



Benemérita Universidad Autónoma de Puebla
Facultad de Ciencias Físico Matemáticas

Implementación de algoritmos de reconstrucción
tomográfica mediante programación paralela (CUDA)

Tesis presentada al

Posgrado en Física Aplicada

como requisito parcial para la obtención del grado de

Doctor en Ciencias

por

Tomás Antonio Valencia Pérez

Asesorado por

Dr. Benito de Celis Alonso
Dr. Eduardo Moreno Barbosa

2 de julio del 2020



Benemérita Universidad Autónoma de Puebla
Facultad de Ciencias Físico Matemáticas

Implementación de algoritmos de reconstrucción
tomográfica mediante programación paralela (CUDA)

Tesis presentada al

Posgrado en Física Aplicada

como requisito parcial para la obtención del grado de

Doctor en Ciencias

por

Tomás Antonio Valencia Pérez

Asesorado por

Dr. Benito de Celis Alonso
Dr. Eduardo Moreno Barbosa

2 de julio del 2020

Título: Implementación de algoritmos de reconstrucción tomográfica mediante programación paralela (CUDA)
Estudiante: Tomás Antonio Valencia Pérez

COMITÉ

Dr. Javier Miguel Hernández López
Presidente

Dr. Mario Iván Martínez Hernández
Secretario

Dr. Mario Rodríguez Cahuantzi
Vocal

Dr. Luis Villaseñor Cendejas
Vocal Externo

Dr. Enrique Varela Carlos
Vocal Externo

Dr. Humberto Salazar Ibargüen
Suplente

Dr. Benito de Celis Alonso
Asesor

Dr. Eduardo Moreno Barbosa
Asesor

Agradecimientos

Muchas gracias a todos los que hicieron posible este trabajo.

Índice general

1. INTRODUCCIÓN	1
1.1 Estado del arte	1
1.2 Trabajo original y esquema de esta tesis	3
2. TOMOGRAFÍA COMPUTARIZADA DE RAYOS X.....	5
2.1 Introducción	5
2.2 Tomografía computarizada.....	5
3. FÍSICA DE LOS RAYOS X	11
3.1 Introducción	11
3.2 Los Rayos X.....	11
3.2.1 Producción de rayos X.....	12
3.3 Interacciones de rayos X con materia.....	14
3.3.1 Dispersión de Rayleigh.....	14
3.3.2 Dispersión Compton	15
3.3.3 Efecto Fotoeléctrico.....	18
3.3.4 Producción de pares.....	21
4. ALGORITMOS DE RECONSTRUCCIÓN DE IMÁGENES MÉDICAS	22
4.1 Introducción	22
4.2 Reconstrucción en general.....	22
5. MÉTRICAS DE CALIDAD PARA IMÁGENES MÉDICAS	28
5.1 Introducción	28
5.2 Parámetros de evaluación de calidad de imagen	29
5.2.1 Selección de métricas de calidad	30
5.2.2 Error cuadrático medio	30
5.2.3 PSNR	30
5.2.4 SSIM	30
5.3 Métricas en la implementación MLEM.....	31
6. IMPLEMENTACIÓN DEL ALGORITMO MLEM EN LA GPU.....	32
6.1 Introducción	32
6.2 Enfoque de reconstrucción de imágenes médicas	32
7. RESULTADOS	36
8. APLICACIONES CLÍNICAS	42
8.1 Estudio clínico.....	42

9. CONCLUSIONES	45
10. PUBLICACIONES	48
11. REFERENCIAS.....	49

Capítulo 1

Introducción

La reconstrucción de imágenes médicas es clave en una amplia gama de tecnologías. Para los sistemas de tomografía computarizada clásica, la cantidad de señales medidas por segundo aumentó exponencialmente en las últimas cuatro décadas, mientras que la complejidad computacional de la mayoría de los algoritmos utilizados no ha cambiado significativamente. Un gran interés y desafío es proporcionar una calidad de imagen óptima con la menor dosis posible de radiación al paciente. Una solución y un campo de investigación activo para resolver ese problema son los métodos iterativos de reconstrucción de imágenes médicas. Su complejidad es múltiple en comparación con los métodos analíticos clásicos que se utilizaron en casi todos los sistemas disponibles comercialmente. En esta tesis se investiga el uso de tarjetas gráficas en el campo de la reconstrucción iterativa de imágenes médicas. Se presentan y evalúan los diferentes enfoques de algoritmos de reconstrucción de imagen acelerados por la GPU (Unidad de Procesamiento Gráfico, por sus siglas en inglés).

1.1 Estado del arte

La tomografía computarizada (TC) es una técnica no destructiva en la que una fuente de rayos X (radiación ionizante) gira en torno de un objeto de interés generando cortes axiales de su interior. La TC es hoy en día una herramienta indispensable en medicina que ayuda al diagnóstico de múltiples enfermedades [1], cómo cáncer de hígado, riñón, páncreas, entre otros, así como afecciones en el cerebro tales como hidrocefalia, etc. Desde su

introducción en 1970, hay cerca de 30,000 escáneres TC en el mundo y su número continúa aumentando de manera exponencial [1]. Incluso si las técnicas de TC representan uno de los mayores desarrollos en el campo de los rayos X en los últimos 50 años [1], aún existe un margen de mejora. Algunas de las líneas de trabajo propuestas para conseguir este objetivo son:

- Reducción del tiempo de exposición del paciente
- Reducción de tiempos de adquisición y procesamiento
- Desarrollo de nuevas técnicas con nuevas funcionalidades
- Reducción de costos [2]

Algunas de las soluciones a estos puntos se han centrado hasta ahora en el desarrollo de hardware de mayor rendimiento que permiten costos más bajos y una entrega de dosis de radiación más baja a los pacientes sin comprometer la efectividad del diagnóstico. También se ha trabajado en el uso de métodos iterativos para la reconstrucción de imágenes, reduciendo los tiempos de procesamiento para la adquisición de imágenes con valor diagnóstico.

Los escáneres TC actuales usan de forma regular para su reconstrucción de imágenes la Retro Proyección Filtrada o FBP por sus siglas en inglés [3,4]. Esta técnica matemáticamente tiene su fundamento en la Transformada de Radon (TR) [5,6]. En ella, la imagen de un objeto se reconstruye a partir de un conjunto de proyecciones paralelas de rayos X sobre el mencionado objeto [7]. Hay muchos algoritmos computacionales que pueden ser utilizados para resolver la TR. El marco teórico de los métodos de reconstrucción se puede consultar en [2,8,9]. Tradicionalmente, se clasifican en dos grupos:

- Métodos de reconstrucción analíticos
- Métodos de reconstrucción iterativos

Para el grupo 1, el uso de los algoritmos FBP ha sido el estándar hasta la fecha [10]. Se trata de un algoritmo de reconstrucción simple y rápido con bajo costo computacional. Sin embargo, es propenso a la aparición de artefactos y en ocasiones presenta baja resolución espacial. Además, requiere de un número elevado de proyecciones sin ruido para que la solución de la reconstrucción sea de calidad. Esto está lejos de las situaciones de la vida real en las que sólo estará disponible un conjunto finito de proyecciones con ruido. Por otro lado, se dispone de métodos de reconstrucción “aproximados” también conocidos como algoritmos iterativos. Son menos sensibles a conjuntos incompletos de datos, el ruido y los artefactos que surgen en el proceso de reconstrucción pueden reducirse o incluso eliminarse [10,11]. Esto se traduce en una mejor calidad en la reconstrucción de las imágenes con respecto a los métodos de reconstrucción analíticos. Además, al requerir menos información para una reconstrucción, pueden usar menos proyecciones/radiación reduciendo las dosis entregadas en un estudio a un paciente. Para una revisión de las matemáticas tras estos métodos ver la siguiente referencia [12]. El más común entre ellos es: *Maximum Likelihood Expectation Maximization* (MLEM) [13]. Sin embargo, la principal limitación que presentan estas técnicas con respecto a los métodos de reconstrucción analíticos es su alto costo computacional.

Una forma de lidiar con la limitación de largos tiempos computacionales es paralelizar el algoritmo de reconstrucción de la imagen. Con este fin, se puede utilizar la *Graphics Processing Unit* (GPU) en combinación con el entorno de programación paralela *Compute*

Unified Device Architecture (CUDA) [14,15,16], mismo que nos permite el manejo y procesamiento de datos en GPUs. Está formada por unidades de ejecución, llamadas *Streaming Multiprocessors* (SM) que a su vez están formadas por núcleos de computación denominados *Streaming Processors* (SP), o núcleos CUDA. Son estos núcleos los que ejecutan las instrucciones, es decir, las operaciones matemáticas o el direccionamiento y transferencia de datos en la memoria. El modelo de programación de estas tarjetas gráficas involucra al mismo tiempo la *Central Processing Unit* (CPU) y la GPU. La parte secuencial de las aplicaciones es ejecutada en la CPU (Host) y la parte de computación intensiva (en informática, es la cualidad de un programa que indica la necesidad de realizar cantidades enormes de operaciones aritméticas con números reales) mediante el uso de la GPU (*Device*) en paralelo. Tienen una aplicación básica dentro de campos tan diversos como: bioinformática, química computacional, imágenes y visión por computadora, clima, análisis numéricos, etc. [17].

El procesamiento de imágenes médicas fue uno de los primeros campos en los que se utilizó esta tecnología [18,19]. Tatarchuk et al. [20] presentó un conjunto de métodos que permiten la exploración interactiva de conjuntos de datos médicos en tiempo real. Flores et al. [21] también presentó el desarrollo de un algoritmo rápido implementado en GPUs para reconstruir imágenes de alta calidad a partir de datos de proyección muestreados y con ruido. Belzunze et al. obtuvo una implementación paralela bajo CUDA con datos de medicina nuclear (SPECT y PET) y se logró un factor de aceleración de hasta 85 veces con respecto a una implementación de CPU de un sólo hilo [22]. Xie et al. propuso una forma de sincronizar la adquisición de datos de TC con el análisis de la GPU reduciendo el tiempo de análisis de computación entre 10 y 16 veces [23]. Finalmente, Xing et al. [24] muestra más aplicaciones sobre el uso de las GPU en esta área.

1.2 Trabajo original y esquema de esta tesis

En esta tesis presentamos la reconstrucción de imágenes médicas TC con un método iterativo que resuelve el algoritmo de reconstrucción MLEM [13]. Mediante el uso de la GPU y CUDA para paralelizar el análisis y reducir sus costos computacionales. Este análisis fue realizado en condiciones no ideales (ruido) y se compararon sus resultados con los de programas comerciales de reconstrucción de imagen médica. Otro objetivo de este estudio es que el esquema de trabajo desarrollado pueda ser implementado en una computadora portátil comercial con una tarjeta gráfica compatible con CUDA y que los requerimientos mínimos nos provean de una plataforma de reconstrucción de imágenes rápida, portátil, simple y de bajo costo.

La tesis está organizada de la siguiente manera:

- En el capítulo [2] se realiza una breve introducción a la tomografía computarizada.
- En el capítulo [3] se discute el proceso de producción de Rayos X y sus principales interacciones en TC.
- En el capítulo se dan a conocer conceptos de programación paralela y las principales características de las GPU.

- En el capítulo 4 ofrecemos una visión general del campo de los algoritmos de reconstrucción y una clasificación de éstos.
- En el capítulo 5 se enumera a un conjunto de métricas de calidad de imagen que posteriormente fueron utilizadas para cuantificar los resultados.
- El capítulo 6 muestra la principal contribución científica en el campo de la reconstrucción de imágenes médicas.
- El capítulo 7 está dedicado a mostrar de manera cuantitativa los resultados obtenidos por las implementaciones del algoritmo MLEM.
- En el capítulo 8 se muestra el comportamiento de aplicar la implementación Parallel-OpT a un estudio clínico.
- Finalmente, en el capítulo 9 se hace un resumen de las conclusiones de la investigación realizada.

Capítulo 2

Tomografía computarizada de Rayos X

2.1 Introducción

La tomografía axial computarizada (TC) fue creada por Godfrey Hounsfield en 1972, conjuntamente con Allan Cormack, quien había desarrollado algunos trabajos previos, fue galardonado con el Premio Nobel de Medicina en 1979. Las aplicaciones de TC de rayos X abarcan varios campos, desde la medicina hasta la industria [25].

En este capítulo se presentan los aspectos básicos de la tomografía computarizada de rayos X, con énfasis en los métodos de reconstrucción de imágenes.

2.2 Tomografía computarizada

La TC permite la adquisición de imágenes bidimensionales en base a la emisión de rayos X correspondientes a delgados cortes axiales a lo largo del cuerpo del paciente. A partir de múltiples cortes axiales consecutivos se puede realizar una reconstrucción del volumen correspondiente al área radiada. Las imágenes TC son capaces de mostrar de forma contrastada distintos tejidos blandos. Su principio básico es que la estructura

bidimensional interna de un objeto puede ser reconstruida a partir de una serie de proyecciones unidimensionales del objeto adquiridas desde distintos ángulos.

Para poder obtener una imagen de una delgada capa de tejido, el haz de rayos X debe ser confinado para obtener, también, un flujo lo más direccional posible. Los detectores, situados de frente al emisor, detectan el número total de rayos que han atravesado al paciente, produciendo una proyección unidimensional. Las intensidades de señal recogidas responden a la distribución de los coeficientes de atenuación a lo largo del corte axial correspondiente a la distribución bidimensional de tejidos. Entonces, el emisor y el receptor repetirán esta operación tras girar un determinado ángulo. Esta operación será repetida hasta conseguir un número suficiente de muestras que permita obtener una imagen con alta resolución espacial. La reconstrucción de la imagen requiere también de un algoritmo de reconstrucción que se explicará posteriormente. Tras la reconstrucción, la imagen TC es un mapa de valores TC (denominado coeficientes de atenuación), de modo que cada píxel de la misma está dotado con el valor TC correspondiente al tejido radiado en esa localización espacial.

La TC de rayos X utiliza mediciones de transmisión para estimar una imagen transversal dentro del cuerpo del paciente. Los rayos X tienen una energía muy alta (en el rango de entre 1 eV a 150 keV) y pueden penetrar el cuerpo del paciente. Sin embargo, no todos los rayos X pueden atravesar el cuerpo del paciente. Algunos rayos X se atenúan o dispersan con otros tejidos dentro del cuerpo, y su energía se debilita. Durante la dispersión de rayos X, un fotón de rayos X interactúa con un electrón dentro del paciente, transfiere parte de su energía a ese electrón y lo desaloja (figura 2.1).

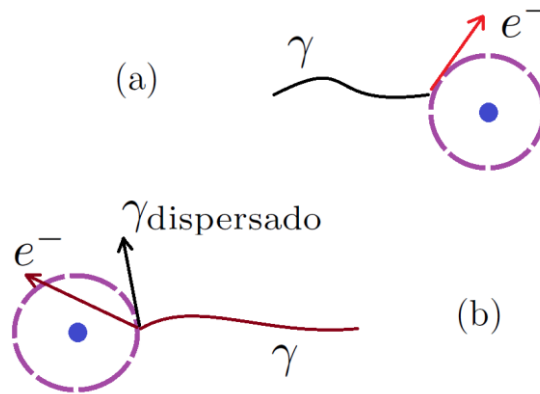


Fig. 2.1 Representación esquemática del (a) efecto fotoeléctrico y de (b) la dispersión de Compton

Los rayos X se desvían a una nueva dirección con energía disminuida. El efecto fotoeléctrico es un proceso en el cual la energía del fotón de rayos X es completamente absorbida por un átomo dentro del paciente. La energía absorbida expulsa un electrón del átomo (figura 2.1). La deposición de energía dentro del cuerpo puede dañar el ADN si la dosis de rayos X es demasiado grande.

Sea I_0 la intensidad de los rayos X antes de ingresar al paciente, y I_d la intensidad que sale del paciente; entonces I_0 e I_d siguen la ley de Beer [9] (figura 2.2):

$$\frac{I_d}{I_0} = \exp(-p) \quad (2.1)$$

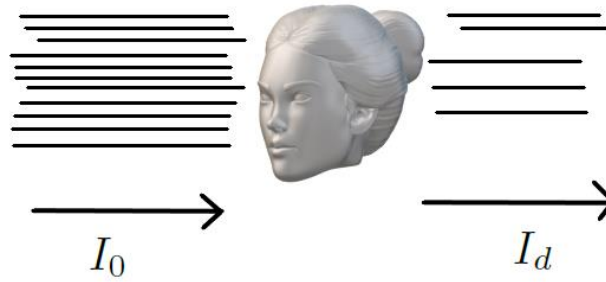


Fig. 2.2 La intensidad de los rayos X se atenúa después de atravesar el objeto.

donde p es la integral de línea de los coeficientes de atenuación lineal a lo largo de la trayectoria de los rayos X.

Una integral de línea de los coeficientes de atenuación se obtiene mediante:

$$I_d = I_0 \exp\left(-\int_L \mu(x)dx\right) \Rightarrow \int_L \mu(x)dx = \ln\left(\frac{I_0}{I_d}\right) \quad (2.2)$$

Por lo que:

$$p = \ln\left(\frac{I_0}{I_d}\right) \quad (2.3)$$

este valor, es el utilizado por los algoritmos de reconstrucción de imágenes.

El objetivo de la TC de rayos X es obtener una imagen de sección transversal de varios coeficientes de atenuación dentro del cuerpo. Una imagen típica de TC de rayos X se muestra en la Figura 2.3. El coeficiente de atenuación (comúnmente denotado por la notación μ) es una propiedad de los materiales y es el logaritmo de la relación de intensidad de entrada/salida por unidad de longitud.

Los huesos tienen valores de μ más altos y los tejidos blandos tienen valores de μ más bajos. El coeficiente de atenuación de un material es inversamente proporcional a la energía de rayos X incidente.

