VNIVERSITAT DE VALÈNCIA

Departamento de Física Atómica, Molecular y Nuclear
Instituto de Física Corpuscular

Programa Oficial de Doctorado en Física

Environment recognition applied to particle detectors

Alberto Corbi Bellot

supervised by Francisco Albiol Colomer and Alberto Albiol Colomer

Valencia – June, 2017

D. Francisco Albiol Colomer y Alberto Albiol Colomer, investigadores del Consejo Superior de Investigaciones Científicas y de la Universidad Politécnica de Valencia, respectivamente,

CERTIFICAN: Que la presente Memoria *Environment recognition applied to particle detectors* ha sido realizada bajo su dirección en el Instituto de Física Corpuscular por D. Alberto Corbi Bellot como Tesis para obtener el grado de Doctor en Física.

Y para que así conste, en cumplimiento de la legislación vigente, presenta ante el Departamento de Física Atómica y Nuclear la referida memoria, firmando el presente certificado en Burjassot (Valencia) a 26 de Junio de 2017.

A mis tres madres: *Mª Laura, Maribel e Isabel.*

A mis tres padres: *Alberto, Celso y Emilio.*

Contents

List of Figures

List of Tables

Resumen en español

Introducción

Los detectores de partículas son dispositivos que registran la radiación ionizante, bien de sistemas activos (rayos X, aceleradores, etc.) o bien de isótopos radiactivos. Para poder realizar medidas de precisión con estos instrumentos, es necesario *modelar geométricamente el entorno*, *contorno* o *escena* bajo estudio. Estas condiciones geométricas se pueden determinar de forma más o menos precisa en algunos experimentos de física de

partículas/nuclear, y en algunos sistemas de imagen, como las tomografías. Sin embargo, este escenario no es necesariamente el habitual. El propósito principal de este trabajo de tesis es desarrollar técnicas e instrumentos que aporten la mencionada *información del entorno* a cualquier sistema de detección de radiación y de *manera general*. Como iremos viendo, estas mejoras tienen lugar mediante la adición de sensores externos (cámaras de video y cámaras de rango, principalmente) capaces de aportar dichos datos sobre el contexto espacial.

Por *escena* o *contorno* se entiende tanto los límites del emplazamiento físico donde se realizan las medidas (habitación, habitáculo, recinto, alrededores, etc.), como el *propio elemento bajo examen* (paciente, objeto contaminado, fuente radioactiva, etc.), incluyendo su posición, giro y volumen relativo al sistema de imagen o a un punto fijo. Tal es el caso de los dispositivos de rayos X de propósito general o los sistemas detectores portátiles usados, por ejemplo, para la medición de radiación ambiental. Como se demuestra a lo largo de este trabajo de tesis, la mencionada *geometría de la escena* puede llegar a complementar o *aumentar* (concepto tomado prestado del mundo de la *visión por ordenador* o *computer vision*) de manera muy significativa la información propia recabada por los sistemas de adquisición utilizados. De manera similar, cuando un dispositivo A *aumenta* un dispositivo B, implica que A provee a B con información espacial relativa a marco de trabajo, de manera que puede derivarse, por ejemplo, información 3D por parte de B, registrar imágenes A+B, etc.

Para alcanzar este objetivo, y como parte de esta investigación, se han explorado técnicas y métodos de reconocimiento del entorno, aplicados a las siguientes áreas:

- aumento de dispositivos de rayos X usados en diagnóstico primario,
- reconstrucción tridimensional de la anatomía de la persona examinada partiendo de radiografías convencionales que luego pueden ser estereográficamente relacionadas,
- obtención de nuevas funciones de transferencia que permitan la generación de imágenes densitométricas a partir de las imágenes de absorción y el volumen del/de la paciente, y
- asignación de coordenadas 3D a fuentes de radiación y a la dosis recibida.

Se ha hecho especial énfasis en los dispositivos de rayos X por su indudable presencia en muchos ámbitos, desde los puramente clínicos hasta los relacionados con la inspección preventiva/forense de objetos. En el contexto de este trabajo, estos sistemas de imagen son aumentados mediante la interacción con dispositivos modernos de posicionamiento, tales como cámaras de video, profundidad, etc. La ventaja de esta *arquitectura de imagen dual* es la posibilidad de determinar geométricamente la escena con precisión y trasladar y superponer esta información al resultado de origen clínico (o al fruto de una inspección relacionada con la gestión de residuos radioactivos, como en el caso de las gamma-cámaras, estudiadas en Capítulo 8).

Además, como parte de los resultados obtenidos en esta tesis, se ha desarrollado una métrica especial (basada en análisis y teoría de la imagen) para cuantificar de manera objetiva la calidad de imágenes radiográficas. Esta técnica es utilizada para estimar la información de las imágenes densitométricas obtenidas mediante los métodos estudiados en este trabajo.

Los rayos X convencionales y sus limitaciones

La modalidad radiológica de rayos X *convencional* es sin duda la más presente y usada en la práctica clínica y ciencias de la salud. Su implantación en todo tipo de centros de salud es muy destacable dada su relativa simplicidad técnica, rapidez y efectividad para diagnosticar muchos tipos de dolencias. La llegada de la *radiografía digital* no ha hecho otra cosa sino profundizar en esta realidad.

Un dispositivo de rayos X consta de un tubo generador de este tipo de radiación instalado dentro de un blindaje, un generador de alta tensión y un chasis o *cassette* que contiene en su interior la película radiográfica o detector digital que integra finalmente la emisión Roentgen que no ha sido absorbida por el/la paciente o el objeto analizado.

A diferencia de otras modalidades como la *tomografía axial computerizada* (TAC), en la modalidad de rayos X *ordinarios* la geometría de la

escena clínica es descrita de manera muy somera. Con enorme frecuencia, el único registro de la misma son sencillas indicaciones relativas a la posición (y sobre todo, orientación) del/de la paciente con respecto a la cubierta protectora del detector de pared vertical y/o mesa horizontal. Es lo que se conoce en literatura como *protocolo* o simplemente, *posicionamiento del paciente*. Estas indicaciones son las que luego se traducen en los conocidos *protocolos de examen* tales como *radiografía postero-anterior*, *antero-posterior, decúbito*, *medio-lateral*, etc. Esta alta variabilidad geométrica proviene del hecho de que en los dispositivos de rayos X para diagnóstico primario existe un desacoplo estructural entre el detector y la fuente de fotones X (el ánodo del tubo). Dicho de otra manera: ambos pueden desplazarse libremente y con plena independencia el uno del otro. Esto se traduce a su vez en una alta *fragilidad* de los parámetros intrínsecos (a diferencia de una cámara fotográfica al uso, donde estos valores permanecen fijos desde el momento de su fabricación). Tanto las mesas de examen como los *estativos verticales* pueden ser fijos, flotantes o semiflotantes e incluso a veces es posible modificar su ángulo con respecto al suelo o pared para realizar exámenes especiales, como los digestivos. En cualquier sistema de imagen, los *parámetros intrínsecos* engloban tanto el punto focal como posibles distorsiones y asimetrías que pueden ser medidas y conocidas.

Un ejemplo que suele resultar llamativo de esta libertad de movimiento en los sistemas de imagen por rayos X es el hecho de que el *punto focal* (distancia desde el ánodo al detector y su posición horizontal y vertical en el plano representado por este) puede llegar a estar situado completamente fuera de la superficie de la imagen. Esto acontece, por ejemplo, en algunos protocolos que exigen proyecciones oblicuas o en ángulos muy *picados* (como las que se muestran en la Figura 2.18 y la Figura 4.5). Nuevamente, esta situación contrasta con la fotografía convencional, donde el punto principal se corresponde normalmente con el pixel central, por ejemplo, en el 640, 540 en el caso de una cámara de video de resolución HD (1920, 1080). Los proyectores de luz (usados comúnmente en presentaciones, arte, etc.) también emplean un punto focal muy desplazado con respecto al centro de la imagen, sin embargo esta sólo se forma con nitidez a una distancia específica y fija (es decir, los parámetros intrínsecos del sistema óptico son nuevamente fijos).

Si bien es cierto que la tecnología y estándares radiológicos están preparados para el registro de ciertas distancias tales como la brecha paciente-

detector (IOD), emisor-detector (SID), etc., estas casi nunca son estimadas, ni medidas y mucho menos inventariadas manual o electrónicamente. Sin embargo, es bien conocido tanto teórica como experimentalmente, así como por la práctica diaria, que estas magnitudes pueden llegar a tener una repercusión no despreciable tanto en la generación de la propia imagen radiográfica y su calidad, así como en la gestión de la dosis recibida por parte del/de la paciente.

Rayos-X aumentados mediante dispositivos de captación de contorno

En este trabajo proponemos una serie de herramientas, metodologías y procedimientos para la determinación del ámbito geométrico en escenarios de diagnóstico basados en sistemas convencionales de rayos X. Estas técnicas se apoyan principalmente en la anexión de un dispositivo de *captación de contorno* o *escena* que permanece rígidamente acoplado al sistema de imagen de rayos X. Los dispositivos de captación de contorno que han sido explorados en este trabajo son cámaras de video y cámaras de profundidad, aunque existen muchas otras alternativas tales como cámaras basadas en tiempo de vuelo (*time-of-flight*), LIDARes (*light detection and ranging*), escáneres 3D láser, sistemas de *visión estereoscópica* con cámaras RGB calibradas, etc. Una cámara *calibrada* (sea del tipo que sea: RGB, profundidad, rayos-X) es aquella de la que se conocen sus parámetros intrínsecos y posición respecto a un punto de referencia externo llamado usualmente *mundo*.

Mediante estas *cámaras adyacentes* y anexionadas de manera rígida es posible la delimitación geométrica de la escena de rayos X, incluidas las distancias anteriormente mencionadas, además de la posición precisa del/de la paciente durante el examen y su volumen. Además, en combinación con una segunda (o más) radiografía(s), es posible aplicar técnicas de *estereoscopía* y reconstrucción 3D y obtener información tridimensional de su anatomía interna, además de otros valiosos datos válidos para complementar el diagnóstico.

En la última década ha acontecido una revolución tecnológica en relación a los dispositivos de captación de contorno, dando lugar a nuevas disciplinas tales como la *detección remota*, la *realidad virtual* o la *realidad aumentada*. Estos nuevos instrumentos conllevan ventajas a las que ya nos hemos ido acostumbrando y se han convertido incluso en cotidianas, tales como la estimación remota de distancias y posiciones, el cálculo de coordenadas, el modelado de superficies, el seguimiento de personas y objetos, la detección barreras y obstáculos, la cartografía y posicionamiento geográfico, entre muchas otras. Los ámbitos de aplicación de los saberes relacionados con la visión por ordenador están ahora al alcance de muchas disciplinas que hasta hace poco se auto-excluían de tales dominios tecnológicos. Entre estas ciencias podemos encontrar a la medicina, la física y otras ciencias básicas.

En lo que concierte a los rayos X, cierto tipo de información geométrica y proyectiva (a excepción del volumen del objeto o persona radiografiada) estaba ya disponible gracias a la intercesión de incómodos y costosos marcos de referencia que contienen *marcadores fiduciarios opacos a la radiación Roentgen*. Esta metodología *heredada* (así como sus *sucesoras* basadas en detectores de contorno que se proponen en este trabajo) radica en el hecho de que un dispositivo de rayos X puede asemejarse a una *cámara pinhole* o *cámara* estenopeica. Una cámara estenopeica es una cámara fotográfica sin lente y que cuenta con un pequeño orificio o *pinhole* por donde entra la luz reflejada por los objetos fotografiados, además un material detector. En el caso de un dispositivo de rayos X, el pinhole es en realidad el *emisor de luz* y coincide estructuralmente con el *ánodo* del tubo de rayos X, que juega también el papel del anteriormente citado punto focal.

El detector en los dispositivos de rayos X *estenopeicos* es la placa radiográfica o el *imaging plate* (en el caso digital). La geometría proyectiva afirma que dados conjuntos de puntos con coordenadas espaciales (3D) y sus correspondientes proyecciones en una imagen, es posible hallar la *ecuación de calibración de cámara* que conecta cualquier otro punto tridimensional en la escena con su localización x, y en la imagen. Es lo que se conoce también con el nombre de *calibración geométrica de cámara*.

El problema con la solución basada en marcos de referencia y fiduciales opacas nombrada anteriormente es que pueden dificultar la movilidad del/de la paciente y/o del sistema, pero sobretodo pueden alterar de ma-

nera significativa la imagen e influir en el diagnóstico alcanzable a partir de la misma. En el Capítulo 3 se estudian y comparan los distintos algoritmos de calibración de cámara pero aplicados al ámbito de los rayos X. Las técnicas propuestas en este trabajo evitan las mencionadas incomodidades para el/la paciente y no interfieren en absoluto en la generación de la placa radiográfica ni en la imagen de absorción final, además de otras ventajas, tales como la posibilidad de guardar registro visual de la escena, adquirir el contorno del/de la paciente o de aplicar protocolos de examen que requieran una gran oblicuidad por parte del sistema de adquisición.

Para combinar geométricamente ambos tipos de dispositivos (sensor de contorno y rayos X) es necesario encontrar con antelación la transformación rígida que los conecta, también conocida como ecuación de la *co-cámara*. Una transformación rígida es una transformación lineal que preserva tamaño y forma, conservando la alineación, el orden y la pertenencia (es decir, las rectas se transforman en rectas y ángulos en ángulos). La búsqueda de esta relación geométrica se detalla en la Sección 4.5 y la Sección 5.4 para el caso de cámaras de visible y de profundidad, respectivamente. En esta fase (y sólo en esta) nos apoyamos en un marco de calibración que incorpora fiduciales detectables por ambos sistemas de imagen (Figura 3.7). Una vez hallada esta matriz de transformación, se dice que ambas cámaras están *registradas*. Tanto en el caso de que la cámara de contorno sea una cámara de video o de profundidad, los marcadores que aparecen en la proyección resultante son fácilmente identificables mediante herramientas de computer vision resumidas en la Section 4.3.3. En el caso de las proyecciones de marcadores opacos a los rayos X, estas son aisladas normalmente de manera manual, aunque es posible aplicar algoritmos de identificación de formas y segmentación sobre la radiografía de calibración. En este trabajo se ha optado por lo primero, aprovechando las mismas herramientas *software* de visualización y diagnóstico del médico-radiólogo. El proceso de hallazgo de la ecuación co-cámara se relata en la Sección 4.3.

Reconstrucción 3D en rayos X

Una vez hallada esta relación de registro entre dispositivos, ya no es necesario el marco de calibración, el cual desaparece de la escena sin perjuicio ni influencia alguna en la(s) radiografía(s) del/de la paciente tal y como se ha anticipado en el párrafo anterior. A partir de este momento, es el detector de contorno el responsable de inferir la geometría de la escena, liberando completamente al sistema de rayos X de esta tarea. Entre los elementos propios de la geometría de la escena que son ahora cómodamente medibles se encuentran, por descontado, las longitudes listadas anteriormente (IOD, SID, etc.). Sin embargo, es posible además inferir otras entidades importantes, tales como el volumen del/de la paciente, sus desplazamientos y los movimientos propios del sistema radiológico entre radiografías consecutivas.

Concretamente, gracias a esta última ventaja (determinación de transformaciones rígidas entre dos desplazamientos) es posible reconstruir tridimensionalmente puntos y distancias internos al/a la paciente mediante técnicas de visión estereoscópica. Para ello sólo son necesarias dos radiografías obtenidas en dos posiciones separadas, ya sea del propio/de la propia paciente o del sistema radiográfico. Esta versatilidad relacionada con los escenarios de aplicación es tratada en la Sección 4.4. Este *seguimiento* o *tracking de la escena* es el que se detalla en el Capítulo 4 y el Capítulo 5 para el caso de que el sensor de contorno sea una cámara RGB y para el caso de una cámara de profundidad, respectivamente. Las cámaras de profundidad consisten en sistemas integrados por una luz láser que es proyectada, formando un patrón conocido, sobre la escena. El reflejo de este patrón es vuelto a ser captado por un sensor CMOS adjunto. A partir de la captura de la deformación del mencionado patrón, es posible determinar información 3D del entorno.

La información 3D obtenida por las cámaras de profundidad es transmitida a otros sistemas informáticos mediante las conocidas *nubes de puntos* o *point clouds*. Una nube de puntos es un conjunto de vértices en un sistema de coordenadas tridimensional. Estos vértices son representaciones de la superficie externa de un objeto (el/la paciente en este caso). Originalmente, las nubes de puntos se utilizaban en la elaboración de modelos tridimensionales en diseño por ordenador (CAD) en la fabricación de piezas, la inspección de calidad en metrología, y muchos otros ámbitos

como animación, y texturización. Desde tiempos recientes han encontrado también un nicho en medicina, como se describe en la Sección 5.1.

En el Capítulo 7 se muestran algunos ejemplos de aplicación de la reconstrucción 3D anatómica en escenarios clínicos reales, tanto con pacientes como con fantomas antropomórficos. En estos ejemplos puede verse claramente cómo es posible reproducir fielmente la longitud de una astilla en el hueso húmero o las distancias entre marcadores fiduciarios emplazados en distintas posiciones dentro de varios de estos fantomas. También se muestra cómo es factible localizar puntos en dos radiografías distintas mediante el trazado de *epipolares acotadas* entre ambas. Los conceptos de línea epipolar y *línea epipolar acotada* se estudian en la Section 4.6.1.

Imágenes densitométricas

En el caso de usar cámaras de profundidad, además de permitir estas el seguimiento e identificación del movimiento en la escena radiológica sin necesidad de ningún tipo de marcador fiducial, también es factible reconstruir volúmenes dentro de la misma. En concreto, es posible dirimir el volumen del/de la paciente si este/esta gira frente al mencionado sensor y se aplican las técnicas y métodos descritos en la Sección 6.1. Esta digitalización del volumen hace uso a su vez del algoritmo KinectFusion (descrito en la Sección 5.5) el cual opera de manera continua sobre las nubes de puntos obtenidas previamente o incluso en tiempo real mediante la ejecución de cálculos en paralelo en la unidad de procesamiento gráfico (GPU). La única complejidad en la aplicación de algoritmo de KinectFusion es la necesidad de eliminar la parte de la nube de puntos referente al fondo de la escena (paredes, decoración, el propio detector, etc.). En la Figura 5.6 se resumen algunos métodos apropiados para ello.

Una vez reconstruido el volumen del/de la paciente, este es trasladado al *punto de vista* del sistema de rayos X. Este hecho permite a su vez la generación de *mapas de longitud recorrida*, es decir, el conjunto formado por todas las distancias recorridas por cada rayo entre el punto en el que

penetran en el/la paciente cuando vienen desde el ánodo y el punto por el que salen del cuerpo hasta alcanzar finalmente un pixel x, y en el detector. La generación de estos mapas se trata en la Sección 6.3.

Con estos mapas de longitud atravesada ya es posible traducir las imágenes de absorción (las típicas obtenidas en la generación de radiografías) por *imágenes de densidad* o *densitométricas* junto con la definición de una nueva *función de transferencia*. La solución más común a este problema era hasta ahora la *absorciometría dual de rayos X*, que consiste en comparar dos imágenes de rayos X tomadas con distinto voltaje. El coste de estos equipos, sin embargo, se incrementa debido a que se requieren dos fuentes de rayos X y/o dos detectores. Otra técnica empleada es el uso de fuentes de rayos X que emitan con al menos dos energías distintas. Estas técnicas reciben el nombre del *imagen de absorciometría de rayos X dual* (*dual-energy X-ray absorptiometry*) o DXA. Un examen DXA es una prueba usualmente indicada para determinar la densidad mineral ósea y diagnosticar, principalmente, desórdenes relacionados con osteoporosis. La técnica de DXA implica el uso de una modalidad radiológica y equipos específicos.

Además de la utilidad inherente a estas imágenes de densidad, en este trabajo también hemos demostrado que estas imágenes contienen objetivamente una mayor calidad y grado de información en comparación con las radiografías de absorción. Las *imágenes densitométricas* cuentan con un rango dinámico más comprimido, lo que se traduce en un realce significativo de los tejidos blandos y en una presencia más balanceada de los tejidos óseo y muscular.

Trabajo en clínica y medida de la calidad de la imagen radiográfica

Respecto a la cuantificación de calidad antes mencionada, el presente trabajo de investigación ha contribuido con una novedosa métrica de cuantificación de la calidad de las imágenes radiológicas basada en los conceptos

de información mutua, entropía, entropía condicional y el filtrado Gabor de imágenes. Un filtro de Gabor consiste en una función gaussiana modulada por una curva sinusoidal a la que se le asigna una determinada frecuencia y dirección, obteniendo una reducción del ruido a la vez que se preserva una dirección de la imagen original. Las funciones de Gabor son importantes en el análisis de texturas, especialmente en la segmentación, ya que diferentes texturas tienden a concentrar su presencia en rangos específicos de frecuencias. Normalmente los filtros de Gabor no se aplican de manera individual a una imagen, sino que se utilizan en grupos de filtros, llamados *bancos*, en los que se permiten diferentes frecuencias y orientaciones.

