una navegación interactiva y fluida en el cliente. Por la razón anterior, el uso de estrategias para acelerar la navegación se ha establecido como un tema vigente de investigación dentro de los sistemas de microscopía virtual. [4, 23].

2.3. Estrategias para Mejorar la Eficiencia en la Navegación de Mega-imágenes

En la sección 2.2, se mencionó que el uso de sistemas de MV usualmente supone la utilización de sistemas cliente/servidor, donde el servidor se encarga de la tarea de almacenamiento de las mega-imágenes y el cliente se encarga de la tarea de visualización. Debido a la utilización de este tipo de arquitectura, y la carga computacional que representa el manejo de la información contenida en las mega-imágenes a través de una red informática, es usual que se usen estrategias de caché/prefetching para optimizar el uso de los recursos computacionales y acelerar la exploración de las mega-imágenes. Una estrategia de almacenamiento en caché se define como un almacenamiento temporal de información utilizada en memoria RAM para su posterior utilización [33]. Este tipo de estrategias usualmente implica la utilización de políticas de reemplazo de la información contenida en la memoria. Uno de los algoritmos más utilizados en microscopía virtual se denomina LRU (Least Recently Used), que consiste en la actualización de la información contenida en el caché mediante el reemplazo de la información que tiene un tiempo más largo sin ser utilizada. En microscopía virtual está vinculada con la suposición de que el patólogo visitará aquellas regiones de la imagen que ha visitado recientemente [34]. Otra técnica conocida consiste en la LFU (Least Frequently Used), en la cual la información cuya frecuencia de uso es menor

es descartada del caché. En el caso de la microscopía virtual este algoritmo promoveria aquellas regiones de la lámina virtual que han recibido un número mayor de visitas [23,34]. Las técnicas de prefetching, por su parte, corresponden a la carga anticipada de información que, se espera, será relevante en un futuro cercano [35]. Este tipo de técnicas ha sido utilizada en la predicción anticipada de movimientos en la navegación de ambientes virtuales [36], y en la navegación de grandes imágenes [35].

La utilización de estrategias de prefetching y de caché en imágenes para reducir el tiempo de respuesta ha sido propuesto por varios autores. Por ejemplo, en [37] se argumenta que el uso de técnicas de caché específicas de la imagen pueden proveer una mejora sobre técnicas de caché mas generalizadas, mediante la recodificación de las imágenes en baja resolución en el caché (técnica conocida como soft-caché). Aunque los métodos de identificación de regiones de interés no han sido directamente utilizadas en estrategias de prefetching/caching, si han resultado efectivas en métodos de compresión selectiva y en la adaptación de imagenes naturales a pantallas pequeñas [38–40], cuyo objetivo común es el mejoramiento de la navegación interactiva de imágenes. En [1] se describe un mecanismo basado en el formato JPEG2000, mediante el cual se logra la reutilización de información en baja resolución de regiones ya visitadas (soft-caché), para aumentar la rapidez en la respuesta del sistema. En [34] se describe una estrategia combinada para la navegación eficiente de mega-imágenes usando una estrategia de caché espacial (basada en peticiones de regiones adjacentes), y una estrategia de soft-caché usando también el formato JPEG2000.

Las estadísticas de navegación sobre las imágenes puede ser información relevante para establecer estrategias de caché o prefetching. En [11] la información de navegaciones previas sobre la imagen es utilizada para definir probabilidades de operaciones de acercamiento/alejamiento en zonas de la imagen y definir así el envío anticipado de regiones de la imagen (server prefetching). Esta información es usada en una estrategia de caché combinada que hace uso de estrategias locales (basado en el algoritmo LRU-least recently used) y globales (basado en estadísticas de navegación sobre la imagen).

La historia de las acciones de navegación del usuario también pueden representar un buen índice para realizar predicciones. En [35] se propone un modelo basado en las acciones de navegación del usuario para anticiparse en los bloques necesarios en la siguiente acción de navegación. Básicamente, se asigna una probabilidad de visita mayor a los bloques que se encuentran en

la dirección de navegación estimada. La utilización de métodos de predicción basados en el modelamiento del perfil del usuario también han sido utilizados exitosamente en la navegación de mega-imágenes histológicas [24]. El perfil de usuario corresponde a una compleja interacción de factores, como el tipo de interfaz (GUI), la mecánica neuromuscular, y la magnificación del sistema de MV [24].

Por otro lado, el uso de esquemas que pretenden modelar la atención visual en problemas similares (transmisión de imágenes a dispositivos móviles) [40–42], hace uso de la información contenida en la imagen. Este tipo de enfoques, localiza sectores u objetos relevantes en la imagen para propiciar operaciones de compresión selectiva o de navegación automática [43] que disminuyen los requerimientos en la transmisión y visualización de la imagen. En este caso, la atención visual en la imagen se define como un conjunto de objetos de atención que contiene información y que "atrapan" la atención del usuario [39].

El impacto de las estrategias de prefetching y caché sobre la navegación en imágenes depende básicamente de las características de la navegación. Descampe et al. realiza un estudio detallado donde se investiga las condiciones bajo las cuales un conocimiento a priori acerca del usuario puede mejorar la reactividad del sistema [44]. Los resultados muestran que el beneficio que reportan este tipo de estrategias en el envío (transmisión) de información está determinado por la capacidad de *anticipación* de las acciones del usuario. Por ello, abordar el problema de anticipación durante la navegación de mega-imágenes, es un ejercicio imprescindible antes de establecer estrategias caché/prefetching efectivas. Más aún si esta navegación corresponde a una tarea de búsqueda particular como es la evaluación diagnóstica de mega-imágenes histopatológicas (ver sección 2.1).

En esta sección se mostraron algunas técnicas de caché/prefetching para la navegación interactiva de navegación. Algunas de ellas, basaban su poder predictivo sobre las acciones de navegación previas del usuario (velocidad y dirección del movimiento del mouse). De esta forma un requisito importante es que la interfaz gráfica restringa los grados de libertad del usuario, exigiendo por ejemplo que la operación sea llevada a cabo mediante un arrastre del mouse, para asegurarse así la información necesaria (posición y velocidad del puntero del mouse). La ventana de tiempo que estos métodos permiten para anticiparse a peticiones de regiones en alta magnificación es relativamente corta, pues el movimiento en cuestión es bastante rápido.

La utilización de estadísticas de navegación sobre la imagen es una solu-

ción sencilla y eficaz, pues es un hecho que los patólogos tienden a visitar regiones de interés comúnes, pues son precisamente estas regiones las que contienen elementos clave para la evaluación diagnóstica [18, 21]. Sin embargo, con un número pequeño de visitas previas no es posible del todo utilizar esta técnica. Además, la información de estadísticas espaciales por sí sola no es generalizable a otras láminas virtuales, pues la disposición espacial de regiones de interés en este tipo de imágenes tiene un caracter aleatorio. Para realizar este tipo de abstracción se requiere de un procesamiento previo de caracterización de las regiones de interés de la mega-imagen.

Las técnicas de soft-caché son adecuadas para las aplicaciones de microscopía virtual, pues es habitual que la información de la lámina virtual se encuentre almacenada en un formato multiresolución, y que la visualización de la lámina generalmente emplee versiones de la mega-imagen en diferente resolución/magnificación (ver sección 2.2). Sin embargo, no se puede desplegar toda la capacidad de esta técnica de caché, si no se conoce de antemano información acerca de las regiones potencialmente más visitadas de la imagen [37], lo cual lleva a la necesidad de desarrollar técnicas para la identificación automática de regiones de interés. Aunque en estos trabajos se ha sugerido la implementación de este tipo de técnicas, no se conoce una implementación de este tipo en estrategias de caché prefetching.

Capítulo 3

Identificación de Regiones de Interés en Tareas de Búsqueda Visual

La identificación de regiones de interés es necesaria dentro de los procesos de búsqueda visual debido a la capacidad limitada de procesamiento del sistema visual humano. El análisis de una escena visual completa se debe realizar mediante un proceso secuencial. Para ello, es importante la existencia de mecanismos que permitan guiar la capacidad de procesamiento hacia regiones de la imagen (regiones de interés) que son susceptibles de contener información relevante para la tarea que realiza el observador [45].

En el capítulo introductorio (ver sección 1.1) se analizó el problema de anticipación de trayectorias diagnósticas en mega-imágenes histopatológicas. La selección de regiones de interés en el exámen diagnóstico de mega-imágenes histopatológicas, permite un análisis secuencial de la mega-imagen cuyo procesamiento es menos exigente que realizar un examen exhautivo. Este comportamiento, es similar al mecanismo de atención selectiva del sistema visual humano. Esta similitud induce a hipotetizar que la identificación previa de regiones de interés diagnóstico es un mecanismo necesario para la anticipación de trayectorias diagnósticas.

En este capítulo se explora el sistema visual humano, las teorías psicofísicas de atención visual, y la implementación computacional de algoritmos para identificar regiones de interés y de atención visual. En términos generales, estos modelos computacionales implementan mecanismos para definir que porciones de la imagen deben ser procesadas por la fóvea, mediante el

concepto de saliencia.

3.1. Organización del Sistema Visual Humano

La visión es frecuentemente catalogada como la modalidad sensorial primaria en primates y humanos. Prueba de ello es la enorme cantidad de información (estimada en el orden de $10^6 - 10^8$ bits por segundo a nivel del nervio óptico [46]) que suministra el sentido visual, la complejidad de las estructuras que proveen la transducción de las ondas luminosas en actividad neuronal, y la importante porción de nuestro cerebro que se encuentra dedicada a tareas relacionadas con la percepción visual [47].

3.1.1. Anatomía del Ojo

El procesamiento visual empieza en los ojos, que son el instrumento sensor en el cual se reciben la luz, y se transforma en impulsos neuronales. La función básica del ojo consiste en recibir las ondas lumínicas y enfocarlas en una estructura laminar fina de células sensoriales receptoras especializadas que se encuentran en la parte posterior del ojo. El globo ocular esta conectado a un elaborado arreglo de musculos que permite que el ojo pueda seguir estimulos objetivo en el ambiente. Los lentes dentro del ojo, cuya función consiste en enfocar la luz, también están conectados a músculos que permiten alterar su forma. Esto permite que se pueda enfocar objetivos a diferente distancia en la parte posterior del ojo, donde finalmente se transforma la energía de las ondas lumínicas en señales nerviosas. Estas señales son modificadas en la retina, de tal forma que se hace enfasis en los cambios y discontinuidades en la iluminación antes de que la señal viaje hacia el cerebro a través del nervio óptico.

El ojo es una esfera de alrededor de 2,5cms de diámetro. La esclera es la capa externa del ojo y se caracteriza por el color blanco de las fibras fuertemente entretejidas que lo componen. En el frente del ojo, la superficie sobresale levemente para formar la cornea, que es transparente y permite la entrada de la luz. Detrás de la cornea se encuentran un conjunto de musculos en forma anular llamado iris. En el centro se encuentra la pupila, y la cantidad de luz que entra en el ojo es determinado por el diámetro de ésta. Los músculos en el iris permiten aumentar o disminuir la apertura de la pupila. Despúes de que la luz pasa a través de la pupila y la camara anterior

del ojo, se encuentra con los lentes del ojo, una estructura biconvexa denominada cristalino. La camara anterior esta llena con un fluido denominado humor acuoso. Este fluido transporta nutrientes y oxigeno a las estructuras cercanas, suplantando la función normalmente ocupada por la sangre, pues ésta obstacularizaria el paso de la luz.

La córnea y el cristalino alteran el paso de la luz de tal forma que se encontrará enfocada en la superficie que se encuentra en la parte posterior del ojo, cubierta por la retina. La forma biconvexa del cristalino invierte la imagen de arriba abajo. Esta inversión no es importante pues las posiciones relativas de los objetos en la imagen son preservadas. Después de pasar a través de los lentes la luz atraviesa el humor vítreo antes de llegar a la retina. El humor vítreo es una sustancia gelatinosa, que a diferencia del humor acuoso no es reemplazada de forma constante.

3.1.2. Organización en la Retina

Se pueden diferenciar tres capas principales en la retina: La capa de células receptoras, la capa bipolar, y la capa de células ganglionares. La capa de células receptoras se encuentra en la parte posterior de la retina, de tal forma que la luz tiene que pasar a través de las capas transparentes sobre ella. Los fotoreceptores forman sinapsis con las células bipolares que también forman sinapsis con las células ganglionares, cuyos axones viajan a través del nervio óptico hasta el cerebro. Todos los axones de las células ganglionares se reunen en un punto llamado disco óptico, desde el cual abandonan la estructura interna del ojo, y se dirigen al cerebro. El disco optico produce un punto ciego en el campo visual, pues no existen fotoreceptores en este punto. No somos conscientes de este punto ciego debido a que nuestro sistema visual llena este "hueco" en algunas áreas de la corteza visual que posteriormente procesan la información visual. La retina también incluye algunas células que integran horizontalmente la información en paralelo (células horizontales y células amacrinas) combinando o sustrayendo mensajes de fotorreceptores adjacentes.

Los fotorreceptores, que son las células en la retina encargadas de la conversión de la información lumínica en patrones de actividad neuronal, se dividen en conos y bastones. Los bastones son capaces de detectar la intensidad del haz luminoso, mientras que los conos son capaces de detectar la longitud de onda (expresada como color). Los conos se encuentran concentrados en una pequeña area de la retina llamada fovea y existen tres clases de

conos. Entre el 5 % y el 10 % del total de la población de conos corresponden a conos que son capaces de detectar las longitudes de onda correpondiente al color azul. El resto de los conos son rojos y verdes y se encuentran en proporción de 2 a 1.