La TC de primera generación, que ya no está en uso, consistía en un único y pequeño detector de centelleo (figura 2.4). La fuente de rayos X y el detector tienen dos

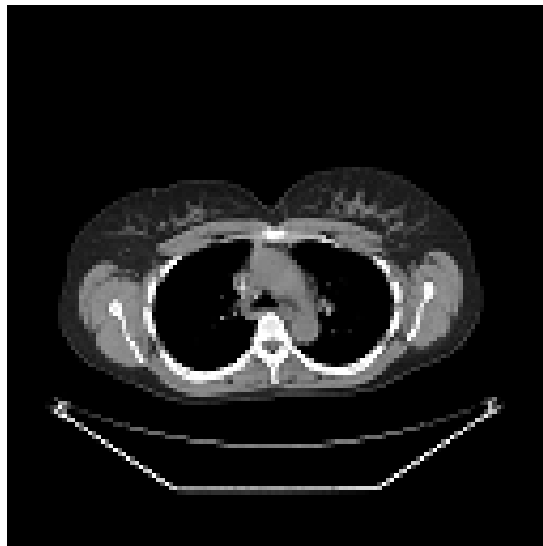


Fig. 2.3 Una imagen TC de rayos X de abdomen [26].

movimientos: traslación lineal y rotación. Los detectores y la fuente de rayos X en primer lugar tienen un movimiento de traslación en el cual se realiza la toma de datos, seguidamente rotan un grado alrededor del cuerpo para comenzar una nueva secuencia de adquisición de datos.

La fuente de rayos X envía un haz estrecho denominado de lápiz, para obtener proyecciones de haz paralelo. El tiempo de exploración era bastante largo (aproximadamente 25 minutos).

La TC de segunda generación utilizó una geometría de haz en abanico, que consistió de 12 detectores de centelleo (figura 2.5).

Al igual que la TC de primera generación, tiene dos movimientos: traslación lineal y rotación. La toma de datos se realiza de manera similar a la TC de primera generación. Puesto que se ha aumentado el número de detectores, el movimiento angular es mayor que junto a la geometría de haz en abanico el tiempo de exploración se acortó a aproximadamente 1 minuto.

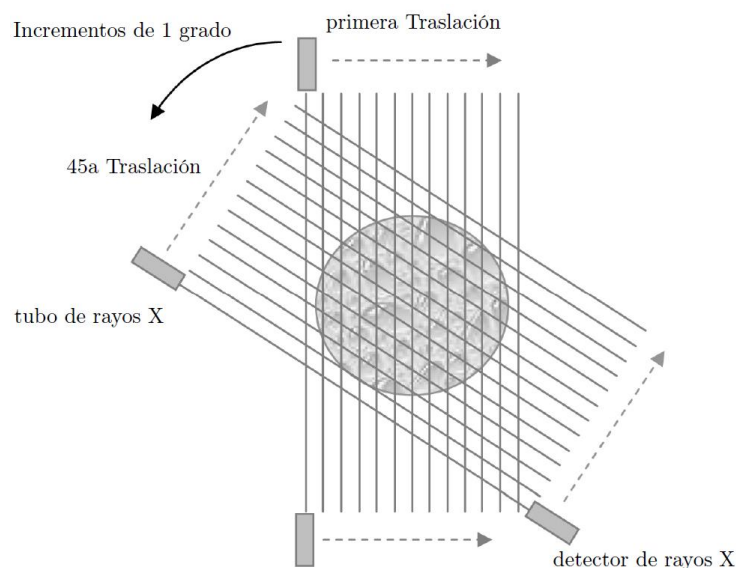


Fig. 2.4 Geometría del escáner TC de primera generación. El tubo de rayos X y el detector se trasladan linealmente para cubrir todo el objeto. Todo el aparato rota 1 grado para repetir la exploración. Figura tomada de [27].

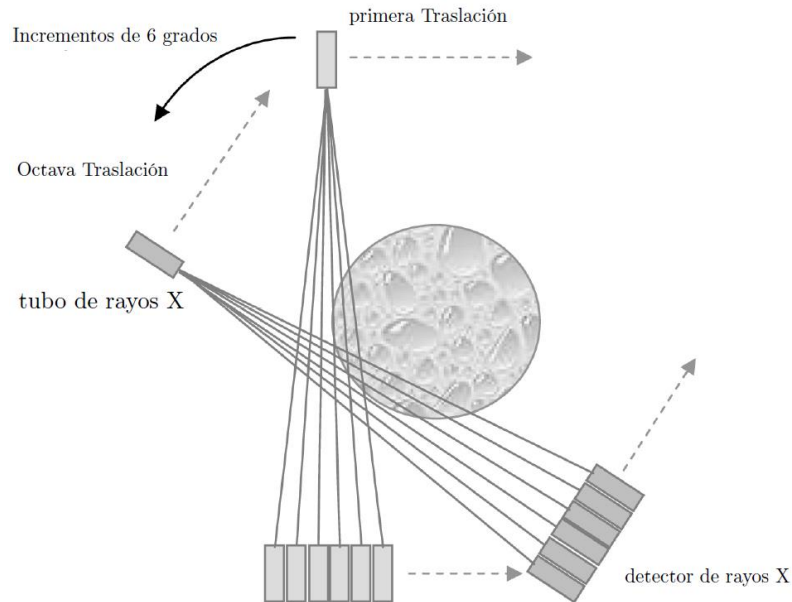


Fig. 2.5 Geometría de escáner TC de segunda generación. Aunque la fuente de rayos X y el detector todavía necesitan ser trasladados linealmente, el tubo de rayos X y el detector pueden rotar cada 6 grados. Figura tomada de [27].

La tercera generación de TC utiliza la geometría de haz en abanico más amplio, que consideraba en promedio 1000 detectores de Gas Xenón o de estado sólido (figura 2.6). No es necesario el movimiento de traslación ya el sistema se basa en una rotación continua de la fuente de rayos X y un arreglo de detectores. La fuente de rayos X produce un haz en abanico y los detectores van tomando los datos de una sola vez por vistas. Con esta geometría el tiempo de exploración se redujo adicionalmente a aproximadamente 0.5 segundos. La TC de tercera generación es actualmente la más usada hoy en día en imágenes médicas.

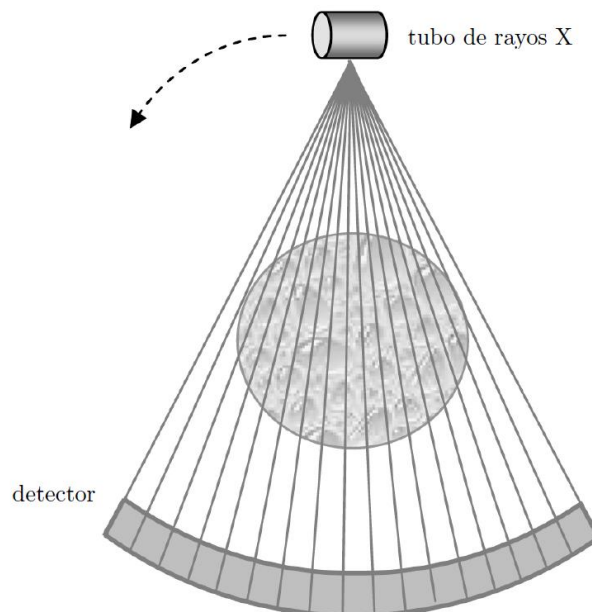


Fig. 2.6 Geometría de escáner TC de tercera generación. El tubo de rayos X y el detector son estacionarios entre sí mientras todo el aparato gira alrededor del paciente. Figura tomada de [27].

La TC de cuarta generación tiene un detector de estado sólido en forma de anillo estacionario. La fuente de rayos X gira alrededor del sujeto (figura 2.7). Este método de escaneo forma una geometría de imágenes de haz de abanico muy rápido; sin embargo, es imposible colimar los rayos X en el detector debido a que los detectores no tienen una

posición fija con respecto a la fuente de rayos X, entonces no se podía utilizar una rejilla (pieza colocada entre cada detector) enfocada para rechazar las radiaciones secundarias por lo que este tipo de geometría produce mucha dispersión.

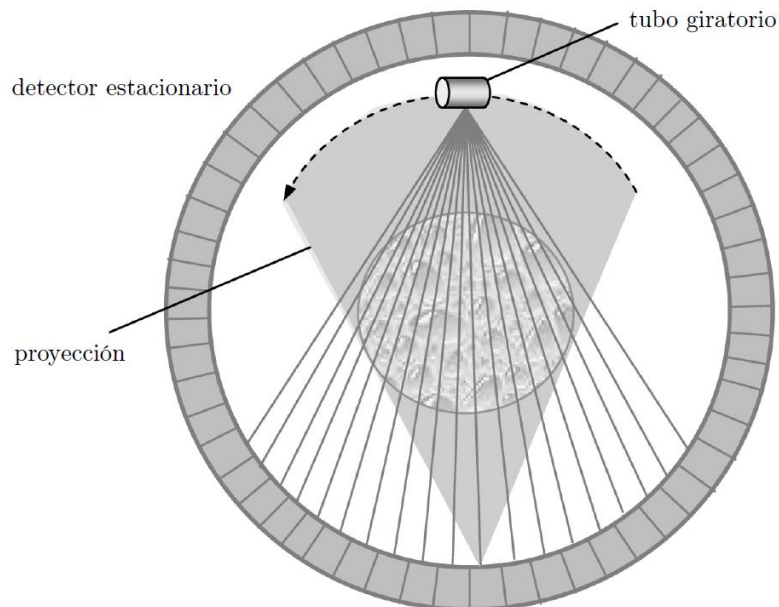


Fig. 2.7 Geometría de un escáner TC de cuarta generación. Una proyección se forma usando muestras de medición de un solo detector a lo largo del tiempo, como se representa por la región sombreada en forma de abanico. Figura tomada de [27].

La TC moderna puede realizar exploraciones helicoidales, que se implementan trasladando la cama del paciente en la dirección axial a medida que la fuente de rayos X y los detectores de estado sólido giran. La TC moderna tiene un detector de múltiples filas 2D y adquiere datos de haz de cono (figura 2.8).

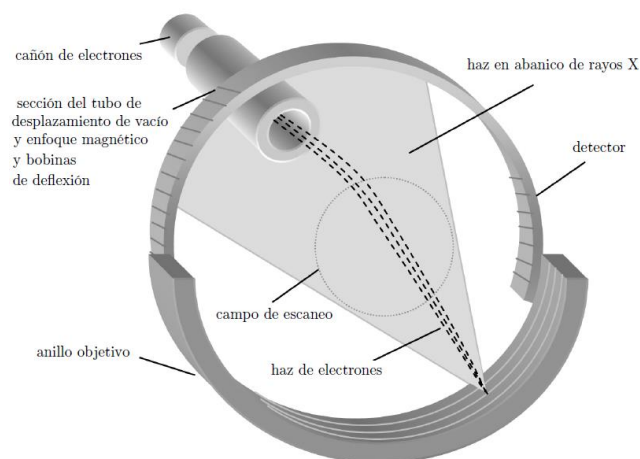


Fig. 2.8 Geometría de un escáner de haz de electrones. El escáner de haz de electrones, a veces llamado escáner de quinta generación, utilizado en tomografía computarizada de haz de electrones. Figura tomada de [27].

Capítulo 3

Física de los Rayos X

3.1 Introducción

En este capítulo se describen los principales componentes del sistema de tomografía de rayos X. El tubo de rayos X utilizado para la producción de éstos, así como la interacción de los rayos X con la materia y el concepto de fluorescencia se presentan en detalle.

3.2 Los Rayos X

Los rayos X, descubiertos por el físico alemán Wilhelm Conrad Von Roentgen en 1895, son una radiación electromagnética de alto poder de penetración. Son fotones de alta energía cuya longitud de onda tienen un rango de 0.01 a 10 nm y tienen aplicaciones médicas e industriales [28].

La figura 3.1 representa a todo el espectro electromagnético respecto a su longitud de onda. Para cada región espectral, se muestra la longitud de onda, las frecuencias y la energía de un fotón. Por ejemplo, ondas de radio (que están en el rango de VHF y UHF) se utilizan para resonancia magnética nuclear y resonancia paramagnética de electrones; rayos X para ionización y rayos γ para efectos nucleares, etc.

El Espectro Electromagnético

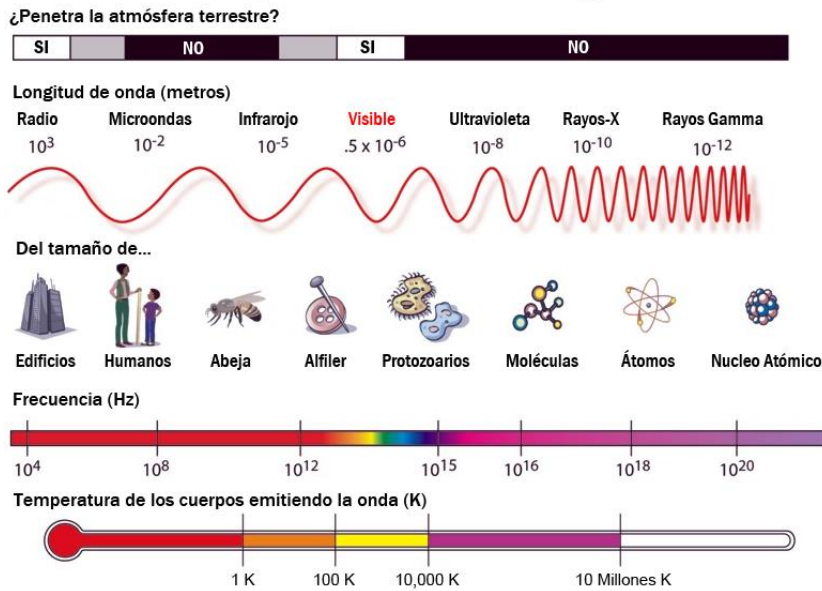


Fig. 3.1 Espectro electromagnético [29].

3.2.1 Producción de rayos X

Los rayos X se producen en un aparato llamado tubo de rayos X. Como se muestra en la figura 3.2, es un tubo de vacío cuyos principales componentes son un cátodo y un ánodo objetivo, con una envoltura de vidrio o metal para mantener un vacío para que los electrones del aire no contribuyan a la corriente de electrones en el interior del tubo de rayos X, lo que interrumpiría el proceso de producción de rayos X y dañaría el tubo [10].

La producción de rayos X, representada en la figura 3.2, implica acelerar un haz de electrones para golpear la superficie de un objetivo metálico.

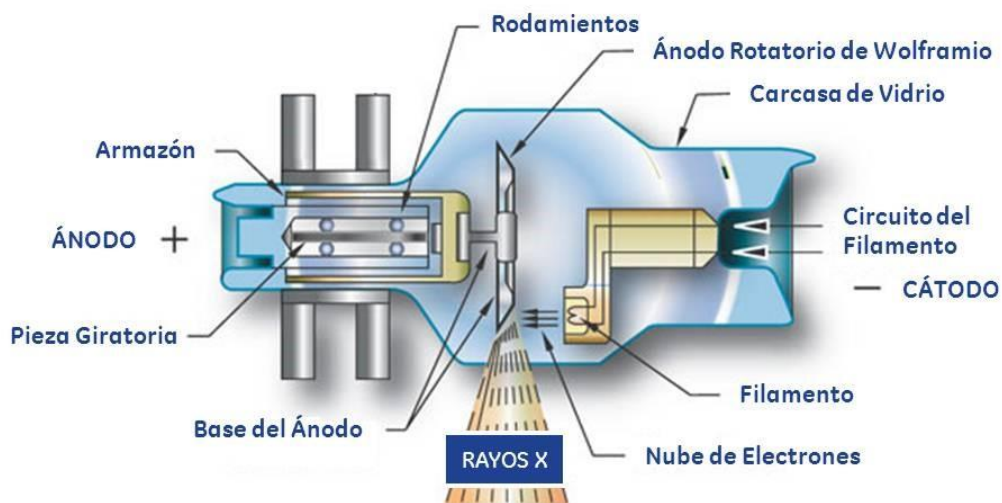


Fig. 3.2 Componentes del tubo de rayos X [30].

Un filamento montado en la ranura del cátodo se calienta mediante un suministro de voltaje relativamente bajo. Los filamentos emiten electrones cuando su energía térmica adquirida supera su energía de unión. Estos electrones se aceleran hacia el ánodo aplicando una diferencia de potencial alta a través del tubo desde la fuente de alimentación de alto voltaje. El objetivo generalmente está hecho de un metal con un número atómico alto como el tungsteno de 74. Los electrones incidentes que golpean el objetivo se acercan al núcleo, pero no penetran debido a la fuerza de la fuerza nuclear objetivo.

Por lo tanto, a cualquier tensión aplicada, se produce una desviación del electrón.

Esta desviación o ralentización provoca una pérdida de energía denominada Bremsstrahlung, una palabra alemana que significa “radiación de frenado”.

Estos fotones de rayos X tienen una nueva dirección y una distribución continua de energías. Cuando el electrón esté cerca del núcleo, se produce mayor energía y menos cantidad de fotones. El otro caso de interacción es cuando la energía del electrón incidente es mayor que la energía de unión del átomo objetivo del ánodo. Por lo tanto, se expulsa un electrón orbital dejando un agujero. Un electrón de capa externa llena la vacante.

Si la brecha de energía entre los niveles es mayor que la del enlace del electrón orbital entonces se libera el fotón de rayos X denominado característico. Como se muestra en la figura 3.3, los dos picos agudos representan las líneas de rayos X característicos y se superponen en el espectro continuo de rayos X de tipo Bremsstrahlung. K_{α} es un rayo X característico producido cuando la transición electrónica es del nivel L al nivel K , mientras que K_{β} es del nivel M al nivel K .

La radiación Bremsstrahlung muestra un rango de valores de energía correspondientes a la pérdida de energía cinética incidente cuando interactúa con el núcleo objetivo.

La eficiencia de los rayos X producidos depende de su energía, que determina su poder de penetración. A medida que aumenta la energía cinética del electrón incidente, aumentan la intensidad y la energía. En general, el 99% de la energía cinética se convierte en energía térmica.