Concretamente, el método propuesto para la asignación de calidad en radiografías (descrito en el Appendix A) computa la información mutua entre las descomposiciones/bancos Gabor de las imágenes en su versión adquirida en alta calidad (típicamente 16 ó 12 bit) y su versión representada en pantalla (8 bit) de menor calidad. La aplicación de la mencionada métrica ha desvelado claramente que las imágenes densitométricas obtenidas en este trabajo contienen más información que las correspondientes radiografías convencionales.

Como se ha comentado previamente, en el Capítulo 7 se detallan experiencias llevadas a cabo en entornos clínicos reales que ejemplifican las técnicas y métodos descritos anteriormente. En lo que concierne a la imagen densitométrica, se muestran varios ejemplos obtenidos a partir de radiografías e información volumétrica de pacientes que se sometían a un examen clínico.

Gestión de la dosis y reconstrucción de la posición 3D de fuentes radioactivas

En las páginas finales de este trabajo de tesis se presentan de manera más somera dos ámbitos de aplicación nuevos de las técnicas de desarrolladas en el mismo.

En el Capítulo 9 se discute una nueva propuesta para la cuantificación de la dosis recibida durante exámenes radiográficos ordinarios. Está metodología sugerida está basada en el cálculo de la *dosis por volumen* real del paciente (lo cual ya es posible mediante las técnicas y métodos descritos en este trabajo).

El otro ámbito de aplicación tratado en el Capítulo 8 es el relativo a la reconstrucción tridimensional de la posición de fuentes radioactivas mediante *cámaras gamma aumentadas* con dispositivos y métodos análogos a los ya trabajados en el caso del diagnóstico por rayos X. Este trabajo se engloba dentro del proyecto *Gamma Unit Advanced Locator Imager* (GUALI) y es el fruto de la colaboración con el Grupo de Espectroscopia Gamma del Instituto de Física Corpuscular. Una cámara gamma o *gammacámara* es un dispositivo de captura de imágenes (*gammagrafías*), comúnmente utilizado en medicina nuclear como instrumento para el estudio de enfermedades, aunque también puede tener aplicación en la caracterización de radioisótopos y en la gestión y clasificación de residuos nucleares. En el caso de su aplicación en medicina, la radiación procede del propio/de la propia paciente a quien se le inyecta, generalmente por vía intravenosa, un trazador radiactivo. En su vertiente relacionada con la gestión de residuos, son los propios elementos contaminados los responsables de la radiación emitida.

Una cámara gamma se encuentra en un espacio de definición y operación intermedio entre una cámara convencional (fotográfica) y los sistemas de rayos X descritos en este trabajo. Por un lado, una cámara gamma no capta luz (fotones) reflejados de una fuente externa (como ocurre en sistemas que hacen uso de la luz visible: video, fotografía, etc.), sino que son las propias fuentes radioactivas las que emiten fotones que son finalmente detectados. Por otro lado, la relación geométrica entre el pinhole (la entrada de *luz gamma*) y el detector permanece fija. Dicho de otra manera: una cámara gamma posee parámetros intrínsecos y estos son invariables (tal y como estamos acostumbrados para las cámaras de visible).

Concretamente, la cámara GUALI ha sido diseñada para detectar los fotones de 662 keV del isómero nuclear metaestable del Ba-137m, el cual es a su vez fruto de la desintegración del Cs-137. Este último isótopo del Cesio es producto principalmente de la fisión. Tiene un periodo de semidesintegración de 30 años, y decae emitiendo partículas β a Bario-137. GUALI es también capaz del resolver los fotones de 1.17 y 1.33 MeV

del Ni-60 que a su vez es el resultado de la desintegración de un isótopo sintético del Cobalto (Co-60).

En lo que concierte a este trabajo de investigación se ha realizado la calibración mutua (registro co-cámara, introducido anteriormente) y se han llevado a cabo varias experiencias tanto en laboratorio (con fuentes de intensidades y posiciones conocidas) como entornos reales relacionados con la gestión de residuos nucleares. Concretamente, se han desarrollado varios ensayos en la Planta Nuclear José Cabrera (Zorita). Esta instalación se encuentra actualmente en proceso de desmantelamiento, operación que conlleva la clasificación de todos los elementos en cuanto a su tipo y grado de contaminación.

Futuros desarrollos

En la última parte de este trabajo (correspondiente al Capítulo 9) se exponen las conclusiones del mismo y se presentan futuras variaciones, mejoras y escenarios de aplicación que exploten de manera más interesante todavía la combinación de sensores de contorno y detectores de radiación.

Una de las propuestas está relacionada con el uso de la combinación de algoritmos RANSAC (*RANdom SAmple Consensus*), SURF (*Speeded up Robust Features*) y SfM (*Structure from Motion*) para realizar la reconstrucción 3D directamente sobre imágenes captadas por la cámara de visible, sin necesidad de emplear marcadores de referencia. RANSAC consiste un método iterativo para calcular los parámetros de un modelo matemático a partir de un conjunto de datos observados que contiene valores atípicos. En el ámbito de la reconstrucción de escenarios tridimensionales, RANSAC es capaz de recrear la escena (incluidos los objetos radioactivos inspeccionados). En ese sentido, puede lograrse un escenario de aplicación muy parecido al descrito para las cámaras de profundidad: obtener referencias espaciales sin ningún tipo de fiducial ni marcador *ad hoc* (la escena es su propio conjunto de fiduciales).

Otra de las propuestas está relacionada con la simplificación de la

obtención del volumen del/de la paciente durante los exámenes radiográficos (para su posterior aplicación en la producción de imágenes densitométricas). A pesar de que la metodología basada en la rotación del/de la paciente y la aplicación del algoritmo KinectFusion resumida anteriormente resultante bastante apropiada, muchos protocolos de examen exigen la ausencia completa de movimiento. En esta propuesta entrarían en juego varias cámaras profundidad emplazadas estratégicamente para capturar el volumen de la persona examinada de manera completa y simultánea a la adquisición radiográfica. La principal complejidad de esta solución radica en la eliminación del fondo de la escena y en el aislamiento correcto del volumen del/de la paciente. La principal ventaja, a parte de la inmediatez en la captura del volumen del/de la paciente, es la posibilidad de generar imágenes densitométricas con protocolos que requieran estativos horizontales (mesas) o incluso inclinados bajo cierto ángulo.

Además de la generación de imágenes densitométricas, la adquisición del volumen del/de la paciente también permite construir imágenes que realzan ciertos tipos de tejidos, sobre todo el óseo. Para ello, se genera una *radiografía virtual* que acaba siendo sustraída de la original. Esta radiografía virtual es *construida* a partir del volumen en agua obtenido mediante las técnicas anteriormente descritas, aplicando la ley de atenuación de Beer-Lambert (que relaciona la absorción de la luz y las propiedades del material por las que transcurre), el coeficiente de atenuación lineal para el agua y un haz monoenergético. El realce de huesos puede contribuir a la estimación de la densidad mineral ósea y en ese sentido puede suponer una alternativa a la modalidad de DXA citada anteriormente. Esta última aplicación se ha explorado también brevemente en la Sección 6.6 y la Sección 7.4.

Conclusiones

Una de los principales resultados de investigación obtenidos en este trabajo de tesis ha consistido en la fusión de la geometría del mencionado *entorno* o *escena* que rodea al estudio radiológico con la propia informa-

ción obtenida fruto el mismo (la gammagrafía, radiografía, etc.). En la investigación desarrollada se ha conseguido que, a partir de dos o más radiografías sea posible: combinar espacialmente y en el mismo sistema de coordenadas, varias de estas imágenes inicialmente *aisladas*, identificar claramente zonas y puntos de interés comunes, calcular distancias, ángulos y dimensiones tanto de órganos como de otros corpúsculos. Asimismo las técnicas expuestas permiten dar cuenta de pequeñas variaciones de la posición del/de la paciente (entre pruebas consecutivas o fechas distintas, por ejemplo), posición del sistema de imagen con respecto al mismo, etc. Una combinación acertada de dos o más radiografías, puede ser tan reveladora desde el punto de vista clínico como lo es un TAC y al mismo tiempo conllevar menos de un 5 % de la dosis aplicada al/a la paciente en comparación con esta modalidad.

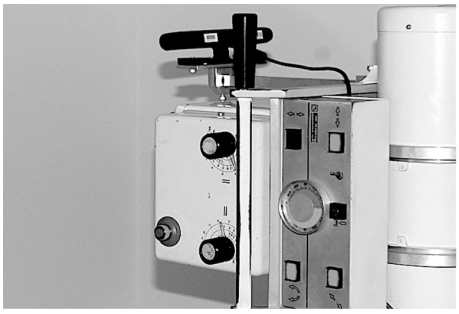
Conocer esta información con detalle y su historia (variabilidad espacial y temporal) puede llegar a ser relevante a la hora de dilucidar y revisar las condiciones de ejecución de una prueba. Las técnicas descritas también pueden aplicarse a otros procesos y sectores en los que estén presentes la inspección y control mediante rayos X (vigilancia, seguridad, calidad, análisis forense, etc.).

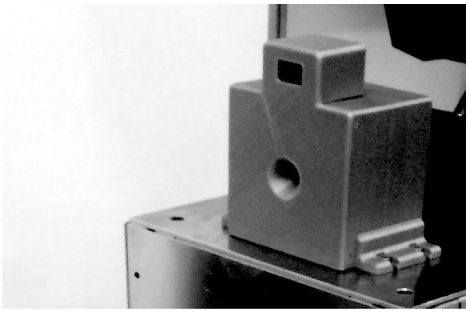
La combinación de información volumétrica y radiográfica del propio/de la propia paciente ha posibilitado también el avance de un nuevo de tipo de imagen radiológica: la imagen de densidad o densitométrica. En la imagen densitométrica, el camino seguido por los rayos X es *corregido* por el volumen real atravesado. El resultado es una *radiografía densitométrica* en lugar de una imagen de absorción que rinde mejor cuenta de cada tipo de tejido y da mayor relevancia a los tejidos blandos. La imagen densitométrica puede llegar a ocupar por tanto un espacio muy útil entre los mecanismos de diagnóstico al alcance del facultativo.

En el ámbito del proyecto GUALI, las técnicas exploradas han permitido la realización de mapas de radiación con una precisión centimétrica. Es decir, por mediación de estos últimos desarrollos no sólo es posible identificar y catalogar fuentes radioactivas, sino que además es posible registrar su posición tridimensional con una mucha exactitud.

1

Motivation





In contrast with other radiological modalities (such as the *computerized axial tomography* or CT) and well-defined experimental setups related to nuclear physics and particle tracking (gamma detectors, accelerators, etc.), many radiation detection devices or *imaging systems* (i.e., those used in medical imaging and/or radiation protection) do not take into account the geometrical information concerning the scene in which they operate. The main goal of this thesis work is the development of the necessary methods and techniques to provide this information to whichever detection device is used and, in a general way. This *augmentation* (concept borrowed from the world of *computer vision*) is achieved through the interplay of an external environment recognition device.

In the context of this thesis work, a *scene* or *environment* entails, not only the outline of the location, room or surroundings where any sort of *photon intensity measurements* or *imaging process* takes place, but also the proper element under examination (person, contaminated object, radioactive source, etc.), including its/his/her position, orientation and volume relative to the imaging system or a fixed point in space. That is the case of *general purpose* X-ray imaging equipment or portable radiation measurement devices used to evaluate, for instance, environmental γ or β emissions. As it is demonstrated in this work, the aforementioned scene's geometry can complement o *augment*, in a very significant way, the inherent information obtained by these imaging systems. Similarly, in the scope of this research project, a device A *augments* another device B when A provides B with accurate spatial references, allowing for instance, 3D reconstruction for B, A+B image overlay (registration), image stitching for B, etc.

In order to achieve the aforementioned goal, some methods and techniques around the determination of the *spatial setting* have been tested and explored within the following areas of application:

- augmentation of primary care X-ray imaging systems,
- three-dimensional reconstruction of the anatomy of the patient under examination using ordinary radiographs,
- derivation of new transfer functions that enable the generation of densitometric images from X-ray absorption ones and the patient's volume,
- assessment of 3D coordinates to radioactive sources and the received dose.

The present research mainly focuses on the augmentation of *conventional* X-ray imaging systems through the interplay of an *external positioning* or *scene-delimitation device* (i.e., a video camera, depth sensor, etc.) like the one shown in the image above. Nevertheless, other particle detection systems (i.e., γ cameras) are also explored. The main advantage of a *dual-camera assembly* is the possibility of geometrically determining the radiographic (or radioactive) scene with accuracy and being able to map both types of information.

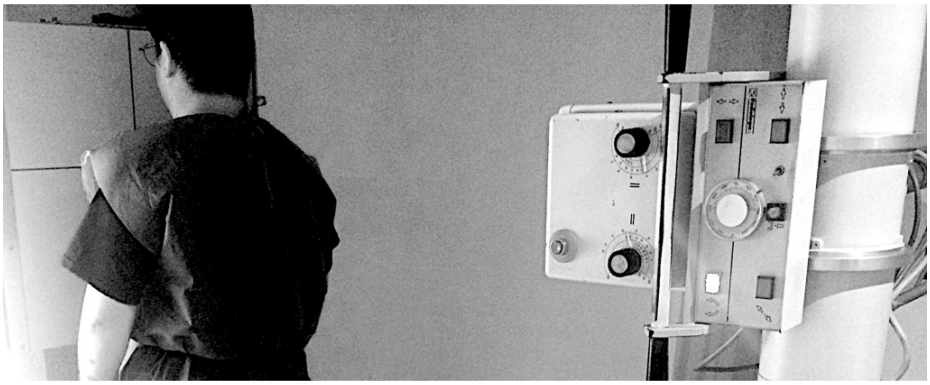
In the context of medical imaging, the reason for choosing this specific equipment type (general purpose X-rays) is their undeniable presence, not only in healthcare, but also industrial domains. The quantification of the available geometrical information surrounding ordinary X-ray examination rooms opens many interesting possibilities which were so far limited to more complex radiological modalities such as CT scanners (i.e., anatomical 3D reconstruction). In contrast with tomographies and as highlighted in Chapter 3, the geometry derived in conventional X-ray imaging can be very variable from session to session and is very rarely (or never) registered, stored or taken into account during the imaging process. More specifically, in this thesis work, we begin by establishing methods and materials (involving the use of the aforementioned external *environment recognition* devices) to account for this, *so far ignored,* scene geometry. This information will later allow us to derive 3D relations from plain radiographs, which is discussed in Chapter 4 and Chapter 5. These 3D reconstruction capabilities are in turn based on an earlier geometrical calibration phase of the imaging system, examined in Chapter 3.

In Chapter 6 we present a technique to obtain densitometric X-ray images from plain radiographs in combination with the patient's volume. The foundations of this technique rely on the theoretical and practical background developed in previous chapters. Next, in Chapter 7 we present some experiments and tests carried out with anthropomorphic phantoms and patients in real clinical setups around the techniques and procedures introduced in this document. In Appendix A, we outline a novel methodology to assess the quality of X-ray images. This metric will enable the objective assessment of the amount of information contained in conventional radiographs and compare it against the new density-based ones.

As stated above, the concept of environment recognition is also applied to particle detectors, notably portable γ-ray cameras. This scope of application is addressed in Chapter 8, where similar steps, methods and tools to those previously applied in the augmentation of X-ray imaging scenarios are now implemented to radiation detection in nuclear waste management locations. Finally, we will also address new approaches for dose assessment based on the measured target's volume.

2

X-ray physics and conventional X-ray settings



An X-ray diagnostic or inspection setting is nothing else but a radiation detection equipment used for clinical or examination/scanning purposes. This type of devices share the main *weakness* highlighted in Chapter 1, that is, they do not take into account the geometrical setting during the imaging process. For this reason, together with the fact of its clinical and industrial relevance, X-ray imaging is one of the main subjects of research in this thesis work.

Conventional X-ray diagnostic has become over the last hundred years a key component of the diagnostic toolbox. These radiological settings may seem clinically *humble* and technically *poor* when compared against

more capable modalities such as computerized tomography, fluoroscopy or tomosynthesis. However, ordinary diagnostic X-ray imaging is still the primary radiological examination required by physicians for many reasons, ranging from the pure budget-related ones to those linked to the limitation of the patient's dose.

In spite of its popularity and unquestionable usefulness, radiographs are the result of radiation absorption operation and as such, all geometrical information is lost in the process. Similarly, a radiograph cannot reveal the intensity of the tissues and materials traversed by the Roentgen radiation.

In this chapter we summarize the physical processes related to conventional X-ray imaging, as well as the historical evolution and basic components of ordinary radiology settings. X-ray procedures and protocols and their geometrical limitations are also discussed. We put a special focus on the absence of *geometry logs* that account for accurate distances, angles, etc., which are ignored during examinations. As next chapters will demonstrate, this information can add significant value to the diagnosis process.

2.1 Overview and limitations of X-ray imaging

X-rays are commonly used in non-destructive analysis since the early 20^{th} century, both in clinical diagnosis and object inspection. They have led to major technological advances in the development of detectors and production methods. The radiographic images can be obtained by placing a natural or artificial source of γ- or X-rays, having them pass through the examined object/patient, and finally detecting them at the other side. Absorption differences due to the nature of the material and material thickness generate an *image of intensities* on the detector. The fundamental difference between X-rays and γ-rays is that the first are derived from a source which generates a continuous spectrum of photons (*polychromatic*), while the latter comes from natural *de-excitation* of atomic nucleus or deep electronic layers in the atom and it is usually *monochromatic*. The use of

radiographic X-ray sources has the advantage that intensity can be modulated and the emission can automatically be cut, while with γ sources it is not possible because of the nature of the process that follows a varying intensity over time:

$$I = I_0 t_0 e^{wt} \tag{2.1}$$

where w is the inverse of the half-life of the isotope, $I_0 t_0$ is the source intensity measured at time t_0 and t is the time where the current measurement is performed. This feature allows applications such as measuring the diffusion of tracers in live systems and allows applications such as the SPECT (cameras that detect the projection of the isotope in a plane) or PET (geometrically paired cameras that detect coincidences of photons from positrons produced by the decay of the nucleus). There are two ways commonly used to obtain the intensity of the source that passes through the scene. In the first, the total intensity is measured without discriminating the individual energy of every photon. In the second, the individual energy of every photon is counted and measured. This latter type of techniques is applied in the aforementioned PET and SPECT applications.

When *diagnostic X-rays* interact with matter they are partly *absorbed* and partly transmitted. The probability of interacting with the material depends on the electron density which is a function that depends on the incident photon energy, and on the elemental composition (Z or atomic number) of the material. This *attenuation principle* is further tackled in Section 2.5.5.

In clinical diagnostic environments, for example, the geometric calibrations are performed by means of mannequins with known densities. Once the equipment is calibrated, the defined parameters are used to highlight different injuries and to analyze the state of bones and tissues, or to locate foreign bodies. In some cases to calibrate the detector response, easily detectable elements with different thicknesses are included in the scene either manually or automatically to establish a link between the measured intensity in the pixel of the detector and the material thickness. Another technique is the use of X-ray sources that emit at two different energies (multi-voltage). These techniques rely on the fact that X-ray absorption in the medium is different and depends on the electron density of the material and the energy of the X-ray beam. The fact of comparing

the images of the same scene acquired at different voltages enables the obtention of *densitometric information* about the examined object/patient. However, this method is inaccurate because the measurements obtained are relative. It also requires extra steps, extra time for the two X-ray snapshots and/or two examination devices devices.

2.2 Brief history of X-rays

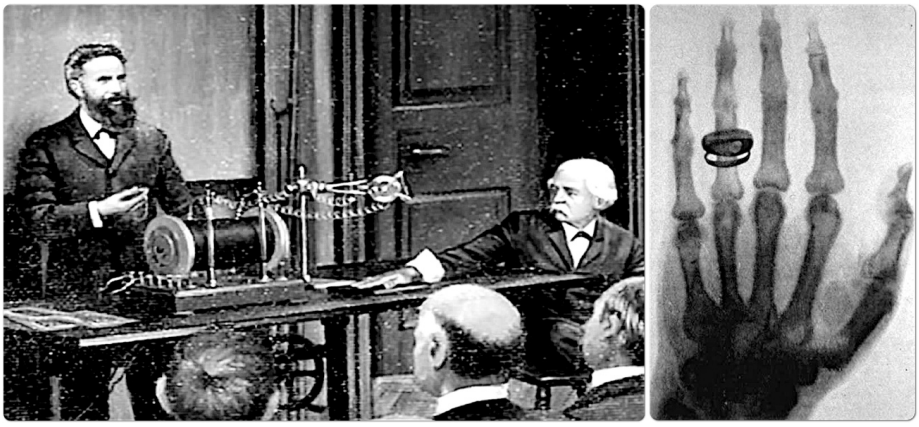
As on other occasions in the history of science, X-rays were were *accidentally discovered*. During the 1870s and 1880s, many physics laboratories were interested in the investigation of cathode rays through evacuated glass tubes known as *Crookes tubes* (after Sir William Crookes). The tube that bears his name was the previous version of modern fluorescent lamps and X-ray equipment. There were many different types of these tubes, and most of them were already capable of producing X-rays. Wilhelm Roentgen was experimenting with a type of Crookes tube when he *fortuitously* discovered X-ray the radiation.