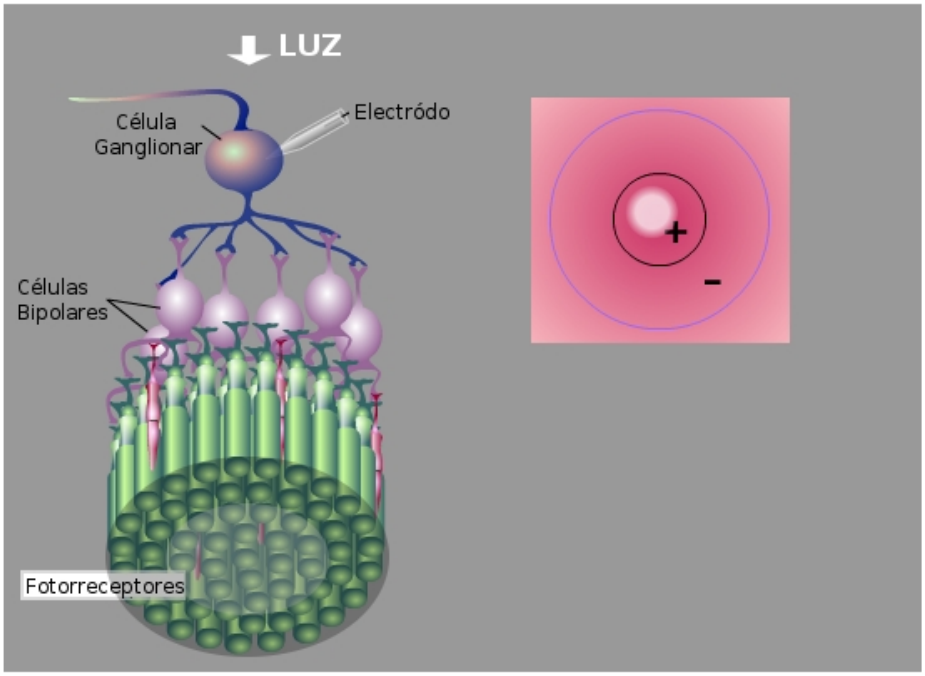

Figura 3.1: Organización celular en la retina. Mecanismo centro-periferia

En este temprana etapa del sistema visual humano, las estructuras envueltas en el proceso exhiben una complejidad sorprendente, y algunos de los mecanismos en la retina no estan plenamente comprendidos. De la retina parten una cantidad cercana a 1 millón de axones hacia el cerebro. Estos axones pertenecen a las células ganglionares de la retina. Existen alrededor de 126 millones de receptores, así que durante la transmisión de la información hacia las capas superiores de la retina existe un proceso de integración de la información. Esta integración tiene como objetivo la detección de patrones de luz que permita determinar la existencia de bordes.

Cada una de las células ganglionares esta conectada a un conjunto de células fotorreceptoras a través de las células bipolares de la retina. Este conjunto de fotorreceptores corresponde a un área en la imagen formada en

la retina y esta área se denomina el campo receptivo de la célula ganglionar. Los fotorreceptores en el campo receptivo estan organizados para permitir un mecanismo *centro-periferia*. Por ejemplo, la luz que excita fotorreceptores en el centro del campo receptivo excita (inhibe) la célula ganglionar, mientras que la luz que excita fotorreceptores en la periferia del campo receptivo inhibe (excita) la célula ganglionar.

3.1.3. Corteza Visual y Áreas Superiores de Procesamiento Visual

Los axones de las células ganglionares se juntan y forman el nervio óptico que abandona el ojo a través del disco óptico. El nervio óptico se proyecta hacia el cerebro en el nucleo lateral geniculado (NLG), que es una zona ubicada en el tálamo cerebral. Ambos hemisferios reciben información de la imagen de los dos ojos, pues los nervios opticos provenientes de los ojos, se intercambian en el quiasma óptico. En este punto los axones correspondientes a las células fotorreceptoras mediales se cruzan, de tal forma que el hemisferio izquierdo (derecho) recibe la porción derecha (izquierda) de la escena visual. El arreglo topográfico de los campos receptivos en las células ganglionares se conserva en el NLG. Las neuronas en el NLG muestran respuestas características similares a las células de las cuales reciben señales eléctricas.

Las neuronas en el NLG proyectan sus axones hacia la corteza visual primaria (V1), también conocida como la corteza estriada. Aquí, la información proveniente de la retina y el NGL es separada y empacada para un analisis mas elaborado en las áreas visuales especializadas del cortex extraestriado. La respuesta característica de las neuronas en V1, es completamente diferentes de las encontradas en la retina o en el NLG. Algunas de ellas incluyen: Sensibilidad a las lineas y barras de diferentes orientaciones y movimientos, especialización de algunas neuronas en el color.

El estudio experimental de la respuesta neuronal de la corteza visual ha determinado que el procesamiento de la información visual para el reconocimiento de objetos se lleva a cabo en forma paralela y jerárquica [48]. Inicialmente, es procesada en áreas dedicadas, que responden a diferentes características: orientación, forma, color, movimiento. Posteriormente, la información procesada es reensamblada en una percepción coherente de nuestro entorno visual en áreas visuales superiores. Así, es notable que a medida que se va avanzando en el sistema visual humano desde el cortex visual (V1) hacia

áreas visuales superiores (como el cortex inferior temporal - AIT), las neuronas responden a estímulos cada vez más complejos, llegando a representar objetos específicos bajo una cantidad variable de condiciones. Un esquema representando está jerarquización se observa en la figura 3.2. Ejemplos de elementos para los cuales las neuronas de un área visual particular presenta respuesta selectiva se muestran dentro de campos receptivos (representados a través de círculos) de diferentes tamaños.

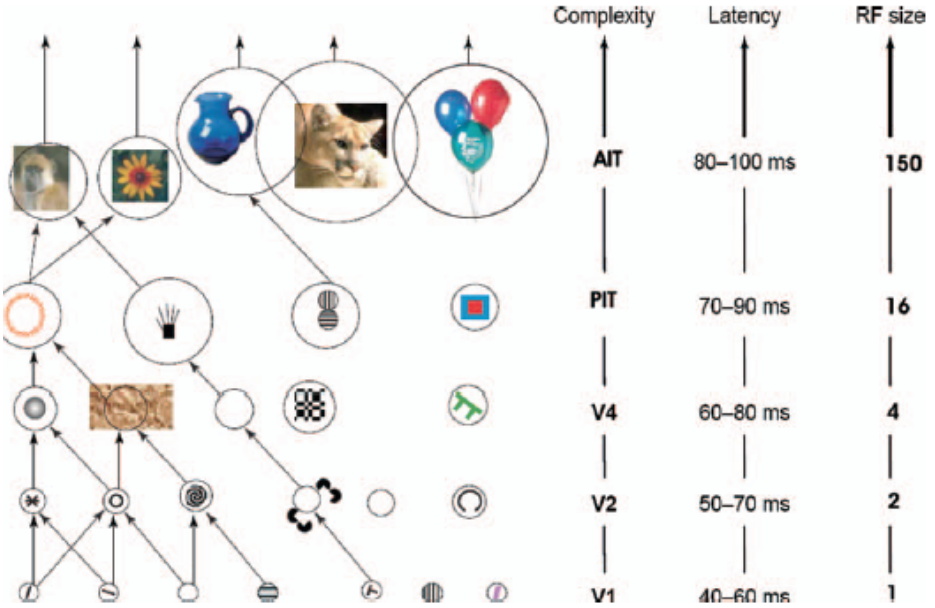

Figura 3.2: Jerarquía del sistema visual humano. Tomado de [47]

Obsérvese que la representación invariante de objetos en las capas superiores se logra mediante la integración progresiva de las respuestas que se obtienen en las áreas de procesamiento visual que se encuentran a un nivel más bajo. Esto también implica que la respuesta en las neuronas de áreas visuales superiores abarca una región mayor del campo visual (campo receptivo). De igual forma el tiempo de respuesta de las áreas visuales superiores es mayor, pues es imprescindible que la información haya sido procesada en áreas visuales que se encuentran más abajo en la jerarquía visual.

La conexión entre las áreas de procesamiento visual no se da solamente en forma unidireccional. Existe un importante mecanismo de retroalimentación desde las áreas superiores que, se supone, modula el comportamiento de las

neuronas en las zonas de procesamiento inicial. De igual manera, áreas de procesamiento pueden influir en otras del mismo nivel de jerarquía mediante conexiónes laterales. En total, se ha determinado la existencia de 32 áreas visuales conectadas por al menos 305 conexiones. A pesar del evidente progreso en la identificación de la funcionalidad de áreas del cerebro en el procesamiento visual, aún no es claro la forma en que interactuan las áreas visuales ni cómo se logra la reconstrucción, representación e identificación de objetos en una escena visual.

3.2. Teorías Psicofísicas de la Atención Visual

La atención se define como el mecanismo de selección que permite el manejo eficiente de la capacidad limitada del cerebro en el procesamiento de la información visual [49]. Este manejo eficiente se logra al enfocar la capacidad de procesamiento visual en un determinado lugar del campo de visión , obligando a que el análisis visual completo sea llevado a cabo mediante una serie consecutiva de movimientos en la escena visual cuyo procesamiento es menos exigente.

El campo visual humano tiene una resolución espacial variable. Por esta razón, el proceso secuencial esta diseñado para permitir que diferentes regiones de la escena visual o imagen sean procesados por la fóvea, que es una pequeña porción de la retina caracterizada por tener la máxima resolución dentro del campo visual. Esta región dispone de una gran cantidad de sensores neuronales en la retina, así como de un importante porcentaje del procesamiento en cada una de las áreas visuales superiores.

Se ha determinado que existen dos mecanismos que guían la percepción visual [50]: uno dado por la información intrínseca en la escena visual o "preatentiva" y otro determinado por la tarea realizada por el observador. En el mecanismo preatentivo características simples son procesadas rápidamente y de forma paralela, mientras que en el segundo un foco de procesamiento especializado, denominado foco de atención, es dirigido a un lugar determinado en el espacio visual, de acuerdo a la tarea que esta desarrollando el observador.

Un modelo psicofísico importante acerca de la atención y búsqueda visual es conocido como el modelo de integración de características propuesto por

Treiman y Gelade [51]. De acuerdo a este modelo, la escena visual es codificada a lo largo de un número de dimensiones separadas: color, orientación, frecuencia espacial, iluminación, dirección de movimiento. Estas características son registradas de forma paralela y automática en un estadío temprano. La atención se constituye en un elemento integrador de las diferentes características que permite la correcta interpretación de un objeto como un todo. Un importante resultado de esta teoría afirma que los elementos u objetos en una tarea de búsqueda visual con una característica distintiva requieren de un tiempo de búsqueda independiente del tamaño del espacio de búsqueda (el efecto pop-up). Por el contrario, si la búsqueda del objeto requiere de la conjunción de varias características (por ejemplo forma y color), el tiempo de búsqueda depende del tamaño del espacio de búsqueda.

La teoría del mapa de saliencia de Koch y Ullman [50] propone la existencia de un mapa que integra las diferentes características visuales en un mapa coherente. Posteriormente un proceso de selección basado en la saliencia local de las características determina regiones candidatas para su análisis atentivo. La existencia de este mapa en la corteza visual, es aún materia de controversia. A pesar de ello, este enfoque ha sido utilizado con éxito en implementaciones computacionales para el análisis de imágenes naturales. Este modelo, representa una excelente aproximación del mecanismo "preattentivo" (bottom-up), pero no ofrece una discusión acerca del mecanismo dependiente de la tarea (top-down).

El mecanismo top-down de la atención visual puede ser explicado mediante la teoría de búsqueda guiada propuesta por Wolfe [52]. De acuerdo a esta teoría un conjunto limitado de atributos, relacionados con aquellas adquiridas con el procesamiento temprano de la visión (nuevamente: color, orientación, frecuencia espacial, iluminación, dirección de movimiento), pueden ser usados para guiar la selección de objetos. De esta forma, la búsqueda de un objeto con determinadas características, por ejemplo una manzana roja, puede ser guiada a través de la modulación en la importancia de ciertos atributos, en este caso el color rojo. El núcleo de esta teoria consiste en el hecho de que las características usadas en el procesamiento visual temprano, sirven para guiar los despliegues de búsqueda selectiva.

3.3. Metodos Computacionales para la Identificación de Regiones de interés y Atención visual

En la sección 2.3 se mencionó que el conocimiento específico de la imagen podría ser suficiente para establecer estrategias de optimización durante la navegación de mega-imágenes. Enfoques clásicos para encontrar regiones de interés en imágenes naturales han consistido en identificar regiones de la imagen con una alta densidad espacial de bordes [53]. En imágenes médicas, la selección de regiones de interés ha sido abordado usando diferentes métodos. Por ejemplo, Karras et al. [54] en imagenes en escala de grises de cancer abdominal, hipotetiza que regiones con una alta densidad de patrones repetitivos son de interés diagnóstico. Estas características fueron clasificados en regiones importantes o no, usando un algoritmo de fuzzy c-means clustering. Este tipo de enfoques difícilmente funcionaría en imágenes histopatológicas, pues éstas se caracterizan por contener regiones con alta concentración de bordes que no poseen sentido clínico [55].

Un exámen diagnóstico en patologia, es el resultado de una serie compleja de actividades que un patólogo experto domina. Teorías psicofisicas clásicas sugieren que estas tareas visuales complejas, incluyendo el exámen histopatológico, involucra altos grados de atención visual [56]. Existe evidencia mostrando que el sistema visual integra las características a bajo nivel de un objeto [51]. Estos hallazgos han inspirado diversos algoritmos computacionales que de alguna manera buscan el conjunto de características a bajo nivel que atraen la atención visual en una determinada imagen [46]. Uno de los modelos más influyentes es el propuesto por Itti et al. [57, 58]. un modelo de atención bottom-up que localiza porciones relevantes de la imagen, basado en un mapa constituido de tres características a bajo nivel: color, intensidad y orientación. A partir de estas consideraciones, el problema de detección de regiones de interés en mega-imágenes histopatológicas ha sido abordado mediante la utilización de un modelo Itti modificado. El ajuste de los parámetros del modelo y una estrategia de aprendizaje sobre imágenes anotadas manualmente, ha resultado ser un enfoque eficaz para la identificación de regiones patológicas en imágenes patológicas de carcinoma basocelular [59]. La utilización de esquemas basados en modelos computacionales de atención visual para la detección de regiones de interés en mega-imágenes histopatológicas es razonable, pues generalmente las patologías ocasionan serias distorsiones en

la estructura de los tejidos. Estas distorsiones hacen que el tejido patológico sea fundamentalmente diferente del tejido normal circundante. Un modelo de atención visual basado en características a bajo nivel tiene una buena probabilidad de captar estas diferencias y por tanto de detectar regiones de interés en algunas megaimágenes-histopatológicas [59]. Sin embargo, este método no puede ser generalizado a todas las patologías pues, tal como se muestra en la sección 2.1, la detección de regiones relevantes requiere de una búsqueda compleja de patrones, que el patólogo aprende a reconocer tras años de entrenamiento médico. Este entrenamiento es necesario, pues los patrones visuales que estan asociados a conceptos diagnóstico no siempre coinciden con regiones que los modelos computacionales definen como salientes en la imagen. La ausencia de información relacionada con la tarea realizada representa una fuerte limitante en la capacidad discriminativa de modelos de atención visual para detectar regiones de interés en mega-imágenes histopatológicas.