Por lo tanto, el material objetivo debe tener un alto punto de fusión y una gran

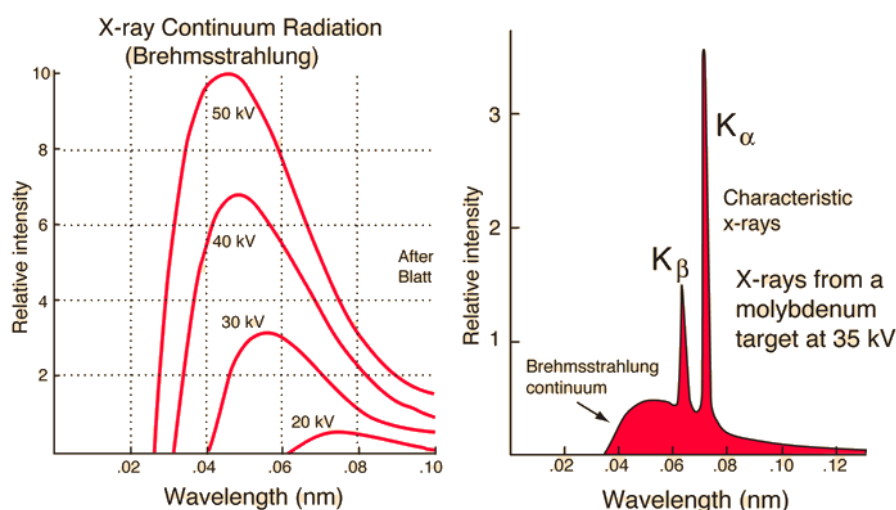


Fig. 3.3 Radiación de Bremsstrahlung, intensidad relativa versus longitud de onda. Figura tomada de [31].

capacidad de disipación de calor. La mayoría de ellos tienen un área focal grande y, en algunos casos, un ánodo giratorio para tener un área efectiva mayor expuesta al haz. La energía de los fotones de rayos X característicos es específica y corresponde a la diferencia de energías entre las capas de transición.

La energía de los fotones de rayos X de Bremsstrahlung varía de cero a una energía equivalente a la energía cinética incidente del electrón del proyectil.

La energía del fotón es definida de la siguiente forma:

$$E = hf_r = \frac{hc}{\lambda} \quad (3.1)$$

donde λ es la longitud de onda, c es la velocidad de la luz que es igual a 3×10^8 m/s en el vacío, f_r es la frecuencia de la radiación y h es la constante de Planck que es igual a 4.1356×10^{15} eVs.

El alto poder de penetración de los rayos X o la alta energía varía con la configuración del tubo. La calidad del haz se define como la habilidad de penetración de la radiación.

Además, depende del voltaje de aceleración (diferencia de voltaje entre el filamento y el ánodo), la corriente del haz y el material objetivo del tubo. Estos factores afectan la intensidad del haz que se define como la cantidad total de energía por unidad de área por unidad de tiempo. Por ejemplo, al aumentar el voltaje del tubo, el haz de rayos X será más energético y, por lo tanto, se producirán fotones más penetrantes. Este control generalmente se expresa en keV. Otro parámetro de la salida del tubo de rayos X es el valor de miliamperios mA. Una corriente más alta dará como resultado un mayor flujo de fotones.

En consecuencia, no hay fotones a energías más altas que al alto voltaje aplicado en el espectro de rayos X. Las bajas energías que se encuentran en el espectro con una longitud de onda más larga se llaman rayos X suaves y generalmente no son deseados, ya que se absorben totalmente en la muestra. Se podría obtener un haz de rayos X más duro mediante la aplicación de un filtro (generalmente una película metálica delgada) que absorbe fotones de baja energía, mientras que los fotones con energías más altas podrán continuar saliendo del filtro.

3.3 Interacciones de rayos X con materia

El rango de energía de los fotones de rayos X generados para la TC esta en el rango de 20 keV a 140 keV. En este rango de energía, hay cuatro formas fundamentales en que los rayos X interactúan con la materia: la dispersión de Rayleigh, el efecto Compton, el efecto fotoeléctrico y la producción de pares [31].

3.3.1 Dispersión de Rayleigh

La dispersión de Rayleigh se da, cuando un fotón incidente interactúa con el átomo total y lo excita, en lugar de los electrones individuales como en la dispersión de Compton o el efecto fotoeléctrico. Esta interacción se produce principalmente con rayos X a energías muy bajas, tal como se utiliza en la mamografía (15 a 30 keV).

Cuando ocurre la dispersión de Rayleigh, una onda electromagnética con un campo eléctrico oscilante provoca que todos los electrones en el átomo oscilen en fase. La nube de electrones del átomo irradia inmediatamente esta energía, emitiendo un fotón de la misma energía, pero en una dirección ligeramente diferente (Fig. 3.4). En esta interacción,

los electrones no son expulsados y por lo tanto no se produce la ionización. En general, el ángulo de dispersión aumenta a medida que la energía de rayos X disminuye.

En la imagen médica, los fotones primarios registrados por el sistema receptor forman la imagen, mientras que los fotones dispersos, pero no absorbidos contribuyen a la formación de una imagen de fondo que degrada la imagen. Sin embargo, este tipo de interacción tiene una baja probabilidad de ocurrencia en el rango de energía de diagnóstico. En el tejido blando, la dispersión de Rayleigh representa menos del 5% de las interacciones de rayos X por arriba de 70 keV y como máximo sólo representa el 12% de las interacciones a aproximadamente 30 keV. Las interacciones de Rayleigh también se conocen como dispersión “coherente” o “clásica”.

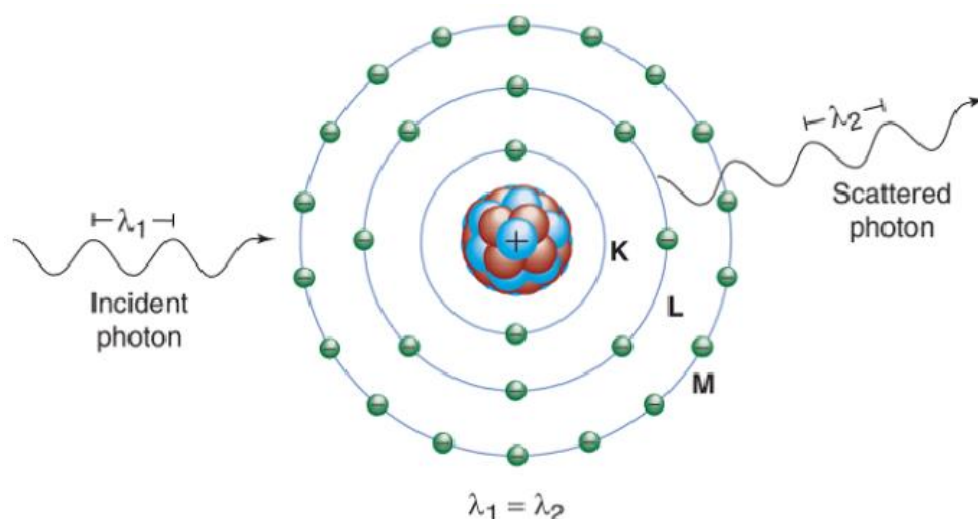


Fig. 3.4 Dispersión de Rayleigh. El diagrama muestra el fotón incidente λ_1 interactuando con un átomo y el fotón dispersado λ_2 está siendo emitido con aproximadamente la misma longitud de onda y energía. Los fotones dispersados de Rayleigh típicamente son emitidos en la dirección de avance bastante cerca de la trayectoria del fotón incidente. K, L y M son capas electrónicas. Figura tomada de [31].

3.3.2 Dispersión Compton

La dispersión Compton (también llamada dispersión inelástica o no clásica) es la interacción predominante de fotones de rayos X y de rayos gamma dentro el rango de energía de diagnóstico con el tejido blando (0.015-15 MeV). De hecho, la dispersión de Compton no sólo predomina en el rango de energía de diagnóstico superior a 26 keV en el tejido blando, sino que continúa predominando mucho más allá de las energías de diagnóstico a aproximadamente 30 MeV. Esta interacción es más probable que ocurra entre los fotones y electrones de la capa exterior (“valencia”) (Fig. 3.5). El electrón es expulsado del átomo, y el fotón es dispersado con alguna reducción en la energía.

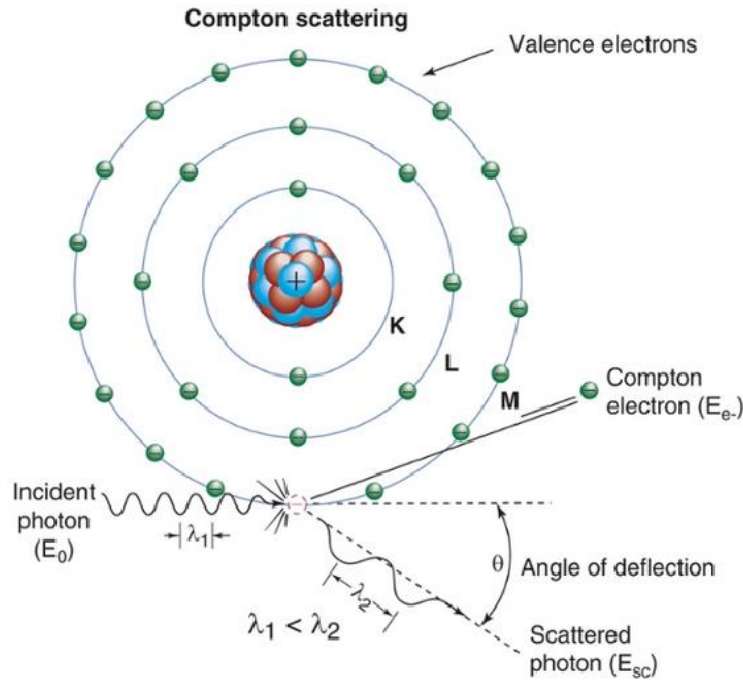


Fig. 3.5 Efecto Compton. Figura tomada de [31].

Como con todos los tipos de interacciones, tanto la energía como el momento deben ser conservados. Así, la energía del fotón incidente (E_0) es igual a la suma de la energía del fotón dispersado (E_{sc}) y la energía cinética del electrón expulsado (E_{e-}), como se muestra en la Ec. 3.2. La energía del enlace del electrón que fue expulsado es relativamente pequeña y puede ser ignorada.

$$E_0 = E_{sc} + E_{e-} \quad (3.2)$$

La dispersión de Compton resulta en la ionización del átomo y una división de la energía del fotón incidente entre el fotón dispersado y el electrón expulsado. El electrón expulsado perderá su energía cinética a través de la excitación y la ionización de los átomos en el material circundante. El fotón dispersado de Compton puede atravesar el medio sin interacción o puede experimentar interacciones subsecuentes tales como dispersión de Compton, la absorción fotoeléctrica, o la dispersión de Rayleigh.

La energía del fotón dispersado puede calcularse a partir de la energía del fotón incidente y el ángulo del fotón dispersado (con respecto a la trayectoria incidente) mediante la siguiente ecuación:

$$E_{sc} = \frac{E_0}{1 + \frac{E_0}{511 \text{ keV}} (1 - \cos \theta)} \quad (3.3)$$

donde E_{sc} es la energía del fotón dispersado, E_0 es la energía del fotón incidente, y θ es el ángulo de fotón dispersado.

En la transmisión de imágenes de rayos X, estos fotones son mucho más propensos a ser detectados por el receptor de imagen, reduciendo así el contraste de imagen. Cuando la dispersión de Compton se produce a energías de rayos X más bajas (Fig. 3.6) utilizadas en el diagnóstico por imágenes (18 a 150 keV), la mayoría de la energía del fotón incidente es transferida al fotón dispersado, que, si es detectado por el receptor de imagen, contribuye a la degradación de la imagen, reduciendo las diferencias de atenuación de fotones primarios de los tejidos. Por ejemplo, siguiendo la interacción de Compton de un fotón de 80 keV, la energía mínima del fotón dispersado es de 61 keV. Por lo tanto, incluso con la pérdida de energía máxima, los fotones dispersados tienen relativamente energías altas y una penetrabilidad del tejido.

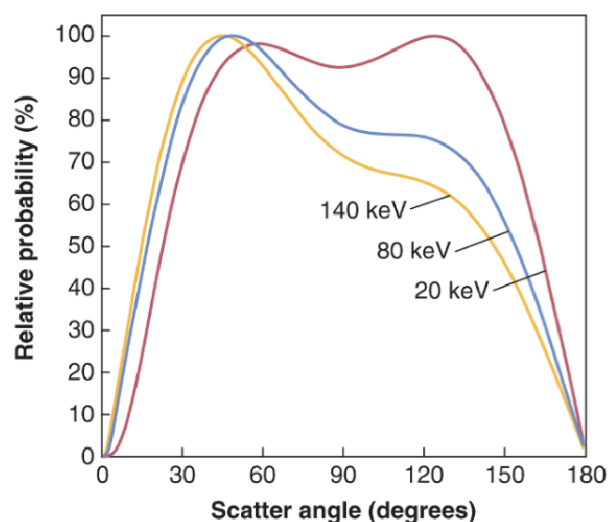


Fig. 3.6 Gráfico del Gráfico que ilustra la probabilidad de Compton relativa como una función del ángulo de dispersión para fotones de 20, 80 y 140 keV en el tejido. Cada curva está normalizada a 100%. Figura tomada de [31].

Las leyes de la conservación de energía y momento establecen límites al ángulo de dispersión y transferencia de energía. Por ejemplo, la transferencia de energía máxima para el electrón Compton (y por lo tanto la reducción máxima de energía del fotón incidente) se produce con una retrodispersión de fotones (desviación en su dirección en ángulos de 180 grados). De hecho, la energía máxima del fotón dispersado está limitada a 511 keV en una dispersión de 90 grados y a 255 keV para un evento de dispersión de 180 grados (retrodispersión). Estos límites a la energía del fotón dispersado se mantienen incluso para fotones con energía extremadamente alta (por ejemplo, el rango de energía terapéutica). El ángulo de dispersión del electrón expulsado no puede exceder de 90 grados, mientras que el del fotón dispersado puede ser cualquier valor incluyendo una retrodispersión de 180 grados. En contraste con el fotón dispersado, la energía del electrón expulsado generalmente es absorbida cerca del sitio de dispersión.

La energía del fotón incidente debe ser sustancialmente mayor que la energía de enlace del electrón para que una interacción de Compton tenga lugar. Por tanto, la probabilidad de una interacción Compton se incrementa, en comparación con la dispersión de Rayleigh o la absorción fotoeléctrica, ya que la energía del fotón incidente aumenta. La probabilidad de la interacción de Compton también depende de la densidad electrónica. Con la excepción del hidrógeno, el número total de electrones/g es bastante constante en el tejido; por lo tanto, la probabilidad de dispersión de Compton por unidad de masa es prácticamente independiente de Z , y la probabilidad de dispersión de Compton por unidad de volumen es aproximadamente proporcional a la densidad del material. En comparación con otros elementos, la ausencia de neutrones en el átomo de hidrógeno resulta en una duplicación

aproximada de la densidad electrónica. Así, los materiales hidrogenados tienen una mayor probabilidad de dispersión de Compton de un material no hidrogenado de igual masa.

3.3.3 Efecto Fotoeléctrico

En el efecto fotoeléctrico, toda la energía del fotón incidente se transfiere a un electrón, que es expulsado del átomo. La energía cinética del fotoelectrón expulsado (E_e) es igual a la energía del fotón incidente (E_0) menos la energía de enlace del electrón orbital (E_b) (Fig. 3.7-Izquierda).

$$E_e = E_0 - E_b \quad (3.4)$$

A fin de que ocurra la absorción fotoeléctrica, la energía del fotón incidente debe ser mayor que o igual a la energía de enlace del electrón que es expulsado. El electrón expulsado es muy probablemente uno cuya energía de enlace es más cercana a, pero inferior a, la energía del fotón incidente. Por ejemplo, para los fotones cuyas energías exceden la energía de enlace de la capa K , las interacciones fotoeléctricas con electrones de la capa K son más probables. Después de una interacción fotoeléctrica, el átomo es ionizado, con una carcasa interior de vacantes de electrones. Esta vacante será ocupada por un electrón de una capa con una menor energía de enlace. Esto crea otra vacante, que, a su vez, es llenada por un electrón de una capa de energía de enlace aún más baja. Por lo tanto, una cascada de electrones desde capas exteriores hasta capas internas tiene lugar. La diferencia en la energía de enlace es liberada ya sea como rayos X característicos o electrones Auger. La probabilidad de la emisión de rayos X característicos decrece a medida mientras el número atómico del material absorbedor disminuye y por lo tanto no ocurre con frecuencia para energía de diagnóstico de interacciones de fotones en el tejido blando.

La probabilidad de absorción fotoeléctrica por unidad de masa es aproximadamente proporcional a Z^3/E^3 , donde Z es el número atómico y E es la energía del fotón incidente. Por ejemplo, la probabilidad de interacción fotoeléctrica en yodo ($Z = 53$) es $(53/20)^3$ o 18.6 veces mayor que en calcio ($Z = 20$) para el fotón de una determinada energía.

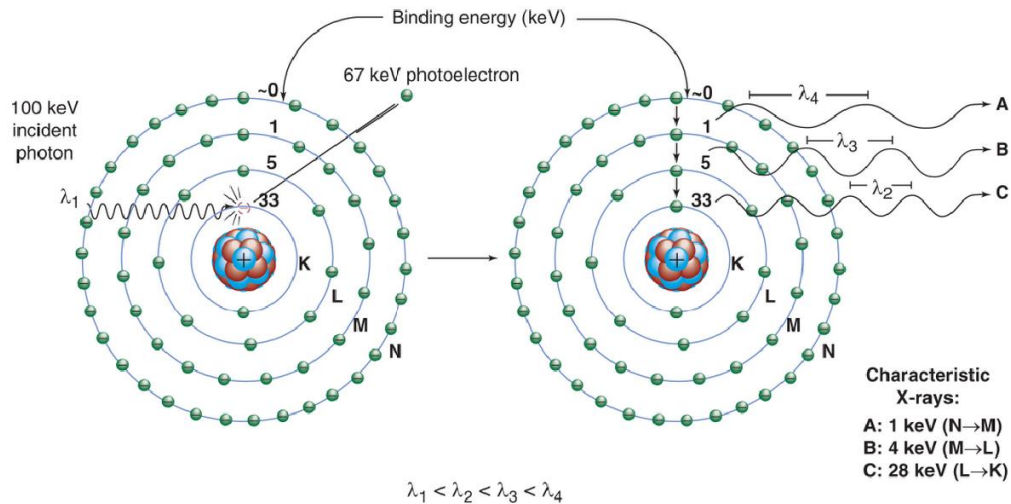


Fig. 3.7 Absorción fotoeléctrica. Izquierda: El diagrama muestra un fotón de 100 keV experimentando absorción fotoeléctrica con un átomo de yodo. En este caso, el electrón de la capa K es expulsado con una energía cinética igual a la diferencia entre la energía del fotón incidente y la energía de enlace de la capa K de 34 o 66 keV. Derecha: La vacante creada en la capa K resulta en la transición de un electrón de la capa L a la capa K. La diferencia en sus energías de enlace, (es decir, 34 y 5 keV), resulta en un rayo X característico de K_{α} de 29 keV. Esta cascada electrónica continuará dando como resultado la producción de otros rayos X característicos de energías más bajas. Tenga en cuenta que la suma de las energías de rayos X característicos es igual a la energía de enlace de los fotoelectrones expulsados. Aunque no se muestra en este diagrama, los electrones Auger de diversas energías podrían ser emitidos en lugar de las emisiones de rayos X característicos. Figura tomada de [31].