On November 8, 1875, Roentgen was working in his physics laboratory at Würzburg University in Germany. He had darkened the room and completely enclosed his Crookes tube with black photographic paper so that he could better visualize the effects of the cathode rays. A plate coated with platinate(2-), tetrakis(cyano-kC), barium (1:1), (SP-4-1), or *barium platinocyanide,* for short (a type of fluorescent material whose molecular formula is C4N4 Pt.Ba) happened to be lying on a bench some meters away from the tube. Because of the enclosing black paper, no visible light was escaping from the setup. However, Roentgen noted that the barium plate was somehow *glowing a mysterious light.*

The intensity of the glow increased as the plate was brought closed to the tube, leaving undoubtedly what was responsible of the glow. Roentgen began investigation this light by interposing several materials: from wood, to aluminum, and even his own hand. His initial investigations were very thorough and intense, allowing him to present his experimental results to

the scientific community in the very same year 1895 (just a few weeks after the first *accidental* discovery). This presentation took place on the 23 January 1895 in the Physical Medical Society. During the presentation, Roentgen took a radiograph of the hand of the renown Albert von Kölliker. At the end of the demo, the crowd was absolutely amazed, claiming they had not seen anything like it. Professor Kölliker was the one who decided that the new type of ray should be name after his discoverer.

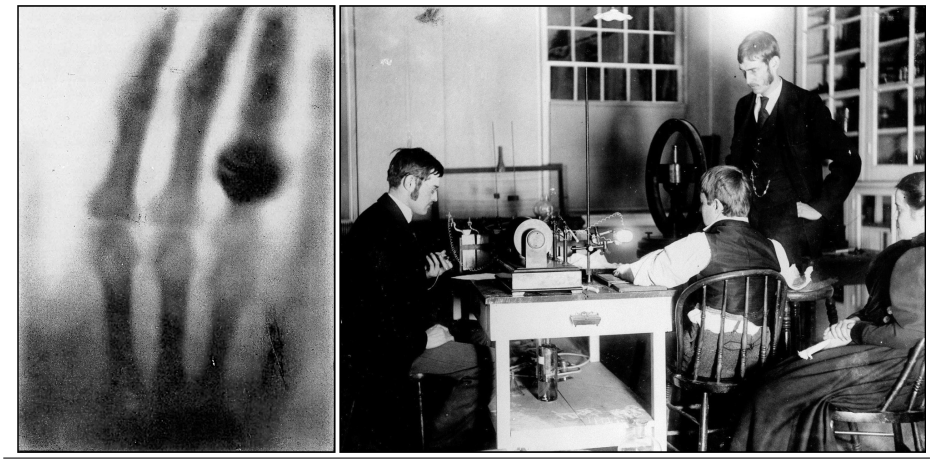
Figure 2.1 First X-ray public demonstration in the Physical Medical Society, including the first public radiograph (Kölliker).



For his research, he received the first Nobel price in physics in 1901. Among his results, he produced the first radiograph in 1876. It was an image of his wife's hand. The first official X-ray examination in the United States took place in 1896, in the physics laboratory at Dartmouth college (Fig. 2.2).

Since then, X-ray-based clinical examinations are omnipresent. Modern radiology consists of mainly five types of X-ray modalities: radiography, fluoroscopy, mammography, positron emission tomography or computerized tomography (represented in Fig. 2.3). Plain radiography uses film or solid-state image receptor and usually and X-ray tube mounted on a track that allows the tube to be moved in any direction. Fluoroscopy, in turn, is conducted with an X-ray tube located under the examination table. The radiologist is then provided with live images on a remote monitor. A computerized tomography uses a rotating X-ray source and detector array. A *volume of data* is acquired so that fixed images can be reconstructed in any plane: sagittal, transverse or oblique.

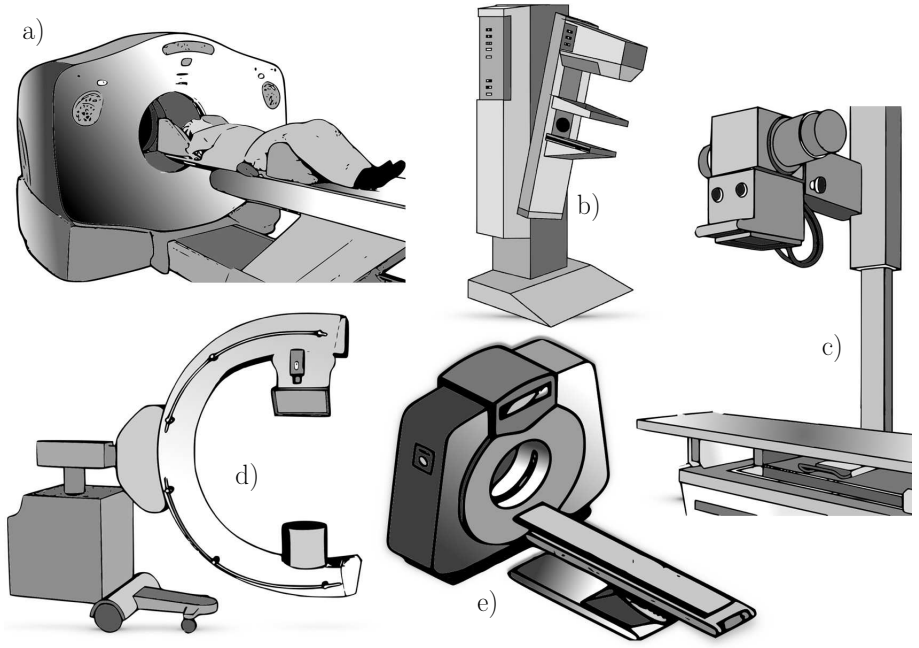
Figure 2.2 First ever produced X-ray images. The hand on the left belongs to Mrs. Roentgen. The photograph on the right records the first medical X-ray examination in the United States. A young patient, Eddie McCarthy broke his wrist while skating and Professor E. B. Frost (from the Dartmouth college) and his brother, Dr. G. D. Frost (medical director at the Mary Hitchcock hospital) examined it with an apparatus assembled by Professor F. G. Austin in this physics laboratory at Reed Hall (Dartmouth college) on February 3, 1896.



Today, voltage and current are supplied to an X-ray tube through a rather complicated set of circuits, but in Roentgen's time, only simple static generators were available. These units could provide currents of only a few mA and voltages up to 50 kVp in contrast with current configurations (i.e., in the order of 1×10^3 mA and 150 kVp).

Other three inventions propelled the use of X-rays into the 20^{th} century. In 1907, H. C. Snook introduced a new high-voltage alternating current power supply. In 1913, William D. Coolidge unveiled his hot-cathode X-ray tube to the medical community. It consisted on a vacuum tube that allowed the independent selection of a beam's intensity and energy with great accuracy. The era of modern X-ray imaging is dated from the matching of the Coolidge tube with the Snook transformer. The third invention corresponds to so-called *bucky*, after Gustav Bucky, who invented the stationary (and a later moving) grid. Simultaneously, the American radiologist Hollis E. Potter patented a similar device, and for this reason it is called the Bucky-Potter grid. This grid is positioned on the opposite

Figure 2.3 Main radiological modalities: PET scanner (a), mammogram scanner (b), ordinary radiography (c), fluoroscopy (d) and computerized tomography (e).



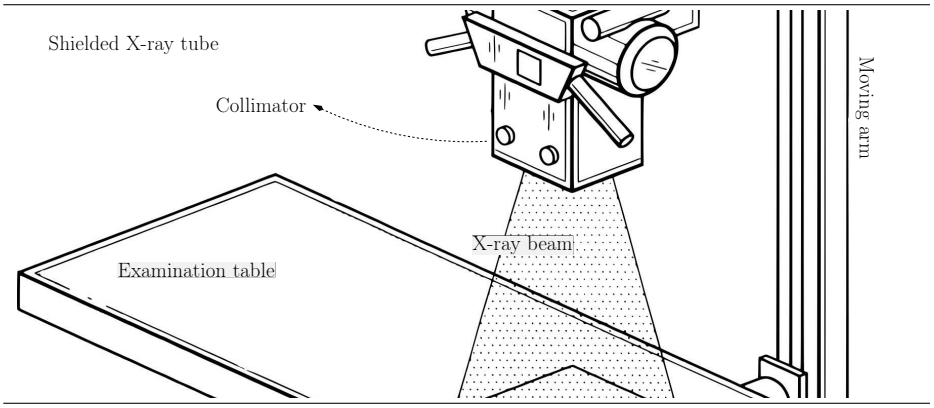
side of the patient from the X-ray tube to reduce the quantity of scattered rays that reach the detector. The Bucky-Potter mechanism allowed the obtention of sharper and higher quality images.

2.3 Main components of conventional X-ray examination rooms

The *general purpose* examination room contains a radiographic imaging system with an X-ray tube, an operating console, a radiation detector and a high-voltage generator. In some types of X-ray imaging systems, such as dental (Fig. 2.7) and portable machines, these components are housed

very compactly.

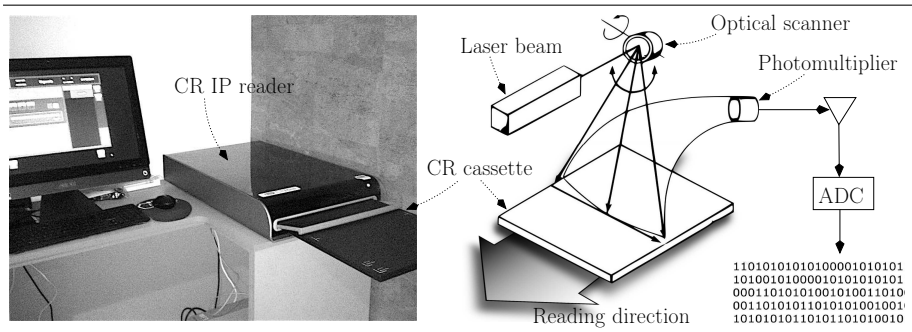
Figure 2.4 Typical X-ray examination setting (for horizontal examination protocols, i.e., *decubitus*).



With most general systems, however, the X-ray tube and the detector –also called *imaging plate* (IP) or *flat panel detector* (FPD)– are located in the examination room (Fig. 2.5). A FPD consists of a thin-film-transistors (TFT) array of detectors which are made of amorphous silicon. These individual detector elements are called *dexels*. Each dexel has a light sensitive area and a hidden area where the electronic components are located. TFT arrays are employed in two types of flat panel detectors: indirect and direct. The indirect detectors use a scintillator to convert the X-rays to visible light and then a photodiode to capture the visible light and produce charge that is read by the TFT array. Certain scintillators can be grown in *columnar crystals*, which act as a pipe for the visible light. These light pipes help reduce the lateral spread of the light, thus preserving spatial resolution. A typical scintillator with light pipes is cesium iodine (CsI). The direct X-ray conversion detectors use a semiconductor material that makes use of electron-hole pairs that are proportional to the incident X-ray intensity. Absorbed energy is directly converted into charge in the detector. Amorphous selenium (a-Se) is the semiconductor most widely used [105]. The a-Se is layered between two surface-area electrodes. Ion pairs are collected under applied voltage across the electrodes. The electric field prevents lateral spread of the charge in the semiconductor, resulting in a high spatial resolution. Due to its low atomic number, the Se layer has to be thick to improve detection efficiency. Digital FPD have several advantages over the conventional screen-film detectors. In

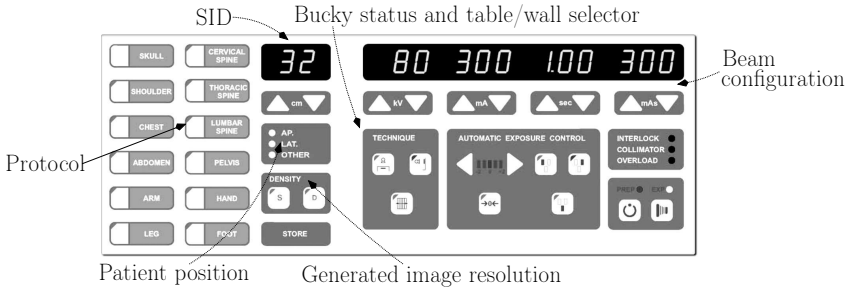
the screen-film radiography, the film type and manufacturing process dictate the required exposure to the technician. A given screen-film has a fixed speed, and if an image is over or underexposed, the image has to be re-exposed. Digital detectors have wider latitude and a variable speed. With this wider dynamic range, digital detectors are more forgiving for an over or underexposed image. Digital images can conveniently be stored in a computer and be assessed and shared at a later time for processing or follow up. Digital images also bypass the need for any chemical processing, which increases the time and cost efficiency. The images are also available for immediate previewing.

Figure 2.5 Example of CR scanner and imaging plate (IP) very commonly used in diagnostic X-ray settings.



The operating console (Fig. 2.6) is usually located in an adjoint room with a protective barrier separating the two. It consists of an on/off control and controls to select kVp, mA, and exposure time.

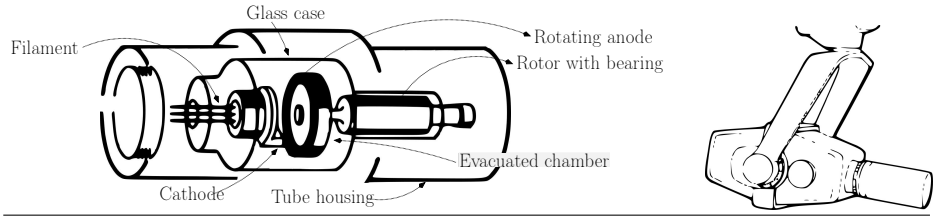
Figure 2.6 Example of an X-ray management/administration console.



The high-voltage generator may be housed in an equipment cabinet positioned against the wall, as close as possible to the X-ray tube. It provides power to the X-ray tube in three possible ways: single-phase, three-phase and high-frequency power.

The X-ray tube is a component of the X-ray imaging system rarely seen. It is contained in a protective housing and is therefore inaccessible. The housing reduces the intensity of *leakage radiation* to less than 1 mGy_a/hr at 1 m. There are two primary parts: the cathode and the anode. Each of these is an electrode, and any electronic tube with a two electrodes is a *diode*. An X-ray tube is therefor a special type of diode. The cathode is the negative side of the X-ray tube and consist of two primary elements, a filament and a focusing cup. The filament emits electrons when it is heated by thermionic emission (i.e., Edison effect). It is usually made of thoriated tungsten. Tungsten provides higher thermionic emission than other metals and the addition of some thorium enhances its efficiency and contributes to prolonging the equipment's life. The focusing cup in which the filament is embedded, confines the beam of electrons to small area of the anode.

Figure 2.7 Left: main components of an X-ray tube. Right: compact dental X-ray tube.



The anode is the positive side of the X-ray tube. There are two types of anodes: rotating and stationary. Stationary anode X-ray tubes are used in dental X-ray imaging systems (Fig. 2.7), some portable imaging systems and other special-purpose units in which high tube current and power are not required.

General purpose X-ray tubes (Fig. 2.8) use the rotating anode because they must be capable of producing a high intensity X-ray beam in a very short amount of time. The anode serves three functions. It is an electrical conductor given that it receives the electrons emitted by the cathode and

conducts them through the tube to the connecting cables and back to the high voltage generator. The anode also provides support for the *target*. The target is the area of the anode *struck* by electrons from the cathode. The target consists of a tungsten alloy embedded in the copper anode. In a rotating tube, a ring of finite area in the disc is the target. Alloying the tungsten with rhenium gives it added mechanical strength to cope with the stresses caused by high speed rotation and repetitive thermal expansion and contraction. Tungsten is the material of choice for the target for general radiography for three main reasons:

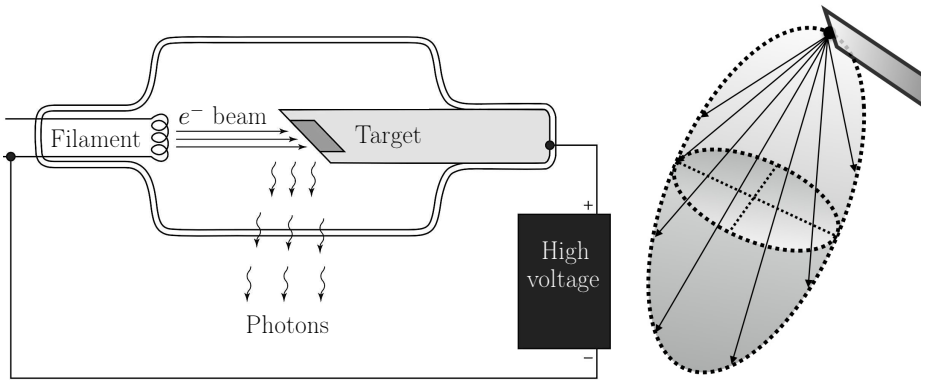
- it has a high atomic number ($Z = 74$) which results in the production of high energy X-rays,
- it has a thermal conductivity almost equal to that of the copper, which makes it very efficient for dissipating the heat produced, and finally,
- it has a high melting point (3400 °C).

The focal spot is the area of the target from which X-rays are emitted. X-ray imaging requires small focal spots because the smaller the focal spot, the better spatial resolution of the radiography. For this reason, the anode is angled. This is primarily due to reduce the focal spot and to limit the self attenuation of the photon beam due to the target. For low electron energies (in the clinical range), bremsstrahlung rays are emitted almost isotropically. As the energy of the incident electrons increases, the distribution angle narrows (Fig. 2.8).

The anode must be a good thermal dissipater too. When the *projectile* electrons from the cathode interact with the anode, more than the 99% of their kinetic energy is converted into heat, which has to be dissipated as quickly as possible. Heat increases (directly) with increasing X-ray tube current (doubling the X-ray tube current doubles the heat produced). It also increases with larger kVp (for diagnostic ranges). For this goal, copper, molybdenum and graphite are the most common ones. In a rotating anode, the electron beam interacts with a much larger target area (500 times more area compared with a stationary anode) and the heating of the anode is not confined to a small and unique spot.

The primary function of the X-ray tube is accelerating electrons in order to produce a consistent X-ray beam, which is something remarkable.

Figure 2.8 Operational representation of an X-ray tube and depiction of the angular distribution of bremsstrahlung photons emitted from the anode.

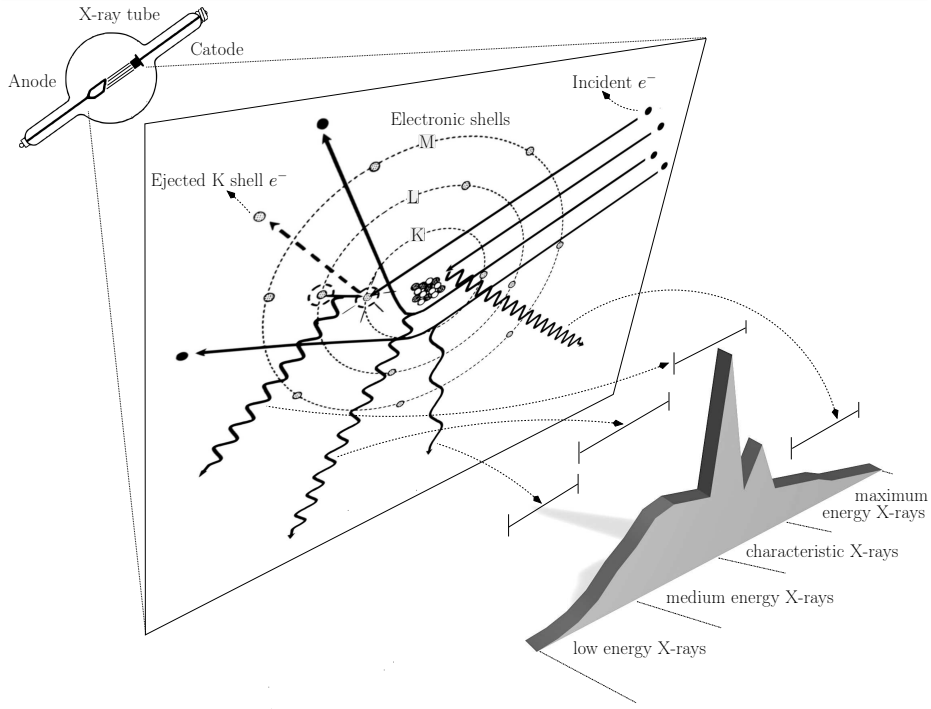


It conveys to the target an enormous number of electrons at a precisely controlled kinetic energy. For instance, at 100 mA, a total of 6×10^{17} electrons travel from the cathode to the anode every second. If the system operates at 70 kVp, each electron arrives at the target with a maximum kinetic energy of 70 keV, which implies speeds faster than half the speed of light. They basically interact with the *outer-shell* electrons but not with enough energy to ionize them. Rather, these outer-shell electrons are raised to an excited state or energy level, from where they drop back to their normal level, emitting infrared light in the meantime. This constant excitation and relaxation or outer shell electrons is the responsible for most of the heat generated in the anode. As state above, only about 1% of the colliding electrons kinetic energy is used for the production of X-ray radiation. This rate is independent of the tube current (regardless of the mA selected, the efficiency of the X-ray production remains the same). In contrast with the tube current, the production of X-rays increases with kVp. For instance, at 60 kVp, only 0.5% of the electron kinetic energy is converted to X-rays. At 100 kVp, this rate reaches approximately 1%. At 20 MeV, almost 70% is converted (of course, this specific case falls outside the common clinical scenario).