3.3.1. Modelos Computacionales de Atención Visual: Modelo de Itti.

El modelo propuesto por Itti [57], se ha constituido en un referente importante en el modelamiento **computacional** del sistema de atención visual.

Este modelo de atención visual se basa en la teoria de un mapa de saliencia [50] descrito anteriormente y modela el comportamiento preatentivo ("bottom-up"). Un diagrama de este modelo se puede apreciar en la figura 3.3.

Inicialmente, el modelo extrae la información de color, intensidad y orientación, y las codifica separadamente en representaciones piramidales multiresolución (multiples escalas). A continuación, se aplican operaciones de normalización y de centro-periferia ("center-surround") sobre la representación piramidal. La idea tras estas operaciones es emular el mecanismo centro-periferia que se observa en el campo receptivo visual observado en las neuronas de la retina, el nucleo geniculado lateral, y la corteza visual primaria (ver sección 3.1.2 y 3.1.3). Este tipo de arquitectura es sensible a discontinuidades locales y permite la localización de regiones que son sustancialmente distintas de su entorno [57]. Tras este proceso se obtiene un conjunto de mapas de características por cada una de los canales iniciales (color, intensidad y orientación). La integración no lineal a lo largo de las escalas (o resoluciones) genera tres mapas que identifican regiones sobresalientes en el

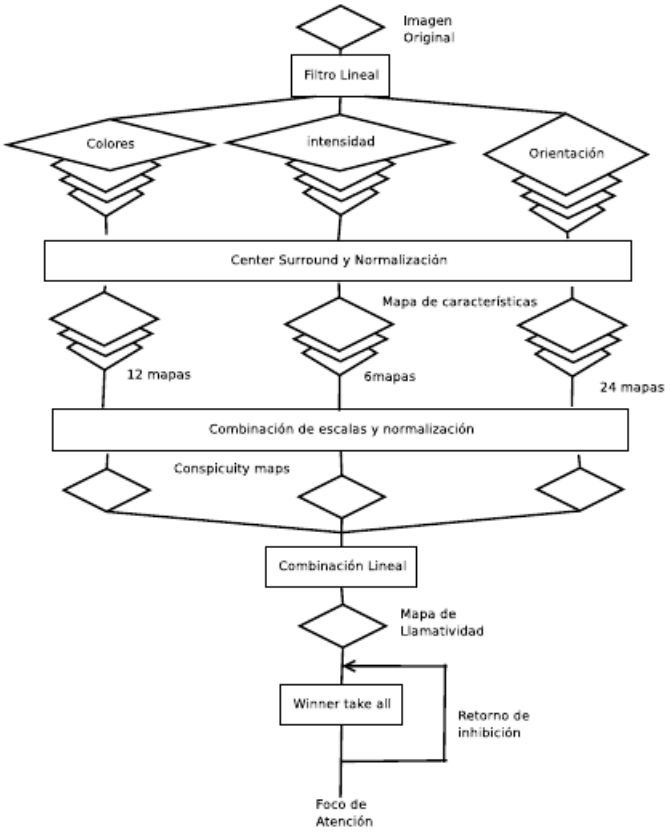

Figura 3.3: Modelo de atención de Itti. Tomado de [60]

dominio del color, la intensidad y la orientación respectivamente. Finalmente, una combinación lineal de los tres mapas produce el mapa de llamatividad o saliencia. Sobre este mapa es posible aplicar un método iterativo para calcular una secuencia de visita sobre regiones relevantes de la imagen original.

A continuación los procesos involucrados en el modelo se explican con mayor profundidad:

Extracción Inicial de Características

De una imagen de color RGB, nueve (9) escalas espaciales son creadas utilizando pirámides gaussianas diádicas, que realizan filtros pasabajos y submuestrean la imagen en sentido vertical y horizontal. Este submuestreo se

realiza desde escala 1:1 hasta 1:256. Para cada una de estas imágenes se obtiene las siguientes características:

- La intensidad de la imagen definida como $I = (r + g + b)/3$
- I es utilizado para crear una pirámide gausiana $I(\sigma)$ con $\sigma \in [0, 8]$.
- Los canales r, g, b, son normalizados en las zonas donde I es mayor a $\frac{I}{10}$ para desacoplar la tonalidad (color) de la intensidad.
- Cuatro canales de colores son creados: $R = r - \frac{g+b}{2}$ $G = g - \frac{r+b}{2}$ $B = b - \frac{r+g}{2}$ $Y = \frac{r+g}{2} - \frac{|r-g|}{2} - b$ Los valores negativos son igualados a 0.

Mapas de Características

Mediante un mecanismo "center-surround" (centro-periferia) se analiza un punto en el espacio. Este procedimiento, se implementa como la diferencia entre dos escalas, el centro es un píxel a escala $c \in 2, 3, 4$ y los alrededores (surround) corresponden al píxel en la escala $s = c + \delta$ con $\delta \in 3, 4$. La diferencia entre dos mapas para todas las escalas se denota con $\ominus$ y es obtenido interpolando cada una de las escalas hasta la escala más fina, y restando punto a punto.

El primer mapa de características es la intensidad, que en los mamíferos es detectada por neuronas sensibles a centros oscuros con vecindarios brillantes o viceversa [48]. Esta característica es simultaneamente calculada en un conjunto de seis (6) mapas, para $c \in 2, 3, 4$ y $\delta \in 3, 4$:

$$I(c, s) = |I(c) \ominus I(s)|$$

En el caso del color, la corteza visual representa esta información utilizando la "doble oposición de color": En el centro del campo, las neuronas son excitadas por un color e inhibido por otro. Este efecto existe entre los colores rojo/verde, azul/amarillo y viceversa. De acuerdo a esto, se generan doce (12) mapas relacionados con el color, seis (6) mapas para la oposición rojo/verde y seis (6) para la oposición azul/amarillo:

$$RG(c, s) = |(R(c) - G(c)) \ominus (G(s) - R(s))|$$

$$BY(c,s) = |(B(c) - Y(c)) \ominus (Y(s) - G(s))|$$

La información local de orientación es obtenida de la información de intensidad aplicando filtros de Gabor $O(\sigma, \theta)$, con $\sigma \in [0, 8]$, para cada escala y $\theta = 0, 45, 90, 135$ como orientaciones preferenciales. Por cada una de las orientaciones se generan seis (6) mapas, para un total de veinticuatro (24) mapas relacionados con la orientación. El mapa de características de orientación se calcula así:

$$O(c,s,\theta) = |O(c,\theta) \ominus O(s,\theta)|$$

Resumiendo, al terminar la aplicación del mecanismo centro-periferia, un total cuarenta y dos (42) mapas de características son calculados, seis (6) para intensidad, doce (12) para color y veinticuatro (24) para orientación.

Mapa de Saliencia

El propósito del mapa de saliencia es colocar en cada lugar del campo visual un valor escalar que guíe el foco de atención. Dado que los cuarenta y dos (42) mapas pertenecen a características y resoluciones diferentes, los objetos llamativos aparecen fuertemente marcados solo en algunos mapas y pueden aparecer como ruido o menos sobresalientes en la mayor parte de los mapas. Puesto que el modelo no posee realimentación "arriba abajo" para integrar los mapas, se utiliza un operador de normalización no lineal N(.) que *promueve globalmente aquellos mapas de características en los cuales un pequeño numero de picos con fuerte actividad se encuentran presentes* [61] Este operador se calcula como sigue:

- Los valores de normalización se encuentran en un rango definido $[0, M]$
- Encontrando el máximo global M y calculando el promedio $\overline{m}$ de todos los otros máximos locales y
- globalmente multiplicando el mapa por $(M - \overline{m}^2)$

Solo la actividad local máxima es considerada, de tal manera que N (.) compara la respuesta asociada a "puntos de activación" en el mapa ignorando las áreas homogéneas. La motivación biológica de N (.) es replicar el mecanismo de inhibición cortical lateral, en donde un vecindario similar es

inhibido por otro, por medio de conexiones anatómicas definidas [61]. Los mapas de características son combinados en tres mapas "principales":

$$\overline{I} = \bigoplus_{c=2}^{4} \bigoplus_{s=c+3}^{c=4} N(I(c,s))$$

$$\overline{C} = \bigoplus_{c=2}^{4} \bigoplus_{s=c+3}^{c=4} [N(RG(c,s)) + N(BY(c,s))]$$

$$\overline{O} = \bigoplus_{\theta \in 0,45,90,135} N(\bigoplus_{c=2}^{4} \bigoplus_{s=c+3}^{c=4} N(O(c,s,\theta)))$$

La motivación para crear estos tres canales y sus normalizaciones individuales parte de la hipótesis que características similares compiten individualmente (preferencialmente).

$$S = \frac{1}{3}(N\overline{I} + N\overline{C} + N\overline{O})$$

En todo momento el máximo del mapa de saliencia define la ubicación de la zona más sobresaliente de la imagen. Mediante la emulación de un proceso conocido como inhibición de retorno, una vez una región es atendida, su saliencia disminuye, permitiendo que el sistema pueda atender una región diferente.

En la figura 3.4 se ilustra el mecanismo de inhibición de retorno sobre una imagen de histopatologia. Después de obtener el mapa de saliencia de la imagen, se procede a seleccionar aquella región de la imagen con un nivel de saliencia más alto mediante la aplicación de una red neuronal winner take all (WTA). En el caso de la figura este corresponde a la región A. Al realizar nuevamente el procedimiento de selección, la red neuronal inhibe la saliencia de aquellos lugares previamente seleccionados. Esto permite que sobresalga una nueva región dentro de la imagen (región B). La aplicación sucesiva de este mecanismo permite la generación de una serie de regiones atendidas, que busca emular el proceso secuencial de análisis de la escena visual. De forma simplificada, el mecanismo de inhibición de retorno funciona como una memoria de aquellos sitios previamente atendidos.

El modelo de Itti logra un comportamiento similar al del componente preatentivo del procesamiento visual. Extensiones del modelo para incluir el

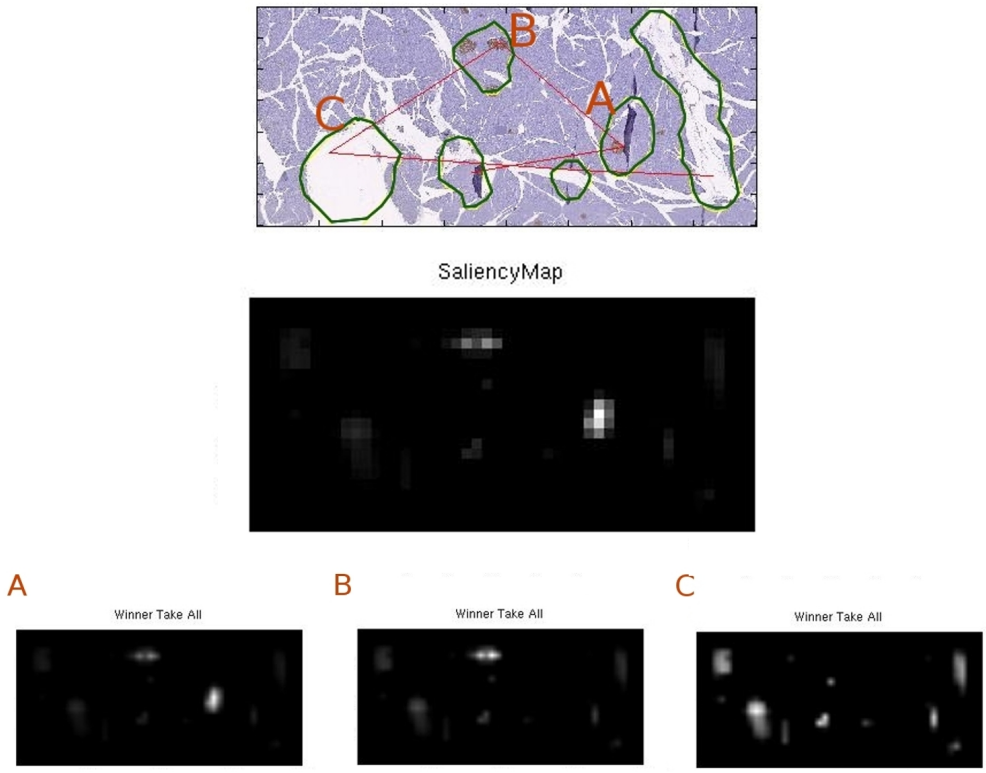

Figura 3.4: Inhibición de retorno.

componente "top-down" se han realizado en trabajos posteriores [62,63], argumentando que la definición de objetivos y distractores de búsqueda visual en una imagen, puede servir de base para la óptima ponderación en los mapas de características, y de esta forma extender el modelo a tareas definidas de búsqueda visual. Finalmente, es interesante mencionar la utilización de plantillas que definen objetos de búsqueda. En [64] se desarrolla un modelo de atención visual para objetos icónicos, en el cual mediante el cálculo de la similitud entre un objeto observado y una representación icónica almacenada se crea un mapa de saliencia. Este proceso usa un criterio de señal a ruido, para decidir si el pico de correlación en el mapa de saliencia es lo suficientemente alto como para definir que el objetivo se encuentra en determinada posición.