El beneficio de la absorción fotoeléctrica en la transmisión de imágenes de rayos X es que no hay fotones no primarios adicionales para degradar la imagen. El hecho de que la probabilidad de interacción fotoeléctrica es proporcional a $1/E^3$ explica, en parte, por qué el contraste de la imagen disminuye cuando se utilizan energías de rayos X más altas en el proceso de formación de imágenes. Si la energía de los fotones es duplicada, la probabilidad de interacción fotoeléctrica se reduce de manera óctuple: $(1/2)^3 = 1/8$.

Aunque la probabilidad de que el efecto fotoeléctrico disminuya, en general, al aumentar la energía del fotón, existe una excepción. Para cada elemento, una gráfica de la probabilidad del efecto fotoeléctrico, como una función de la energía fotónica, presenta

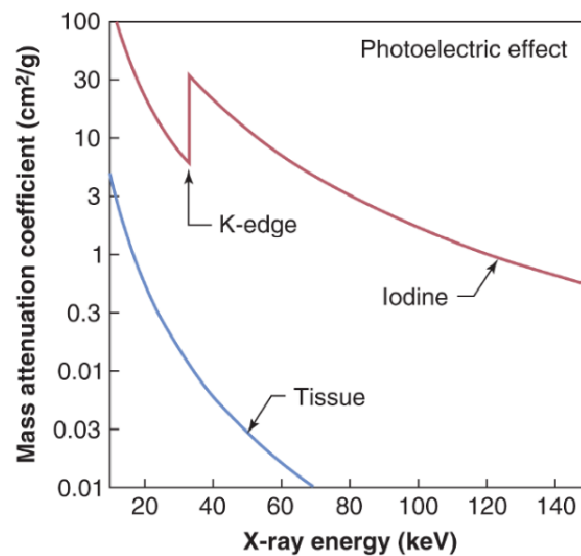


Fig. 3.8 Coeficientes de atenuación de masa fotoeléctricos para el tejido ($Z = 7$), yodo ($Z = 53$), y bario ($Z = 56$) como una función de la energía. El abrupto incremento de los coeficientes de atenuación llamado “bordes de absorción” se producen debido al incremento de la probabilidad de absorción fotoeléctrica cuando la energía del fotón sólo excede la energía de enlace de los electrones de la capa interior (por ejemplo, K, L, M...) por lo tanto aumentando el número de electrones disponibles para la interacción. Este proceso es muy importante para materiales con alto Z , como yodo y bario, en el rango de energía de diagnóstico. Figura tomada de [31].

discontinuidades agudas llamadas bordes de absorción (Fig. 3.9). La probabilidad de la interacción para los fotones de energía justo por encima de un borde de absorción es mucho mayor que la de los fotones de energía ligeramente por debajo del borde. Por ejemplo, un

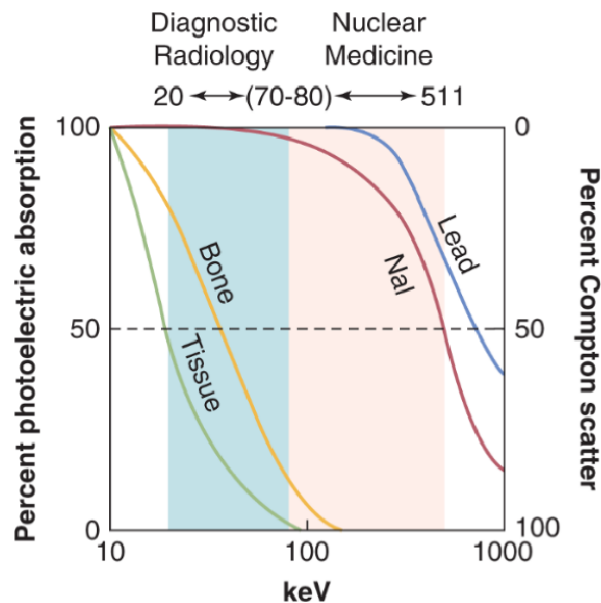


Fig. 3.9 Gráfica del porcentaje de contribución de los procesos de atenuación fotoeléctrico (escala izquierda) y Compton (escala derecha) para diferentes materiales como una función de la energía. Cuando los fotones de energía de diagnóstico (es decir, energía efectiva de rayos X de diagnóstico de 20 a 80 keV; fotones de formación de imágenes de medicina nuclear de 70 a 511 keV), interactúan con materiales de bajo número atómico (por ejemplo, tejido blando), el proceso Compton domina. Figura tomada de [31].

fotón de rayos X de 33.2 keV es casi seis veces más probable de tener una interacción fotoeléctrica con un átomo de yodo que un fotón 33.1 keV.

Como se mencionó anteriormente, un fotón no puede experimentar una interacción fotoeléctrica con un electrón en una capa atómica en particular o subcapa si la energía del fotón es menor que la energía de enlace de esa capa o subcapa. Esto provoca la dramática disminución en la probabilidad de absorción fotoeléctrica para los fotones cuyas energías están justo por debajo de la energía de enlace de una capa. Por lo tanto, la energía del fotón correspondiente a un borde de absorción es la energía de enlace de los electrones en esa capa o subcapa particular. Un borde de absorción es designado por una letra, que representa la capa atómica de los electrones, seguido de un número que denota la subcapa (por ejemplo, K, L1, L2, L3, etc).

La energía de los fotones correspondiente a un borde de absorción en particular se incrementa con el número atómico (Z) del elemento. Por ejemplo, los elementos primarios que comprenden el tejido blando (H, C, N, y O) tienen bordes de absorción por debajo de 1 keV. Los elementos de yodo ($Z=53$) y bario ($Z=56$), usados comúnmente en agentes de contraste radiográfico para proporcionar una mayor atenuación de los rayos X, tienen K bordes de absorción de 33.2 y 37.4 keV respectivamente (Fig. 3.8). La energía del K-borde de plomo ($Z=82$) es de 88 keV.

Para energías de fotones por debajo de 50 keV, el efecto fotoeléctrico juega un papel importante en la formación de imágenes de tejidos blandos. El proceso de absorción fotoeléctrico puede ser utilizado para amplificar las diferencias en la atenuación entre los tejidos con números atómicos ligeramente diferentes, mejorando así contraste de la imagen. Esta absorción diferencial es explotada para mejorar el contraste de la imagen en la selección de material del blanco del tubo de rayos X y filtros en la mamografía y en el uso de fósforo que contienen elementos de tierras poco comunes (por ejemplo, lantano y

gadolinio) en la intensificación de las pantallas. El proceso fotoeléctrico predomina cuando los fotones de menor energía interactúan con materiales con un Z alto (Fig. 3.9). De hecho, la absorción fotoeléctrica es el principal modo de interacción de los rayos X de diagnóstico con pantallas de fósforo, materiales de contraste radiográfico, y hueso. Por el contrario, la dispersión Compton predominará en la mayoría de las energías de diagnóstico de los fotones en los materiales con un número atómico inferior, tal como el tejido y el aire.

3.3.4 Producción de pares

La producción de pares sólo puede ocurrir cuando las energías de los rayos X exceden los 1.02 MeV. En la producción de pares, un rayo X interactúa con el campo eléctrico del núcleo de un átomo. La energía del fotón es transformada en un par electrón-positrón (Fig. 3.10-A). La energía de la masa en reposo equivalente de cada electrón es 0.511 MeV y por eso el umbral de energía para esta reacción es 1.02 MeV. La energía del fotón por encima de este umbral es impartida a los electrones como energía cinética. El electrón y el positrón pierden su energía cinética vía la excitación e ionización. Como se discutió previamente, cuando el positrón queda en reposo, interactúa con un electrón cargado negativamente, resultando en la formación de dos fotones de aniquilación con direcciones opuestas de 0.511 MeV (Fig. 3.10-B).

La producción de pares no tiene ninguna consecuencia en las imágenes de rayos X de diagnóstico debido a las extremadamente altas energías necesarias para que esto ocurra. De hecho, la producción de pares no se hace significativa a menos que las energías de fotones superen ampliamente el umbral de energía de 1.02 MeV.

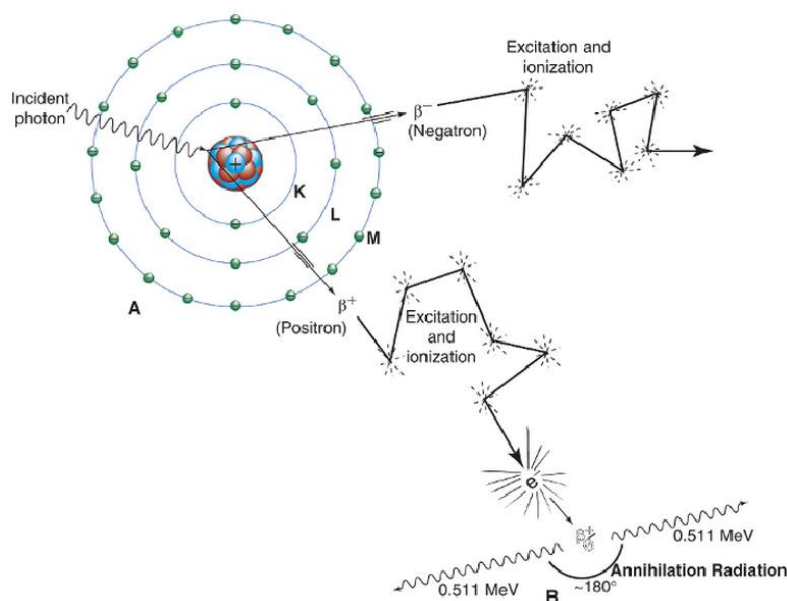


Fig. 3.10 La producción de pares. A: Diagrama ilustra el proceso de producción de pares en el que un fotón de alta energía incidente, bajo la influencia del núcleo atómico, se convierte en el par de materia y antimateria. Ambos electrones (positrón y negatrón) gastan su energía cinética por la excitación y la ionización en la materia que atraviesan. B: Sin embargo, cuando el positrón queda en reposo, se combina con un electrón produciendo los dos fotones de radiación de aniquilación de 511 keV. K, L y M son capas electrónicas. Figura tomada de [31].

Capítulo 4

Algoritmos de reconstrucción de imágenes médicas

4.1 Introducción

La tecnología de imágenes médicas no solo se basa en dispositivos mecánicos y sistemas electrónicos como detectores de estado sólido y emisores de rayos X. Durante décadas, el procesamiento de datos y específicamente los algoritmos de reconstrucción han estado en el centro de la investigación y el desarrollo aún en curso. En este capítulo se presenta un panorama general del campo de los algoritmos de reconstrucción y una posible clasificación en diferentes grupos.

4.2 Reconstrucción en general

En TC, la información sobre la imagen original se da en forma de un conjunto de proyecciones tomadas bajo diferentes ángulos, que se conoce como la transformada de Radon de la imagen, y la inversa de la transformada, representa la imagen. Considérese una rebanada axial que se divide en vóxeles con resolución espacial Δx , Δy , Δz , donde a cada vóxel puede asignarse una atenuación efectiva μ . Considérese también un rayo de intensidad I_0 , que penetra un objeto a lo largo de una trayectoria L , en línea recta, pasando

por cada vóxel (o región discretizada del objeto) con una distribución no homogénea de atenuaciones $\mu(x)$. La intensidad del rayo que alcanza el detector $I(x)$ depende no sólo de la distancia atravesada x sino también de la atenuación $\mu(x)$ de cada punto en su trayectoria, obedeciendo la ley de Beer-Lambert:

$$I(x) = I_0 \exp\left(-\int_L \mu(x) dx\right) \quad (4.1)$$

Dado que es posible medir tanto I_0 como la intensidad $I(x)$ en el detector del tomógrafo, resulta conveniente reescribir a la ecuación 4.1 de la siguiente forma:

$$p(x) = -\ln\left(\frac{I(x)}{I_0}\right) = -\int_L \mu(x) dx \quad (4.2)$$

El resultado de esta ecuación provee la proyección $p(x)$, y tiene implicaciones importantes. La primera, que el detector registra la integral de línea y esta depende de las atenuaciones en cada región del objeto en la trayectoria del rayo (regiones que fueron divididas en vóxels). La segunda, que, aunque se usa información volumétrica (vóxels), el detector registra la proyección $p(x)$, que es una señal unidimensional para cada ángulo θ (Fig. 4.2-a y 4.2-b). Dichas proyecciones usualmente se guardan en una matriz $p(x, \theta)$, que constituye el sinograma (Fig. 4.2-c).

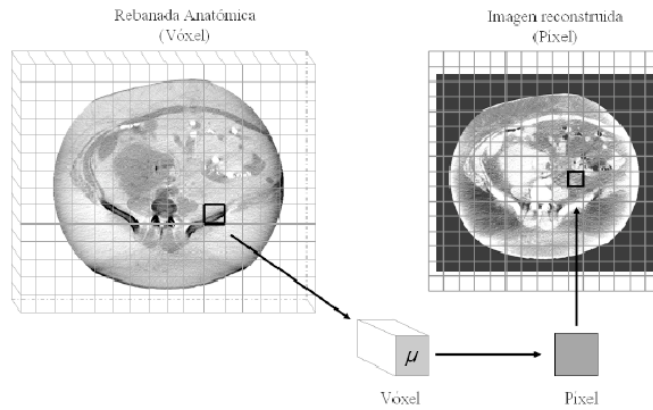


Fig. 4.1 Proceso de adquisición de la imagen. Una rebanada tridimensional, artificialmente discretizada en voxels, es proyectada a través de integrales de línea (una dimensión). Un proceso de reconstrucción proyecta los resultados en una imagen bidimensional, representada por píxeles que contienen la atenuación efectiva de cada vóxel evaluado [31].

Tras la reconstrucción, se obtiene una imagen bidimensional, donde cada píxel tiene por intensidad el valor estimado de su atenuación μ . Finalmente, nótese que se asume que cada vóxel contiene una atenuación uniforme (que corresponde a un tejido específico), lo

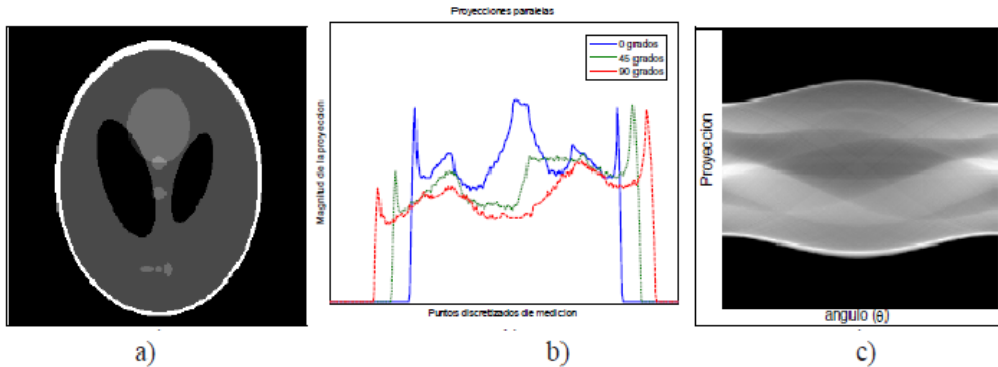


Fig. 4.2 Simulación del proceso de adquisición de imágenes por tomografía utilizando Matlab\textregistered 2008. (a) Shepp-Logan phantom de 256×256 píxeles. (b) perfiles de las proyecciones sobre 0° , 45° y 90° . (c) sinograma resultante de la proyección del objeto con pasos $\Delta \theta = 0.5^\circ$ con θ variando desde 0° a 180° [31].

cual no es necesariamente cierto, ya que es muy probable que algunos vóxeles contengan dos, o incluso más, materiales simultáneamente, especialmente en los bordes o interfaces entre tejidos. Este fenómeno es llamado el efecto del volumen parcial, que en algunas aplicaciones específicas (que no se discutirán aquí), debe ser corregido.

El problema de la reconstrucción de la imagen consiste en asignar la atenuación μ adecuada para cada vóxel que se utilizó para discretizar el objeto, dadas las proyecciones $p(x, \theta)$.

Dar una visión general de los algoritmos de reconstrucción de imágenes médicas es una tarea bastante compleja ya que el campo de investigación en el procesamiento de imágenes médicas siempre está en desarrollo y año tras año se introducen algoritmos nuevos o modificados. Por lo tanto, nos enfocamos en métodos bien conocidos y/o de uso comercial. Los algoritmos de reconstrucción son divididos en tres grupos diferentes donde los dos últimos grupos pertenecen a una clase diferente de algoritmos llamados métodos iterativos. La descripción general también se ilustra en la Figura 4.3.

El primer grupo cubre los métodos de reconstrucción analíticos. La retroproyección filtrada (FBP) es una aplicación directa de la teoría de las transformaciones integrales de radón en el caso de la TC de primera generación que proporciona una geometría de haz paralelo en 2-D. El teorema de la rebanada de Fourier [7] como un caso especial de la transformación integral de Radon es la base de muchos métodos de reconstrucción analítica.

Los métodos de reconstrucción analíticos están muy extendidos y todavía se usan ampliamente en muchos campos debido a sus beneficios de rendimiento computacional. Sin embargo, también tienen algunos inconvenientes, por ejemplo, necesitan una colección de datos completa, lo que no siempre es posible y no proporcionan una reconstrucción óptima cuando la señal contiene ruido [34].

La medición del ruido generalmente se ignora en la formulación del problema de los métodos analíticos. Con el fin de tratar los problemas relacionados con el ruido, a menudo se aplican diferentes tipos de filtrado previo y posterior. Al igual que el Teorema tradicional de Fourier, las formulaciones analíticas asumen mediciones continuas y proporcionan una solución de forma cerrada utilizando integrales continuas. Al discretizar estas soluciones, los problemas de muestreo, por ejemplo, en el espacio de Fourier, se tratan y compensan mediante el procesamiento previo y posterior o durante la interpolación en el detector. Otro inconveniente es que los métodos analíticos requieren ciertas geometrías estándar, por ejemplo, una adquisición circular con muestreo completo en dirección radial y ángulos equidistantes.