2.4 Production of X-rays

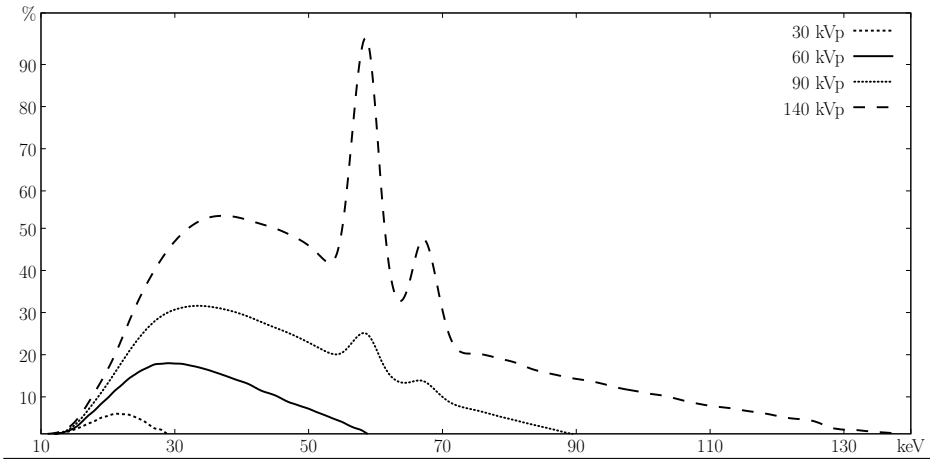
When the electrons produced and accelerated in the X-ray tube collide with the atoms in the anode, several physical process can happen. These processes are graphically summarized in Fig. 2.9. An X-ray beam (in ordinary radiography production) consists in a bundle of photons with an associated electromagnetic field and energy in the order of the 0 to 160 keV and a wavelength from 0.01 to 10 nm (generally smaller than the diameter of an atom) and a frequency from 3×10^{16} to 3×10^{19} Hz.

Figure 2.9 Production of X-rays. Four electrons are presented, along with their collisions with the target, which are translated to X-rays of different energies.



At the end, taking into account the sum of all the processes reviewed below, we obtain a characteristic X-ray energy spectrum, which depends on the peak energy of the colliding electrons (Fig. 2.10).

Figure 2.10 Characteristic X-ray spectra at different electron kVp.



2.4.1 Rayleigh Scattering

In Rayleigh scattering (also referred to as *coherent* or *classical* scattering), the incident photon interacts with the total atom, as opposed to individual electrons as in Compton scattering (tackled in Section 2.5.1) or the photoelectric effect (discussed in Section 2.5.2). This interaction occurs mainly with very low energy X-rays (15 to 30 keV).

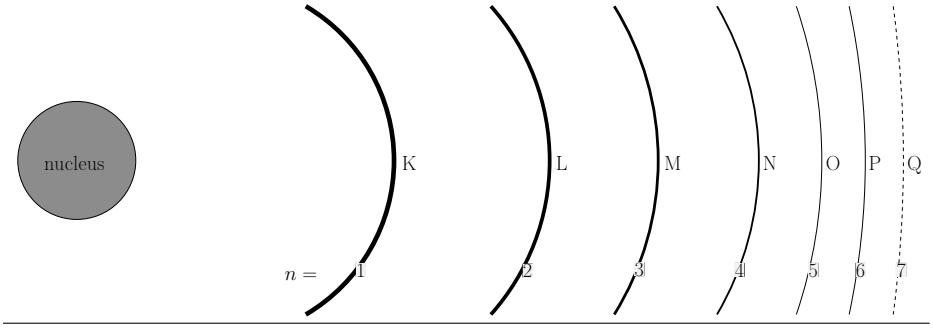
During Rayleigh scattering the electrons in the scattering atom oscillate as *a whole* and in phase. The electron cloud subsequently radiates this energy, emitting a photon of the same energy but with a different angle. The atom is not ionized (electrons are not ejected). The average scattering angle decreases inversely proportional to the incident X-ray energy. In X-ray imaging, detection of the scattered photons will not contribute to image quality, quite the contrary. Fortunately, Rayleigh scattering only accounts for less than 5% of the X-ray interactions in the range of 70 – 150 keV and 10% for interactions at 30 keV or under.

2.4.2 Characteristic radiation

If the projectile electrons interact with inner shell electrons of the target atom rather than with outer shell ones, characteristic X-rays can be

produced. Characteristic X-rays result when the interaction is sufficiently *violent* to ionize the target atom through total removal of an inner shell electron. In this scenario, a temporary *electron void* is produced in that specific inner shell. This is a highly unnatural state for the anode atom and nature corrects this situation by making an outer shell electron to fall into and fill the *electric vacant* in the inner shell. The main electronic shells and vacancies are represented in Fig. 2.11.

Figure 2.11 Electron shell vacancies that contribute to characteristic X-ray radiation.



The transition of an orbital electron from an outer shell to an inner shell is accompanied by the emission of an X-ray. This X-ray has the energy equal to the difference in the binding energies of the orbital electrons involved. The only transitions that have a relevance in diagnostic X-rays are those from outer shells to the K-shell. The K-shell is the electronic orbital with two electrons and with a principal quantum number equal to 1, an azimuthal quantum number equal to 0 and two possible spin values ($+1/2$ and $-1/2$). The rest of shell transitions represent a range of energies too low to be of any use for diagnostic and are usually filtered away. Because the electron binding energy for every element is different, the energy of characteristic X-rays produced is also different. The effective energy of characteristic X-rays increases with increasing atomic number of the target element.

$$\mathrm{E_{Xray}} = \mathrm{E_{vacant\,shell}} - \mathrm{E_{original\,shell}} \tag{2.2}$$

The angular distribution of the characteristic radiation has been studied by [88].

2.4.3 Bremsstrahlung radiation.

Apart from the characteristic radiation, another type of interaction consists in the electron losing its kinetic energy when interacting with the nuclear field of a target atom when passing relatively close to its nucleus. The closer the electron gets to the nucleus, the more it is influenced by its electric field. This field is very strong given that it contains many protons. As the electron passes near the nucleus, it is slowed down and changes its course with less kinetic energy.

This loss of kinetic energy is translated to Roentgen radiation. These types of X-rays are called *bremsstrahlung* X-rays. They can be considered as radiation that results from the braking of electrons by the nucleus. An electron can lose all, none or any intermediate level of its initial kinetic energy. In the diagnostic range, most X-rays are bremsstrahlung X-rays. For instance, at 100 kVp, only 15% of the X-ray beam corresponds to characteristic rays. In contrast with characteristic X-rays whose originating electrons need a specific minimal kinetic energy, bremsstrahlung X-rays can be produced at any electron energy and their range of energies form a continuous spectrum. The bremsstrahlung energies range from zero to a peak and back to zero (maximum energy will be equal to the selected kVp of operation). This is referred to as the *continuous X-ray spectrum*. The majority of the *clinically useful* X-rays are part of this continuous spectrum.

2.4.4 Other factors affecting the X-ray emission spectrum

The number of X-rays emitted from an X-ray tube can be determined by adding together the number of X-rays emitted at each energy over the entire spectrum. The general shape of an emission spectrum is always the same, but its relative position along the energy axis can change.

- If we double the current, twice as many electrons will flow from the cathode to the anode. This operation will produce twice as many X-rays at every energy. In other words, the X-ray emission spectrum will change in amplitude but not in shape.
- If the kVp is raised, the area under the curve increases to an area ap-

proximating the square of the factor by which kVp is increased. The relative distribution of the emitted X-ray energy shifts in average to the right, but the maximum emission energy remains numerically equal to the chosen kVp. A change in kVp affects both the amplitude and the position of the X-ray emission spectrum. In the diagnostic range, a 15% increase in kVp is equivalent to doubling the number of projectile electrons.

- X-ray number is proportional to the selected current.
- X-ray number is proportional to the selected peak voltage.
- X-ray quantity is inversely proportional to the square of the distance from the source (SID)
- When SID is increased, current must be increased by SID^2 to maintain constant exposure to the image receptor.

2.4.5 Effect of filtration

Adding filtration reduces the X-ray beam intensity (less X-ray photons) while increasing its average energy. Filtration absorbs more efficiently low energy rays. Therefore, the bremsstrahlung X-ray emission spectrum is reduced further on the left than on the right. Filtration is often called *hardening* because of the increase in the average energy. Filtration also reduces patient dose. At least, every X-ray equipment has an inherent filtration which is related to the X-ray tube design (glass window, housing, dielectric oil bath, etc.).

2.4.6 Effect of the target material

The atomic number of the target affects both the number and the energy of X-rays. As the atomic number of the target material increases, the efficiency of the production of bremsstrahlung radiation increases and high energy X-rays increase in number to a greater extent than low energy X-rays. The characteristic spectrum is shifted to the right. This phenomenon

is a direct result of the higher electron binding energies associated with bigger atomic numbers.

2.5 X-ray interaction with matter

X-rays have very short wavelengths (10^{-8} to 10^{-9} m). The higher the energy of an X-ray, the shorter is its wavelength. Consequently, low energy X-rays tend to interact with whole atoms, which have diameters of approximately 10^{-9} to 10^{-10} m. Moderate energy X-rays generally interact with electrons and high energy electrons usually interact with the nuclei. X-rays interact at these various structural levels through five mechanisms: coherent scattering, Compton scattering, photoelectric effect, pair production and photo-disintegration. Photoelectric effect and Compton scattering are of special relevance in diagnostic X-rays and are tackled in this section.

Radiography is performed with an X-ray source on one side of the patient and a Roentgen radiation detector on the other side. A half second pulse is emitted by the X-ray tube. A fraction of this beam interacts with the patient. However some of the X-rays pass through the patient and reach the detector, where a radiography is finally formed. The initial distribution of rays entering the patient is severely altered. A fraction of X-rays is removed from the beam (i.e., attenuated) by scattering and absorption processes within the atoms of the patient. The contribution to the attenuation of tissues such as bone, fat, and air inside the patient are substantially different, resulting in a new distribution of X-rays *emerging from the patient*. The radiographic image is nothing else but a picture of this X-ray distribution (or the absence of it).

2.5.1 Compton scattering

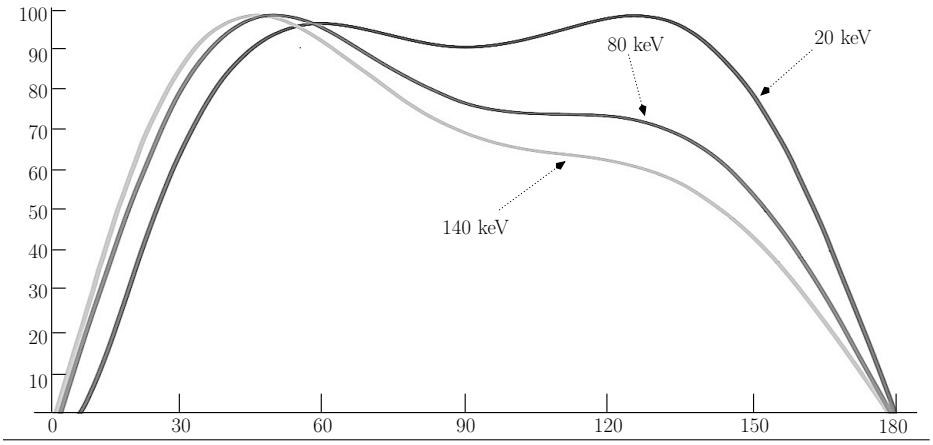
When X-rays interact with the outer shell electrons they are not only scattered, but its energy is also reduced and the target atom is ionized as

well. This type of interaction is called *Compton scattering*, also known as *inelastic* or *nonclassical scattering*. The energy of the Compton-scattered photons is equal to the difference between the energy of the incident X-ray and the energy of the scattered electron. In turn, the energy of the ejected electron is equal to its binding energy in addition to the kinetic energy with which it leaves the atom. During Compton scattering, the energy is divided between the scattered X-ray and the Compton electron. Both the scattered electron and the X-ray may still have sufficient energy to produce additional ionizing processes before they lose all their energy.

$$E_{\text{compton}} = \frac{E_{\text{original}}}{1 + \frac{E_{\text{original}}}{E_e}(1 - \cos\theta)}$$

Ultimately, the scattered X-ray is *photoelectrically absorbed*. The Compton electron loses all of its kinetic energy and drops into a vacancy available in an orbital.

Figure 2.12 Relative Compton scatter probability as a function of the scattering angle.



When Compton scattering occurs at the diagnostic imaging energies ($15 \leftrightarrow 150$ keV), the greatest part incident photon energy goes to the scattered photon. X-rays can be deflected in any direction, even in the back direction. In this case it is called *backscatter radiation*. The probability of Compton scattering is inversely proportional to X-ray energy and independent of the atomic number. In other words, an X-ray photon is almost

as likely to undergo Compton scattering with an atom of soft tissue as with an atom of bone.

Scattered X-rays from Compton effect *provide no useful information* in a radiograph and may contribute to *extra dose damage*. Even with a huge energy loss, a Compton scattered photons have a relatively high momentum and tissue penetrability. The accidental detection of scattered photons by the image receptors results in a degradation of image contrast and an increase in *unwanted noise*. This noise can be reduced thanks to the interleaving of a static of moving *Bucky layer* (Section 2.2).

If a Compton interaction event is to take place, the incident photon energy must be significantly greater than the electron's binding energy. Thus, the relative probability of a Compton interaction increases as the incident photon energy increases. The probability of Compton interaction also depends on the electron density, which with the exception of hydrogen, is fairly constant in all types of tissues.

2.5.2 Photoelectric effect

Diagnostic X-rays also undergo ionizing interactions with inner shell electrons. During this type of interactions the X-ray photon is absorbed completely. The electron removed from the atom (called photoelectron) escapes with kinetic energy equal to the difference between the energy of the incident X-ray and the binding energy of the electron. This entails and total X-ray absorption.

$$\mathrm{E}_{\mathrm{electron}} = \mathrm{E}_{\mathrm{incident\,photon}} - \mathrm{E}_{\mathrm{orbital\,binding}} \tag{2.3}$$

For low atomic number atoms (soft tissues), the binding energy of shell electrons (even those residing in the K-shell) is very low. Therefore, the electron is released with almost equal energy to the incident X-ray photon. For higher atomic numbers (such as the calcium found in bones), binding energies are higher and therefore, the kinetic energy is lower.

Secondary characteristic X-rays may be produced if the ejected electron corresponds to that of a K-shell. In this case, an outer shell electron (for instance, from the L-shell) drops into the vacant orbital and and a

characteristic X-ray is emitted. The energy of this secondary X-ray is equal to the difference between binding energies of the two orbitals. These subsidiary photons can be considered as scattered rays and their contribution to producing X-ray images is meaningless. The photoelectric effect does normally occur with valence shell electrons. In this case, no inner shell electron cascade can take place and no subsequent characteristic X-rays are produced.

The probability of photoelectric effect is directly proportional to the third power of the Z of the absorbing atom and to third power of the incident photon energy (Z^3/E^3). The great benefit of photoelectric absorption in diagnostic X-ray imaging is that it avoids scattered photons degrading the image. The fact that the probability of photoelectric interaction is proportional to $1/E^3$ explains, why image contrast decreases when higher X-ray energies are used in the imaging process

2.5.3 Differential absorption

More important than interaction of the X-ray by Compton scattering or photoelectric effect is the X-rays transmitted through the body without interaction. As stated before, *Compton scattered X-rays contributes with no useful information* to the image formation. On the contrary, X-rays that undergo photoelectric effect do provide diagnostic information because of the fact of not reaching the detector. These missing X-rays are representative of anatomical X-ray-opaque structures. The photoelectric absorption of X-rays produces the light areas in a radio graph, such as those corresponding to the bones. Other X-rays pass through the body and reach the detector with no interaction whatsoever. They produce dark areas in a radiograph. The structures and tissues traversed are transparent to X-rays.

An X-ray image results from the difference between those X-rays absorbed photoelectrically in the patient and those transmitted to the image receptor. This difference in X-ray interaction is called differential absorption. Approximately 1% of the X-rays incident on a patient reach the image receptor. Fewer than half of those interact to form an image. Thus, a radiograph results from just 0.5% of the X-rays emitted by the X-ray tube.

Producing good quality X-ray images requires the adequate selection of the kVp do that as to obtain the maximum differential absorption. Unfortunately, if we reduce kVp, apart from increasing differential absorption (and also image contrast), we also increase patient dose. As a consequence, differential absorption controls the contrast of the image.

2.5.4 Dependence on mass density

Mass density is the quantity of matter per unit volume, specified in $\mathrm{kg\,m^{-3}}$. The interaction of X-rays with tissue is proportional the mass density of the tissue, regardless of the type of interaction, i.e., when the mass density is doubled, the probability of X-ray interaction is doubled.

2.5.5 Exponential attenuation

The relative frequency of each Roentgen interaction mechanism depends on the atomic number, the mass density and the own X-ray energy. Absorption is a critical condition fox X-ray interaction. If the X-ray is only partially absorbed (like in the Compton scattering) do not contribute to the image formation process. The total reduction in the number of X-rays in the original beam after they penetrate through a given thickness is called *attenuation*. Attenuation is therefore the product of absorption and scattering and it takes place exponentially, following the well-known Beers-Lambert law.

The fraction of photons removed from a mono-energetic beam of X-rays or gamma rays per unit thickness of material is called the linear attenuation coefficient ($\mu(E)$), typically expressed in $\mathrm{cm^{-1}}$. The number of photons removed from the beam traversing a very small thickness can be expressed as:

$$n = \mu(E) N \Delta l \tag{2.4}$$

where n is the number of photons removed from the beam and N is the number of original photons incident on the material.

For a mono-energetic beam (with energy E) of photons incident upon either thick or thin slabs of material, an exponential relationship exists between the number of incident photons (N_0) and those that are transmitted (N) through a thickness l without interaction:

$$N = N_0 e^{-\mu(E)l} \tag{2.5}$$

For a given material and thickness, the probability of interaction is proportional to the number of atoms per volume. This dependency can be overcome by normalizing the linear attenuation coefficient for the density of the material. The linear attenuation coefficient, normalized to unit density, is called the *mass attenuation coefficient* τ (Table 2.1). This coefficient depends on the type of interaction, as shown in Fig. 2.13. This figure illustrates the addition of different X-ray interaction's mass attenuation coefficients as well as the total mass attenuation coefficient for soft-tissue.

Figure 2.13 Mass attenuation coefficients for soft tissue. The data for this plot comes from National Institute of Standards and Technology XCOM photon cross-sections database.

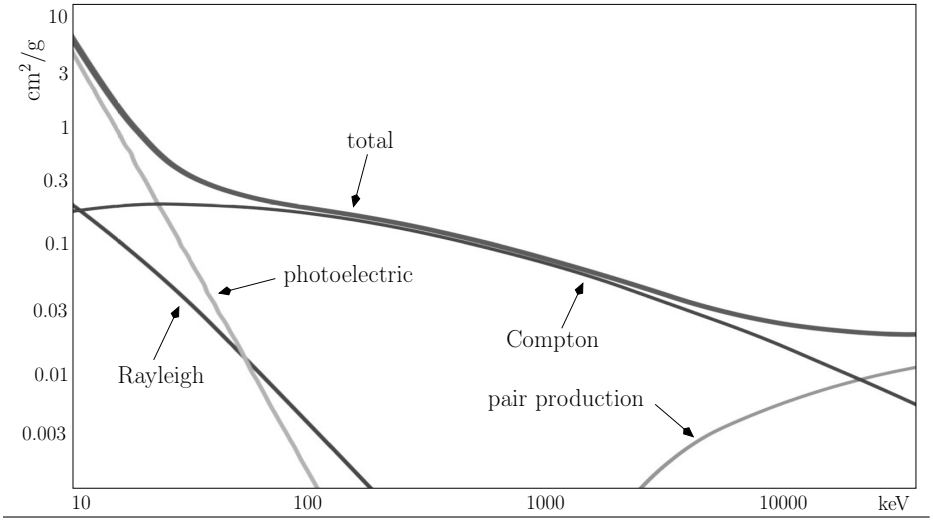


Table 2.1 Some mass attenuation coefficients (cm^2/g).