3.4. Modelamiento de Movimientos Oculares

Como se ha dicho anteriormente, la atención es un proceso secuencial, cuyo objetivo consiste en facilitar que la gran cantidad de información sea procesada por una cantidad limitada de recursos neuronales. Usualmente este proceso se realiza mediante las sacadas o movimientos angulares cortos de los ojos seguidas de periodos de inmovilidad ocular llamadas fijaciones. Si bien, estos movimientos durante la búsqueda visual se caracterizan por ser idiosincráticos, existe una cantidad considerable de estudios que enlazan la actividad del movimiento ocular (lugar hacia el cual esta dirigido la fóvea) con la actividad realizada por el observador, y es ampliamente reconocido que dicha actividad es indicativa de las operaciones cognitivas subyacentes del observador [25, 65]. En el contexto de evaluaciones diagnósticas médicas esta aseveración se conserva: Durante el análisis de imágenes radiológicas existe correlación entre las fijaciones de radiologos, pues a pesar de la enorme diferencia entre la secuencia espaciotemporal de movimientos, las fijaciones coinciden en algunos casos espacialmente y en otros en las características extraídas de las imágenes [66].

Una aproximación para emular este comportamiento secuencial se basa en la identificación de regiones relevantes a través de un conjunto de algoritmos de procesamiento de imágenes (relacionados con simetria, orientación, bordes, entropia, contraste, operaciones center-surround, y descomposiciones en el espacio frecuencial), tal como fue propuesto por Privitera y Stark [67]. Posterior a la identificación de regiones de interés, el algoritmo construye clusters de las zonas identificadas y define secuencias de visitas del foco de atención sobres estas regiones de acuerdo al valor obtenido por cada una de estas regiones.

La mayor parte de los modelos computacionales de la atención visual, producen un mapa de saliencia. En este mapa las regiones con un mayor valor de saliencia, atraen el foco de atención dentro de la escena visual. Este comportamiento usualmente se modela mediante el uso de una red neuronal, en la cual el máximo global es identificado. Debido a que solo existe un máximo dentro del mapa de saliencia, es común además la implementación de mecanismos de inhibición de retorno, que permiten emular los cambios en el foco de atención. De acuerdo a este mecanismo, las regiones visitadas por el punto de atención (esto es, la fovea) son transitoriamente inhibidas permitiendo que el foco de atención pase a la siguiente región con más intensidad en el mapa de saliencia [58].

En el proceso de navegación automática de imágenes, propuesta como una solución a la adaptación de imágenes en dispositivos de visualización pequeños, también se utiliza un mapa de saliencia para definir las regiones de interés en la imagen. Esta saliencia es obtenida de acuerdo al valor definido, de forma ponderada, por algoritmos de reconocimiento de caras, reconocimiento de texto y características visuales de bajo nivel. Después de definir qué objetos son relevantes en la imagen, el algoritmo procede a generar secuencias entre estas regiones siguiendo conceptos observados en patrones de búsqueda de alimentos en animales: Maximización del consumo, y minimización del recorrido [43].

Para simular el proceso de cambio del punto de fijación, también se han utilizado algoritmos de caminata aleatoria (random walk). En este caso el mapa de saliencia de la imagen es usada como representación del espacio sobre el cual un algoritmo es implementado [68]. La aleatoriedad de la caminata está influenciada directamente por la saliencia, pues la distribución de probabilidad de los saltos del foco de atención esta directamente relacionada al valor de saliencia que se haya en un punto en particular.

Otros mecanismos para la generación de movimientos sacádicos se basan en la hipótesis de independencia entre los procesos de localización y reconocimiento del objeto. Mediante la representación icónica de objetos, comprendida como la respuesta a filtros spacio-cromáticos a múltiples escalas es posible realizar comparaciones iterativa desde una baja resolución a una alta resolución. Teniendo en cuenta esto, y aplicando también heurísticas de búsqueda (por ejemplo, el objeto tiende a estar en el centro de la imagen) es posible calcular el siguiente punto de fijación, sin haber calculado de forma completa el mapa de saliencia. Esta estrategia permite encontrar distribuciones de fijación en arreglos fotométricos de objetos que tienen similitud con los observados en humanos [64].

3.5. Reconocimiento y Atención Visual

En este capitulo se revisó el modelamiento computacional para la atención visual, particularmente mediante el cálculo de la saliencia intrínseca de la imagen (bottom-up). Algunas características comunes a este tipo de modelamiento incluyen la utilización de características a bajo nivel, la utilización de operaciones centro-periferia para capturar diferencias a lo largo de diferentes niveles de resolución, la computación de un mapa de saliencia

único, y la generación de scan-paths basadas en mecanismos de inhibición de retorno. Si bien este tipo de modelamiento describe de forma eficiente el despliegue de los mecanismos atencionales durante los primeros cientos de milisegundos después de la presentación de una nueva escena, el desarrollo de mecanismos dependientes de la tarea del observador no han sido incluidos en este tipo de modelos. Así, un importante campo de desarrollo consiste en la integración de la información característica de la imagen, y la guiada por la tarea desarrollada por el observador.

Una primera aproximación consiste en la modificación de los modelos bottom-up (como el modelo Itti) para ponderar de forma más fuerte las características que discriminan de forma efectiva las regiones interesantes de la imagen de aquellas que se consideran como background o distractores [62].

Intentos más elaborados por caracterizar un subsistema de alto nivel se basan en la independencia de procesos como memorización, búsqueda y reconocimiento. La memorización es concebida como un proceso mediante el cual se almacena tanto las fijaciones, expresadas como características invariantes (Asociada a lo que se está buscando-**qué**), así como la secuencia ordenada de las fijaciones (Asociada a donde se esta buscando-**dónde**). La operación de búsqueda, utilizada tanto en la memorización como en el reconocimiento, consiste en un exámen secuencial de la imagen basado en un algoritmo de búsqueda. Este algoritmo determina el salto entre dos fijaciones teniendo en consideración puntos en su vecindad que presenten un cambio brusco en intensidad, la distancia de estos puntos al punto de fijación actual, y la 'novedad' del punto de fijación (inhibición de retorno). En el proceso de reconocimiento, se ejecuta la búsqueda y se busca fijaciones que logren parecerse a las características de las fijaciones almacenadas. Este modelo puede reconocer complejas escenas y rostros con invarianza a variables como traslación y rotación [69].

El modelo propuesto por Deco y Zihl, usa la información de saliencia extraída a partir de características de bajo nivel en una resolución baja. Se seleccionan regiones candidatas correspondientes a regiones con una alta saliencia. Estas regiones son analizadas de forma secuencial en una resolución cada vez mayor hasta que un puntaje de reconocimiento, con respecto a un objeto almacenado en memoria, lo suficientemente alto es obtenido. Este modelo ha sido aplicado de forma exitosa en experimentos psicofisicos que muestran la validez del mejoramiento progresivo de la resolución espacial [70].

La atención visual es un mecanismo selectivo que permite detectar regiones de la imagen hacia las cuales se debe dirigir el grueso del limitado poder

computacional del cerebro. Posterior al proceso de atención visual aparece el reconocimiento de objetos que consiste en la identificación y asociación de la escena visual a determinado(s) objeto(s) que se encuentra almacenado(s) en la memoria. El proceso biológico sobre el cual se desarrolla el reconocimiento de objetos aún no ha sido completamente entendido, pero se reconoce que la atención es un requisito imprescindible para llevar a cabo este proceso.

Capítulo 4

Método para Predecir Patrones de Navegación en Mega-imágenes Histopatológicas

En el capítulo introductorio se describió el problema de investigación abordado en este trabajo. Se mostró que al asumir el problema de la anticipación de la trayectoria diagnóstica como una búsqueda visual, éste se podía subdividir en dos problemas: La identificación de regiones relevantes y el mecanismo de transición entre estas regiones. Siguiendo los lineamientos generales que fueron expuestos anteriormente, en este capítulo se describe en detalle la solución propuesta.

En la sección 4.1 se presenta la estrategia de detección de relevancia utilizada para definir regiones de interés en la mega-imágen. Esta estrategia basa su funcionamiento en la introducción de conocimiento específico del experto mediante la selección de regiones de ejemplo objetivo (ejemplos positivos) y/o distractores (ejemplos negativos) de las mega-imágenes histopatológicas. El uso de esta configuración, está motivada por investigaciones recientes que muestran que durante tareas de búsqueda visual, incrementar la ponderación de la respuesta de neuronas cuyas características están relacionadas con un distractor es una estrategia óptima [62]. La similitud/disimilitud de cada una de las regiones en la mega-imágen con respecto a estas imágenes de ejemplo es el punto de partida para la construcción de un mapa único de relevancia, cuyo objetivo es similar al mapa de saliencia utilizado en los modelos com-

putacionales de atención visual(ver sección 3.3). El valor escalar en cada uno de los puntos determina la relevancia del punto de la imagen espacialmente asociado. Este mapa de relevancia representa una definición "difusa" de regiones de interés, pues no consiste en una clasificación binaria (relevante/no relevante). Para evaluar el desempeño de este método se realizó una binarización mediante la selección de un umbral, pero este paso no es necesario en el método general.

Una vez encontrado el mapa de relevancia de la imagen, se procede a modelar la transición entre regiones de interés, es decir, establecer un criterio que permita determinar cuál será la próxima región requerida en alta resolución por el patólogo conociendo únicamente algunos datos como la posición actual y la velocidad del dispositivos de entrada(*mouse*). Obsérvese que este modelamiento esta íntimamente ligado a las características de la predicción necesaria en el sistema, así como de la interfaz gráfica utilizada. En la sección 4.2 se analizan dos posibles casos: en uno de ellos el sistema es considerablemente lento de tal forma que no es posible usar la información de velocidad para realizar una predicción en la trayectoria diagnóstica. El segundo caso incluye la información de velocidad del *mouse*.

4.1. Estrategia de Detección de Relevancia

La estrategia de detección de relevancia es un método mediante el cual se asigna un valor escalar a cada una de las regiones de la imagen de acuerdo a su relevancia. El proceso empieza dividiendo la imagen usando una partición rectangular. Se escogió este tipo de partición porque es la más sencilla y no requiere de ningún otro tipo de conocimiento sobre las regiones contenidas en la imagen histopatológica. Al aplicar la partición, se obtiene un número considerable de subimágenes. Cada una de ellas tiene una dimensión escogida de forma arbitraria, y es representada por características relacionadas con color, intensidad, orientación y textura. Es recomendable que la selección del tamaño de bloque no implique un número excesivo de subimágenes (pues la complejidad del algoritmo es exponencial), y que a la vez, la región contenida en cada uno de los bloques presente cierta uniformidad en sus características. El siguiente paso consiste en la selección de ejemplos positivos y negativos. Éstos son manualmente escogidos, basándose en las navegaciones reales realizadas por los patólogos. A continuación, se construyen diferentes mapas de distancia, mediante el cálculo de la distancia de cada subimagen a los

ejemplos positivos y negativos empleados. Este cálculo se realiza tomando un par métrica-característica específico. Finalmente, un mecanismo de integración de información es utilizado para fusionar la información proveniente de diferentes pares métrica-característica.

En la figura 4.1 una representación esquemática de los procesos el modelo es visualizado.

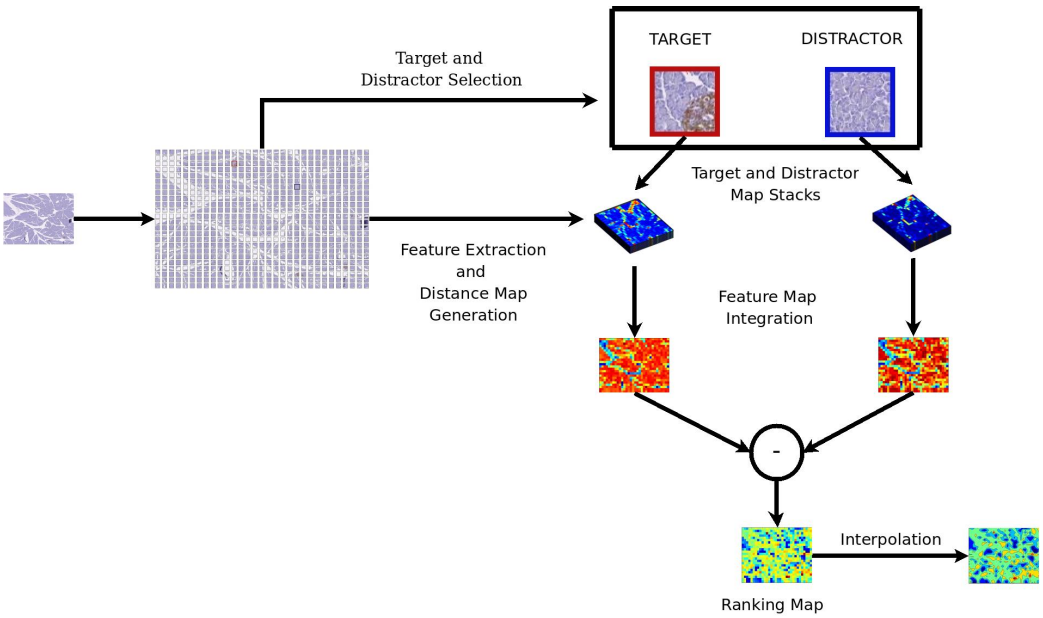

Figura 4.1: Diagrama de flujo del método propuesto.

A continuación se describen cada uno de los pasos con mayor detalle:

4.1.1. Extracción de Características y Métricas

Las características relacionadas con intensidad, color, orientación y textura han sido usadas en sistemas de recuperación de información por contenido para imágenes histopatológicas [71, 72]. Este tipo de características son importantes dentro de las mega-imágenes histopatológicas porque permiten la discriminación de las estructuras patológicas e histológicas dentro del tejido. La información contenida en el color es importante en las imágenes histopatológicas, pues el uso de tinciones precisamente delimita estructuras de interés histológico y patológico [73]. Un ejemplo claro, es el uso de la tinción H&E

(Hematoxilina y Eosina), donde el primer componente (Hematoxilina) marca con un color azul oscuro aquellas estructuras con un alto contenido de ácidos, como el núcleo (que contiene Acido desoxirribonucleico-DNA) y la matriz del cartílago; mientras que la Eosina marca otros componentes citoplasmáticos y el colágeno de color rosado [14]. Las características asociadas a textura y orientación también resultan relevantes dentro del contexto histopatológico, pues la aparición de ciertas patologias (por ejemplo, cancerígenas) involucra una fuerte distorsión en la arquitectura de los tejidos [74].