Esto motiva a la segunda clase de algoritmos de reconstrucción, a saber, métodos iterativos, que pueden superar estos inconvenientes. Dividimos la segunda clase en dos grupos diferentes. Primero se presenta el grupo de métodos de reconstrucción algebraica y segundo los métodos de reconstrucción estadística.

El método más antiguo de reconstrucción algebraica es la Técnica de Reconstrucción Algebraica (ART) [35], que también fue utilizada por Hounsfield para el primer prototipo de TC [36].

El ART fue propuesto por Gordon, Bender y Herman en [35] como un método para la reconstrucción de objetos tridimensionales a partir de fotografías de rayos X o escaneos microscópicos electrónicos. En lugar de describir un sistema continuo y luego discretizarlo para su aplicación, el ART define el problema como una tarea de álgebra lineal. Lo más simplificado se puede describir como una matriz A y un vector conocido d_μ , donde se debe calcular el vector desconocido μ , de modo que:

$$A \cdot \mu = d_\mu \quad (4.3)$$

Para matrices A de tamaño moderado (del orden de $A_{100 \times 100}$) e invertible, el sistema se puede resolver fácilmente utilizando métodos directos. Sin embargo, debido al enorme tamaño de la matriz del sistema en TC, los métodos directos no son aplicables.

En 1972 Gilbert et al. introdujeron la técnica de reconstrucción iterativa simultánea (SIRT) [37]. En lugar de corregir μ con respecto a cierta fila de la matriz A y una sola medición d_μ , su algoritmo realiza una corrección de μ para todas las mediciones d_μ en un solo paso de iteración. Por lo tanto, en cada iteración, el término de corrección para todas las mediciones debe calcularse y aplicarse en μ . Las correcciones realizadas son más pequeñas que para SART y aseguran el comportamiento de convergencia. Por lo tanto, este método proporciona una convergencia más lenta.

La siguiente mejora en la reconstrucción algebraica fue publicada por Andersen y Kak en 1984 con la introducción de la Técnica de Reconstrucción Algebraica Simultánea (SART) [38]. Sugirieron corregir iterativamente μ con respecto a mediciones múltiples d_μ en un paso de iteración. Los autores argumentaron que este método podría suprimir los artefactos de rayas que existían en las reconstrucciones usando ART.

Mientras que la técnica de reconstrucción iterativa simultánea, publicada en 1972 por Gilbert et al. - también puede suprimir estos artefactos, el método SART aún proporcionó

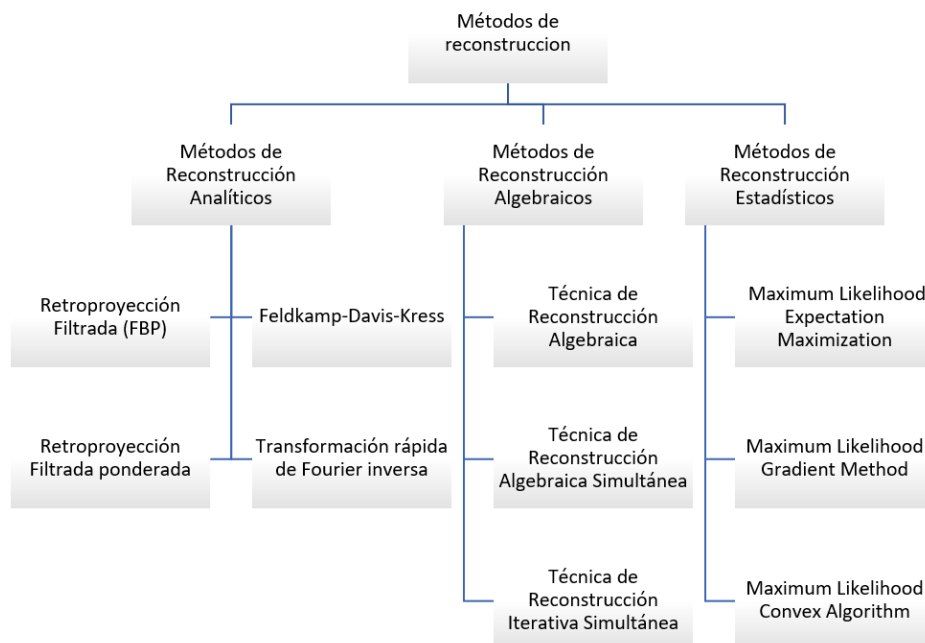


Fig. 4.3 Descripción general del campo de los métodos de reconstrucción en imágenes médicas con enfoque en tomografía computarizada [33].

algunas mejoras debido a una convergencia más rápida. Por lo tanto, la velocidad de convergencia de SART en general está justo entre el ART más rápido y el SIRT más lento.

La idea de realizar una corrección con respecto a un cierto número de ecuaciones, respectivamente, filas de A ya se usa en el método SART, pero también se puede mover al siguiente paso. Por ejemplo, en lugar de corregir μ para todas las mediciones de una proyección, se puede corregir para algunas proyecciones n al mismo tiempo. Para más detalles sobre los métodos de reconstrucción algebraica, consulte la literatura [39].

Los métodos de reconstrucción estadística como el segundo grupo en la clase de métodos iterativos están motivados modelando la física de la creación, absorción y detección de rayos X en un sistema matemático. Este enfoque estadístico permite superar las condiciones de adquisición subóptimas, por ejemplo, debido a un ángulo de adquisición limitado o datos muy ruidosos por dosis de radiación bajas.

Los datos limitados y los datos muy ruidosos son característicos de la tomografía por emisión, donde la reconstrucción estadística, por ejemplo, la reconstrucción de Máxima Verosimilitud (ML) [13] por Shepp et al, es ampliamente utilizada.

Lange et al. introdujo los algoritmos de reconstrucción de Maximización de Expectativas (EM) para la tomografía de emisión y transmisión [40] en 1984. Si bien se introdujeron muchos métodos de reconstrucción estadística para la tomografía de transmisión en las últimas décadas, todavía no se utilizan en este campo debido a su complejidad computacional.

Una introducción completa y una comparación de los tres métodos de reconstrucción estadística mencionados en la Figura 4.3, se pueden encontrar en el artículo [41] de Lange y Fessler sobre algoritmos de reconstrucción globalmente convergentes.

Esto completa la visión general de los métodos de reconstrucción médica en tres grupos diferentes, de los cuales los algoritmos seleccionados se presentan con más detalle en las siguientes secciones. El último grupo introducido de algoritmos de reconstrucción pertenece a la clase de métodos de reconstrucción iterativos.

Los métodos de reconstrucción estadística proporcionan diferentes beneficios, como modelar efectos físicos o incorporar estadísticas de medición. El inconveniente de estos métodos es que son más complicados y tienen altos costos computacionales. Por lo tanto, se utilizan principalmente en aplicaciones donde otros métodos no pueden cumplir con los requisitos clínicos. Los métodos de reconstrucción estadística se utilizan ampliamente en la tomografía por emisión y rara vez en la tomografía de transmisión. En comparación con la tomografía de transmisión, la tomografía de emisión generalmente se limita a una menor cantidad de datos de medición que está corrompida por características de ruido más altas.

En este trabajo nos centramos en la reconstrucción de máxima probabilidad para la tomografía de transmisión, este método es llamado MLEM.

En el trabajo de Lange et al. [42], se encuentra el algoritmo de reconstrucción del MLEM para la tomografía de emisión y transmisión. Los diferentes pasos analíticos o procesos seguidos en este algoritmo se presentan en la siguiente sección (ver Fig. 4.4) y podrían expresarse matemáticamente mediante la siguiente fórmula:

$$x_j^{r+1} = \frac{x_j^r}{\sum_{i=1}^M a_{ij}} \sum_{i=1}^M \frac{a_{ij} y_i}{\sum_{l=1}^N a_{il} x_l^r} \quad (4.4)$$

Aquí x_j^r es el píxel j de la imagen x en la iteración r -ésima, y_i es el intervalo i de la proyección medida dada por el escáner TC, a_{ij} es la matriz del sistema $M \times N$ cuyos coeficientes conectan los valores x_j de la imagen con las proyecciones y_i y describe la probabilidad de detectar un fotón en el píxel j por el bin i .

Todos los algoritmos desarrollados en este estudio requirieron de una matriz de sistema a_{ij} que se usa en la resolución del algoritmo MLEM dado en 4.4. La creación de la matriz del sistema a_{ij} depende del número de líneas de proyección y_i , del número de ángulos de proyección y del tamaño de la imagen a reconstruir $n \times n$. Esta matriz puede ser calculada de diferentes maneras y en este proyecto se utilizaron los desarrollos de [43,44,45] para generar y simularla en MatLab.



Fig. 4.4 Pasos del algoritmo MLEM [46].

Capítulo 5

Métricas de calidad para imágenes médicas

5.1 Introducción

En el desarrollo de algoritmos de procesamiento de imágenes, la medición de calidad de imagen juega un papel importante para evaluar el rendimiento de la imagen procesada.

La calidad de imagen se define como una característica de la imagen que mide su degradación comparándola con una imagen ideal.

Las técnicas de evaluación de la calidad de la imagen se dividen en técnicas subjetivas y técnicas objetivas [47]. La evaluación por técnicas subjetivas es la que presenta una mejor correlación con el sistema visual humano. Sin embargo, es un método basado en el análisis estadístico sobre la apreciación que diferentes observadores realizan sobre un conjunto a evaluar, por lo que es un método con un procesamiento lento, costoso en tiempo y en recursos.

La evaluación mediante técnicas objetivas pretende desarrollar fórmulas matemáticas que permitan obtener un valor numérico sobre la degradación obtenida, y que éste presente una elevada correlación con la valoración visual humana.

Dada la simplicidad de este método y las ventajas respecto a los métodos subjetivos en cuanto a mayor rapidez y menor necesidad de recursos, el desarrollo de técnicas objetivas que proporcionen una adecuada valoración según el sistema visual humano es un aspecto muy demandado.

Estas métricas pueden clasificarse en función de la disponibilidad de la imagen de referencia, como referencia completa, referencia parcial y sin referencia. La referencia

completa indica que la imagen original está disponible para calcular la calidad de la imagen degradada y reconstruida, mientras que la referencia parcial representa la disponibilidad de información parcial sobre la imagen original.

Del mismo modo, ninguna categoría de referencia significa métricas de calidad que pueden evaluarse sin la imagen original. Esta categoría también se conoce como tipo de referencia ciega de métricas de calidad.

Las métricas de calidad de referencia completa también se clasifican en diferentes clases como:

- métricas basadas en diferencia de píxeles,
- métricas basadas en similitud estructural,
- métricas basadas en correlación,
- métricas basadas en contenido espectral y
- métricas basadas en el sistema visual humano.

La evaluación de la calidad de las imágenes médicas, las imágenes de rayos X, las imágenes de ultrasonido, las imágenes de resonancia magnética, etc. es un trabajo crucial en comparación con el de las imágenes fotográficas generales.

En el caso de las imágenes de rayos X, se debe dar alta preferencia a la información de borde al decidir la calidad de esas imágenes.

Las imágenes de ultrasonido están corrompidas principalmente por el ruido moteado (propiedad inherente de la ecografía médica y debido a esto, la resolución de la imagen y el contraste se reducen, lo que afecta el valor diagnóstico de esta modalidad de imagen [48]), que oculta las lesiones y otra información estructural importante en la imagen. La mayoría de los algoritmos de eliminación de ruido eliminan el ruido moteado de las imágenes de ultrasonido a costa del suavizado. A veces, los médicos prefieren una imagen de ultrasonido ruidosa que una imagen suavizada.

Entonces, determinar la calidad de las imágenes de ultrasonido es en sí mismo un trabajo difícil.

5.2 Parámetros de evaluación de calidad de imagen

Algunas métricas de calidad básicas populares son el error cuadrático medio (MSE), la señal pico a la relación de ruido (PSNR), la relación señal a ruido (SNR), etc. Podemos agregar esta lista con error absoluto normalizado (NAE), diferencia promedio (AD), diferencia máxima (MD), etc.

Todas estas métricas de calidad básicas podrían estar basadas en la diferencia de píxeles. Estos son más populares debido a su simplicidad matemática. MSE y PSNR son proporcionales a la energía de la distorsión/ruido y se basan en los valores digitales de las imágenes que en la luminancia física real [49]. Debido a esta razón, estas métricas de calidad difieren más de la percepción humana.

Para superar las limitaciones de las métricas de calidad basadas en la diferencia de píxeles, se introdujeron métricas de calidad basadas en la evaluación estructural, como el índice de similitud estructural (SSIM) [50].

5.2.1 Selección de métricas de calidad

En el caso de las imágenes de rayos X debido a su proceso de formación contiene ruido de Poisson que depende de la señal por lo que hemos seleccionado las siguientes métricas de calidad para el análisis de las imágenes reconstruidas.

5.2.2 Error cuadrático medio

La métrica de calidad más simple es el Error cuadrático medio (MSE por sus siglas en inglés) [51]. La definición matemática de la métrica de calidad se da a continuación:

$$MSE = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N \left(I_{Est}(i, j) - I_{Ref}(i, j) \right)^2 \quad (5.1)$$

dónde I_{Ref} denota la imagen de referencia de dimensión MN y la imagen estimada se denomina I_{Est} . El valor de MSE mide la diferencia entre una imagen procesada y una imagen de referencia. Un valor grande de MSE significa que la imagen es de baja calidad y el valor más pequeño del MSE representa el mejor resultado.

5.2.3 PSNR

La Proporción Máxima de Señal a Ruido o PSNR (por sus siglas en inglés) [52] es uno de los parámetros métricos más utilizados. Describe la relación entre la potencia máxima posible de una señal y la potencia del ruido corruptor y normalmente se representa en una escala de decibelios. PSNR se puede expresar de la siguiente manera:

$$PSNR = 10 \log_{10} \left(\frac{I_{Max}^2}{MSE} \right) \quad (5.2)$$

dónde I_{Max} es el valor de intensidad máxima presente en la imagen de referencia y MSE es el error cuadrático medio entre la imagen de referencia y la estimada. SNR y PSNR generalmente se miden en decibelios (dB).

5.2.4 SSIM

El Índice de similitud estructural (SSIM por sus siglas en inglés) [53] es una métrica de calidad de imagen que evalúa el impacto visual de tres características de una imagen: luminancia (l), contraste (c) y correlación estructural (s). Por lo que SSIM es una combinación multiplicativa de estos tres términos.

$$SSIM(x, y) = [l(x, y)]^\alpha \cdot [c(x, y)]^\beta \cdot [s(x, y)]^\gamma \quad (5.3)$$

dónde,

$$\begin{aligned} l(x, y) &= \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1}, \\ c(x, y) &= \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2}, \\ s(x, y) &= \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3} \end{aligned} \quad (5.4)$$

dónde,

μ_x es la media de la imagen de referencia, μ_y es la media de la imagen estimada, σ_x es la desviación estándar de la imagen de referencia, σ_y es la desviación estándar de la imagen estimada y σ_{xy} es la correlación cruzada entre la imagen de referencia y la estimada.

5.3 Métricas en la implementación MLEM

A fin de evaluar la calidad de las imágenes reconstruidas se utilizaron distintos parámetros de calidad. Los parámetros estudiados fueron: La relación Contraste-Ruido (CNR por sus siglas en inglés) y relación Señal-Ruido (SNR por sus siglas en inglés) [54]. Estas indican qué tan bien podrían diferenciarse dos tejidos vecinos (CNR) y qué tan fuerte era una señal con respecto al ruido de fondo (SNR). Otros parámetros de calidad de imagen evaluados fueron: MSE, PSNR y SSIM. La MSE proporciona una medición de las diferencias entre una imagen reconstruida en comparación con la imagen de referencia. Cuanto menor sea el valor de MSE, mejor será el resultado. PSNR describe la relación de la potencia máxima posible de una señal con la potencia de la corrupción del ruido. Se representa en una escala de decibels y un alto valor de PSNR significa que la reconstrucción es de alta calidad. Por último, SSIM se usa para medir la similitud entre dos imágenes (imagen TC original y las imágenes reconstruidas). Se normaliza entre 0 y 1, siendo 1 la situación en la que ambas imágenes son iguales.

Capítulo 6

Implementación del algoritmo MLEM en la GPU

6.1 Introducción

Teniendo en cuenta los antecedentes técnicos y teóricos de nuestra investigación en los capítulos y 4, en este capítulo se muestra la principal contribución de este trabajo en el campo de la reconstrucción de imágenes médicas.

6.2 Enfoque de reconstrucción de imágenes médicas

La dificultad en la programación de GPU que usa CUDA por un lado es la paralelización del problema, ya que el algoritmo debe dividirse en muchos sub-algoritmos, en el mejor de los casos independientes, con cálculos similares. El desafío por el otro lado es comprender el hardware subyacente y modificar o seleccionar un esquema de paralelización de tal manera que se puedan lograr altas tasas de cómputo (elevado número de cálculos/segundo), y por tanto, un alto rendimiento.

Para un esquema de realización de alto rendimiento, es esencial diseccionar de manera óptima este algoritmo, de modo que los cálculos se ejecuten de manera paralela con un uso óptimo de los recursos computacionales.

Revisando la regla de actualización del algoritmo MLEM en la ecuación 4.4:

$$x_j^{r+1} = \frac{x_j^r}{\sum_{i=1}^M a_{ij}} \sum_{i=1}^M \frac{a_{ij} y_i}{\sum_{l=1}^N a_{il} x_l^r} \quad (6.1)$$

muestra los cálculos requeridos a realizarse.