Material		Air	Water	Muscle	Bone	Fat
Density		0.001	1.0	1.04	1.65	0.92
incident energy	10	4.91	5.07	5.07	19.79	3.08
	15	1.52	1.57	1.57	6.19	1.00
	20	0.73	0.76	0.76	2.75	0.53
	30	0.33	0.36	0.36	0.95	0.29
	40	0.24	0.26	0.26	0.50	0.23
	50	0.20	0.22	0.22	0.34	0.21
	60	0.18	0.20	0.20	0.27	0.19
	70	0.16	0.18	0.18	0.21	0.17
	80	0.15	0.17	0.17	0.18	0.16

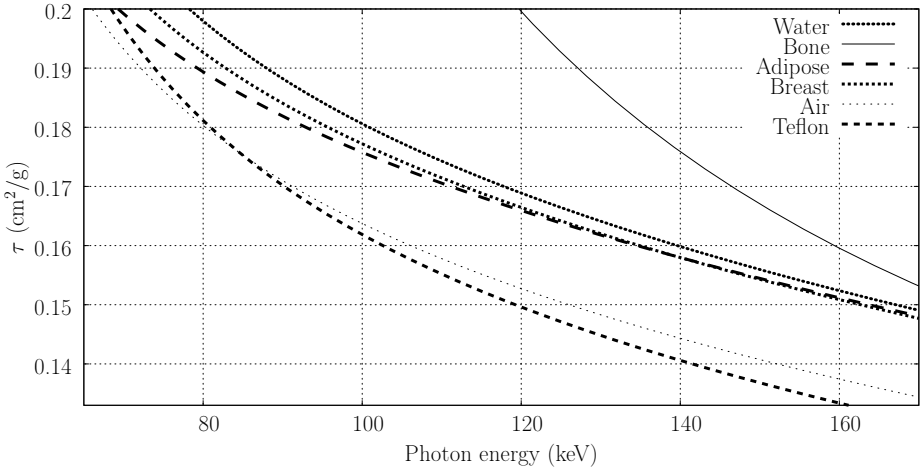
2.6 Radiographic image

A pixel with intensity $\mathcal{I}(x, y)$ in the final X-ray image is the result of *chaining* Eq. (2.5) for each traversed element i and accounting (adding or better, integrating) for each possible energy dE in the original spectrum (Fig. 2.10), which leads to the following *transfer function*:

$$\mathcal{I}(x, y) = \int_0^{\text{kVp}} N_0(E) \cdot \prod_i e^{-\mu_i(E) \cdot l_i(x,y)} dE \tag{2.6}$$

where $l_i(x, y)$ is the traversed length in an element i of a X-ray whose *final destination* is the pixel x, y in the detector, $N_0(E)$ is the initial number of photons of energy E and kVp is the maximum E. This number will depend on the chosen exposure time and current. Of course, each pixel will output a different value for Eq. (2.6). Besides, the integrated intensity of each $\mathcal{I}(x, y)$ is finally resampled to an integer value between 0 and 4095 in the case of 12-bit images or 0 and 65535 in the case of 16-bit images. Of course, this final conversion from energy to gray level also depends on the detector efficiency, nature, a even specific software configuration. Consumer computer screens can usually display just 8-bit (0 to 255 levels). Each of these values is a grayscale value represented by screening software. Recent studies have demonstrated that a human observer can perceive up to 900 shades of gray [123]. Finally, some X-ray management software

Figure 2.14 X-Ray mass attenuation coefficients for some materials and tissues present in X-ray rooms and for typical energy ranges used in diagnostic imaging.



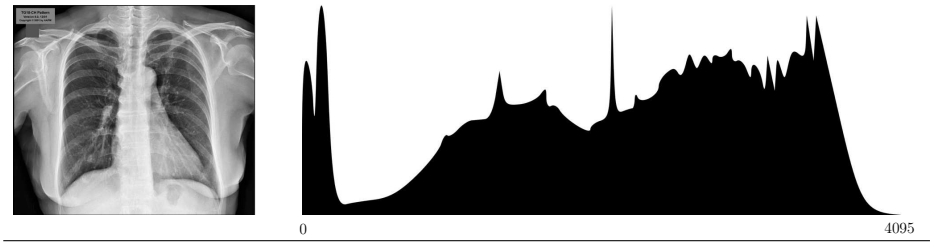
may apply specific correction algorithms that remove artifacts or enhance contrast, concept tackled next.

2.6.1 Image Contrast

After the X-ray image is generated, we are done with the physics layer and we end up with a grayscale image representing the transmitted radiation that was not absorbed by the examined patient/object. A grayscale image can be assigned an intensity histogram. An image histogram is a graphical representation of the tonal distribution in a digital image. It plots the number of pixels for each gray value. The horizontal axis of the graph represents the tonal variations, while the vertical axis represents the number of pixels in that particular level. The left side of the horizontal axis represents the black and dark areas, the middle represents medium grey and the right hand side represents light and white areas. The histogram for a very dark image will have the majority of its data points on the left side of the graph. Conversely, the histogram for a very bright image with few dark areas and/or shadows will have most of its data points on the right side and center of the graph.

Contrast is the difference in grayscale values in an image. An image with just one grayscale level has no contrast. On the contrary, an image with transitions between dark gray levels and light gray pixels shows high contrast. The contrast of an X-ray image (Fig. 2.15) is related to the different photon fluencies that reach the detector after interacting with the patient's tissues through different physical phenomena. The dynamic range of an X-ray imaging system is the ratio of the largest and smallest input X-ray intensities that can be imaged. It can be also be defined as the total number of distinctive pixel values that occur in the image. Dynamic range is limited by the number of bits per pixel used to represent the image.

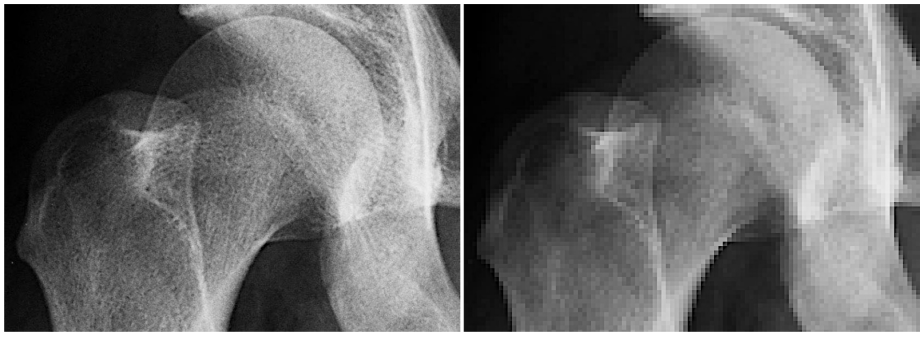
Figure 2.15 Example of a 12-bit X-ray image gray level/intensity histogram.



2.6.2 Spatial resolution

Spatial resolution refers to the ability of an X-ray system to see small detail (Fig. 2.16). An X-ray detector has higher spatial resolution if it can resolve the presence of small objects in the image. The limiting spatial resolution is the size of the smallest resolvable object. The wavelength of the electromagnetic energy used to probe the patient is a fundamental limitation of the spatial resolution of the X-ray imaging modality.

Figure 2.16 Different image resolutions in conventional X-ray images.



The frequency (and thus, the wavelength) of X-rays depends on the beam energy, but even the lowest energy X-ray (involving wavelength of about 1 nm). This is far from the actual resolution by digital or legacy X-ray sensor, but it does represent the theoretical limit on the spatial resolution of diagnostic X-rays equipment. This spatial resolution is usually a tenth of the millimeter for fully digital and CR detectors, but can reach the hundredth of a millimeter in the case of legacy screen film radiography. However, such limits begin to enter into conflict with the own spatial resolution of the human eye [115].

2.7 Conventional X-ray settings for primary diagnostic

When patients undergo X-ray examinations, they are usually offered a comfortable gown to put on instead of their own clothes. They are then guided through the positioning process, which can take place in two main positions in the case of a trunk radiography: standing erect or lying. X-ray examination rooms usually have separate detector holders for each type of exam. A detector is also placed accordingly, except for the cases of fully digital equipment, where flat panel detector as usually welded to the supporting structure. The positioning of the patient, the detector and the

X-ray tube and the distances between them are always roughly stablished and no precise measurements are registered.

Once the patient is positioned, the technician or radiologist triggers the activation of the X-ray beam. The patient can be advised, just a few instants before, to hold his/her breath (or not) if it is clinically necessary. Once the exposure has ended, the detector information is read, either by manually inserting it in a CR scanner or by automatically launching a readout process. The process can be repeated several times more if it is clinically convenient. All these steps are graphically summarized in Fig. 2.17.

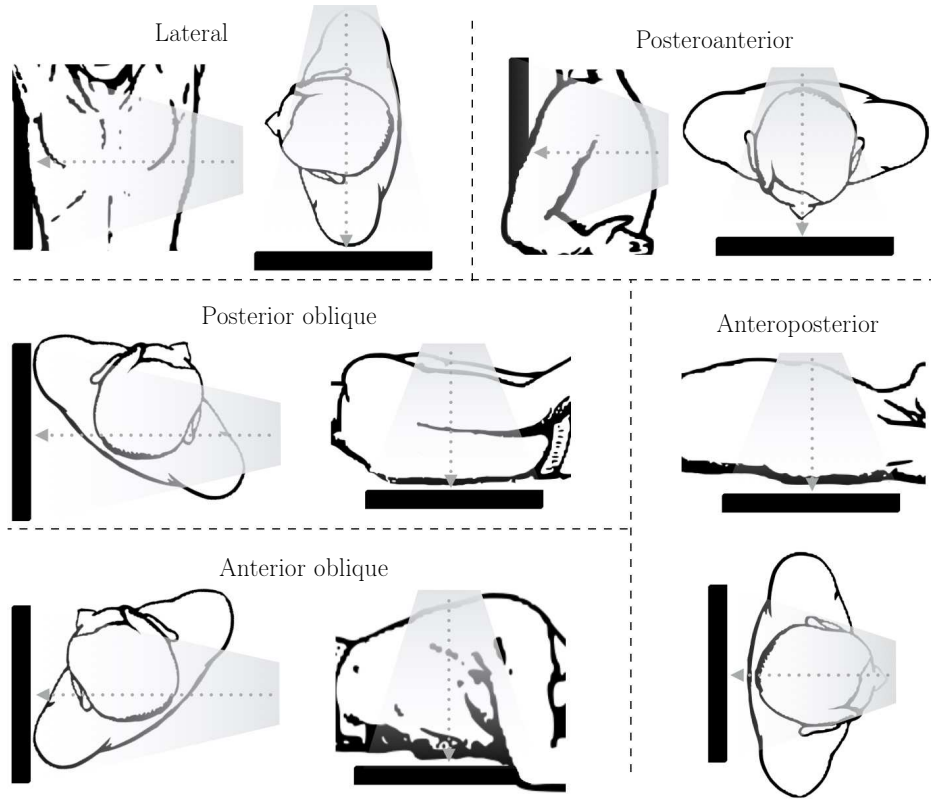
Finally, the radiologist or doctor diagnoses the acquired images and informs the patient about his/her health status and gives counseling about future actions.

Figure 2.17 Typical steps during an X-ray exam: a) patient guidance during positioning, b) radiograph acquisition, c) imaging plate digitalization, d) radiograph visualization and diagnosis and e) patient counseling.

2.7.1 Patient positioning

During an X-ray examination, the patient is asked to adopt any of the following positions, always expressed relative to the detector. They are also graphically summarized in Fig. 2.18.

Figure 2.18 Most common positions/protocols in conventional X-ray imaging.



Decubitus position is obtained with the patient faces towards the cassette while lying in decubitus position.

Posteroanterior when the patient faces towards the cassette.

Anteroposterior when the patient's back is towards the cassette.

Anterior oblique is obtained with the right front of the patient against the imaging plate. The patient is turned approximately 45 degrees

toward the right side, placing the patient's right shoulder in contact with the grid device and left hand on their hip. This position demonstrates the maximum area of the left lung field.

Posterior oblique is obtained with the right back of the patient against the cassette. The patient is rotated 45 degrees with left posterior shoulder against the imaging plate. It is comparable to the anterior oblique view in demonstrating the maximum area of the left lung field.

These positions are usually never measured nor logged for a posteriori processing. Patient dimensions and exact orientation relative to the imaging system is also very frequently ignored for diagnosis and/or post-processing.

2.7.2 Common distances in X-ray examinations

Distances in conventional X-ray imaging are seldom or never log or taken into account and they vary from snapshot to snapshot, as Fig. 2.19 shows. They are also difficult to measure. For instance, the patient may have an heterogeneous surface and the real position of the anode may be difficult to establish (as it is housed inside a shielding structure). The basic lengths in a conventional X-ray room are:

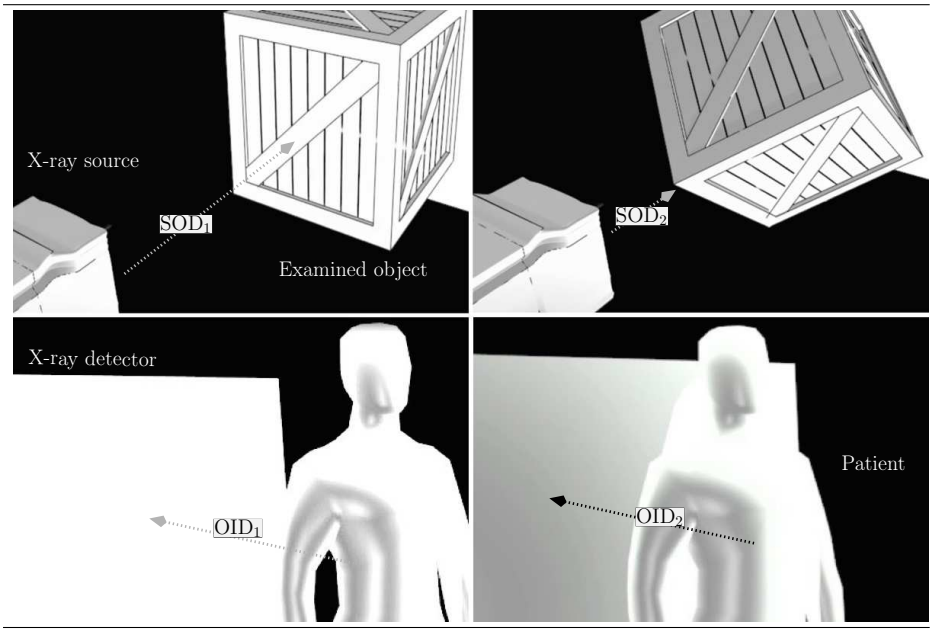
Source to Image Distance or is a measurement of the distance between the radiation source (anode in our case) and the radiation detector (a FPD, for instance).

Source to Object Distance or distance from the anode to the patient of which a radiographic image is to be generated.

Object to image receptor distance or between the patient and the radiographic image receptor.

These parameters can also be stored as part of the Information Object Definition (IOD) of each radiograph according to the Digital Imaging and Communications in Medicine (DICOM) [157]. The DICOM standard was designed for handling, storing, printing, and transmitting information in medical imaging. It includes, besides a file format definition and a network

Figure 2.19 Variable object-to-image (OID) and source-to-object (SOD) distances, along with different possible orientations.



communications protocol, a tagging mechanism for all sorts of data, such as the parameters aforementioned described:

DistanceSourceToDetector (0018, 1110) Distance in mm from source to detector center. This value is equivalent to the SID.

DistanceSourceToPatient (0018, 1111) Distance in mm from source to the table, support or bucky side that is closest to the subject, as measured along the central ray of the X-Ray beam. This definition is less useful in terms of estimating geometric magnification than a measurement to a defined point within the patient, but accounts for what is realistically measurable in an automated fashion in a clinical setting. This measurement does not take into account any air gap between the patient and the front of the table or bucky. If the detector is not mounted in a table or bucky, then the actual position relative to the patient is implementation or operator defined. This value is equivalent to the SOD.

DistanceSourceToEntrance (0040, 0306) Distance in mm from the

source to the surface of the patient closest to the source during the acquisition of the image. It may be an estimated value based on assumptions about the patient's body size.

These parameters have both a name in English and a tuple 8-digits code (i.e., PatientsName or 0010, 0010). As with the patient positioning, the above listed distances are seldom measured.

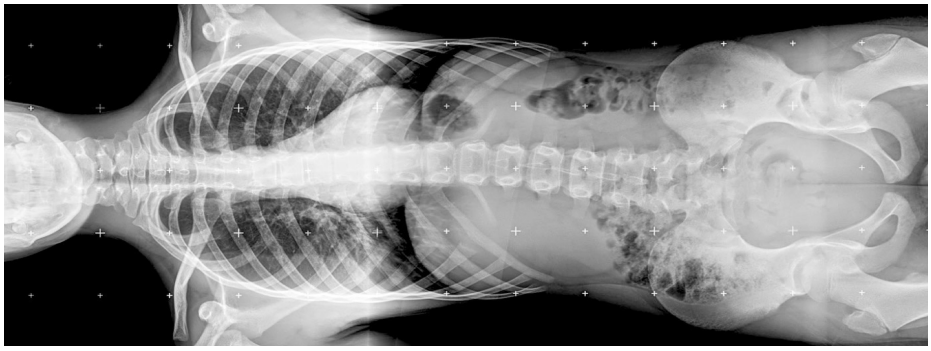
2.8 Conclusions

In this chapter we have summarized the basic and well stablished physical processes regarding the production of X-rays and their interaction with matter. The implications of the attenuation and filtration have also been presented. Next, conventional X-ray setups used in primary diagnostic, their role in patient diagnosis and their limitations regarding the determination of lengths and distances (such as SID, OID and SAD) have been exposed. These distances and angles can have a significant impact in the image properties (like those tackled in Section 2.6) and even in dose assessment (as we will later see through Monte Carlo and deterministic simulations in Chapter 9).

In the next chapters, methods and materials that address the aforementioned barriers will be discussed. These proposed methodologies will require the augmentation of ordinary X-ray equipment with external positioning devices (Chapter 4 and Chapter 5) or the use of structures that hold X-ray-opaque markers, as will be reviewed in the next chapter.

3

Geometrical calibration of X-ray equipment



As stated in the previous Chapter, in conventional X-ray imaging, an absorption image is generated from the examined object/patient. This type of imaging technique *ignores* the geometry of the environment, and thus, 3D information cannot be derived. This spatial information may turn out very useful, as will be emphasized in the rest of the text. In this chapter we discuss, among other things, the most basic approaches to address these limitations. These techniques are, in turn, based on the adoption of X-ray-opaque fiducial systems and their projection onto X-ray images.

X-ray systems can be operatively simplified and geometrically modeled as pinhole cameras. This fact enables the application of modern projective geometry methodologies and the measurement, in X-ray settings, of the same lengths and geometry relations discussed in Section 2.7.2. The combination of X-ray images and spatial information also facilitate, for instance, multiple radiograph stitching, as the example shown in the image above.

In this chapter, we also explore three different alternatives for obtaining intrinsic and extrinsic parameters in conventional diagnostic X-ray frameworks: the direct linear transform (DLT), the Zhang method, and the Tsai approach. We analyze and describe the computational, operational, and mathematical background differences for these algorithms when they are applied to ordinary radiograph acquisition. We also highlight the extreme variability of the intrinsic parameters in ordinary X-ray imaging setups.

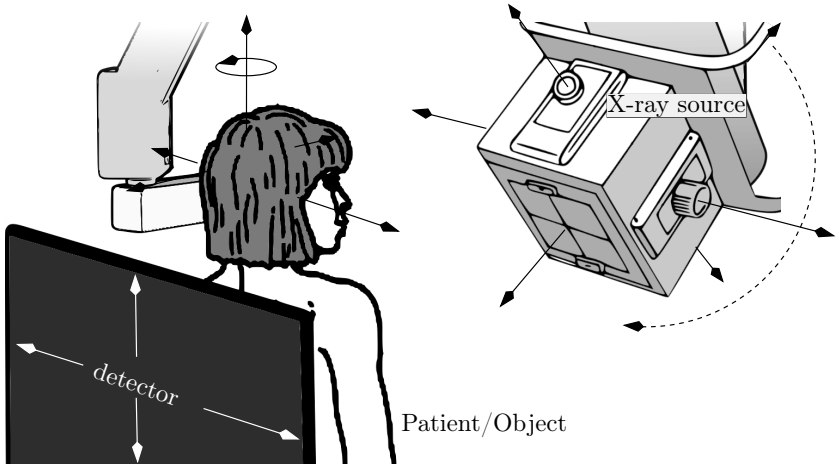
For our study, we developed an initial 3D calibration frame with tin cross-shaped fiducials at specific locations. The three studied methods enable the derivation of projection matrices from 3D to 2D point correlations. We propose a set of metrics to compare the efficiency of each technique. The results show a clear superiority of the DLT approach, both in accuracy and operational suitability. We paid special attention to the Zhang calibration method. Although this technique has been extensively implemented in the field of computer vision, it has rarely been tested in depth in common radiograph production scenarios. Zhang's approach can operate on much simpler and more affordable 2D calibration frames, which were also tested in this chapter. All in all, in this Chapter we expand the research tasks carried out in [13].

3.1 X-ray devices as pinhole cameras

Camera calibration is an important preprocessing step in computer vision applications and it also has its relevance in daily diagnostic X-ray imaging

scenarios like the one exemplified in Fig. 3.1.