En el modelo propuesto, estas características son utilizadas para representar cada uno de los bloques de la mega-imágen. La asignación de relevancia se lleva a cabo calculando la similitud entre cada uno de los bloques y un conjunto de bloques de ejemplo (positivo y negativo). Este proceso en cierta forma define que características son relevantes, asignandole mayor (menor) ponderación a aquellas características de bajo nivel, vinculadas a las imágenes de ejemplo positivo (negativo). Este mecanismo esta inspirado en la teoría de búsqueda guiada de Wolfe (ver sección 3.2), que expresa que las características de bajo nivel usadas en el procesamiento visual temprano, sirven para guiar los despliegues de búsqueda selectiva.

Las características usadas se describen a continuación:

- Histograma de escala de grises y de color:

  Para la escala de grises, se uso un histograma de 256 bins. Histogramas son representados usando partición del espacio de color, tal como se describe en [75]. El espacio RGB de color fue particionado y cuantizado en $8 \times 8 \times 8 = 512$ bins.

- Histograma de partición local binaria:

  Esta característica de textura se obtiene de la siguiente forma: Para cada pixel P en la imagen, se construye una cadena binaria de 8 posiciones. Cada una de las posiciones de la cadena identifica si un pixel en su 8-vecindad contiene un valor de intensidad mayor (1) o menor (0) que el pixel P. Esta cadena puede tomar 256 valores diferentes, y mediante el análisis de todos los pixeles en la imagen se construye un histograma de 256 bins.

- Histograma de texturas de Tamura:

  Las texturas de tamura estan fuertemente correlacionadas con la percepción visual humana. Tres características de Tamura se utilizaron:

rugosidad, contraste y direccionalidad. El espacio generado por estas características es particionado, de forma similar al espacio de colores, en $8 \times 8 \times 8 = 512$ bins.

- Histograma de Sobel:

  El operador de Sobel es uno de los más populares en el campo de procesamiento de imágenes para la detección de bordes. El operador, calcula la diferencia de intensidad en el vecindario de un pixel en las direcciones vertical y horizontal. Esta diferencia puede ser interpretada como la derivada de la función representadondo la imagen en ese punto. En este caso, se utilizo un filtro con una ventana de 3×3 para analizar la 8-vecindad de cada uno de los pixeles en la imagen. Un histograma con 512 bins es calculado.

Como las características a bajo nivel utilizadas son histogramas, se requieren métricas para evaluar las diferencias entre las distribuciones de probabilidad. Las características seleccionadas son histogramas cuyo rango varía entre los 256 y 512 valores posibles. La similitud entre histogramas de diferentes imágenes puede ser evaluada usando diferentes funciones de distancia, cada una de ellas valorando diferentes particularidades de los histogramas. La selección de estas métricas esta vinculada a su uso extensivo en sistemas de recuperación por contenido [71, 75].

Para las siguientes definiciones, H y H' son histogramas, ambos con un número de bins M. En adición, H_m correponde al emésimo bin en el histograma H:

- Distancia Euclidiana:

$$L_2(H, H') = \sqrt{\sum_{m=1}^{M} (H_m - H'_m)^2} \tag{4.1}$$

- Intersección Histogramas:

$$D_I(H, H') = \sum_{m=1}^{M} min(H_m, H'_m) \tag{4.2}$$

- Divergencia Jensen-Shannon:

$$D_{JSD}(H, H') = \sum_{m=1}^{M} \left(H_m log \left(\frac{2H_m}{H_m + H'_m} \right) + H'_m log \left(\frac{2H'_m}{H'_m + H_m} \right) \right) \tag{4.3}$$

- Desviación relativa entre bins:

$$D_{RBD}(H, H') = \frac{1}{m} \sum_{m=1}^{M} \frac{|H_m - H'_m|}{H_m + H'_m} \tag{4.4}$$

- Desviación relativa:

$$D_{RD}(H, H') = \frac{\sqrt{\sum_{m=1}^{M} (H_m - H'_m)^2}}{\sqrt{\sqrt{\sum_{m=1}^{M} (H_m)^2} + \sqrt{\sum_{m=1}^{M} (H'_m)^2}}} \tag{4.5}$$

4.1.2. Integración de los Mapas de Distancia

Como se mostró en la subsección anterior, un total de seis (6) diferentes características, y cinco (5) métricas para histogramas han sido presentadas para la representación del contenido de imágenes histopatológicas y el cómputo de la similitud entre imágenes. Aplicando cada uno de los pares métrica-característica para encontrar la distancia de cada bloque a la imagen de ejemplo nos da un total de 30 mapas de distancia. Como se estan usando al menos dos imágenes (ejemplo positivo y negativo), el número de mapas crece proporcionalmente a 60. Es probable, que el uso de algunas de las características y métricas sea suficiente para obtener una representación adecuada para una imágen específica; pero si queremos generalizar el método y abarcar un número mayor de tipo de imágenes histopatológicas, es necesario el uso de una vasta selección de metricas y características que sean capaces de valorar diferentes propiedades de la imagen en cuestión. Este sistema debe ser capaz de ponderar la importancia de cada uno de los pares métrica-característica y dar una mayor prioridad a aquella que se ajuste mejor a los parámetros de búsqueda, en este caso los ejemplos positivos y negativos.

Los mapas generados proceden de diferentes características y son evaluados usando diferentes métricas, lo cual plantea un dilema ante la necesidad de combinar esta información para llegar a una decisión. Para resolver un problema similar, en [57,61] se propone la utilización de un operador de normalización no lineal que *promueve aquellos mapas de características en los cuales un número pequeño de picos presentan un actividad fuerte* [61], o en otras palabras, que tienen una serie de máximos definidos. El uso de este operador es resumido de la siguiente forma:

- Reasignación del valor minimo en el mapa al segundo valor mínimo. Al pertenecer la imagen de ejemplo a la colección de imágenes, se espera que sea dominante en la colección. Se propone este paso, para evitar una promoción excesiva de la imagen de ejemplo en la colección. Si la imagen de ejemplo no esta contenida en la colección (mega-imagen), este paso puede ser ignorado.

- Normalización de los mapas en el rango [0, M] donde 0 corresponde al valor máximo de distancia y M al mínimo. De esta forma convertimos los mapas de distancia en mapas de similitud con rango dinámico idéntico. En el caso de la métrica de intersección de histogramas este paso es ignorado.

- Para cada uno de los mapas, se encuentra un máximo global M y se calcula el promedio de los máximos locales $\overline{m}$ (sobre una vecindad definida $n = 3 \times 3$).

- Multiplicar cada uno de los elementos de el mapa por $(M - \overline{m})^2$. De esta forma, si un mapa de relevancia contiene un pico global que sea sustancialmente mayor que los máximos locales en toda la imagen, es ponderado de una forma fuerte.

- Sumar todos los mapas para obtener solo uno.

Después de el proceso de integración, tenemos un mapa relacionado con el ejemplo positivo y otro con el negativo. Para obtener un mapa de valor de interés, simplemente restamos estos dos mapas e interpolamos para obtener un mapa continuo.

El mapa resultante puede ser expresado de la siguiente forma:

$$R_T = \sum_{m=1}^{N} (\alpha_m \cdot T_m - \beta_m \cdot D_m) \tag{4.6}$$

donde N denota el número total de mapas dado por los pares métrica-características, α_m y β_m denotan un factor de ponderación dado por el proceso de normalización, T_m y D_m corresponden al mapa de similitud normalizado del par métrica-característica m.

4.2. Predicción en una Estrategia de Prefetching

El segundo componente dentro del método desarrollado para anticipar la trayectoria de observación diagnóstica consiste en el modelamiento de la transición entre regiones de interés, definidas por la estrategia de detección de relevancia. Para ello se hace uso de dos hipótesis, una de ellas diseñada para sistemas de MV con una baja reactividad y el otro para sistemas con alta reactividad.

Estos dos casos se describen a continuación:

4.2.1. Hipótesis de Navegación - Baja Reactividad

En este modelo no existe información acerca del proceso de transición entre regiones de interés, por lo cual la información existente corresponde a la posición y la dimensión de la región observada por el patólogo en la ventana de alta magnificación. La hipótesis utilizada en este caso se basa en la siguiente premisa: Los siguientes bloques visitados corresponden a aquellos que tienen una relevancia alta y esten en un vecindario arbitrariamente determinado.

Este modelo corresponde a situaciones en las que es necesario realizar una anticipación con mucho tiempo, y no es posible realizar predicciones durante el proceso de transición de la ventana de alta magnificación entre regiones de interés.

En la figura 4.2 se puede observar con detalle esta operación. Si la ventana de observación es el área en rojo, el area de búsqueda de potenciales zonas que serán accesadas se limita al vecindario (en el esquema en azul claro). Los bloques que tengan una relevancia más alta dentro de esta zona

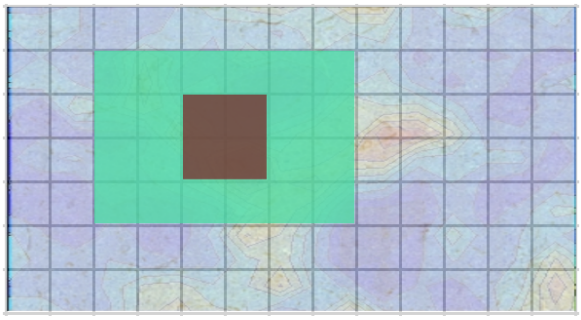

Figura 4.2: Modelo de navegación en un sistema de baja reactividad

tendrán mayor probabilidad de ser accesados, y por esa razón deberían ser seleccionados en una estrategia de prefetching.

4.2.2. Hipótesis de Navegación - Alta Reactividad

En este segundo caso, las peticiones en alta resolución que realiza el experto los movimientos del usuario sobre la imagen estan restringidos a operaciones de "dragging" y "dropping" y se tiene información de velocidad y posición del mouse durante la transición de regiones de interés. En este caso se recolectan estadísticas del movimiento del sistema para definir la dirección del movimiento. Junto con la información de velocidad, es posible calcular los bloques que tienen mayor probabilidad de ser accesados en un tiempo determinado.

En la figura 4.3 se representa en rojo oscuro la región actual de la ventana de observación. En rojo claro se observan las posiciones previas de la ventana. Con base en estadísticas y la velocidad que lleva el movimiento es posible determinar una zona con alta probabilidad de ser accesada.

La integración con la información proveniente del mapa de relevancia es llevada a cabo mediante la aplicación de la siguiente formula:

$$T_s = a * N_{score} + (1 - a) * R_{score}$$

donde T_s corresponde al score total de cada uno de los bloques, N_{score} corresponde al score del bloque proporcionado por el modelo de navegación y

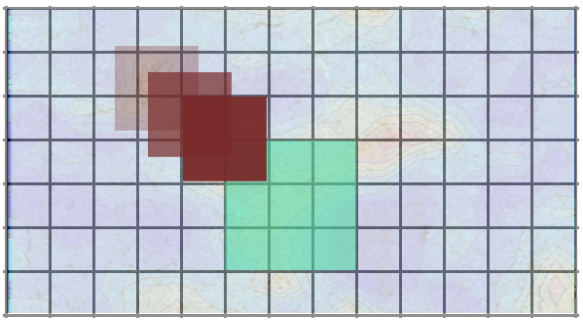

Figura 4.3: Modelo de navegación en un sistema de alta reactividad

R_{score} corresponde al score del bloque proporcionado por el modelo de asignación de relevancia.

4.3. Experimentación y Resultados

4.3.1. Evaluación de la Estrategia de Detección de Relevancia

Un total de siete láminas histológicas fueron digitalizadas y ocho imágenes de alta resolución fueron contruidas [34]. Los espécimenes consisten en un pancreas normal de ratón procedente de un estudio no relacionado, un especimen de tumor neuroendocrino de la tiroide, un tercer especimen de un adenoma de tiroides atipico. Adicionalmente, cuatro especimenes diferentes con tinción de hematoxilina-eosina con diagnosis de apendicitis, hiperplasia folicular reactiva de un nodulo linfático, leiomioma y una tiroides normal. Las imágenes de alta resolución fueron construidas usando registro automático usando la correlación cruzada como medida de similitud [34] y fueron almacenadas en formato JPEG2000 [76] para su posterior acceso y navegación. Las regiones de interés de control y los bloques de interés y de distracción se calculan usando navegaciones reales de patólogos sobre las imágenes construidas. Para esto tres patólogos expertos, cada uno de ellos certificados con al menos cinco años de experiencia, fueron previamente entrenados en el uso de un prototipo de microscopio virtual.

Para la construcción del mapa de relevancia se utilizaron versiones en

mediana resolución de las mega-imágenes. La partición rectangular utilizada para subdividir las imágenes fue de 70×70 pixeles. Si bien la selección de la dimensión de los bloques es arbitraria, se trató de seleccionar un tamaño de bloque que no implicara un número excesivo de subimágenes (pues la complejidad del algoritmo es exponencial), y que a la imagen contenida en cada uno de los bloques obtenidos presentara cierta uniformidad en sus características.