Una iteración de la reconstrucción MLEM se realiza en los siguientes pasos:

- **Datos de Entrada:** En un primer paso, se carga los archivos proporcionados por el escáner TC. Esto incluye la matriz del sistema a_{ij} y las proyecciones medidas y_i . Estas estructuras se denominaron Matrix y Vector respectivamente. La matriz a_{ij} y el vector y_i proporcionaron información como el número de filas y columnas de la matriz del sistema o el número de elementos del vector proyección y_i .
- **Estimación Inicial de Imagen.** Para reducir el número de iteraciones, se utilizó una estimación inicial de la imagen: x_l^0 . Fue creada ya sea con todos sus valores igual a la media de los datos de proyección, o mediante la fijación de sus valores a una intensidad de entre 1 y 255 (escala de grises).
- **Proyección hacia Adelante.** La primera iteración se realizó en la imagen con una estimación inicial que simulaba las proyecciones de un x_j^0 dado. Es requerido el cálculo de $y_i^{simulated} = \sum_l a_{il} x_l^0$ para determinar los valores de proyección correspondientes.
- **Comparación con los Datos Medidos.** El paso de comparación se realizó dividiendo los datos de proyección originales y_i con los datos de proyección simulados $y_i^{simulated}$, obteniendo una relación $Y_i = \frac{y_i}{y_i^{simulated}}$. Aquí Y_i formó un factor de corrección multiplicativo para cada proyección.
- **Proceso de Retroproyección.** La Y_i calculada fue retro proyectada para obtener un factor de corrección X_j para la imagen inicial. Esto se hizo con el cálculo de un producto del sistema $X_j = \sum_{i=1}^M a_{ij} Y_i$.
- **Normalización.** El factor de corrección X_j se normalizó antes de multiplicarlo por x_j^0 . Se dividió por un término de ponderación basado en el modelo del sistema para aplicar la intensidad deseada de cada factor de corrección de la imagen. Esto produjo $\epsilon_j = X_j / \sum_{i=1}^M a_{ij}$, donde ϵ_j representa el factor de corrección normalizado.
- **Actualización de la Imagen.** La estimación inicial de la imagen x_j^0 fue multiplicado por el factor de corrección normalizado ϵ_j obteniendo la nueva estimación de imagen x_j^1 que luego se reingresó al algoritmo en el paso 1.

La siguiente fase fue las implementaciones seriales del algoritmo MLEM en los lenguajes de programación C y MatLab (Serial-C y Serial-MatLab, respectivamente) para establecer:

- Los procesos que requieren una mayor cantidad de recursos del CPU.
- Procesos adecuados para paralelizar.
- Analizar la reconstrucción de las imágenes médicas en ambas implementaciones del algoritmo MLEM.

A partir de la ejecución de estas implementaciones seriales y la medición del rendimiento para cada uno de los pasos de reconstrucción se encontró lo siguiente:

- Los procesos que consumieron más recursos eran los pasos de Forward-Projection y Back-Projection. Entre ambos, utilizan aproximadamente el 75% del proceso computacional.
- Forward-Projection y Back-Projection resultaron ser los procesos más adecuados para paralelizar.
- Las reconstrucciones obtenidas por ambas implementaciones son prácticamente idénticas.

A partir de estos resultados preliminares fueron desarrolladas las rutinas Forward-Projection y Back-Projection en paralelo.

Se desarrollaron dos algoritmos: Paralelo usando GPU (Parallel-GPU) y Paralelo usando GPU, pero optimizando la transferencia de datos (Parallel-Opt). Ambos algoritmos resolvieron la ecuación 4.4, usando tanto C como CUDA. CUDA se empleó para desarrollar exclusivamente los procesos de Forward-Projection y Back-Projection, conocidos por ser los que más recursos consumen.

Las operaciones realizadas sobre las estructuras de datos fueron manejadas en C por cuatro kernels (Back-Projection, Forward-Projection, Normalization, Image Update), que distribuyeron los datos a los bloques de threads. Esto se hizo en semejanza al caso serial.

Los diferentes procesos de los algoritmos paralelos programados se describen completamente aquí:

Rutinas Forward-Projection y Back-Projection paralelas del algoritmo MLEM

Rutina Forward-Projection. Para este fin, se aplicó una rutina Forward-Projection basada en la biblioteca cuBLAS [55] de CUDA. La matriz a_{ij} fue cargada al Host y transferida al Device; Vector x_l^0 fue generado sobre el Host y también transferido al Device. La matriz a_{ij} fue particionada en grids de bloques de threads. Cada threads lee una fila de la matriz a_{ij} y realiza el cálculo con una columna del vector x_l^0 para obtener un vector resultante en el Device. El vector resultante $y_i^{simulated}$ es transferido de nuevo al Host.

Rutina Back-Projection. Como en el proceso anterior, la rutina de Back-Projection es lanzada sobre el Device. Esta rutina requirió la mayor parte de los recursos computacionales para todo el proceso de reconstrucción, ya que consistió en dos operaciones: Cálculo de la transpuesta de a_{ij} y su producto por el vector Y_l .

Optimización del tiempo de cómputo de la GPU. Las principales transferencias de datos entre el Host y el Device se realizaron cuando se ejecutaron los procesos de Proyección hacia adelante y de Retroproyección. Para mejorar el rendimiento de los algoritmos paralelos, la transferencia de datos entre la memoria RAM de la CPU y la memoria global de la GPU se manejó uniendo las rutinas de Forward-Projection, Comparación con los datos medidos y Back-Projection en una función, manteniendo los cálculos de datos en la GPU. Con ello se creó la versión Parallel optimized (Parallel-Opt).

Estos procesos se ejecutaron iterativamente hasta que el error relativo entre la imagen estimada y la real alcanzó un valor inferior al 5%.

En este trabajo para comparar las implementaciones del algoritmo MLEM y dos algoritmos de TIGRE se reconstruyeron dos imágenes clínicas TC junto con una imagen de 512×512 del fantasma Shepp-Logan (SLP) [56], (que consiste en una elipse grande que incluye elipses más pequeñas, representación del cerebro y sus características). La primera imagen clínica fue un corte axial del cerebro que cubre el cerebelo, así como las cavidades nasales y auditivas [57]. La segunda imagen clínica representa un corte axial abdominal a través de los pulmones, el corazón, el páncreas y el hígado [57]. Las matrices de ambas imágenes TC eran de 512×512 con un rango dinámico de 8 bits por píxel y una matriz

de sistema de 23944×262144 y vector proyección de 23944×1 para un ángulo de barrido de 5 grados.

La medición de los parámetros de calidad de imagen se completó para todas las imágenes TC y para las cuatro implementaciones del algoritmo que desarrollamos, así como las dos versiones de TIGRE. Adicionalmente consideramos dos escenarios en los que añadimos un 5% de ruido de Poisson a las imágenes bajo estudio a fin de evaluar el impacto del ruido en el tiempo y la calidad de la reconstrucción.

Con objeto de comparar las implementaciones desarrolladas en este trabajo para el algoritmo MLEM con los de un código externo (comercial), se consideró la biblioteca MatLab TIGRE [58]. A partir de ésta se utilizaron dos toolboxes de reconstrucción tomográfica iterativa basadas en GPUs: Primero el algoritmo OS-SART que se utiliza para resolver reconstrucciones de imágenes de TC de haz cónico mediante técnicas de reconstrucción algebraica simultánea con subconjuntos ordenados. Segundo, el algoritmo ASD-POCS que se basa en la proyección de descenso más pronunciado adaptativa en subconjuntos convexos proporcionando reconstrucciones de la imagen sobre conjuntos incompletos de datos de proyección [59].

Las implementaciones serial y paralela se desarrollaron y ejecutaron en una computadora portátil con un procesador Intel Core i7-5500U corriendo a 2.40 GHz y equipada con una tarjeta gráfica NVIDIA GeForce 840M con una memoria global de 2048 MB y 384 núcleos CUDA sobre un sistema operativo Ubuntu 15.04 [60]. Se empleó MATLAB R2017a [43] y GCC 4.9.2 [61] para la parte serial. CUDA C 7.5 [55] fue utilizado para la parte paralela. Para el cálculo de las mediciones de calidad de imagen se empleó la biblioteca Open Source Computer Vision (OpenCV 3.2.0) [62,63].

Capítulo 7

Resultados

En las Figuras 7.1 a 7.4 se muestran los resultados obtenidos para las cuatro implementaciones del algoritmo desarrollados en esta tesis (versión Serial-MatLab, Serial-C y 2 versiones paralelas) así como el software de prueba externa (dos versiones).

Las imágenes de la Fig. 7.2 corresponden a una reconstrucción con Parallel-OpT con un máximo de 1000 iteraciones. En la primera fila son presentados SLP y dos imágenes TC de 512×512 píxeles. En la siguiente fila son presentadas 4 Regiones de Interés (ROI, por sus siglas en inglés) de 128×128 píxeles de las imágenes anteriores. Estas regiones fueron seleccionadas al azar. En la primera columna de la Fig. 7.2 se muestra la imagen original de SLP (primera columna).



Fig. 7.1: Imágenes originales TC. (a) SLP [56]. (b) Cabeza TC. (c) Abdomen TC [57].

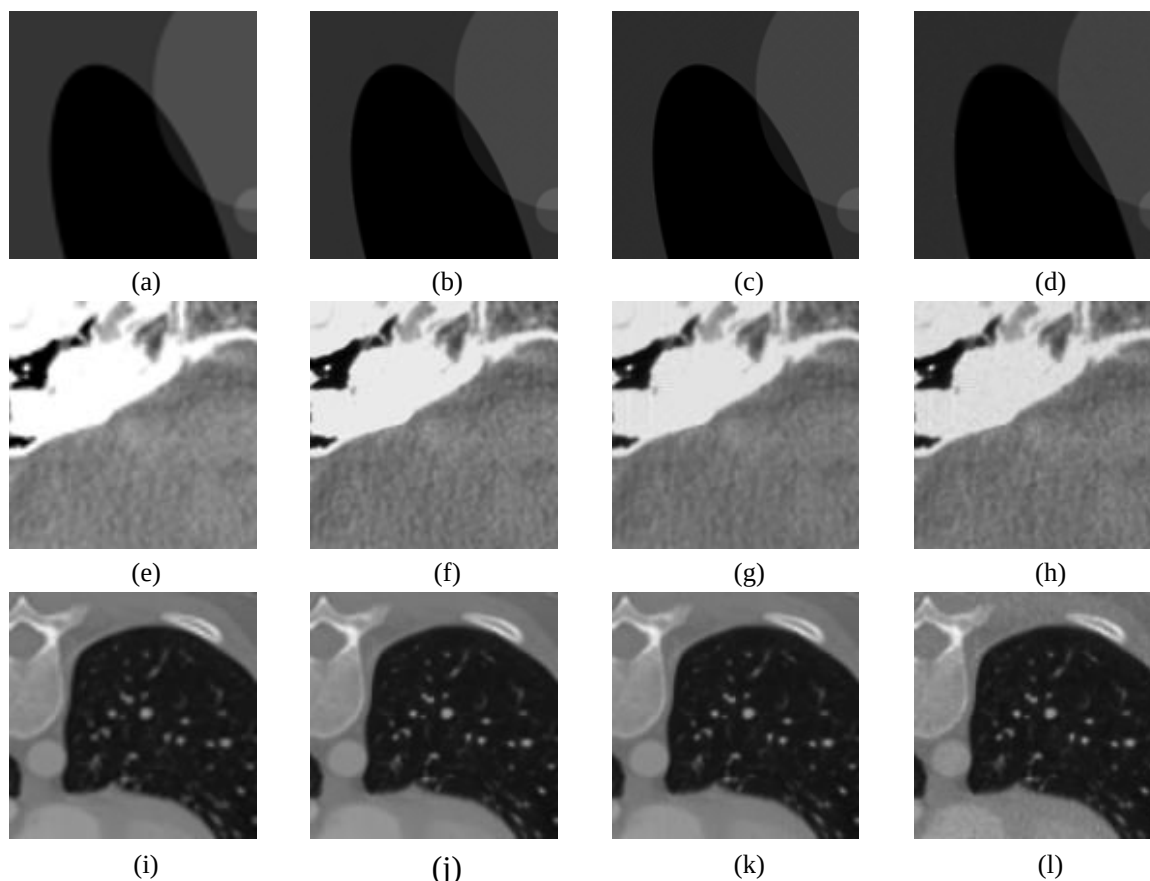


Fig. 7.2. Comparación de imágenes TC originales y reconstruidas utilizando el algoritmo MLEM, Parallel-OpT, SLP. (a) ROI Imagen original TC. (b) ROI La mejor. (c) ROI Funcional. (d) ROI Funcional con ruido de Poisson. Cabeza TC (e) ROI Imagen original TC. (f) ROI La mejor. (g) ROI Funcional. (h) ROI Funcional con ruido de Poisson. Abdomen TC (i) ROI Imagen original TC. (j) ROI La mejor. (k) ROI Funcional. (l) ROI Funcional con ruido de Poisson.

En la segunda columna, se muestra la imagen reconstruida “iteración 64”, que se denominará a partir de este momento como “the best” y denominada iteración 64. El número 64 corresponde al número de iteraciones necesarias para alcanzar el valor de SSIM más alto. En la tercera columna, se muestra la imagen reconstruida denominada “functional”. En este caso se llama “iteración 35”, y se caracteriza por el hecho de que las diferencias en sus parámetros de calidad con los de “the best” son inferiores al 5%.

En la última columna se muestra la imagen “functional” con ruido de Poisson. Las imágenes TC de la fila central corresponden al corte axial del cerebro mientras que en la última fila de la Fig. 7.2 se presenta el corte axial abdominal. Los datos de las 2 últimas filas se presentan en el mismo esquema que para la imagen de SLP.

Los valores de los diferentes parámetros de calidad de cada imagen son presentados en la Tabla 1. También es presentado el número de iteraciones mínimas necesarias para obtener el valor de calidad de imagen dado, así como el tiempo necesario para lograrlo. En la Fig. 7.3 son presentados las medias de los valores para los parámetros de calidad de las imágenes reconstruidas de la Fig. 7.2: MSE, PSNR y SSIM vs. número de iteraciones. La Tabla 2 presenta la cuantificación de estos valores.

Las imágenes obtenidas de la reconstrucción fueron matrices de 512×512 formadas a partir de proyecciones de TC que variaron entre 0 y 180 grados en pasos (vistas) de 5 grados. Inicialmente, los valores de calidad de las imágenes reconstruidas se calcularon para todas las implementaciones en serie y paralelas; sin embargo, las diferencias en los

TABLA 1. Parámetros de calidad de imagen de la Fig. 7.2.

	Sheep-Logan						
	CNR	SNR	SSIM	PSNR	MSE	Iteraciones	Tiempo (s)
Original	0.98	5.17					
La Mejor		5.10	0.99	31.49	46.1	64	408.31
Funcional		4.90	0.99	28.85	84.8	35	55.99
Funcional + ruido			0.98	28.60	88.0	35	54.96
	Cabeza						
	CNR	SNR	SSIM	PSNR	MSE	Iteraciones	Tiempo (s)
Original	0.35	3.54					
La Mejor		3.51	0.98	27.84	107.0	46	295.01
Funcional		3.60	0.97	25.43	186.2	28	44.93
Funcional + ruido			0.96	24.52	229.6	28	47.24
	Abdomen						
	CNR	SNR	SSIM	PSNR	MSE	Iteraciones	Tiempo (s)
Original	0.98	4.63					
La Mejor		4.79	0.98	41.65	4.45	24	155.62
Funcional		4.71	0.98	41.50	4.60	24	39.04
Funcional + ruido			0.96	39.36	7.53	24	39.18

valores absolutos de estos parámetros de calidad fueron menores al 0.001%. Por esta razón, sólo se muestran los resultados de Parallel-OpT, y para su comparación, los de la biblioteca TIGRE: OS-SART renombrado como TIGRE 1 y ASD-POCS como TIGRE 2.

TABLA 2. Cuantificación de los datos presentados en la Fig. 7.3.

	MSE		PSNR		SSIM	
	Mín. Iteración	Tiempo (s)	Máx. Iteración	Tiempo (s)	Máx. Iteración	Tiempo (s)
Shepp-Logan	32.2	1000	1548	33	1000	1548
Shepp-Logan + ruido	58.95	1000	1595	30.42	1000	1595
TC Abdominal	0.842	146	227.4	48.8	146	227.4
TC Abdominal + ruido	7.46	22	35.98	39.4	22	35.98
TC Cabeza	100	1000	1586	28	1000	1586
TC Cabeza + ruido	139.1	1000	1612	26.69	1000	1612
Shepp-Logan con OS_SART	262.2	30	26.8	23.94	30	26.8
Shepp-Logan con ASD_POCS	272.6	40	795.8	23.77	40	795.8

Adicionalmente, fue estudiado la dependencia entre el número óptimo de vistas/proyecciones para la reconstrucción de la imagen en las implementaciones. Estos resultados son presentados como material complementario, (Fig. A.1). En este análisis, los parámetros de calidad de imagen, así como el tiempo de cálculo se estudiaron en función del número de iteraciones y del número vistas. Se encontró que el uso pasos de 5 grados por vista, para cubrir todo el campo de visión, así como de 50 a 100 iteraciones, produjo resultados confiables sin extender el tiempo de cálculo y obteniendo valores de parámetros de imagen de alta calidad.

Cabe mencionar que cuando el campo visual fue cubierto con proyecciones que variaban entre 0-180 grados o 0-360 grados, las reconstrucciones de los diferentes programas produjeron resultados muy similares (en cuanto a parámetros de calidad de imagen). Debido a esto y desde este punto, sólo se tomaron en cuenta reconstrucciones con una matriz del sistema que varió de 0 a 180 grados. La Fig. 7.4 presenta las medias del tiempo de cómputo para las cuatro implementaciones/programas desarrollados en esta tesis, así como para las dos versiones usadas sobre TIGRE. También muestra los resultados con ruido de Poisson añadido. Los datos se presentan en función del número de iteraciones.