Figure 3.1 A conventional X-ray setting.



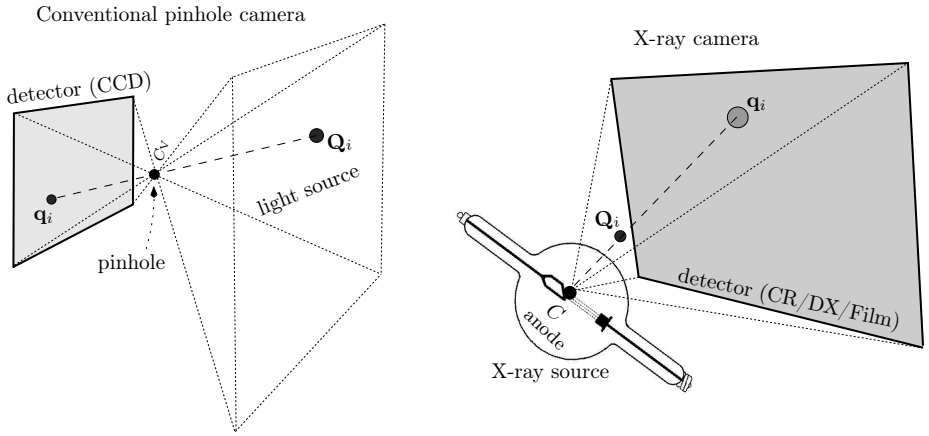
X-ray devices –if modeled as pinhole cameras– are composed of a Röntgen radiation source and a disengaged sensitive surface. In Fig. 3.2 we graphically juxtapose a conventional pinhole device and X-ray equipment. One of the key differences is that, in the case of X-rays, a projected point $\mathbf{Q}_i$ is located *between* the anode C –which plays the role of optical center– and the detector. In other words, every $\mathbf{Q}_i$ point is projected to a 2D *shadow* in a specific coordinate $\mathbf{q}_i$ in the sensor.

C is also the *origin* of the photon stream, whereas in a conventional camera, the *photon source* is the *photographed scene* itself, which *radiates* reflected light (Fig. 3.2). This diffused light *enters the camera* through the pinhole and *hits* the detector. A slide projector (commonly used in presentations) is the source of the light stream as well, however, its focal length is fixed (we only get a neatly formed image at a specific distance) and it is possible to derive it [19].

The sensor part of an X-ray deployment is usually a photographic film or an array of *dots* (*imaging plate* or IP) sensitive to this type of radiation. Depending on how the information is readout, we mainly have DX or CR modalities. Both divide the detector surface in *sensing pixels* whose resolution λ may vary between producers and between clinical protocols. λ is usually independent of the orientation ($\lambda_x = \lambda_y$) and is normally

referred as the *linear resolution* λ_o provided by the manufacturer.

Figure 3.2 X-rays (right) vs. pinhole cameras (left). In both devices, $\mathbf{Q}_i$ is a 3D point whose projection in the image is $\mathbf{q}_i$.



The pinhole camera metaphor has already been applied to X-rays in many previous works [143, 207]. In spite of this increasing and creditable contribution of literature on the subject, very few authors explicitly combine ordinary radiograph generation and modern computer vision camera calibration techniques. An adaptation of Tsai's approach developed in [245] is examined by [159] in order to obtain measurements from planar and non planar targets. Authors in [163] benefit from the Direct Linear Transform (DLT) calibration process by augmenting X-ray systems with laser rangefinders. DLT also appears in [217] where authors make use of a 3D phantom applied to orthopedics. Some research lines claim not to use any specific calibration method but specially devised techniques, involving nonlinear optimizations. For instance, [222] applies these procedures to cylinder-shaped frames. In addition, phantom grids take part in the work carried out by [68] along with the minimization of the retro-projection error.

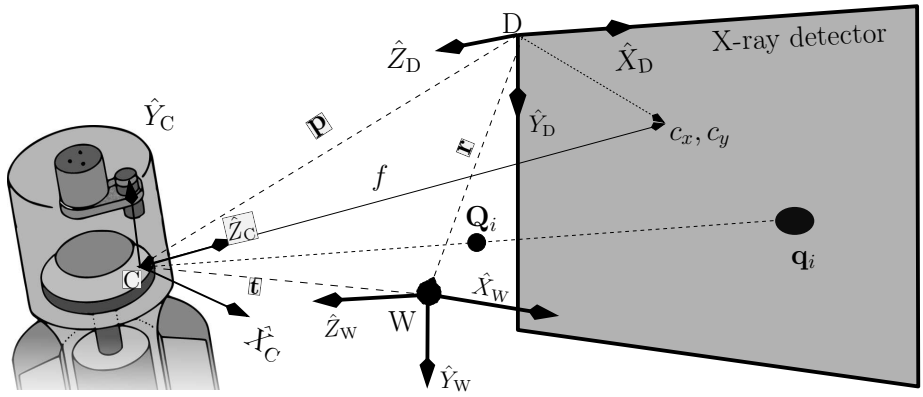
3.2 Background on the geometry of X-ray imaging systems

The fact of identifying X-ray imaging systems as pinhole devices entails some considerations, both geometrical and operational, that makes them particularly different from conventional systems. We review below the most relevant.

3.2.1 World, anode and detector reference frames

In the specific case of X-ray pinhole cameras, it turns out appropriate to work with different reference frames, all represented in Fig. 3.3. The first and most intuitive one is the *world coordinate system* $\mathrm{W} = (\hat{X}_\mathrm{W}, \hat{Y}_\mathrm{W}, \hat{Z}_\mathrm{W})$, usually placed at a known 3D point in the radiographed scene, typically over a frame or DUT. Next is the coordinate frame *attached* to detector plate itself $\mathrm{D} = (\hat{X}_\mathrm{D}, \hat{Y}_\mathrm{D}, \hat{Z}_\mathrm{D})$ whose origin is normally coincident with the upper-leftmost pixel. The requirement of D comes from the physical fact of dealing with a radiation responsive layer not *tied* to C and from the possibility of moving the anode and/or detector, with complete freedom, around the 3D scene. Finally, the coordinate system $\mathrm{C} = (\hat{X}_\mathrm{C}, \hat{Y}_\mathrm{C}, \hat{Z}_\mathrm{C})$ has its origin in C itself.

Figure 3.3 World (W), detector (D) and anode (C) reference frames.



Any point $\mathbf{X}$ in the 3D space can be associated with any of the aforementioned reference frames. For instance, if $\mathbf{X}_{\mathrm{C}}$ represents the coordinates of $\mathbf{X}$ relative to C and $\mathbf{X}_{\mathrm{D}}$ is its representation relative to D, then, from Fig. 3.3, it can be verifiable that:

$$\mathbf{X}_{\mathrm{D}} = \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}}_{\pi \text{ rotation about } \hat{X}_{\mathrm{C}}} \cdot \mathbf{X}_{\mathrm{C}} + \mathbf{p} \tag{3.1}$$

where $\mathbf{p} = (c_x, c_y, f)$ is known as the *principal point* and represents the coordinates of C in D. The line that passes through C and is perpendicular to the detector plane is the *principal axis*, and intersects the image plane at the point $(c_x, c_y, 0)$. The distance between C and the detector plane is the focal length f. The anode C can also be expressed in W coordinates by means of vector $\mathbf{t}$ that starts at the origin of W and ends at C. This vector can be interpreted as *the position of the world origin in camera coordinates*. This is often counter-intuitive, because we usually want to specify how the camera is transformed relative to the world (*pose*).

Finally, if $\mathbf{t}$ and $\mathbf{p}$ are known, it then becomes possible to derive vector $\mathbf{r}$ connecting both W and D:

$$\mathbf{r} = \mathbf{p} - \mathbf{t} \tag{3.2}$$

3.2.2 Camera calibration and projection matrices

As discussed in the first paragraph of this chapter, a prerequisite for any application in the field of computer vision is the calibration of the camera. This step is necessary to determine the pose between the imaging system and real world objects. The *pose* of an object or camera is the combination of position and orientation. It involves the calculation of 5 intrinsic (internal) and 6 extrinsic (external) parameters, which can be grouped in a 3×4 matrix called the *camera projection/calibration matrix* P.

Mathematically, P maps 3D points –expressed in W coordinates– to 2D points, using the expression: $\hat{\mathbf{q}}_i = \mathrm{P} \cdot \hat{\mathbf{Q}}_i$, where each $\hat{\mathbf{q}}_i$ is an image point and $\hat{\mathbf{Q}}_i$ is a W-referenced point, both in homogeneous coordinates.

As mentioned above, P can be decomposed in two blocks of intrinsic (K) and extrinsic parameters, respectively. K projects 3D points, expressed in the D reference frame, to their corresponding image coordinates:

$$\mathrm{K} = \begin{pmatrix} \alpha_x & s & x_0 \\ 0 & \alpha_y & y_0 \\ 0 & 0 & 1 \end{pmatrix} = \underbrace{\begin{pmatrix} \lambda_x & 0 & 0 \\ 0 & \lambda_y & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{\lambda} \cdot \begin{pmatrix} f & \sigma & c_x \\ 0 & f & c_y \\ 0 & 0 & 1 \end{pmatrix} \tag{3.3}$$

where λ is a *resolution matrix* –tackled above– expressing the *number of pixels per unit of length*, for both x and y axes, and $\alpha_x = f\lambda_x$ and $\alpha_y = f\lambda_y$ represent the focal lengths in pixel units. Similarly, x_0 and y_0 are the image units counterparts of c_x and c_y. The parameter s is the skewness of the camera and defines the angle between the x and y axes. However, as stated above, we assume that pixels are square, which allows us to simplify Eq. (3.3) as:

$$\mathrm{K} = \begin{pmatrix} \alpha & 0 & x_0 \\ 0 & \alpha & y_0 \\ 0 & 0 & 1 \end{pmatrix} \tag{3.4}$$

where we have also deliberately *forced* both focal lengths to be equal to α. This assumption makes sense in the field to which our research is intended to contribute, i.e., X-ray diagnostic imaging and item scanning. However, this hypothesis should be taken with caution in the case of CR plates, where, for instance, the IP scanning process has its own optical oddities and limitations [209, 72].

The extrinsic parameters describe a rigid transformation mapping points in space between W and C frames. This matrix can also be decomposed in a rotation matrix R (accounting for angles $\theta_x, \theta_y, \theta_z$) and the aforestated translation vector $\mathbf{t}$:

$$\overbrace{[\mathrm{R} \mid \mathbf{t}]}^{\text{extrinsic matrix}} = \overbrace{\underbrace{\left(\begin{array}{c|c} \mathrm{I} & \mathbf{t} \end{array}\right)}_{\text{3D translation}} \cdot \underbrace{\left(\begin{array}{c|c} \mathrm{R} & 0 \\ \hline 0 & 1 \end{array}\right)}_{\text{3D rotation}}}^{\text{extrinsic matrix}} \tag{3.5}$$

It is common to see a version of Eq. (3.5) with extra row of (0, 0, 0, 1)

added to the bottom. This makes the matrix square, which allows us to further decompose this matrix into the aforementioned rotation and translation

Given P, we can decompose it into its intrinsic/extrinsic parts using a RQ decomposition [99]. We can do this given the fact that R is orthogonal and K has an upper-triangular shape:

$$\mathrm{P} = \overbrace{\mathrm{K}}^{\text{intrinsic matrix}} \cdot \overbrace{[\mathrm{R} \mid \mathbf{t}]}^{\text{extrinsic matrix}} \tag{3.6}$$

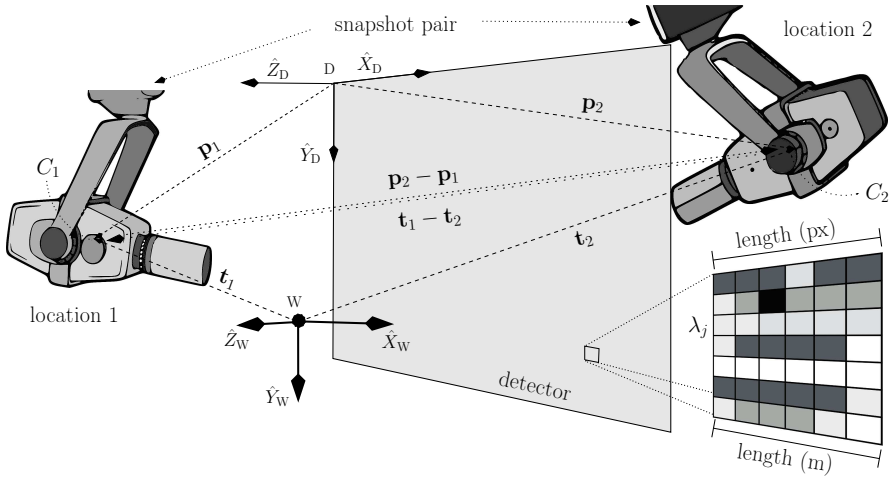
3.2.3 Estimation of the detector resolution

Even though we usually have access to the detector resolution specified by the manufacturer (λ_o), this parameter can also be experimentally determined (λ) from two X-ray images produced at different geometrical configurations. If we take a look at Fig. 3.4, we can easily notice that the distances represented by the vector relations $|\mathbf{p}_2 - \mathbf{p}_1|$ and $|\mathbf{t}_2 - \mathbf{t}_1|$account for the same spatial gap. However, there exists an important difference between the two: the former is expressed in pixel units and the latter is specified in physical units. This fact enables the calculation of λ with the expression:

$$\lambda \cdot |\mathbf{t}_2 - \mathbf{t}_1| = |\mathbf{p}_2 - \mathbf{p}_1| \tag{3.7}$$

where the tuples $\mathbf{p}_2, \mathbf{p}_1$ and $\mathbf{t}_2, \mathbf{t}_1$ represent the anode coordinates in D and W reference frames, for X-ray source locations 1 and 2, respectively. That is, we experimentally resolve how many pixels per meter λ holds our X-ray detector (for a given *stereo snapshot*), which should be *a priori* close to factory specs $\lambda \approx \lambda_o$.

Figure 3.4 Derivation of detector linear resolution (λ). The X-ray tube (a dental one in this specific case) is activated at two different locations (1 and 2) and two images are generated at each system pose. Taking into account the pair of vectors connecting the anode and the world ($\mathbf{t}_1$ and $\mathbf{t}_2$) and the pair of vectors linking the anode and the detector ($\mathbf{p}_1$ and $\mathbf{p}_2$), it is possible to derive λ.



This method for the estimation of the linear resolution of a detector represents on its own a very interesting tool which can be integrated into the quality control protocols applied during the ordinary inspection of X-ray equipment. As stated above, the $\mathbf{t}$ vectors have natural units (m, mm, etc.) and the $\mathbf{p}$ vectors have pixel units so that when divided, we obtain a unit for linear resolution (i.e., pixels per meter). Both $\mathbf{p}$ and $\mathbf{t}$ vector pairs are obtained through the RQ decomposition of each P_1 and P_2 representing the camera calibration at each of the two stereo anode poses:

$$\mathrm{P} \xrightarrow[\text{decomposition}]{\text{RQ}} \underbrace{\begin{bmatrix} \alpha & s & x_0 \\ 0 & \alpha & y_0 \\ 0 & 0 & 1 \end{bmatrix}}_{\mathbf{p}=(x_0,y_0,\alpha)} \cdot [\mathrm{R} \mid \mathbf{t}] \tag{3.8}$$

where each vector $\mathbf{p}$ is built exclusively from the same intrinsic parameters already expressed in Eq. (3.4), and $\mathbf{t}$ is nothing else but the *translation*

vector, part of the extrinsic side of the camera matrix equation.

Finally, Eq. (3.7) and Eq. (3.8) already point at the fact that in X-ray systems, extrinsic and intrinsic parameters are closely related, which will be further developed in the next section.

3.2.4 Pose-dependent intrinsic matrices on X-ray systems

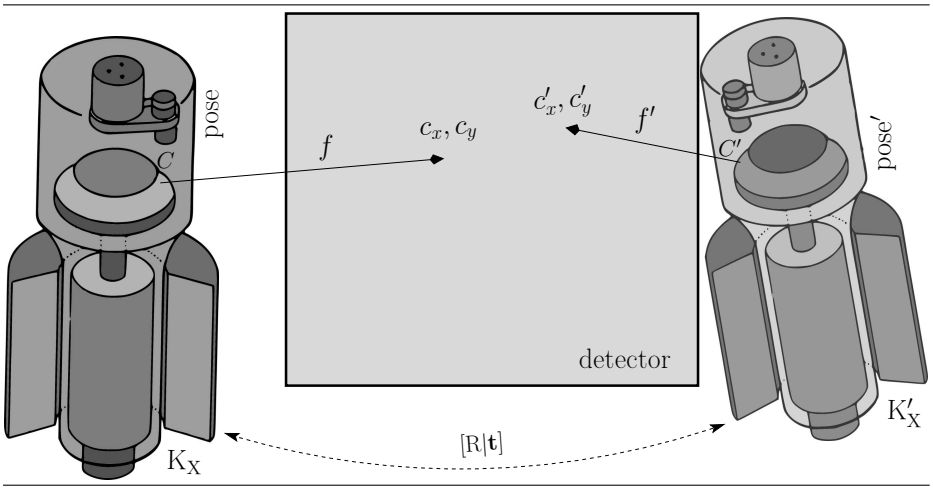
It is worth noting that in the specific scenario of X-ray imaging a controversy may arise when qualifying K as *intrinsic*. In conventional pinhole cameras, intrinsics do not change if the device is repositioned in the scene. However, in X-ray frameworks, these parameters may vary significantly between consecutive snapshots if either the sensor or anode are shifted and/or rotated relative to each other, as illustrated in Fig. 3.5. This variability lies in the fact that detector surface and anode are detached –and structurally independent– one from the other. The term *pseudointrinsic* might be coined here.

This fact contrasts with conventional camera systems, where the sensor (usually a CCD/CMOS array) is *architecturally* fixed to the optical center. In X-ray imaging, we face a very interesting scenario in which the intrinsic and extrinsic parts of P are *tied*. Therefore, for two anode/detector locations C and C', it is not unreasonable to think that $\mathrm{K} \neq \mathrm{K}'$. This is the reason why the D reference frame is required when describing the geometry accompanying X-ray settings (as already discussed in Section 3.2.1).

3.3 Outline of current calibration Algorithms

We now outline the three most used methodologies for the estimation of P. These same approaches will later be applied (and compared) to X-ray imaging systems. Several authors have proposed solutions to the problem of camera calibration. Among the most popular are Tsai's algorithm [247],

Figure 3.5 Intrinsic parameters variation when relocating the anode from an original pose to a different one (the same would happen if the detector plate was shifted and the X beam source remained anchored at a fixed location).



DLT [98], and Zhang's method [262], which we will now summarize. Further details about these methods can be found in [201] and [264] addressed to conventional cameras.

3.3.1 Tsai's Method

Tsai's camera calibration method –presented in [247]– by Roger Tsai in 1986 is one of the most famous –and probably one of the very first– *modern* algorithms for camera calibration. An updated description of the Tsai's algorithm can be found in [90] and [149].

The algorithm recovers the camera parameters using the relationship between points in a three-dimensional calibration mold and their projections in the image plane. A key aspect of this algorithm is that x_0 and y_0 (pixel coordinates of the principal point c_x, c_y) must be passed as input parameters. The pixel coordinates of the principal point are normally known, fixed and provided by the detector manufacturer in the case of conventional cameras. Unfortunately, this approach is not applicable in X-ray systems, where sensor and emitter are detached one from the other,

as discussed above.

Tsai's method entails two stages. The first one determines the extrinsic parameters and the focal length. This is achieved by solving a system of linear equations whose inputs are the 3D/2D coordinates of points in a calibration frame, both in image and real world. The second phase involves a non-linear minimization process where the radial distortion factor is determined and all other parameters are further refined.

3.3.2 DLT

DLT is a simple algorithm used to obtain the projection matrix given a sufficient set of point correspondences. It was originally devised by Abdel Y. I. Aziz and H. M. Karara [27] and is updated by [98].

DLT estimates P using a projective transformation presented in Section 3.2.2 and a set of point correspondences. However, since points are in homogeneous coordinates, two points will be equal if their coordinates are proportional. For this reason, each point pair $\hat{\mathbf{q}}_i, \hat{\mathbf{Q}}_i$ introduces a restriction which is better written using the vector cross product: $\hat{\mathbf{q}}_i \times \mathrm{P}\hat{\mathbf{Q}}_i = 0$. This restriction generates two independent equations. Since the number of independent unknowns is 11 (P is determined up to a scale factor in homogeneous coordinates), a simple linear solution for P can be derived with 6 correspondences.