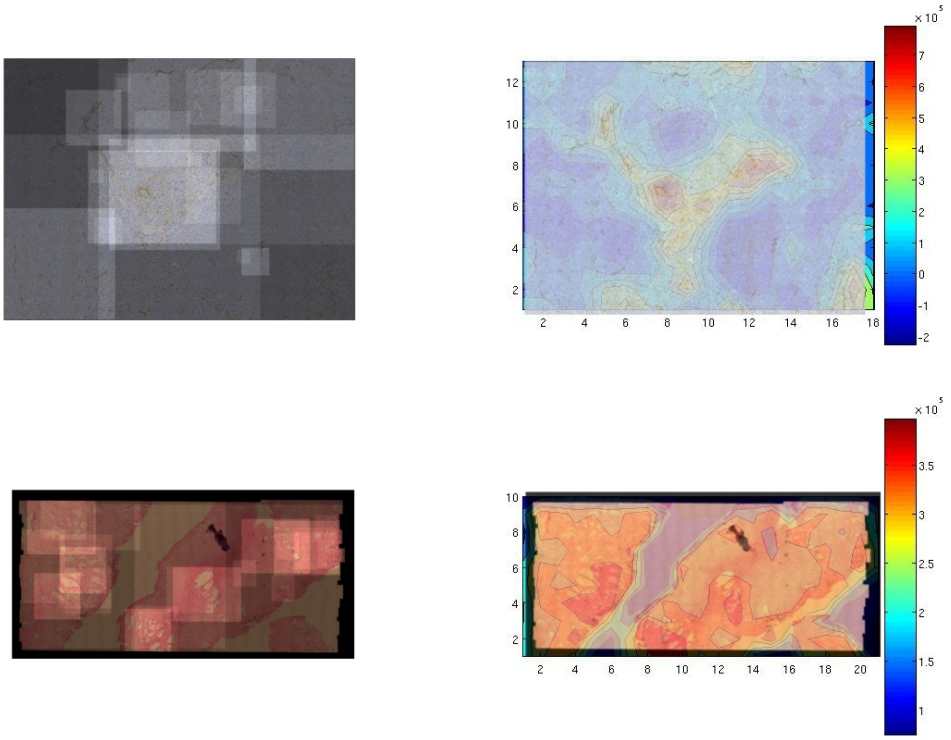

Figura 4.4: Aplicación estrategia de relevancia. A la izquierda se observa una lámina virtual en baja resolución. La frecuencia de visita de los patólogos corresponde a la iluminación. A la derecha se observa el mapa de relevancia obtenido a través del método propuesto. Las áreas en rojo corresponden a un valor más alto de relevancia.

En la columna izquierda de las figuras 4.4 y 4.5 se puede observar una imagen en baja magnificación de algunas láminas virtuales. Sobre las imágenes se ha superimpuesto la trayectoria de observación, representada por aquellas zonas con una intensidad más alta. La frecuencia de visita es directamente

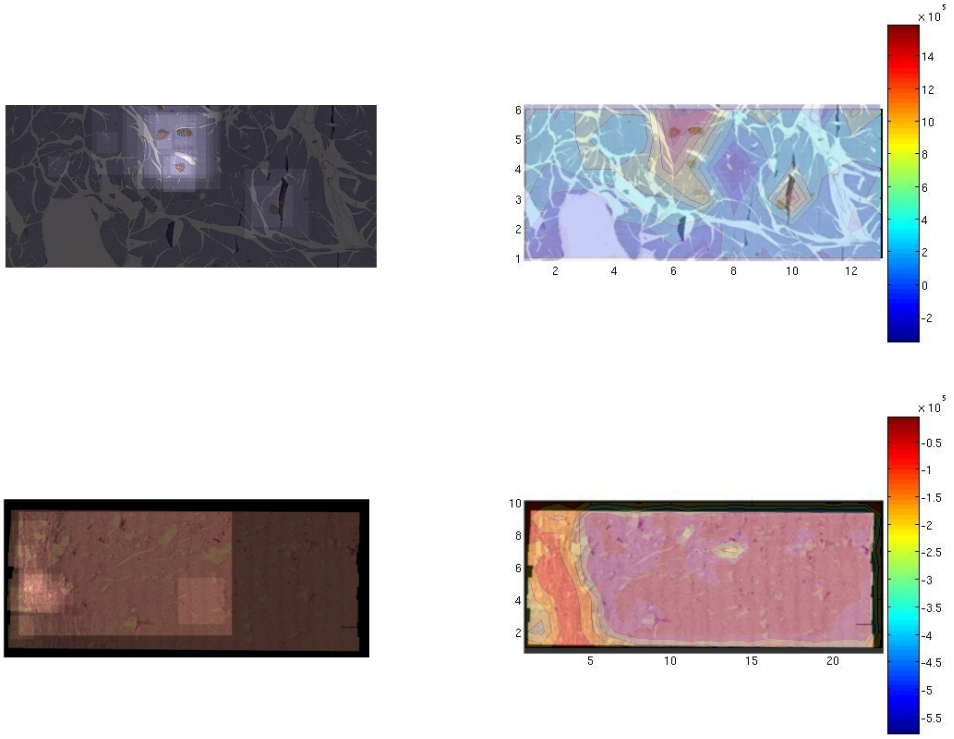

Figura 4.5: Aplicación estrategia de relevancia. A la izquierda se observan láminas virtuales en baja resolución. La frecuencia de visita de los patólogos corresponde a la iluminación. A la derecha se observa el mapa de relevancia obtenido a través del método propuesto. Las áreas en rojo corresponden a un valor más alto de relevancia.

proporcional a la iluminación en la porción de la imagen. En la columna derecha se muestra el mapa de relevancia obtenido al aplicar la estrategia de detección de relevancia sobre las imágenes en consideración. En esta escala aparecen en rojo las areas más relevantes y en azul las menos relevantes. De forma aproximada se puede decir que la relevancia coincide con las zonas visitadas por los patólogos en la mayoría de los casos, aún si tenemos en cuenta que las imágenes corresponden a especímenes histopatológicos completamente diferentes.

En la parte superior de la figura 4.4 se observa una imagen de un páncreas obtenida a través de un proceso de inmunotinción, y en la parte inferior un

tejido correspondiente a la próstata en la cual se aplicó una tinción de hematoxilina y eosina. El pancreás tenía muy pocas estructuras relevantes para explorar y los patólogos dedicaron su tiempo a explorar la imagen en el centro, básicamente porque en esa región se localizaban los islotes de langerhans, que son una estructura funcional importante. En el especimen de prostata se observa un patrón más distribuido, pero que es igualmente emulado por el método propuesto.

El principal objetivo de la evaluación del método de detección de relevancia consiste en medir la correlación entre los bloques con un valor alto de relevancia y los bloques pertenencientes a la trayectoria de observación en navegaciones reales. Los bloques pertenecientes a la trayectoria de observación son definidos como interesantes **I**, cuando el número de visitas es mayor que el 50 % de las visitas recibidas por el bloque más visitado. Estos bloques son comparados con los bloques estimados **E** por el método.

El experimento realizado consiste en la asignación de un nivel binario de relevancia para los bloques en la trayectoria de observación y la medición del nivel de coincidencia con los bloques más relevantes (En este caso los bloques cuyo valor de interés se encontraba en el 10 % más alto). De forma similar, se procedió con los mapas de relevancia obtenidos con el modelo de Itti [57], particionando el mapa de saliencia o relevancia en una grilla idéntica a la utilizada en el método propuesto. Por último, se midió repetidamente precisión y recall, medidas usadas ampliamente en la comunidad dedicada a la recuperación de información, contra una muestra escogida de forma uniformemente aleatoria de bloques de la imagen. Los resultados se encuentran descritos en la tabla 4.1.

Con relación a los bloques definidos como interesantes (I), y estimados (E), la medida de precisión y recall se definen de la siguiente manera:

$$Prec. = \frac{I \cap E}{E}$$

$$Rec. = \frac{I \cap E}{I}$$

Los resultados expuestos en la tabla 4.1 indican que la introducción indirecta de conocimiento de expertos a través del método propuesto, tuvo un impacto positivo en la identificación de regiones relevantes en imágenes histopatológicas. De forma general, las medidas de precisión y recall son consistentemente más altas en la estrategia propuesta con una precisión me-

Tabla 4.1: Precisión y Recall

	Propuesto		**IttiMod**		**Aleatorio**	
Imagen	*Prec.*	*Recall*	*Prec.*	*Recall*	*Prec.*	*Recall*
1	0.12	0.23	0.013	0.03	0.05	0.09
2	0.43	0.16	0.30	0.11	0.25	0.09
3	1	0.26	0.09	0.02	0.35	0.09
4	0.57	0.67	0.14	0.17	0.08	0.08
5	0.57	0.2	0.1	0.03	0.28	0.1
6	0.48	0.73	0.17	0.27	0.06	0.09
7	0.5	0.42	0.31	0.26	0.11	0.09
8	0.69	0.32	0.05	0.08	0.21	0.1
Promedio	0.55	0.38	0.14	0.12	0.17	0.09

dia del 0.55 y un medida recall media de 0.38, que son consideradas como adecuadas por la comunidad dedicada la recuperación de información. A excepción de la primera y la tercera imagen que tuvieron medidas de precisión extremas, la medida de precisión varió entre el 40 % y el 70 %. En general, las medidas de recall fueron bajas, pues las regiones solicitadas por los patólogos, abarcaban la zona de interés y el área circundante, además de que es posible que se hayan fijado por regiones de interés diagnóstico que no fueron correctamente caracterizadas.

El modelo Itti tuvo un desempeño variable en el conjunto de datos analizados, obteniendo un rendimiento bueno en algunas imágenes (particularmente la 2 y la 7 que corresponden a una lámina virtuales con páncreas, y vesicula biliar respectivamente), en algunos otros similar a una selección aleatoria y en otros un desempeño pobre. Esta variabilidad se debe a que mientras en algunos casos las regiones identificadas por el modelo Itti se correlacionan con las regiones de interés, en otros casos se identifican regiones que no tienen ningun sentido clínico.

4.3.2. Evaluación de la Predicción en una Estrategia de Prefetching

Un sistema para la navegación remota de imágenes, requiere una configuración cliente servidor similar a la que se muestra en la figura 4.6. Además de los elementos esenciales, como la base de datos, el servidor, el cliente, y el

canal de comunicación, se pueden observar elementos auxiliares que buscan maximizar la eficiencia de la navegación mediante la utilización de tecnicas de cache y de prefetching, similares a las descritas en la sección 2.3.

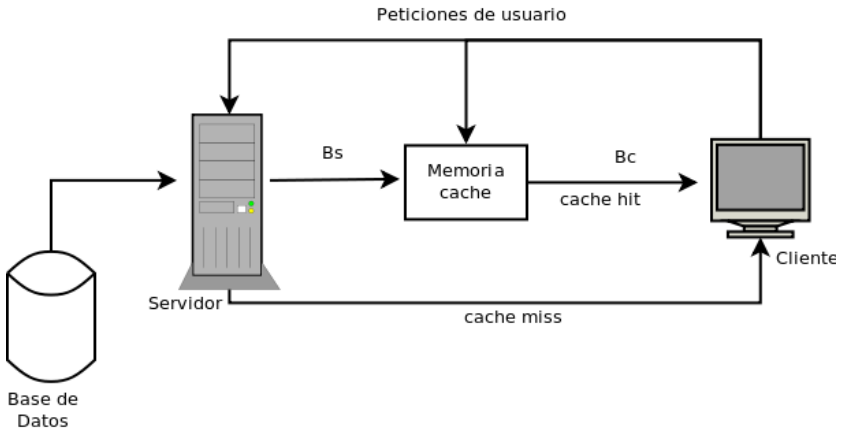

Figura 4.6: Configuración cliente/servidor

Para la experimentación se asumirá que el ancho de banda del canal de comunicación es ilimitado, de esta forma despreciamos los efectos que pueda llegar a tener el tiempo en que el cliente realiza peticiones. Aunque este factor es determinante en el rendimiento general del sistema, no se constituye en un elemento esencial a la hora de comparar diferentes modelos de predicción.

La información de las mega-imágenes, es representada como un conjunto de bloques o subregiones obtenidos a partir de una partición regular de la megaimágen. Así, una petición para visualizar una determinada región o ventana de interés (WOI por sus siglas en inglés-Window of Interest), es entendida en el modelo como una petición de un número determinado de bloques correspondientes a esa región.

La memoria caché es una entidad que puede almacenar una cantidad limitada de bloques. Si alguno de los bloques almacenados corresponde a una petición del cliente, tenemos un evento denominado cache hit. En caso de que la petición no se encuentre tenemos un cache miss, y es necesario que la petición del usuario se envié al servidor.

Para ello utilizaremos una medida conocida como Cache Hit Ratio (C_{HR}):

$$C_{HR} = \frac{C_{hit}}{N}$$

donde C_{hit} corresponde al total de bloques almacenados en el caché que cumple con las peticiones del usuario, y N corresponde al numero total de peticiones.

Con el objetivo de conocer el impacto que un modelo de predicción puede tener en la calidad de la navegación de imágenes histopatológicas durante el proceso de diagnóstico, se analizaron las dos hipótesis de navegación para un sistema de alta reactividad y uno de baja reactividad descritas anteriormente.

Para el caso del sistema con alta reactividad se utilizaron los siguientes modelos:

- Modelo Estático: En este modelo, los bloques contenidos en la región contenida por la ventana de interés actual del patólogo, son almacenados directamente en el caché. Se supone que los mismos serán posteriormente requeridos

- Modelo Navegación: En este modelo se construye una serie de estadísticas sobre la navegación del mouse para cada uno de los usuarios basadas en navegaciones previas. Estas estadísticas son posteriormente utilizadas para predecir hacia donde se dirige el experto. Corresponde a un ajuste del parametro a=0, en la ecuación descrita en la subsección 4.2.2.

- Modelo Navegación+Mapas Relevancia: Este modelo combina el modelo de navegación anterior y los mapas de relevancia obtenidos con el método de asignación de relevancia mediante el método descrito anteriormente en la subsección 4.2.2. El parámetro de combinación a, fue ajustado tras pruebas sucesivas en a=0.9.

- Modelo Navegación+Histograma de visitas: Corresponde al mismo descrito en el item anterior. Sin embargo el mapa de relevancia corresponde al histograma de visitas de los bloques obtenidas a partir del conjunto de navegaciones reales. Este método se considera el groundtruth.

Para observar el comportamiento de los modelos se recrearon las navegaciones realizadas por 3 patólogos sobre 8 Megaimágenes histopatológicas provenientes de diferentes especímenes histopatológicos, y que se encuentran descritas en la sección 4.3.1. Utilizando cada uno de los modelos expuestos anteriormente se calcula la medida de cache-hit ratio para cada una de las navegaciones. Finalmente se obtiene un promedio del desempeño en todas las navegaciones para cada uno de los patólogos, variando el tamaño del

cache (1 %, 3 %, 5 %, 7 %, 10 %, 15 % y 20 % del número total de bloques que constituyen la mega-imágen).