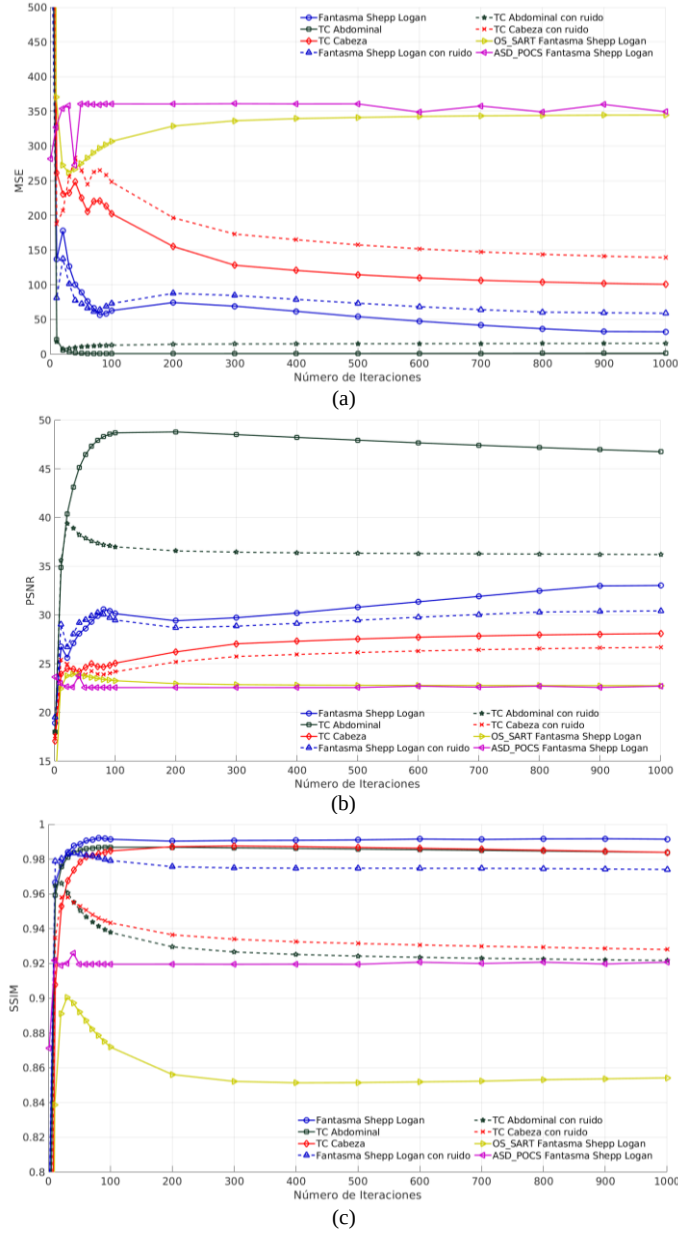


Fig. 7.3 Parámetros de calidad de imagen MSE, PSNR y SSIM vs. número de iteraciones. Parámetros de imagen obtenidos después de la reconstrucción con el algoritmo Parallel-OpT. (a) MSE (b) PSNR (c) SSIM.

Se puede observar que las diferencias en tiempo de cómputo entre las implementaciones Parallel-OpT y Serial-C fueron del orden de 12 veces más rápido para la solución paralela. El Parallel-OpT fue 170 veces más rápido que Serial-MatLab. Las implementaciones Parallel-OpT y Parallel-GPU difirieron en un factor de 2 cuando se consideró el tiempo de procesamiento de la computadora. Se observó que los tiempos de cómputo de TIGRE 2 fueron comparables a los del Serial-C y aproximadamente 10 veces más lentos que Parallel-OpT. En contraste, TIGRE 1 fue el más rápido de todos, siendo 2 veces más rápido que la implementación Parallel-OpT en CUDA. Cuando se analizaron los datos con ruido agregado, no se encontraron diferencias en el tiempo de cómputo con respecto a Parallel-OpT.

Esto fue así para la reconstrucción de las tres imágenes estudiadas. Finalmente, vale la pena señalar que el tiempo computacional para la reconstrucción de las diferentes implementaciones en CUDA del algoritmo MLEM, fue estable y constante entre

ejecuciones, y no dependió del número de iteraciones completadas o la adición de ruido de Poisson a las imágenes para reconstruir.

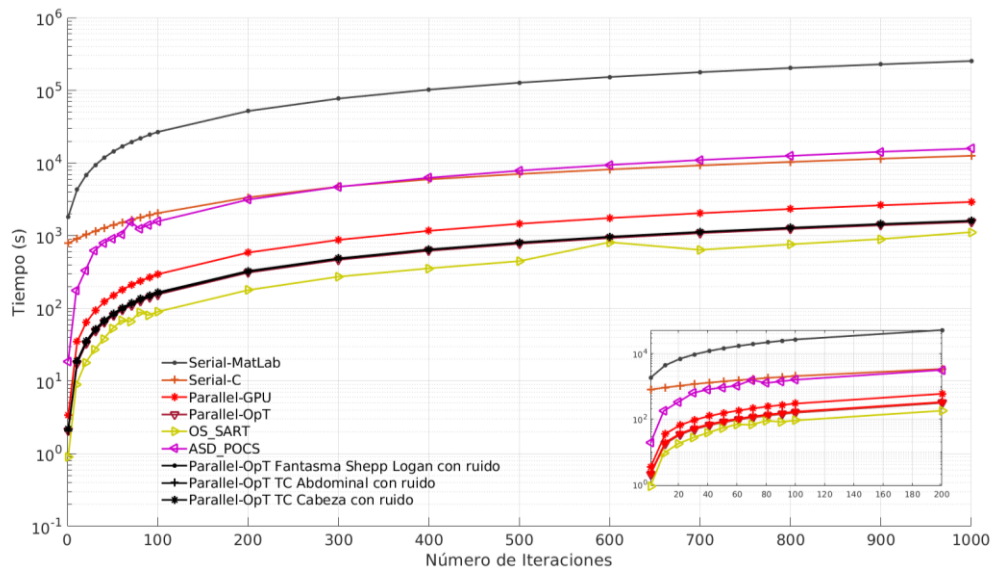


Fig. 7.4. Tiempos de los cálculos para diferentes algoritmos desarrollados en este estudio y algoritmos de TIGRE. En la esquina inferior derecha es presentada una imagen ampliada de las primeras 200 iteraciones.

Como ya se mencionó anteriormente y como se observa en las Figuras 7.3 y 7.4, Parallel-OpT fue hasta 13 veces más rápido que TIGRE 2 y 7% mejor en calidad de imagen. TIGRE 1 fue, en contraste, un 50% más rápido, pero con una calidad de resultados de imagen inferior. El MSE fue 720% peor, el PSNR fue 7.5% peor y el SSIM fue casi 14% más bajo. Estas diferencias son resaltadas en la Fig. 7.5.

En la Fig. 7.5, es mostrado la comparación general entre Parallel-OpT y los dos elegidos de la biblioteca TIGRE. Puede ser apreciada la mejor relación calidad/tiempo desde la primera implementación.

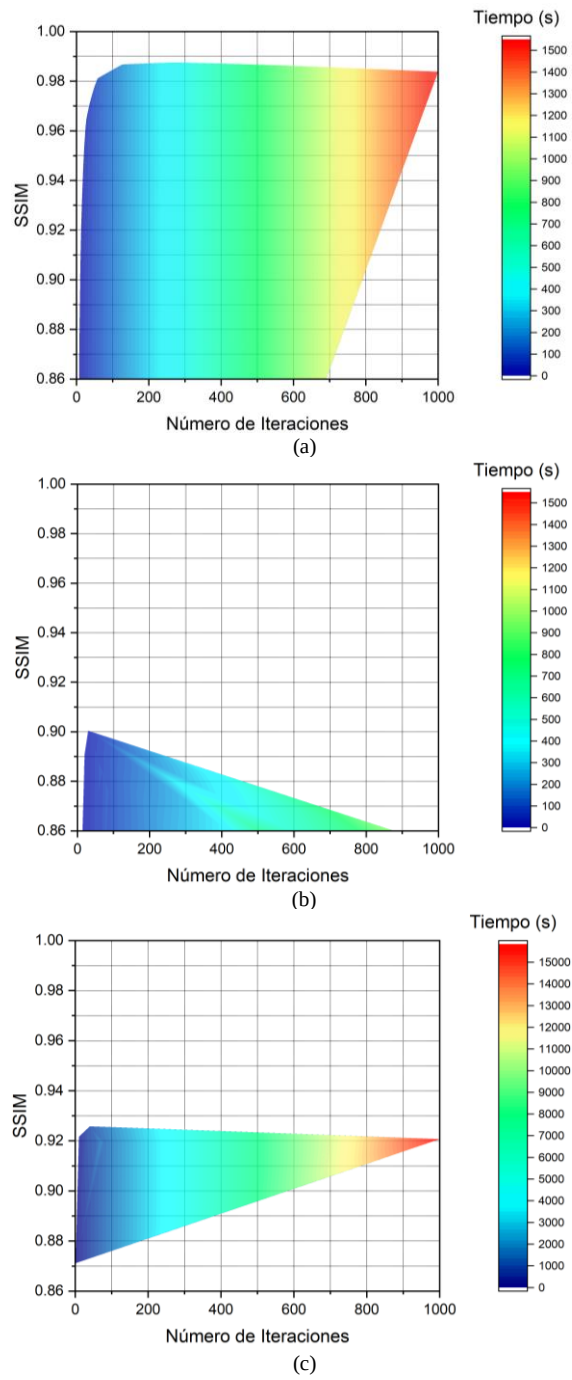


Fig. 7.5: SSIM vs. número de iteraciones vs. Tiempo. (a) Parallel-OpT (b) TIGRE 1 y (c) TIGRE 2. Incrustado en la imagen, el código de color indica el efecto del tiempo sobre los valores SSIM.

Capítulo 8

Aplicaciones clínicas

Dada la base técnica y teórica explicada en los capítulos anteriores, las aplicaciones clínicas son el foco principal de este capítulo.

Para enfatizar el impacto y la relevancia de esta tesis, mostramos la aplicación de nuestra implementación Parallel-OpT a un estudio clínico que incluyen aspectos de comparación de rendimiento y calidad de imagen con TIGRE 1 y 2.

8.1 Estudio clínico

The Cancer Imaging Archive (TCIA) [57] es un servicio que aloja una gran cantidad de imágenes médicas de cáncer accesibles para descarga pública y acceso abierto para apoyar la investigación, el desarrollo y las iniciativas educativas que utilizan imágenes médicas avanzadas del cáncer. Los datos están organizados como “colecciones”. Por lo general, las imágenes de los pacientes están relacionadas con una enfermedad común (por ejemplo, cáncer de pulmón), modalidad o tipo de imagen (MRI, TC, etc.). Los datos relacionados con las imágenes, como los resultados del paciente, los detalles del tratamiento, la genómica y los análisis de expertos también se proporcionan cuando están disponibles.

El estudio clínico utilizado fue tomado de la colección The Cancer Genome Atlas Sarcoma (TCGA-SARC) [26], un conjunto de 101 imágenes TC, correspondientes a un corte axial de una pelvis distinta para cada voluntario y por lo tanto, no fue necesario permiso ético para realizar este estudio de nuestra parte.

Primero, el intervalo de iteración utilizado para obtener la mejor imagen está entre 50 y 100. Aquí que se estableció en 100 para los algoritmos Parallel-OpT, TIGRE 1 y TIGRE 2. Segundo, el siguiente paso para este análisis fue la creación de una matriz de sistema utilizada para reconstruir imágenes.

Las matrices de sistema construidas aquí, había la misma cantidad de filas y columnas que rayos de proyección y píxeles respectivamente. En general, una matriz de sistema podría presentar grandes dimensiones dependiendo del número de líneas de proyección, el número de ángulos de proyección y el tamaño de la imagen. Para obtener más detalles sobre cómo se construyó la matriz del sistema, consulte la metodología de [43,44,45].

En nuestro caso, se utilizaron imágenes de 512×512 píxeles con una matriz del sistema de 23944×262144 y un vector de proyección de 23944×1 para un ángulo

de barrido de 5 grados. En cada imagen, SSIM fue calculado como en las secciones anteriores. Los datos de SSIM correspondientes a Parallel-OpT son presentados en la Fig. 8.1-(a) (fue utilizado MatLab Distribution Fitter app software). Los valores de SSIM alcanzaron un máximo de 0.96. Los datos presentaron una distribución normal de 0.902 ± 0.019 (media $\pm$ desviación estándar). Esto fue confirmado por una prueba de Shapiro-Wilks en el software SPSS ($p < 0.05$). Todos los datos, excepto uno, superaron el

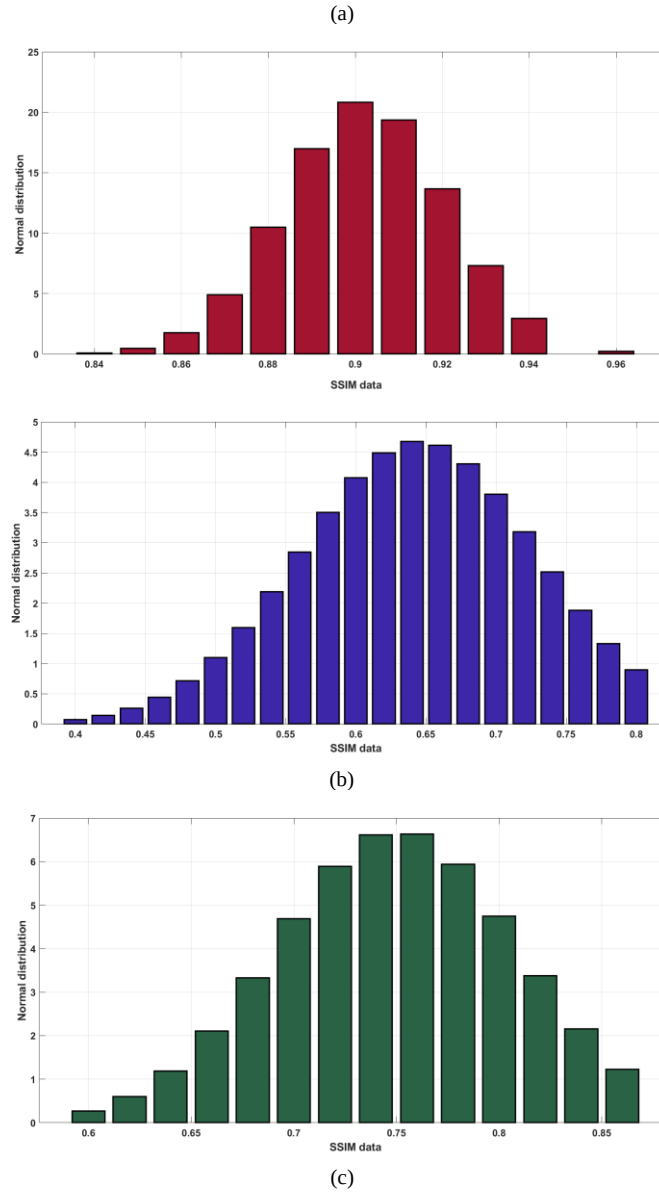


Fig. 8.1 Histograma de valores SSIM de imágenes reconstruidas a partir de un estudio clínico. Los datos presentan una distribución de valores igualmente normalizada. Para (a) Parallel-OpT (b) TIGRE 1 (c) TIGRE 2.

valor umbral de 0.85 SSIM, un límite estándar para una buena calidad de imagen.

En segundo lugar, los datos SSIM correspondientes a TIGRE 1 se presentan en la Fig. 8.1-(b). Los valores de SSIM alcanzaron un máximo de 0.78. Los datos presentaron una distribución normal 0.645 ± 0.085 (media $\pm$ desviación estándar). Ningún dato alcanzó el valor umbral. Finalmente, los datos SSIM correspondientes a TIGRE 2 se presentan en la Fig. 8.2-(c). Los valores de SSIM alcanzaron un máximo de 0.85. Los datos presentaron una distribución normal 0.750 ± 0.059 (media $\pm$ desviación estándar). Solo 2 datos excedieron el valor umbral de 0.85, 22 establecidos entre 0.8 y 0.85. Los datos restantes del estudio estaban por debajo de 0.8 SSIM.

Capítulo 9

Conclusiones

En este trabajo se muestra la implementación en paralelo del algoritmo de reconstrucción MLEM, utilizando las capacidades de tarjetas gráficas. Las soluciones de software desarrolladas aquí se compararon con reconstrucciones de imágenes clínicas utilizando varios parámetros de calidad de imagen (CNR, SSIM, MSE, etc.).

El tiempo de computo utilizado para la reconstrucción de imágenes en programas basados en CUDA fue menor que el asociado con la solución serial por factores que varían entre 2 y 170 veces

En comparación con el software TIGRE, las imágenes se obtuvieron de una calidad similar, pero en algunos casos y para algunos factores de calidad de imagen específicos, las imágenes fueron hasta 7 veces mejores con los programas desarrollados en CUDA.

La adición de un 5% de ruido a las señales no aumentó el tiempo de reconstrucción proporcionalmente, de hecho, lo dejó casi inalterado. Aún así, la calidad de la imagen disminuyó en general en un factor del 5% (a excepción del parámetro PSNR que perdió el 30% de su valor máximo en algunos escenarios). Adicionalmente, los resultados del estudio clínico mostraron que los cálculos fueron consistentes y reproducibles. Además, la calidad de la reconstrucción de la imagen se conservó incluso con tiempos de reconstrucción cortos. Todos estos resultados implican que las implementaciones de MLEM en CUDA usando las capacidades de GPU presentadas aquí, fueron confiables y rápidas.

Con respecto a la adquisición de imágenes, se demostró que, para generar la matriz de sistema, solo se necesitaban proyecciones en un rango de 0 a 180 grados. Esto en contraste con la adquisición de proyecciones de campo completo (de 0 a 360 grados). Este es un hecho bien establecido en el campo, por lo que no se presentaron datos o imágenes para respaldar este hecho en este trabajo. En general, las adquisiciones de ángulos pequeños (1,2,3 grados) a 1000 iteraciones, tomaron hasta 5000 s más que las reconstrucciones en pasos de 5 grados (400% veces más) o mayores.

Las mediciones con ángulos más grandes se describieron en la Figura complementaria A1. Allí se pudo apreciar que las mediciones de 5 grados presentaban un compromiso perfecto entre la calidad de la reconstrucción y el tiempo de reconstrucción. Por lo tanto, este valor fue el seleccionado para la reconstrucción en este estudio.

La Fig. 7.3 mostró cómo las diferentes métricas de calidad de imagen (MSE, PSNR, SSIM) disminuyeron primero y luego aumentaron rápidamente a medida que aumentó el número de iteraciones. Esto permitió deducir un rango “ideal” de iteraciones para la reconstrucción de imágenes para esta implementación de software, que se estableció entre 50 y 100.

Por lo tanto, como se puede ver en la Fig. 7.3b, el parámetro PSNR para el SLP y la imagen abdominal, alcanzó un valor máximo local dentro del rango de 50 a 100 iteraciones. Para la imagen de la cabeza, este efecto fue similar pero menos prominente. Obviamente, cuando se realizó un mayor número de iteraciones, los valores de este último parámetro fueron mayores y, por lo tanto, mejores. Sin embargo, el tiempo utilizado para mejorar el 5% de este parámetro podría ser fácilmente de 5 a 10 veces mayor. Eso hizo que este enfoque de iteraciones más grandes fuera menos atractivo para fines prácticos. Los valores máximos de la métrica SSIM para las imágenes SLP, cabeza y abdomen fueron: 0.992, 0.987 y 0.986, respectivamente. Estos valores máximos se alcanzaron dentro del rango de iteraciones propuesto (ver Fig. 7.3c). Después del máximo, los valores se mantuvieron estables y no se vieron afectados por un mayor número de iteraciones. Como se indicó anteriormente en el texto, los valores de SSIM superiores a 0.85 representaban imágenes de buena calidad clínica.