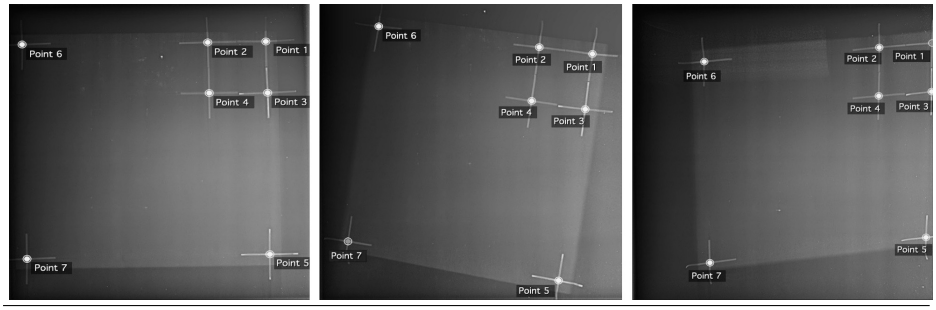
Inasmuch as point coordinates are always measured with some error, the linear equation system used to obtain P is normally solved using the SVD method. Unfortunately, the results of SVD depend on the origin and scale of the coordinate system in the image, which makes the algorithm unstable. To deal with this issue, [98] advocates for a former normalization of each $\hat{\mathbf{q}}_i$ that scales and centers all points.

3.3.3 Zhang's Method

This method was originally devised by Zhengyou Zhang in [262] and requires N_{p} projections of one or more planar calibration targets $N_{\boldsymbol{p}}$,

each with its own fiducial set $\mathbf{Q}_{i\cdots N_\mathrm{f}}$ as shown in Fig. 3.6. Conventionally in computer vision, $N_{\boldsymbol{p}} = 1$, i.e., just one calibration frame photographed/beamed N_p times at different poses ($N_\mathrm{p} = 1 \times N_\mathrm{p}$). However, we can also use several $N_{\boldsymbol{p}}$ frames, each portrayed once ($N_\mathrm{p} = N_{\boldsymbol{p}} \times 1$). For each projection j, a 2D homography can be estimated from the acquired image. In order to compute these homographies, more than 4 non-collinear points are needed. Zhang calculates then a series of projective transformations $\{\mathrm{P}_1, \mathrm{P}_2, \cdots, \mathrm{P}_j, \cdots, \mathrm{P}_{N_\mathrm{p}}\}$ to points $\mathbf{q}_{i\cdots N_\mathrm{f} \times N_\mathrm{p}}$ in the N_p bitmaps, up to a scale factor. It is important to denote here that each projection j used in the scope of Zhang's method will return its own independent $[\mathrm{R} \mid \mathbf{t}]_j$ set, however, they all share the same K. This means that we can only reproject points related to each projection j and we miss the opportunity to locate a *shared* and unique world reference frame W unless one of the plane frames is radiographed at a well-known and traceable location in the 3D scene. The minimization phase helps refining all previously derived parameters, whose amount is proportional to N_p, that is, 3 intrinsic $(c_x,\, c_y,\, f)$ + $N_\mathrm{p} \times (\theta_x, \theta_y, \theta_z, t_x, t_y, t_z)$ extrinsics.

Figure 3.6 Zhang planes applied in the X-ray spectrum.



3.3.4 Non-linear refinement

All the previous algorithms can be improved if the retrieved results are refined using a non-linear cost function $g(\ldots)$. The most common one is the *geometric distance* (or *transfer error*), which measures the Euclidean distance between the projection of a N_points set of 3D world spots $\mathbf{Q}_i$ and their *observed* correspondences $\mathbf{q}_i$ in the image:

$$g(\mathbf{q}_{1\cdots N_{\text{points}}}, \mathbf{Q}_{1\cdots N_{\text{points}}}) = \sum_{i}^{N_{\text{points}}} \|\mathbf{q}_i, \mathrm{P}\mathbf{Q}_i\|^2 \tag{3.9}$$

In the case of the Zhang method discussed above, each plane has its own projection j and its own P_j. Each $\mathbf{Q}_{j,i}$ (fiducial i on beamed frame j) must be then projected using the corresponding P_j. In this scenario, the cost function depends on $M = N_{\text{p}} \times N_{\text{points}}$ total amount of parameters. Eq. (3.9) should be then rewritten as suggested by [152]:

$$g(\mathbf{q}_{1\cdots M}, \mathbf{Q}_{1\cdots M}) = \sum_{j=1}^{N_{\text{p}}} \sum_{i=1}^{N_{\text{points}}} \|\mathbf{q}_{j,i} - \mathrm{P}_j\mathbf{Q}_{j,i}\|^2 \tag{3.10}$$

where $\mathbf{q}_{j,i}$ is the observed projection of a coplanar point i to imaged plane j. However, if plane projections correspond to one single physical flat frame ($N_{\boldsymbol{p}} = 1$) with a common set of fiducials $\mathbf{Q}_{1,\cdots,i,\cdots,N_{\text{points}}}$ arbitrarily shifted/rotated between snapshots –as we explore in Section 3.8– Eq. (3.10) can be simplified as:

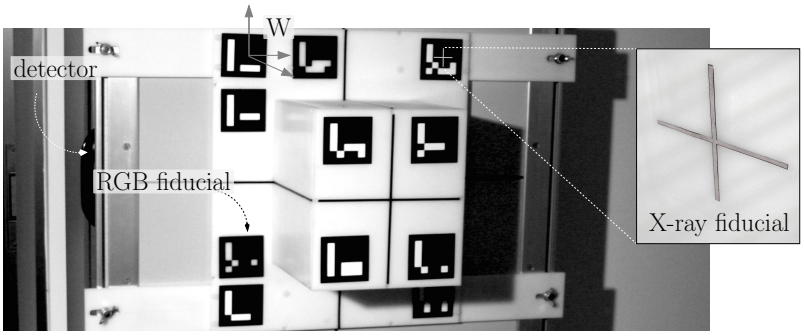
$$g(\mathbf{q}_{1\cdots M}, \mathbf{Q}_{1\cdots N_{\text{points}}}) = \sum_{j=1}^{N_{\text{p}}} \sum_{i=1}^{N_{\text{points}}} \|\mathbf{q}_{j,i} - \mathrm{P}_j\mathbf{Q}_i\|^2 \tag{3.11}$$

3.4 X-ray imaging setup and calibration frame

In order to appropriately describe an X-ray diagnostic system like the one represented in Fig. 3.1, it is initially necessary to establish a reference frame located at the detector. The chosen one is displayed in Fig. 3.3 and Fig. 3.8.

Also, an initial polytetrafluoroethylene 3D calibration frame or DUT (Fig. 3.7), outlined in Fig. 3.8 has been specially built. The DUT has

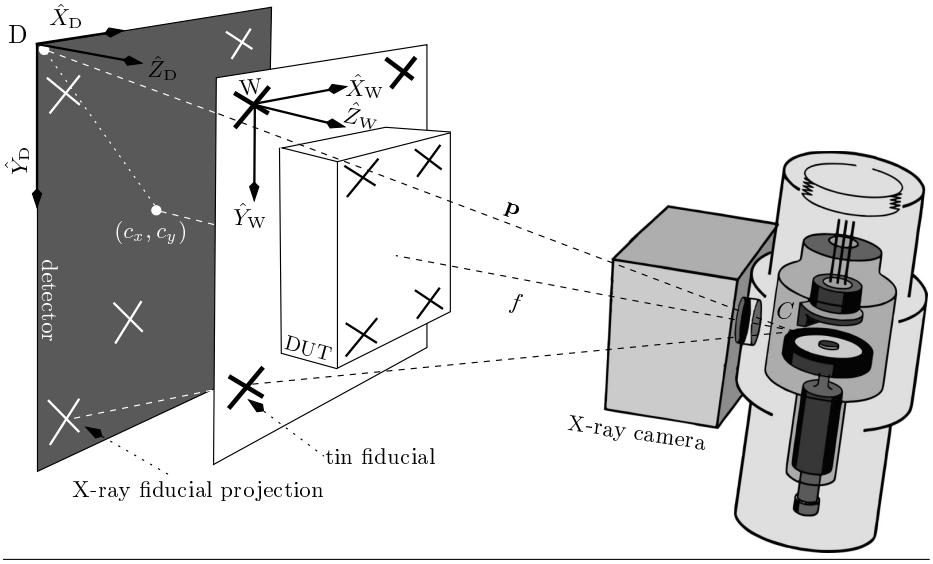
Figure 3.7 Polytetrafluoroethylene 3D calibration frame with hidden X-ray fiducials.



a simple rectangular shape in order to be manageable and, if necessary, geometrically checked by third party adopters.

In our set up, this DUT remains fixed and can be imaged from many poses.

Figure 3.8 World and detector coordinate systems.



Additionally, a set of $N_\mathrm{f} = 13$ cross-shaped tin markers opaque to Röntgen rays were distributed on two levels of the DUT. The bottom level accommodates 9 markers and the top one just 4 of them. The DUT

is also responsible for establishing the W reference frame.

Besides the N_{f} cross-shaped markers, our DUT also shelters 9 lead spherules (4 mm diameter) inside its Teflon cavity (Fig. 4.19-left). These *birdshots* will play a role later when verifying the goodness of each method and they are ignored during the calibration.

The DUT is radiographed from many angles and beam positions (57 snapshots), imitating the AP vertical examination. In our specific setup, the IP surface is roughly the same size as the calibration frame in order to capture all possible fiducial marker projections. The pixel centers of all cross-shaped traces have been manually identified.

3.5 Comparison metrics

In order to objectively compare the three methodologies outlined in Section 3.3, several metrics are defined. *Reprojection RMSE* and *intrinsic parameters consistency* are applied on a per projection basis. The other two (*detector resolution* and *distance to epipolars*) imply the use of pairs of stereo radiographs as input.

3.5.1 Reprojection root-mean-square difference

Or, how well each algorithm performed when projecting again into the image those same 3D N_{fiducial} points that were used in the calibration:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N_{\mathrm{f}}} \|\mathbf{q}_i - \mathbf{P}\mathbf{Q}_i\|^2}{N_{\mathrm{f}}}} \tag{3.12}$$

3.5.2 IP resolution estimation

Knowing the linear resolution of the IP, provided by the manufacturer ($\lambda_o = 10^4$ px/m) it is possible to use it as a screening/test parameter. Based on Eq. (3.7) and the arguments put forward in Section 3.2.3, we can write:

$$\lambda_j = |\mathbf{p}_{j,2} - \mathbf{p}_{j,1}| / |\mathbf{t}_{j,2} - \mathbf{t}_{j,1}| \tag{3.13}$$

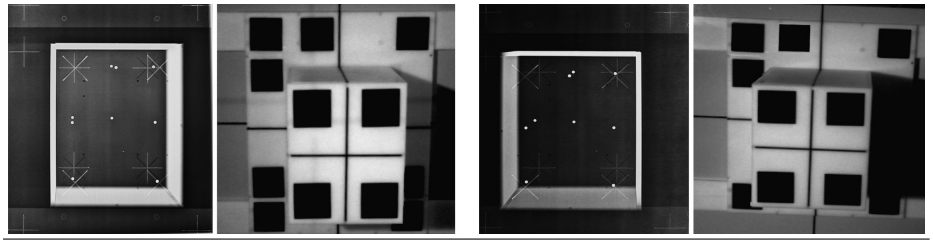
where λ_j is the computed linear resolution for a stereo snapshot j. Each snapshot obtained through Eq. (3.13) represents two stereo X-ray tube positions comprising two separated images. If we gather several $\lambda_j, \lambda_k, \ldots \lambda_n$ we can compare the calculated (mean) linear resolution with that provided by the manufacturer

Over $N_{\text{comb}} = 1540$ pair combinations from 56 radiographs were made to estimate λ_j, like the one exemplified in Fig. 3.9. Then for each calibration algorithm, we obtain the overall mean value:

$$\lambda_{\text{method}} = \sum_{j=1}^{N_{\text{comb}}} \lambda_j / N_{\text{comb}} \tag{3.14}$$

for *method* being one of Tsai, DLT or Zhang.

Figure 3.9 Example of a snapshot pair used to derive IP's λ_j from two X-ray source poses (photographs were taken, for informative purposes, with a consumer digital camera near C and roughly pointing at the same direction as the collimator).



3.5.3 2D distance between epipolar lines and spherules

A key step in stereo imaging entails finding point correspondences in two images. Using epipolar geometry, the search for a corresponding point can be reduced from scrutinizing a whole image to just looking across a specific line within it, called the *epipolar line*. In other words, a point in one image is a line in its *stereo partner*. The left camera *sees* it as a point because it is directly *in line* with that camera's center of projection. However, the right camera *discerns* this ray as an infinite straight segment in its image plane. The same situation arises in X-rays too, where the beam source is shifted between two positions with a rigid transformation.

In order to produce these epipolars, we first compute the *fundamental matrix* F. This entity can be obtained by several means, one of them being from the interplay of two camera projection matrices P_1 and P_2, as described by [22, 138].

As mentioned in Section 3.4, a set of $N_\mathrm{s} = 9$ extra lead markers were placed in the DUT plastic receptacle at specific 3D locations thanks to a foam scaffold. Using the same snapshot pair analogy we applied for the detector resolution step, we can compute the epipolar lines and their distances $\Delta_{s,j}^{\text{method}}$ to the projections $\mathbf{q}_s$ of each spherule $\mathbf{Q}_s$, again, for each snapshot pair j and for each calibration method. Epipolar geometry establishes that the epipolar line $\mathbf{l}_s^2$ on 2nd image linked to the projection of point $\mathbf{Q}_s$ on 1st image $\mathbf{q}_s^1$ can be calculated with:

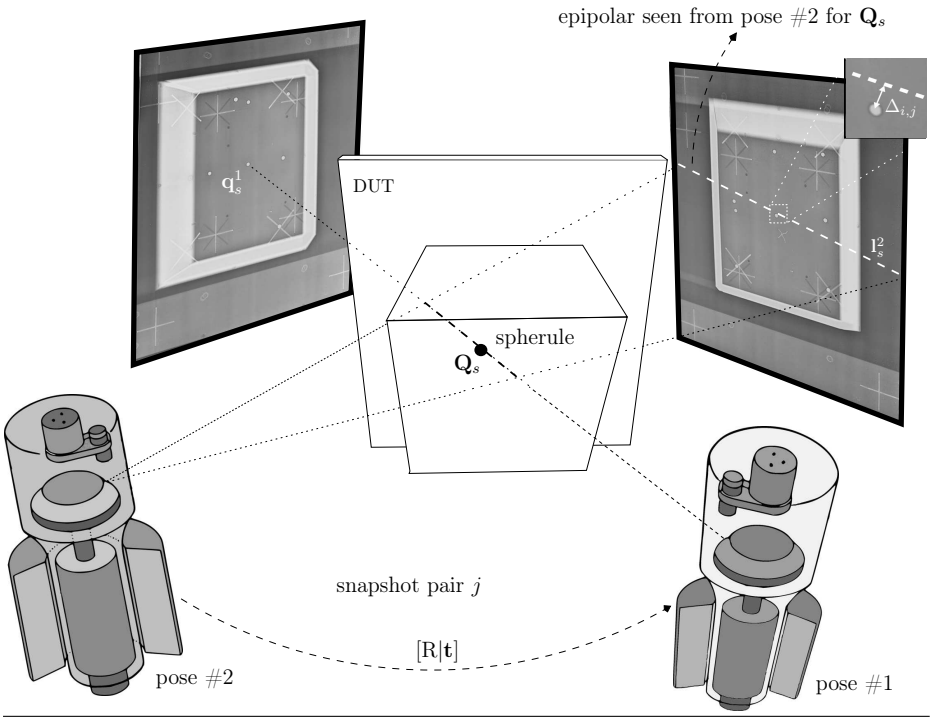
$$\mathbf{l}_s^2 = \mathrm{F}_{\mathrm{P}_2}^{\mathrm{P}_1} \cdot \mathbf{q}_s^1 \tag{3.15}$$

The whole idea is graphically summarized in Fig. 3.10. The mean value of all $\Delta_{s,j}^{\text{method}}$ is given as a result for the Tsai ($\Delta_{s,j}^{\text{Tsai}}$) and DLT ($\Delta_{s,j}^{\text{DLT}}$) calibration techniques:

$$\Delta_{s,j}^{\text{method}} = \frac{\sum_{j=1}^{N_\text{comb}} \sum_{s=1}^{N_\text{s}} \Delta_{s,j}}{N_\text{s} \times N_\text{comb}} \tag{3.16}$$

When deriving $\Delta_{s,j}^{\text{Zhang}}$ and regarding Eq. (3.15), P_1 and P_2 should represent the corresponding projection matrices to the same Zhang plane (#1, ..., #13 in Fig. 3.12) containing each given projected spherule from

Figure 3.10 Epipolar geometry in stereo X-ray imaging.



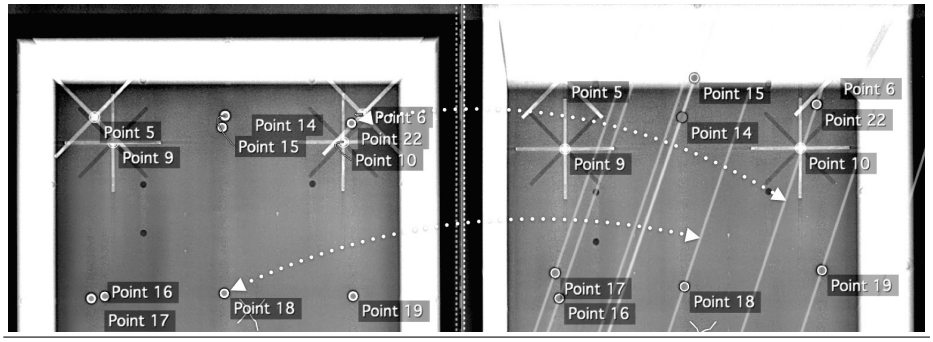
each stereo location.

In this occasion, a subset of 43 X-ray images was used (those with spherules, later added) and $N_{\text{comb}} = 903$ snapshot pair combinations. In Fig. 3.11 we show some real examples of these stereo pairs and epipolars.

3.5.4 Consistent intrinsic parameters

A final test might consist on analyzing whether the intrinsic parameters inferred at the RQ decomposition stage are physically consistent, i.e., if they represent cohesive physical dimensions and distances between the X-ray emitter and the detector. More specifically, we will pay attention to the focal length f of each X-ray system pose. From here we check whether we can derive a coherent spatial distance between anode and IP. In order to verify this aspect, the focal distance obtained with the DLT approach

Figure 3.11 Two stereo snapshot pairs and the epipolar line from each lead spherule as seen from *a paired* anode location.



will be taken as a *near true* distance and Tsai and Zhang methodologies will be compared against it (Δf_{DLT}). From experience DLT almost always delivers consistent intrinsic sections so it seems a good idea to compare the other two methodologies against this one.

This metric is not intended to serve as a comparison point for the three calibration methods but as a *final health check* of the consistency of each calculated P. A projection matrix can perform very good in all three aforementioned metrics. However, this fact does not necessarily mean that intrinsics (i.e., f) are well derived, neither that we end up with well formed and physically relevant $\mathrm{K}_{\mathrm{Tsai}}$ and $\mathrm{K}_{\mathrm{Zhang}}$ matrices.

3.6 Practical considerations

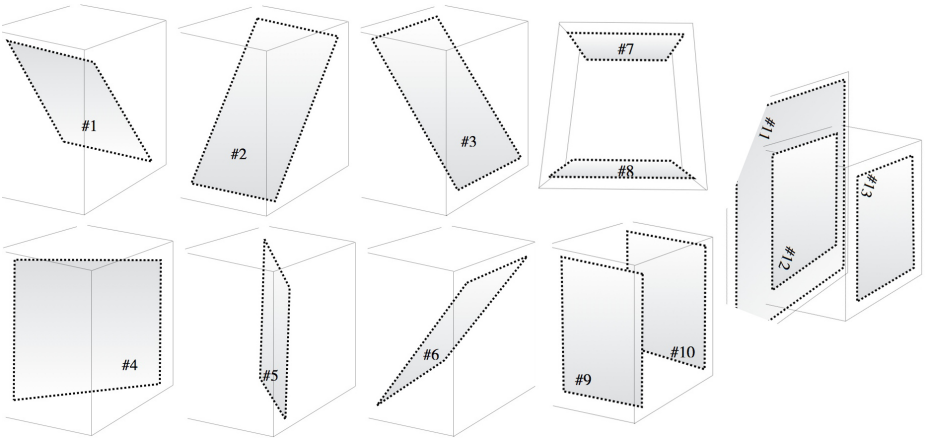
The most straightforward method is DLT, whose application starts with a set of correspondences between *world points* N_{f} and image points; a linear solution can be then computed. Tsai is quite similar but it compulsory needs an initial guess of x_0 and y_0 (usually the image center), which later has to be fine-tuned with a least-squares stage.

The Zhang method differs significantly from the other two, as it operates over 2D frames and their planar projections. As with DLT and Tsai,

with Zhang we initially assumed that the accuracy of the camera calibration process would be in proportional relationship with the amount of 3D-2D point correspondences, which in turn implies more planar projections (as discussed in Section 3.3.3). In this context, we initially conceived a total of $N_{\boldsymbol{p}} = 13$ *virtual planes* using all possible fiducial marker combinations in the DUT. These ideal planes are shown in Fig. 3.12 and a sample subset of their X-ray projections in Fig. 3.13.