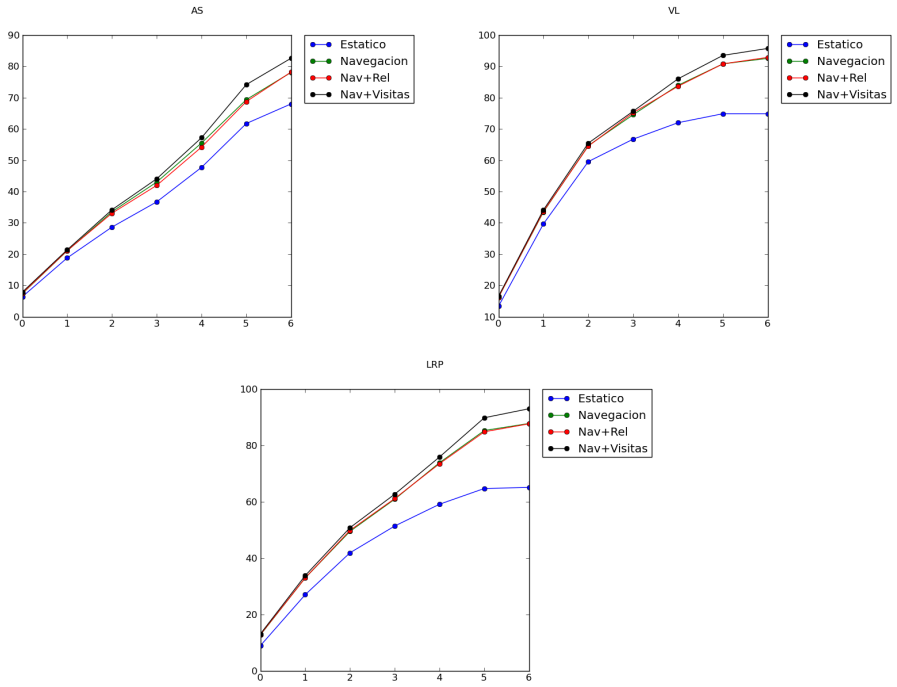

Figura 4.7: Resultados con hipotesis de alta reactividad

De acuerdo a las figura 4.7, se observa que la información predictiva suministrada por el modelo de navegación resulta más precisa y determinante que la suministrada por el modelo de asignación de relevancia. Solo el ground truth tuvo un desempeño superior, indicando que es necesaria una aproximación mayor al histograma de visitas para obtener una ganancia en el desempeño.

En el caso de baja reactividad se implementan los siguientes esquemas de predicción:

- Modelo LRU: Modelo que almacena en memoria únicamente los últimos bloques visitados. Los bloques menos visitados son susceptibles de ser reemplazados durante la siguiente petición del cliente.
- Modelo Vecindad: En este modelo se almacena en memoria cache los bloques que constituyen la vecindad de la ventana de interés actual. La

vecindad de la región de interés se define como aquellos bloques que limitan con los bloques más externos de los bloques requeridos en la petición actual.

- Modelo Uniforme: A diferencia del modelo anterior, el proceso de selección al azar se realiza en una área de la mega-imagen que contiene la zona visualizada en un determinado momento. En este caso particular, la zona en cuestión corresponde a 1/4 del total del área de la mega-imagen.

- Modelo basado en Mapa de Relevancia: En este modelo la memoria caché es actualizada en cada operación de visualización. Los bloques almacenados en el cache corresponden a los que tengan mayor relevancia(de acuerdo al método para asignación de relevancia descrito en la sección anterior) en un área limitada de la mega-imagen que contiene la posición actual de la ventana en observación. En este caso particular, la zona en cuestion corresponde a 1/4 del total del área de la mega-imagen.

De forma similar al caso de alta reactividad, se varia el tamaño del caché (1 %, 3 %, 5 %, 7 %, 10 %, 15 % y 20 % del número total de bloques que constituyen la mega-imágen), y se procede a medir el número de cache hits de cada uno de los modelos que se obtiene al recrear cada una de las navegaciones realizadas previamente por los patólogos.

Los resultados obtenidos (ver figuras 4.8 y 4.9) muestran que el método desarrollado presenta una ganancia en la capacidad de predicción con respecto a una selección de bloques al azar sobre un área determinada de la mega-imágen (Método Uniforme). A la vez presenta un desempeño similar al de métodos basados en las ventanas de interés actuales (vecindad y LRU), y con tamaños de caché lo suficientemente grandes se logra mejorar el desempeño. En tamaños menores de caché es evidente que estrategias sencillas de selección más sencillas (específicamente LRU) tienen más impacto, presumiblemente porque las peticiones de visualización presentan una redundancia espacial considerablemente alta.

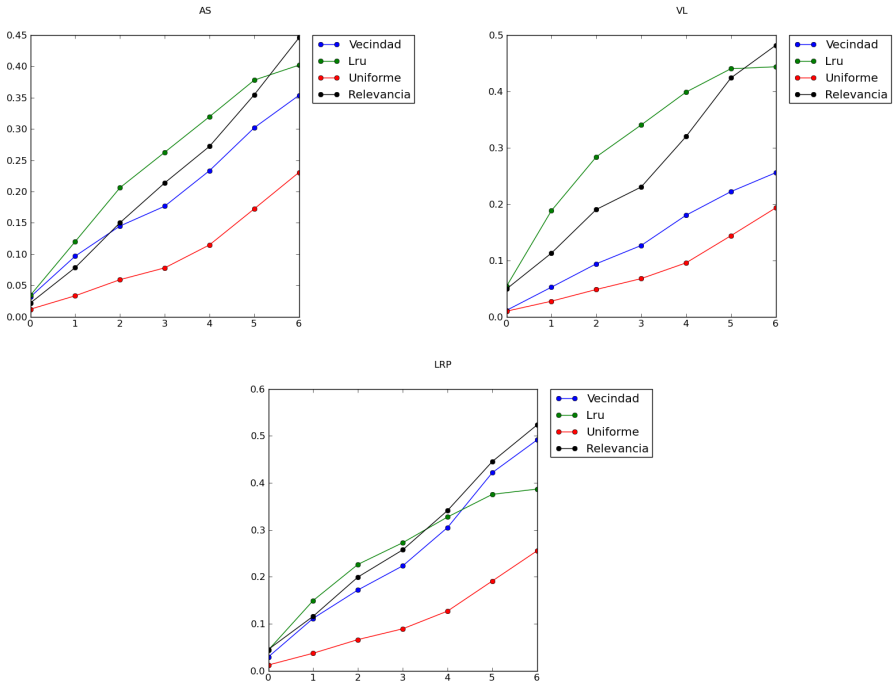

Figura 4.8: Resultados con hipótesis de baja reactividad

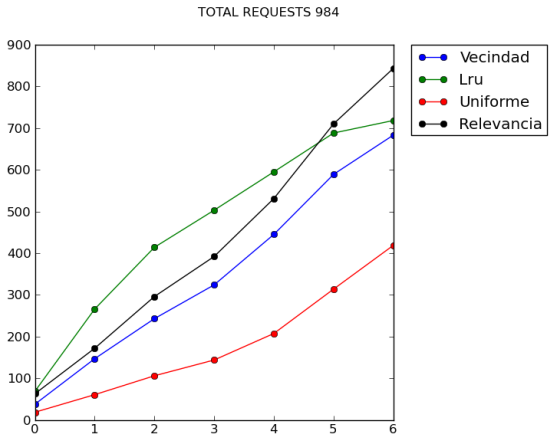

Figura 4.9: Resultados con hipótesis de baja reactividad

Capítulo 5

Conclusiones y Discusión

5.1. Conclusiones del Método Propuesto

El método desarrollado a lo largo de este trabajo presenta un paradigma novedoso en el problema de anticipación de trayectoria de observación durante la tarea de evaluación diagnóstica de mega-imágenes histopatológicas. Al dividir el problema de investigación en la identificación de regiones de interés y el modelamiento de la transición entre regiones de interés, fue posible introducir conocimiento específico del campo histopatológico.

El método de detección de relevancia presentado en este trabajo esta basado en la selección de ejemplos para extraer regiones de interés diagnóstico en imágenes histopatológicas. La información es integrada a partir de diferentes características de bajo nivel que son extraídas de la lámina virtual. Los resultados indican que la introducción de conocimiento del experto, a través de ejemplos positivos (negativos) tiene un impacto positivo en la identificación de estas regiones. De igual forma los experimentos demostraron que los modelos computacionales de atención visual, basados únicamente en información de bajo nivel de las imágenes, son inadecuados para resolver este problema, probablemente porque las regiones estimados con estos modelos no siempre estan relacionados con los patrones visuales que identifican las patologías.

Al analizar el impacto de la introducción del conocimiento "a priori", representado como un mapa de relevancia de la imagen, en estrategias de caché/prefetching se reconoció que las características del sistema de interfaz (particularmente la reactividad del sistema) juegan un papel determinante. En un sistema con reactividad alta, es decir una respuesta rápida a las peti-

ciones del usuario de la interfaz y del sistema de caché/prefetching, la información proveniente del mapa de relevancia tiene un impacto casi nulo. Esto se debe a que la información de posición y velocidad del mouse durante la transición de regiones de la mega-imágen es muy confiable y permite realizar una predicción precisa. En sistemas con reactividad baja, la formulación de políticas de prefetching basadas en el conocimiento específico de la imagen resulta competitivo con técnicas más naturales(tradicionales) como el caché LRU (Least Recently Used) y prefetching de bloques vecinos. Los resultados en la sección 4.2 muestran que el desempeño puede superar estos métodos tradicionales, si se dispone de un caché lo suficientemente grande.

Una conclusión importante de este trabajo es que la utilización de diversas características y métricas permite generalizar el método a diferentes tipos de muestras histopatológicas obteniendo resultados consistentes. Es importante notar que las regiones de interés dependen del tipo de espécimen histopatológico y que son sustancialmente diferentes entre sí. La generalidad del método esta basada en la propiedad de representar cada una de las imágenes mediante diferentes características y métricas, y en la posterior integración de éstas. Esta integración se realiza de tal forma que se promueven de forma más fuerte aquellas características que poseen un mayor poder discriminativo.

5.2. Trabajo Futuro

El trabajo desarrollado explora la utilización de conocimiento del experto en la formulación de estrategias de caché-prefetching. El enfoque utilizado se basa en el uso de varios subprocesos que llevan a cabo la identificación de regiones de interés y el modelamiento de transición entre estas regiones. Cada uno de ellos es susceptible de ser analizado y mejorado.

El principal componente del método es la estrategia de detección de relevancia. Las variables de esta estrategia estan asociadas al tamaño de la partición de la lámina virtual, la selección de bloques de ejemplo y el proceso de integración de características. Con respecto al primer elemento, es cuestionable si el tamaño utilizado captura realmente las características principales de una región de interés. La selección de bloques de ejemplo es función de la experticia del patólogo, pero en algunos casos podría extenderse a la utilización de un conjunto de bloques. Por esta razón sería interesante desarrollar una representación compacta de los bloques de ejemplo, pues con cada bloque nuevo la cantidad de procesamiento necesaria aumenta exponen-

cialmente. En cuanto a las estrategias de integración es pertinente continuar estudiando estrategias que permitan una selección de características con alto poder discriminativo.

El segundo proceso en el método consiste en establecer mecanismos que modelen el cambio entre las regiones de interés. Para modelar esta acción, se utilizó una sencilla hipótesis: El cambio se realiza hacia los bloques de la imagen que tengan una relevancia alta y esten en un vecindario arbitrariamente determinado. Si bien este supuesto es razonable dentro de la tarea de búsqueda visual, no tiene en cuenta elementos importantes como la dimensión de la ventana de observación y la revisita de regiones. Un paradigma interesante podría ser la utilización de modelos generadores de secuencias, similares a los de cambio en el foco de atención durante los movimientos sacádicos, para generar histogramas de visitas teniendo en cuenta las dimensiones de la ventana de visualización (alta magnificación).

Bibliografía

[1] J. Li and H. H. Sun. On interactive browsing of large images. *IEEE Transactions on Multimedia*, 5:581–590, 2003.

[2] D. Capel. *Image Mosaicing and Super-Resolution*. Springer, 2004.

[3] R. Szeliski. Image mosaicing for tele-reality applications. In *Proceedings of the Second IEEE Workshop on applications of Computer Vision*, 1994.

[4] M. Iregui, F. Gomez, and E. Romero. Virtual microscopy in medical image: a survey. *Modern Research and Educational Topics in Microscopy*, 2:996–1006, 2007.

[5] D. C. Allen. *Histopathology specimens: clinical, pathological and laboratory aspects*. Springer Verlag, 2004.

[6] R. Ferreira, B. Moon, J. Humphries, A. Sussman, J. Saltz, R. Miller, and A. Demarzo. The virtual microscope. In *American Medical Informatics Association, 1997 Annual Fall Symposium.*, pages 449–453, 1997.

[7] A. Afework, M. D. Beynon, F. Bustamante, A. Demarzo, R. Ferreira, R. Miller, M. Silberman, J. Saltz, A. Sussman, and H. Tsang. Digital dynamic telepathology-the virtual microscope. In *AMIA Annual Fall Symposium*, 1998.

[8] P. Fontelo, E. DiNino, J. Krista, A. Khan, and M. Ackerman. Virtual microscopy: Potential applications in medical education and telemedicine in countries with developing economies. In *Proceedings of the 38th Hawaii International Conference on System Sciences*, 2005.

[9] R. S. Crowley and O. Medvedeva. An intelligent tutoring system for visual classification problem solving. *Artificial Intelligence in Medicine*, 36:85–117, 2006.