Una vez que fue establecido el intervalo de iteración, así como los grados entre las proyecciones utilizadas para adquirir las imágenes, se compararon los tiempos de procesamiento de las diferentes implementaciones desarrolladas (ver Fig. 7.4)

Se descubrió que la diferencia de tiempo entre Serial-C y Parallel-OpT para lograr el mismo resultado era de alrededor de 1000 segundos a 50 iteraciones (12 veces más) y 9000 segundos en 1000 iteraciones (170 veces más).

Estos resultados mostraron que las implementaciones del algoritmo MLEM basado en CUDA, podrían realizar buenas reconstrucciones de imágenes en tiempos más cortos que los métodos tradicionales basados en CPU.

Sorprendentemente, la adición de ruido de Poisson no afectó los tiempos de reconstrucción de la implementación Parallel-OpT.

De hecho, se necesitaba el mismo número de iteraciones para lograr las mejores imágenes funcionales, independientemente de eso. En comparación con los resultados de las implementaciones de TIGRE, se lograron rendimientos de tiempo iguales o mejores, siempre con una mejor reconstrucción de la calidad de imagen.

Una de las principales limitaciones de nuestro estudio fue no tener acceso a un escáner TC clínico ni a sus datos sin procesar. Idealmente, compararíamos los tiempos de computación y la calidad de reconstrucción de nuestros algoritmos con los implementados por las compañías de salud. Esto daría una mejor idea de la fortaleza de nuestra metodología en un entorno clínico. También tener acceso a imágenes adicionales nos habría permitido llevar a cabo análisis más grandes y confiables.

Sin embargo, este último punto debe considerarse con cuidado, ya que las imágenes de escáneres TC comerciales generalmente se presentan a los usuarios después de manipulaciones severas (es decir, filtrado). Debido a esta calidad, los parámetros de las imágenes clínicas pueden verse mejor de lo que realmente son. Otra limitación que vale la pena mencionar fue la potencia de cómputo de la máquina portátil y la GPU utilizada que se mantuvo al mínimo para que cualquiera pudiera tener acceso a este tipo de tecnología. Sin embargo, esto provocó que el tamaño de la memoria de la GPU resultara insuficiente para procesar imágenes directamente. Esto estaba claro en la etapa de construcción de la matriz del sistema. Debido a esta limitación, durante este análisis, las imágenes tuvieron que dividirse antes de la reconstrucción. Los bloques se construyeron con una relación tamaño de imagen/bloque de 16. De esta manera, no hubo un cuello de botella informático

al enviar en serie cada bloque de la CPU a la GPU. Esta transferencia de datos estaba limitada por el ancho de banda del bus de las GPU y su latencia.

Además, debe tenerse en cuenta que nuestra implementación no es el primer y único software a considerar cuando se trata de la reconstrucción tomográfica. La mayoría de los usuarios pueden preferir paquetes de software comerciales, ya que proporcionan una interfaz gráfica de usuario fácil de usar.

Por lo tanto, podría compararse con otros softwares como The ASTRA Toolbox [64] para tener una medición más clara de su eficiencia y calidad de reconstrucción de imágenes clínicas y mejorar la interactividad con el usuario a través de una interfaz gráfica.

Para utilizar el software presentado aquí, el usuario requiere conocimientos de lenguajes de programación y la habilidad para analizarlos y procesarlos en las capas de CUDA C.

Además, no incluye los algoritmos típicos de preprocesamiento (por ejemplo, Geometría TC) y posprocesamiento (por ejemplo, Corrección de desenfoque). Estas desventajas limitan el objetivo de nuestro software principalmente a investigadores y usuarios con experiencia en informática.

Los lectores deben considerar que una de las ideas básicas detrás de este trabajo fue implementar el algoritmo MLEM de una manera novedosa utilizando las capacidades de GPU en sistemas informáticos comerciales de baja capacidad y esto fue logrado. La solución para el análisis de imágenes presentada aquí podría usarse en instituciones o sistemas nacionales de salud con recursos financieros limitados, de manera confiable, conservando la calidad de la imagen y manteniendo los tiempos de reconstrucción lo más bajos posible.

Capítulo 10

Publicaciones

Los resultados de esta tesis han dado lugar a los artículos donde se describen las implementaciones utilizadas.

- Valencia Pérez, T. A., Hernández López, J. M., Moreno-Barbosa, E., de Celis Alonso, B., Palomino Merino, M. R., & Castaño Meneses, V. M. (2020). Efficient CT Image Reconstruction in a GPU Parallel Environment. *Tomography* (Ann Arbor, Mich.), 6(1), 44–53. <https://doi.org/10.18383/j.tom.2020.00011>
- T. A. Valencia-Pérez, J. M. Hernández-López, E. Moreno-Barbosa, \& B. de Celis-Alonso. (2020). Study of CT Images Processing with the Implementation of MLEM Algorithm using CUDA on NVIDIA’S GPU Framework. *Journal of Nuclear Physics, Material Sciences, Radiation and Applications*, 7(2), 165-171. <https://doi.org/10.15415/jnp.2020.72021>
- T. A. Valencia Pérez, J. M. Hernández López, E. Moreno Barbosa, M. I. Martínez Hernández, G. Tejeda Muñoz, and B. de Celis Alonso, Parallel approach to tomographic reconstruction algorithm using a nvidia gpu, *AIP Conference Proceedings*, vol. 2090, no. 1, p. 050015, 2019. <https://doi.org/10.1063/1.5095930>

Referencias

- [1] B. SC., Manual de radiología para técnicos: Física, biología y protección radiológica, 2010.
- [2] T. Buzug, Computed Tomography: From Photon Statistics to Modern Cone-Beam CT, 01 2008.
- [3] L. A. Feldkamp, L. C. Davis, and J. W. Kress, “Practical cone-beam algorithm,” J. Opt. Soc. Am. A, vol. 1, no. 6, pp. 612–619, Jun 1984. [Online]. Available: <http://josaa.osa.org/abstract.cfm?URI=josaa-1-6-612>
- [4] A. C. Kak, M. Slaney, and G. Wang, “Principles of computerized tomographic imaging,” Medical Physics, vol. 29, no. 1, pp. 107–107, 2002. [Online]. Available: <https://aapm.onlinelibrary.wiley.com/doi/abs/10.1118/1.1455742>
- [5] S. R. Deans, The Radon transform and some of its applications, 1983.
- [6] J. Radon, Akad. Wiss., pp. 262–277.
- [7] F. Natterer, The Mathematics of Computerized Tomography. USA: Society for Industrial and Applied Mathematics, 2001.
- [8] B. F. Hutton, J. Nuyts, and H. Zaidi, Iterative Reconstruction Methods. Boston, MA: Springer US, 2006, pp. 107–140. [Online]. Available: https://doi.org/10.1007/0-387-25444-7_4
- [9] G. Zeng, Medical image reconstruction: A conceptual tutorial, 01 2010.
- [10] J. Hsieh, B. Nett, Z. Yu, K. Sauer, J.-B. Thibault, and C. A. Bouman, Recent Advances in CT Image Reconstruction, 03 2013. [Online]. Available: <https://doi.org/10.1007/s40134-012-0003-7>
- [11] S. Vandenberghe, Y. D’Asseler, R. Van de Walle, T. Kauppinen, M. Koole, L. Bouwens, K. Van Laere, I. Lemahieu, and R. Dierckx, “Iterative reconstruction algorithms in nuclear medicine,” Computerized Medical Imaging and Graphics, vol. 25, no. 2, p. 105–111, 2001. [Online]. Available: [https://doi.org/10.1016/s0895-6111\(00\)00060-4](https://doi.org/10.1016/s0895-6111(00)00060-4)
- [12] L. L. Geyer, U. J. Schoepf, F. G. Meinel, J. W. Nance, G. Bastarrika, J. A. Leipsic, N. S. Paul, M. Rengo, A. Laghi, and C. N. De Cecco, “State of the art: Iterative ct reconstruction techniques,” Radiology, vol. 276, no. 2, pp. 339–357, 2015, pMID: 26203706. [Online]. Available: <https://doi.org/10.1148/radiol.2015132766>
- [13] L. A. Shepp and Y. Vardi, “Maximum likelihood reconstruction for emission tomography,” IEEE Transactions on Medical Imaging, vol. 1, no. 2, pp. 113–122, 1982.
- [14] S. Cook, CUDA Programming: A Developer’s Guide to Parallel Computing with GPUs, 1st ed. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2012.
- [15] J. Sanders and E. Kandrot, CUDA by Example: An Introduction to General-Purpose GPU Programming, 1st ed. Addison-Wesley Professional, 2010.
- [16] M. Schellmann, S. Gorlatch, and Meil, “Parallel medical image reconstruction: from graphics processing units (gpu) to grids.”

- [17] NVIDIA, “Gpu-accelerated applications,” 2019. [Online]. Available: <https://www.nvidia.com/en-us/data-center/gpu-accelerated-applications/catalog/>
- [18] An., “Medical image processing on the gpu – past, present and future,” *Medical Image Analysis*, vol. 17, no. 8, pp. 1073 – 1094, 2013. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S1361841513000820>
- [19] K. T, S. P, and S. K, “Survey of using gpu cuda programming model in medical image analysis,” *Informatics in Medicine Unlocked*, vol. 9, pp. 133–144, 2017. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S235291481730045X>
- [20] T. Natalya, S. Jeremy, and D. Christopher, “Advanced interactive medical visualization on the gpu,” *Journal of Parallel and Distributed Computing*, vol. 68, no. 10, pp. 1319–1328, 2008.
- [21] F. L. A., V. Vicent, M. Patricia, R. Francisco, and V. Gumersindo, “Ct image reconstruction based on gpus,” *Procedia Computer Science*, vol. 18, pp. 1412–1420, 2013.
- [22] B. V and V. C, “Cuda parallel implementation of image reconstruction algorithm for positron emission tomography,” *Open Medical Imaging Journal*, pp. 108–118, 2012. [Online]. Available: <http://www.citeulike.org/group/15740/article/12092940>
- [23] X. Lizhe, H. Yining, Y. Bin, W. Lin, Y. Benqiang, L. Wenyuan, Z. Libo, L. Limin, S. Huazhong, and C. Yang, “An effective cuda parallelization of projection in iterative tomography reconstruction,” *PLoS ONE*, vol. 10, no. 11, pp. 1–17, 2015.
- [24] P. Guillem and X. Lei, “Gpu computing in medical physics: A review,” *Medical Physics*, vol. 38, no. 5, pp. 2685–2697, 2011. [Online]. Available: <https://aapm.onlinelibrary.wiley.com/doi/abs/10.1118/1.3578605>
- [25] V. Cnudde and M. Boone, “High-resolution x-ray computed tomography in geosciences: A review of the current technology and applications,” *Earth-Science Reviews*, vol. 123, pp. 1 – 17, 2013. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S001282521300069X>
- [26] C. Roche, E. Bonaccio, and J. Filippini, “Radiology data from the cancer genome atlas sarcoma [tcga-sarc] collection. the cancer imaging archive,” 2016. [Online]. Available: <http://doi.org/10.7937/K9/TCIA.2016.CX6YLSUX>
- [27] J. Hsieh, *Computed Tomography Principles, Design, Artifacts, and Recent Advances*, 2nd ed. John Wiley and Sons Inc, 10 2009.
- [28] L. D. Chiffre], S. Carmignato, J.-P. Kruth, R. Schmitt, and A. Weckenmann, “Industrial applications of computed tomography,” *CIRP Annals*, vol. 63, no. 2, pp. 655 – 677, 2014. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0007850614001930>
- [29] W. Commons. (2008) Diagrama del espectro electromagnético. [Online]. Available: https://commons.wikimedia.org/wiki/File:EM_Spectrum_Properties_es.svg
- [30] Diagrama del tubo de rayos x. [Online]. Available: <https://commons.wikimedia.org/wiki/File:Roentgen-Roehre.svg>
- [31] J. T. Bushberg, J. A. Seibert, E. M. L. Jr., and J. M. Boone, *The Essential Physics of Medical Imaging*, 2nd ed. Lippincott Williams and Wilkins, 2002.
- [32] Nvidia-Corporation. (2017) Aplicaciones de la gpu. [Online]. Available: <https://developer.nvidia.com/>
- [33] B. Keck, “High performance iterative x-ray ct with application in 3-d mammography and interventional c-arm imaging systems,” Ph.D. dissertation, 11 2014.
- [34] A. Macovski, *Medical Imaging Systems*, ser. Prentice-Hall information and system sciences series. Prentice-Hall, 1983. [Online]. Available: <https://books.google.com.mx/books?id=G74eAQAAIAAJ>
- [35] R. Gordon, R. Bender, and G. T. Herman, “Algebraic reconstruction techniques (art) for three-dimensional electron microscopy and x-ray photography,” *Journal of Theoretical*

- Biology, vol. 29, no. 3, pp. 471 – 481, 1970. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/0022519370901098>
- [36] C. Cavedon and S. Rudin, Cardiovascular and Neurovascular Imaging: Physics and Technology. CRC Press, 2015.
- [37] P. Gilbert, “Iterative methods for the three-dimensional reconstruction of an object from projections,” Journal of Theoretical Biology, vol. 36, no. 1, pp. 105 – 117, 1972. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/0022519372901804>
- [38] H. A. Andersen and C. A. Kak, “Simultaneous algebraic reconstruction technique (sart): A superior implementation of the art algorithm,” Ultrasonic Imaging, pp. 81–94, 1984.
- [39] R. Yagel and K. Mueller, “Fast and accurate three-dimensional reconstruction from cone-beam projection data using algebraic methods,” 1998.
- [40] K. Lange and R. F. Carson, “Em reconstruction algorithms for emission and transmission tomography.” Journal of computer assisted tomography, vol. 8 2, pp. 306–16, 1984.
- [41] K.
- [42] K. Lange, M. Bahn, and R. Little, “A theoretical study of some maximum likelihood algorithms for emission and transmission tomography,” IEEE Transactions on Medical Imaging, vol. 6, no. 2, pp. 106–114, 1987.
- [43] E. Gopi, Digital Signal Processing for Medical Imaging Using Matlab. Springer Science and Business Media, 2012.
- [44] W. S. Van Hemelryck Tessa, G. Maggie, B. K. Joost, and S. Jan, “The implementation of iterative reconstruction algorithms in matlab,” 2007.
- [45] H. L. (2017) Tools for 2-d tomographic reconstruction. [Online]. Available: <https://github.com/phymhan/matlab-tomo-2d>
- [46] T. Valencia-Pérez. (2020) Pasos del algoritmo mlem.
- [47] V. Vora, A. Suthar, Y. Makwana, and S. Davda, “Analysis of compressed image quality assessments,” International Journal of Advanced Engineering and Application, vol. 10, 01 2010.
- [48] J. W. Goodman, “Some fundamental properties of speckle,” JOSA, vol. 66, no. 11, pp. 1145–1150, 1976.
- [49] D. M. Chandler and S. S. Hemami, “Vsnr: A wavelet -based visual signal-to-noise ratio for natural images,” IEEE Transactions on Image processing, vol. 16, no. 9, pp. 2284–2298, 2007.
- [50] Z. Wang, A. C. Bobik, H. R. Sheikh, and E. P. Simoncelli, “Image quality assessment: From error visibility to structural similarity,” IEEE Transactions on Image processing, vol. 13, no. 4, pp. 600–612, 2004.
- [51] Z. Wang and A. C. Bovik, “Mean squared error: Love it or leave it? a new look at signal fidelity measures,” IEEE signal processing magazine, vol. 26, no. 1, pp. 98–117, 2009.
- [52] D. Salomon, Data compression: the complete reference. Springer Science and Business Media, 2004.
- [53] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, “Image quality assessment: from error visibility to structural similarity,” IEEE transactions on image processing, vol. 13, no. 4, pp. 600–612, 2004.
- [54] G. Dougherty, Digital image processing for medical applications. Cambridge University Press, 2009.
- [55] NVIDIA, “Cuda cublas library,” 2018. [Online]. Available: <https://docs.nvidia.com/cuda/cublas/index.html>

- [56] L. A. Shepp and B. F. Logan, "The fourier reconstruction of a head section," IEEE Transactions on nuclear science, vol. 21, no. 3, pp. 21–43, 1974.
- [57] K. Clark, B. Vendt, K. Smith, J. Freymann, J. Kirby, P. Koppel, S. Moore, S. Phillips, D. Maffitt, M. Pringle et al., "The cancer imaging archive (tcia): maintaining and operating a public information repository," Journal of digital imaging, vol. 26, no. 6, pp. 1045–1057, 2013.
- [58] A. Biguri, M. Dosanjh, S. Hancock, and M. Soleimani, "Tigre: a matlab-gpu toolbox for cbct image reconstruction," Biomedical Physics and Engineering Express, vol. 2, no. 5, p. 055010, 2016.
- [59] E. Y. Sidky and X. Pan, "Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization," Physics in Medicine and Biology, vol. 53, no. 17, p. 4777, 2008.
- [60] M. Helmke, E. K. Joseph, J. A. Rey, P. Ballew, and B. M. Hill, The official Ubuntu book. Pearson Education, 2014.
- [61] "Gcc, the gnu compiler collection gnu project free software foundation [fsf]." [Online]. Available: <https://gcc.gnu.org/>
- [62] I. Corporation and W. G. Itseez, "Open source computer vision," 2000. [Online]. Available: <https://opencv.org/>
- [63] A. Kaehler and G. Bradski, Learning OpenCV 3: computer vision in C++ with the OpenCV library. O'Reilly Media, Inc., 2016.
- [64] W. van Aarle, W. J. Palenstijn, J. D. Beenhouwer, T. Altantzis, S. Bals, K. J. Batenburg, and J. Sijbers, "The astra toolbox: A platform for advanced algorithm development in electron tomography," Ultramicroscopy, vol. 157, pp. 35 – 47, 2015. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0304399115001060>