[10] T. A. Drake, J. Braun, A. Marchevsky, I. S. Kohane, C. Fletcher, H. Chueh, B. Beckwith, D. Berkowicz, F. Kuo, Q. T. Zeng, U. Balis, A. Holzbach, A. McMurry, C. E. Gee, C. J. McDonald, G. Schadow, M. Davis, E. M. Hattab, L. Blevins, J. Hook, M. Becich, R. S. Crowley, S. E. Taube, and J. Berman. A system for sharing routine surgical pathology specimens across institutions: the shared pathology informatics network. *Human Pathology*, 38:1212–1225, August 2007.

[11] C. Lin and Y. F. Zheng. Fast browsing of large-scale images using server prefetching and client caching techniques. In Andrew G. Tescher., editor, *Applications of Digital*

Image Processing XXII., volume 3808 of *Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference.*, pages 376–387, 1999.

[12] A. Descampe, C. De Vleeschouwer, M. Iregui, B. Macq, and F. Marques. Pre-fetching and caching strategies for remote and interactive browsing of jpeg2000 images. Technical report, Laboratoire de Telecommunications et Teledetection. Universite catholique de Louvain, 2007.

[13] T. Schrader, S. Niepage, T. Leuthold, K. Saeger, K. Schluns, P Hufnagl, K. Kayser, and M. Dietel. The diagnostic path, a useful visualisation tool in virtual microscopy. *Diagnostic Pathology*, 1(1):40, 2006.

[14] L. C. Junqueira and J. Carneiro. *Basic Histology.* McGraw-Hill, 10th edition, 2003.

[15] R. S. Crowley, G. J. Naus, and C. P. Friedman. Development of visual diagnostic expertise in pathology. In *Proceedings of the 2001 AMIA Symposium*, volume 125, page 9, 2001.

[16] R. S. Weinstein, M. R. Descour, C. Liang, A. K. Bhattacharyya, A. R. Graham, J. R. Davis, K. M. Scott, L. Richter, E. A. Krupinski, J. Szymus, K. Kayser, and B. E. Dunn. Telepathology overview: From concept to implementation. *Human Pathology*, 32(12):1283 – 1299, 2001.

[17] R. S. Weinstein, A. R. Graham, L. C. Richter, G. P. Barker, E. A. Krupinski, A. M. Lopez, K. A. Erps, A. K. Bhattacharyya, Y. Yagi, and J. R. Gilbertson. Overview of telepathology, virtual microscopy, and whole slide imaging: prospects for the future. *Human Pathology*, 40:8, 2009.

[18] R. S. Crowley, G. J. Naus, J. Stewart III, and C. P. Friedman. Development of visual diagnostic expertise in pathology-an information-processing study. *Journal of the American Medical Informatics Association*, 10:39, 2003.

[19] G. Bussolati. Dissecting the pathologist's brain: mental processes that lead to pathological diagnoses. In *Virchows Archiv*, pages 739–743, 2006.

[20] J. R. Pani, J. H. Chariker, and R. D. Fell. Toward a theory of qualitative visual reasoning in microanatomy. In *Proceedings of the eighteenth International workshop on qualitative reasoning.*, 2004.

[21] E. A. Krupinski, A. A. Tillack, L. Richter, J. T. Henderson, A. K. Bhattacharyya, K. M. Scott, A. R. Graham, M. R. Descour, J. R. Davis, and R. S. Weinstein. Eye-movement study and human performance using telepathology virtual slides. implications for medical education and differences with experience. *Human pathology*, 37(12):1543–1556, 2006.

[22] L. Roa-Pena, F. Gomez, and E. Romero. An experimental study of pathologist's navigation patterns in virtual microscopy. *Diagnostic Pathology*, 5(1):71, 2010.

[23] S. Wienert, M. Beil, K. Saeger, P. Hufnagl, and T. Schrader. Integration and acceleration of virtual microscopy as the key to successful implementation into the routine diagnostic process. *Diagnostic Pathology*, 4(1):3, 2009.

[24] F. Gomez, M. Iregui, and E. Romero. Prediction of pathologist navigation patterns in virtual microscopy based on a soft-computing approach. In *Human Computer Interaction*, 2008.

[25] S. P. Liversedge and J. M. Findlay. Saccadic eye movements and cognition. *Trends in cognitive sciences*, 4(1):6–14, 2000.

[26] J. Gu and R. W. Ogilvie. *Virtual Microscopy and Virtual Slides in Teaching, Diagnosis, and Research*. CRC Press, 2005.

[27] M. H. Kim, Y. Park, D. Seo, Y. J. Lim, D. I. Kim, C. W. Kim, and W. H. Kim. Virtual microscopy as a practical alternative to conventional microscopy in pathology education. *Basic and Applied Pathology*, 1:46–48, 2008.

[28] R. K. Kumar, G. M. Velan, S. O. Korell, M. Kandara, F. R. Dee, and D. Wakefield. Virtual microscopy for learning and assessment in pathology. *The Journal of Pathology*, 204:613–618, 2004.

[29] D. Romer and S. Suster. Use of virtual microscopy for didactic live-audience presentation in anatomic pathology. *Annals of Diagnostic Pathology*, 7:67 – 72, 2003.

[30] J. Gilbertson, J. Ho, L. Anthony, D. Jukic, Y. Yagi, and A. Parwani. Primary histologic diagnosis using automated whole slide imaging: a validation study. *BMC Clinical Pathology*, 6(1):4, 2006.

[31] S. Saltz. Virtual microscope: Databases and system software for multiscale problems. In *E-poster presented in Advancing Pathology Informatics, Imaging, and the Internet, APIII.*, 1999.

[32] U. Catalyurek, M.D. Beynon, C. Chialin, T. Kurc, A. Sussman, and J. Saltz. The virtual microscope. *IEEE Transactions on Information Technology in Biomedicine.*, 7:230–248, 2003.

[33] B. D. Davison. A survey of proxy cache evaluation techniques. In *Proceedings of the 4th International Web Caching Workshop*, pages 67–77, 1999.

[34] M. Iregui, F. Gomez, and E. Romero. Strategies for efficient virtual microscopy in pathological samples using jpeg2000. *Micron*, 38:700–713, October 2007.

[35] A. Descampe, J. Ou, P. Chevalier, and B. Macq. Data prefetching for smooth navigation of large scale jpeg 2000 images. In *IEEE International Conference on Multimedia and Expo, 2005. ICME 2005.*, pages 908–911, 2005.

[36] A. Chang, R. Lau, and B. Ng. Motion prediction for caching and prefetching in mouse-driven dve navigation. *Transactions on Internet Technology (TOIT)*, 5:70–91, 2005.

[37] A. Ortega, F. Carignano, S. Ayer, and M. Vetterli. Soft caching: web cache management techniques for images. In *IEEE First Workshop on Multimedia Signal Processing.*, pages 475–480, 1997.

[38] G. K. Anastassopoulos and A. N. Skodras. Jpeg 2000 roi coding in medical imaging applications. In *Proceedings of the 2nd IASTED International Conference on visualisation, imaging and image processing (VIIP2002)*, 2002.

[39] L. Q. Chen, X. Xie, X. Fan, W. Y. Ma, H. J. Zhang, and H. Q. Zhou. A visual attention model for adapting images on small displays. *Multimedia Systems*, 9:353–364, 2003.

[40] Y. Hu, X. Xie, Z. Chen, and W. Y. Ma. Attention model based progressive image transmission. In *IEEE International Conference on Multimedia and Expo, ICME'04*, volume 2, 2004.

[41] L. Q. Chen, X. Xie, W. Y. Ma, H. J. Zhang, and H. Q. Zhou. Image adaptation based on attention model for small form factor devices. In *Proceeding of the 9th International Conference on Multimedia Modeling*, pages 421–439, 2003.

[42] X. Fan, X. Xie, W. Y. Ma, H. J. Zhang, and H. Q. Zhou. Visual attention based image browsing on mobile devices. In *Proceedings of the International Conference on Multimedia and Expo, ICME'03.*, volume 1, 2003.

[43] H. Liu, X. Xie, W. Y. Ma, and H. J. Zhang. Automatic browsing of large pictures on mobile devices. In *Proceedings of the eleventh ACM international conference on Multimedia*, pages 148–155, 2003.

[44] A. Descampe, C. De Vleeschouwer, M. Iregui, B. Macq, and F. Marques. Pre-fetching and caching strategies for remote and interactive browsing of jpeg2000 images. *IEEE Transactions on Image Processing.*, 16:1339–1354, 2007.

[45] J. M. Wolfe. *Attention*, volume 1, chapter Visual search, pages 13–73. University College London Press, 1998.

[46] L. Itti and C. Koch. Computational modelling of visual attention. *Nature Reviews Neuroscience*, 2(3):194–204, 2001.

[47] M. J. Tovée. *An introduction to the visual system.* Cambridge Univ Press, 1996.

[48] S. Zeki. *Una visión del cerebro.* Editorial Ariel S.A, 1995.

[49] R. Desimone and J. Duncan. Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, 18(1):193–222, 1995.

[50] C. Koch and S. Ullman. Shifts in selective visual attention: towards the underlying neural circuitry. *Human neurobiology*, 4(4):219–227, 1985.

[51] A. M. Treisman and G. Gelade. A feature-integration theory of attention. *Cognitive psychology*, 12(1):97–136, 1980.

[52] J. M. Wolfe. Guided search 4.0: Current progress with a model of visual search. *Integrated models of cognitive systems*, 25:9479–9487, 2007.

[53] H. Watanabe and T. Ogita. Transcoding by automatic roi extraction from jpeg2000 bitstream. In *Proceedings of Picture Coding Symposium*, 2003.

[54] D. A. Karras, S. A. Karkanis, and D. E. Maroulis. Efficient image compression of medical images using the wavelet transform and fuzzy c-means clustering on regions of interest. In *Proceedings of The 26th EUROMICRO Conference.*, 2000.

[55] B. Lessmann, T. W. Nattkemperb, V. H. Hansc, and A. Degenharda. A method for linking computed image features to histological semantics in neuropathology. *Journal of Biomedical Informatics*, 40:631–641, 2007.

[56] G. P. Pena and J. Andrade-Filho. How does a pathologist make a diagnosis? *Archives of Pathology & Laboratory Medicine*, 133:124–132, 2009.

[57] L. Itti, C. Koch, and E. Niebur. A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on pattern analysis and machine intelligence*, 20(11):1254–1259, 1998.

[58] L. Itti and C. Koch. A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision research*, 40(10-12):1489–1506, 2000.

[59] F. Gómez, J. Villalón, R. Gutierrez, and E. Romero. Finding regions of interest in pathological images: an attentional model approach. In *Proceedings of SPIE*, volume 7260, page 72603G, 2009.

[60] R. Gutiérrez. Compresión selectiva en imágenes digitales de microscopio en patología. Master's thesis, Universidad Nacional de Colombia, 2009.

[61] L. Itti and C. Koch. Feature combination strategies for saliency-based visual attention. *Systems Journal of Electronic Imaging*, 10:161–169, 2001.

[62] V. Navalpakkam and L. Itti. Search goal tunes visual features optimally. *Neuron*, 53(4):605–617, 2007.

[63] V. Navalpakkam and L. Itti. Optimal cue selection strategy. *Advances in neural information processing systems*, 18:987, 2006.

[64] R. Rao, G. J. Zelinsky, M. M. Hayhoe, and D. H. Ballard. Eye movements in iconic visual search. *Vision Research*, 42(11):1447–1463, 2002.

[65] M. DeAngelus and J. Pelz. Top-down control of eye movements: Yarbus revisited. *Visual Cognition*, 17:790 – 811, 2009.

[66] L. Dempere-Marco, X. Hu, S. M. Ellis, D. M. Hansell, and G. Z. Yang. Analysis of visual search patterns with emd metric in normalized anatomical space. *IEEE Transactions on Medical Imaging*, 25(8):1011–1021, 2006.

[67] C. M. Privitera and L. W. Stark. Algorithms for defining visual regions-of-interest: Comparison with eye fixations. *IEEE Transactions on pattern analysis and machine intelligence*, 22(9):970–982, 2000.

[68] G. Boccignone and M. Ferraro. Modelling gaze shift as a constrained random walk. *Physica A: Statistical Mechanics and its Applications*, 331(1-2):207–218, 2004.

[69] I. A. Rybak, V. I. Gusakova, A. V. Golovan, L. N. Podladchikova, and N. A. Shevtsova. A model of attention-guided visual perception and recognition. *Vision Research*, 38:2387– 2400, 1998.

[70] G. Deco and J. Zihl. A neurodynamical model of visual attention: Feedback enhancement of spatial resolution in a hierarchical system. *Journal of Computational Neuroscience*, 10:231–253, 2001.

[71] J. Caicedo. A prototype system to archive and retrieve histopathology images by content. Master's thesis, Universidad Nacional de Colombia, 2008.

[72] J. C. Caicedo, F. A. Gonzalez, and E. Romero. A semantic content-based retrieval method for histopathology images. In *Proceedings of the Asia Information Retrieval Symposium*, 2008.

[73] F. Willemse. *A colored view on quantitative pathology : aspects of true color image analysis in routine pathology.* PhD thesis, University of Groningen, 1996.

[74] C. Demir and B. Yener. Automated cancer diagnosis based on histopathological images: A systematic survey. Technical Report Technical Report TR-05-09., Computer Science Department at Rensselaer Polytechnic Institute., 2005.

[75] T. Deselaers. *Features for Image Retrieval.* PhD thesis, RWTH Aachen University of Technology, 2003.

[76] ISO/IEC JTC1/SC29 WG1. Jpeg2000 part i final committee draft version 1.0. Working Draft, 2003.

Apéndice A

Productos

Ponencia en conferencia internacional

Romo, D., Romero, E., González, F.: *Análisis de relevancia usando patrones de navegación en mega-imágenes histopatológicas*. IV Seminario Internacional de Procesamiento y Análisis de Imágenes Médicas, SIPAIM 2009. Universidad de los Andes - Universidad Nacional de Colombia. 2009

Artículo en conferencia international

Romo, D., Romero, E., González, F.: *Learning Regions of Interest from Low-Level Maps in Virtual Microscopy*. 10th European Congress on Telepathology and 4th International Congress on Virtual Microscopy. 2010. Sometido a revisión para ser publicado en el journal electrónico: *Diagnostic Pathology*[1].

[1]El artículo se anexa a continuación