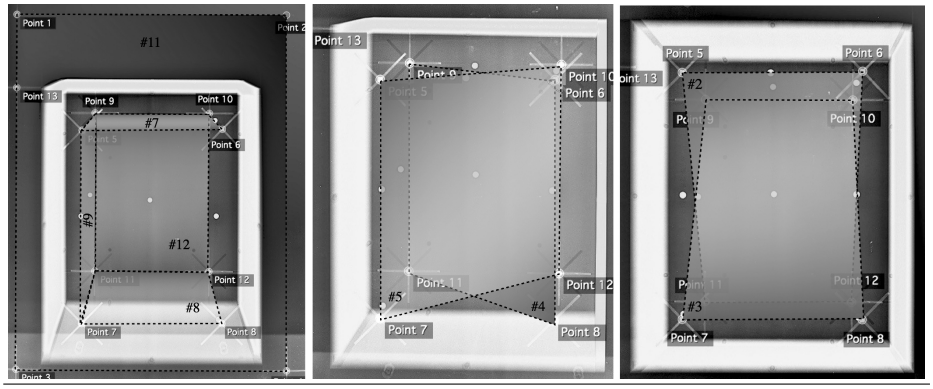
Each of these planes contains 4 fiducial points, except for instances #11 and #12, which can be *expanded* with the extra antisymmetry copper landmark (whose X-ray projection is labeled as *point 13* in Fig. 3.13).

Figure 3.12 6 of the total of 13 *made-up* internal planar structures inside the 3D DUT used for the Zhang calibration method.



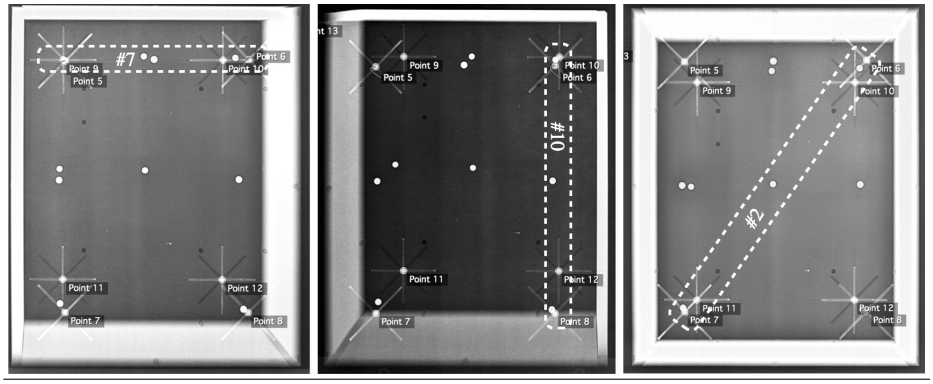
Lead spherules described in Section 3.4 have been situated as coplanar as possible to some of the *built* Zhang planes from #1 to #10 shown in Fig. 3.12, though some have been left out the calculations if their geometrical distance exceeded 5 mm from a given plane. In the case of the Zhang method, each plane is related to its own projection matrix $\mathrm{P}_{1\cdots13}$, and thus, it is only possible to project 3D traits that lie on each specific plane only. This consideration has been taken into account when applying the comparison metric described in Section 3.5.3.

Figure 3.13 Some X-ray projections of *synthetic* Zhang planes *buildable* with fiducials present in the DUT.



After running some preliminary tests, we found that using all the 13 planes for calibration produced unstable results in many cases. A more careful look showed that fluctuating outcomes were obtained when some of the virtual planes turned out to lie very perpendicularly relative to the image surface and/or anode. When our DUT was radiographed, these steep planes projected an almost negligible area. In Fig. 3.14 we show some examples of this type of *problematic* traces.

Figure 3.14 X-ray projections of some *problematic* Zhang virtual planes (inside the dotted line) which affect results.



For this reason, we have devised a pre-step to the Zhang algorithm that easily discards planes that generate unstable results based on their projected area, in number of pixels. Section 3.7 explores this topic.

Besides, in order to be able to apply the test metric described in Section 3.5.2, we need to select the projection matrix P_{11} linked to plane #11 (Fig. 3.12) on each stereo snapshot. This plane contains the fiducial cross marker (upper-left corner) that defines a shared W reference frame. The matrix P_{11} linked to this plane, when RQ-decomposed, reveals the translation vector $\mathbf{t}$ that allows us to derive the detector resolution λ_j for a specific radiograph stereo pair j, using Eq. (3.13).

As discussed in Section 3.3, all calibration methods allow a further optimization through non-linear minimization. We have chosen the *geometric distance* as the non-linear cost function.

3.7 Results and discussion

Table 3.1 compares all calibration algorithms using the metrics introduced in Section 3.5. We provide the results *before* and *after* the refinement process (denoted with suffix /R). As introduced in Section 3.5, σ is the reprojection error –using Eq. (3.12)– of the same fiducial markers taken into account for calibration. Metric Δ is the mean distance from each spherule image location to the epipolar of that same spherule –obtained with (Eq. (3.15))– seen from another paired anode. Δ is calculated with Eq. (3.16). It should be brought to mind that, in the case of the Zhang algorithm, each spherule is projected using the particular P_j related to the plane which best embeds it.

The mean resolution λ for the detector is estimated in pixels/meter and as a result of Eq. (3.13). λ is then compared with factory indications λ_o for each calibration method. We recall here that in order to compute λ in the case of the Zhang method, the projection matrix P_{11} for plane #11 and its translation vector $\mathbf{t}$ have been used.

We also draw some attention to the quality of f provided by Tsai and Zhang methods by comparing it with the focal length derived with DLT (Δf_{DLT}).

Table 3.1 Test metrics and results for the three calibration methods. The /R suffix means non-linear least squares refined.

		Tsai	DLT	Zhang	Tsai/R	DLT/R	Zhang/R
σ	px	36	7	270	7	7	19
std(σ)		15	5	136	5	5	11
Δ	px	15	3	69	3	3	12
std(Δ)		19	6	49	7	6	8
λ	$\frac{\text{px}}{\text{m}}$	10071	9993	10114	9991	9991	10040
std(λ)		287	192	237	181	190	278
Δf_{DLT}	m	0.10		0.12	0.00		0.01
std(Δf_{DLT})		0.06		0.02	0.00		0.06

In general, all methods report good results and show capabilities to be used for X-ray calibration. σ is very low, λ is virtually identical to λ_o and Δ is nearly *zero*.

DLT reports best results even without a minimization phase. This entails that this time-consuming stage can be ignored. On the contrary, Zhang and Tsai need undergoing a non-linear minimization step in order to deliver accurate results.

Although Tsai's approach is quite precise, and even operates better when refined through a minimization episode, it must be initialized with the principal point coordinates, which are, a priori, unknown in X-ray environments. If the X-ray emitter is located more or less perpendicularly to the examined object, as in setup a) in Fig. 3.15, x_0 and y_0 can be assigned the coordinates of the image center w.l.o.g and still get satisfying results. However, at poses approaching the limits of the X-ray system, this central landmark can have a distant coordinate from the center, as exemplified in Fig. 3.5.

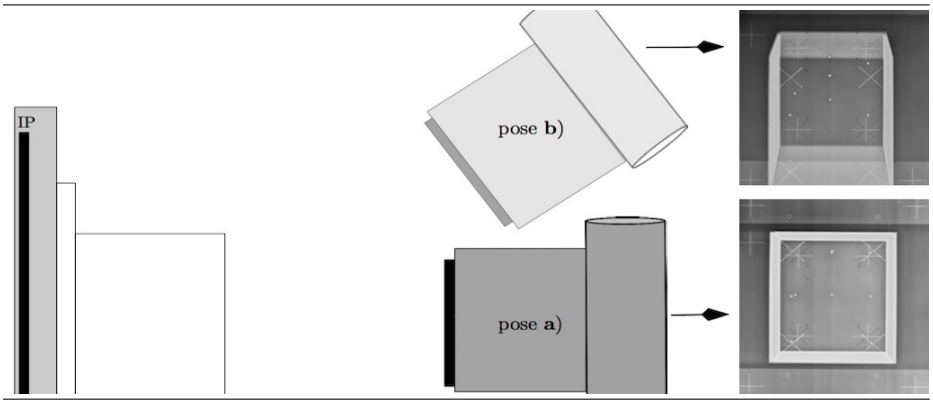
As detailed in Section 3.6, In the case of Zhang's method, we have had to first reject some *degenerate* planes with the simple technique of requiring a minimum projected area belonging to each virtual plane. Most accurate results for this method are recovered when a minimal area of 10^3

px for each plane is set as a threshold. However, these results are not as accurate as those obtained with DLT/Tsai. An explanation for this may reside in the fact that planes should be populated with a denser fiducial grid, as [214] also observed.

In order to better visualize the influence of the minimum required projected area in Zhang's algorithm, we have generated Fig. 3.17. We there have plotted the mean difference (Δf_{DLT}) of the Zhang's derived focal lengths f_{Zhang} to those obtained using DLT (f_{DLT}).

From these plots we can infer that the best performance seems to be achieved with the combination of 4 *large* projected planes: the three widest parallel ones (#11, #12 and #13 in Fig. 3.12) plus a fourth extra one. In other words, the Zhang method can be improved by adding more planes, but it is of greater importance which planes are selected based on their projected area. Too many parallel planes or too perpendicular can interfere the calibration process. This outcome agrees with the experimental results obtained by [262]. We also found that, Zhang's algorithm also delivered better results when the DUT was radiographed from wider angles –configuration b) in Fig. 3.15– that originated larger projected zones.

Figure 3.15 Geometrical configuration of the X-ray system that works better with Zhang's approach (b) and frontal placement that suits better Tsai's algorithm (a).



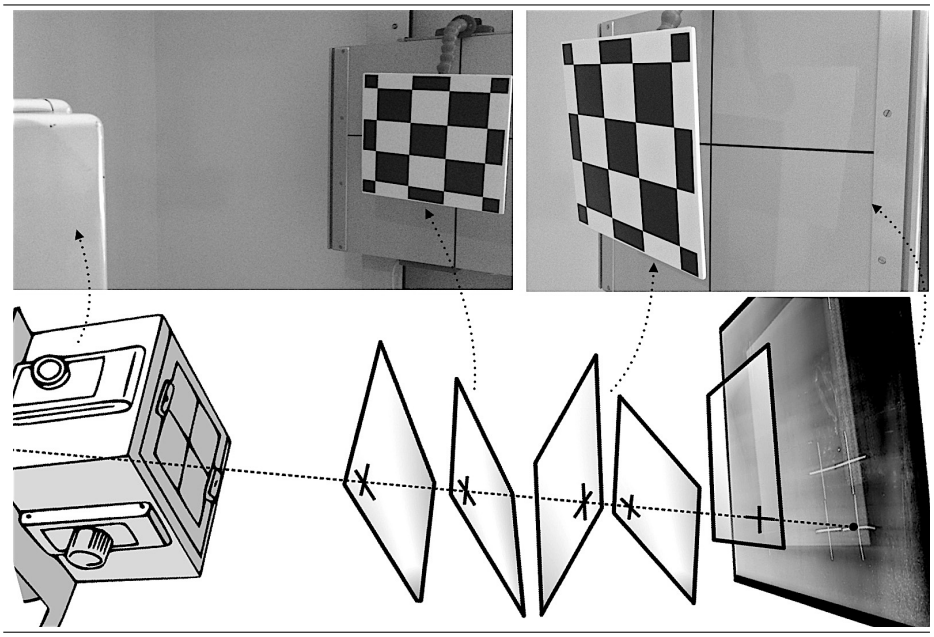
Zhang's method has also been evaluated in its original conception for computer vision: using a single 2D calibration frame ($N_{\boldsymbol{p}} = 1$) projected at alternating poses (discussed in Section 3.3.3). This parallel study is

detailed below.

3.8 Alternate Zhang calibration experiment with a 2D calibration frame

As a side experiment, we have radiographed a separate 2D squared DUT –equipped with 16 radiation-opaque fiducials plus 4 extra to break symmetry– at 27 arbitrary orientations and distances (Fig. 3.16 and Fig. 3.17-b). We have designed a special holding structure consisting of an adjustable

Figure 3.16 Visual representation –for the Zhang method– of the experimental setup with a single calibration flat panel adjusted over a strand of flexible plastic segments which allows perfect positioning. The panel, equipped with 16 hidden X-ray-opaque markers, is then beamed at random poses and f_{Zhang} is then derived.



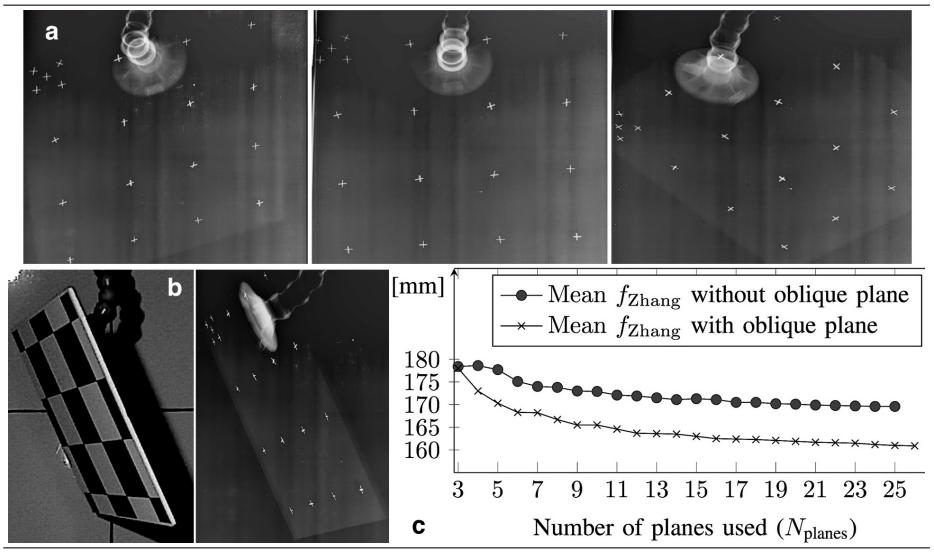
support system with a panoramic head that allowed us to arbitrarily move around and re-orientate this structure before been radiographed.

The focal length f_{Zhang} has been estimated with the help of this device while the X-ray imaging system has remained still throughout the measurements. We have calculated the mean focal length (Fig. 3.17-c) for each number of plane combinations (i.e., 2300 combinations for 3 planes, 12650 for 4 planes, 53130 for 5 planes, and so on).

These results clearly demonstrate how Zhang's method, together with a humbler 2D frame, can represent a good alternative to calibrate any X-ray device. As expected, the more projected planes involved in the calculation, the more accurate –or convergence towards a stable– focal length. Nevertheless, some combinations of 3 or 4 planes, already offered a very precise solution. As an example, Fig. 3.17-a shows a combination of just three planes from which we can already derive the same f_{Zhang} achievable with the remaining 24 projections.

In our tests, we also discovered a degenerate projection instance (tackled in Section 3.6). In Fig. 3.17 we show such *conflictive* plane whose pose is very oblique and projected area very small. It is important to identify these planes and taken them out of calculations in order to speed up a stable and consistent calibration. Their inclusion can dramatically affect results, as Fig. 3.17-c clearly demonstrates (this time correct focal length is known by construction).

Figure 3.17 (a) combination of three random Zhang planes that produce a very accurate focal length. (b) Problematic configuration of a Zhang plane. (c) Convergence of the mean f_{Zhang} determined using the random positioning of DUT (with and without oblique planes) towards the stable mean focal length due to an increase in the number of planes used.

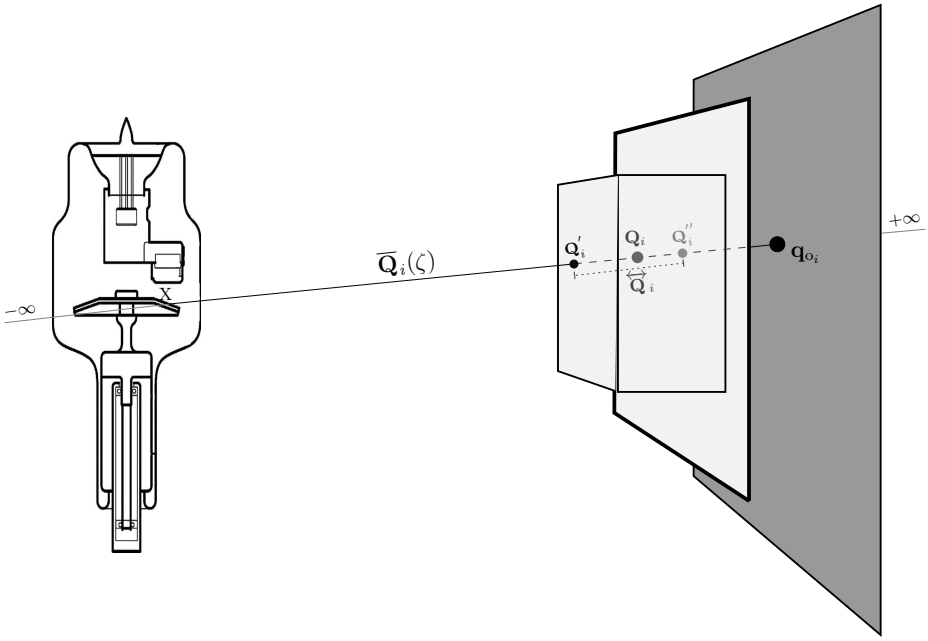


3.9 Conclusions

We have applied three well-known camera calibration methods to primary diagnostic X-ray environments. All calibration methods share the same principle: finding a projection matrix that connects scene 3D points and their projections in the image. A well calibrated camera allows us to *back-project* points in the image (a radiograph, in this case) to their corresponding location in the 3D space (as shown in Fig. 3.18). This fact will enable the registration between different camera systems, as explained in Section 4.6.3 and Section 8.1.5 for visual and X-ray images and for visual and γ images, respectively.

In order to study these calibration methods, two 3D adjustment frames outfitted with fiducial markers and a set of metrics has been devised. The main and most immediate outcome is that all three methods can perform nicely on these type of X-ray settings. In other words: all three methodologies, with a similar calibration frame to ours, will deliver functional

Figure 3.18 Back-projection of an image j 2D point into the 3D scene in the form of an infinite ray $\overline{\mathbf{Q}_{j,i}}$ (expressed in the W reference frame). A finite segment of this ray is also shown $\overleftrightarrow{\mathbf{Q}_{j,i}}$ and roughly spans the length of the DUT.



projection matrices and accurate intrinsics, before or after a non-linear least squares refinement. However, there are some meaningful operational differences and particularities related to each technique that have been analyzed as well.

We have paid special attention to the Zhang method, massively used for computer vision but not in regular X-ray contexts. We show it can be a reasonable alternative to heavier 3D phantoms. Within a Zhang framework, three radiographs from a simpler and single planar device can suffice, however, very oblique projections can significantly alter calculations and should be previously identified and discarded. From the experiences reported in this text, fixed 3D calibration structures are more recommendable if working under DLT or Tsai schemes. For this reason we chose DLT for the rest of the calibration operations with X-ray equipment performed in this work.

4

Calibration and augmentation of X-ray settings with visual information

In this chapter, we present a methodology to recover the geometrical calibration of conventional X-ray settings with the help of an ordinary video camera and visible fiducials that are present in the scene. In contrast with what was tackled in Chapter 3 and following the goals elaborated in the introduction of this thesis work, we now resolve the scene's geometry thanks to the interplay of an external environment recognition device.

After calibration, equivalent points of interest can be easily identifiable with the help of the epipolar geometry. The same procedure also allows the measurement of real anatomic lengths and angles and obtains

accurate 3D locations from image points. Our approach completely eliminates the need for X-ray-opaque reference marks (and necessary supporting frames) which can sometimes be invasive for the patient, occlude the radiographic picture, and end up projected outside the imaging sensor area in oblique protocols.

Two possible application frameworks are envisioned: a spatially shifting X-ray anode around the patient/object and a moving patient that moves/rotates while the imaging system remains fixed. As a proof of concept, experiences with a device under test (DUT) have been carried out. The results show that it is possible to identify common points with a proper level of accuracy and retrieve three-dimensional locations, lengths and shapes with a millimetric level of precision.

The presented approach is simple and compatible with both current and legacy widespread diagnostic X-ray imaging deployments and it can represent a good and inexpensive alternative to other radiological modalities like CT. In the context of this thesis work, this entails an evolution from what was addressed in [10] and [58].

4.1 3D information from ordinary radiographs

Recovering the geometrical information from multiple X-ray snapshots of a same object/patient generated from different angles and positions has become of increasing relevance in medicine, industry and surveillance. More specifically, interest has grown in relation to the identification of common points or areas of interest in several radiographs, and the derivation of useful 3D information (distances, angles, etc.) from a sparse set of images produced in conventional and primary diagnostic X-ray imaging settings. In this context, distances are usually very poorly estimated, and, in many cases, simple X-ray-opaque objects (like coins) are used as reference landmarks. In this chapter, a new methodology that represents a step forward towards a better assessment of anatomical distances with standard/legacy diagnostic equipment, fiducial-less radiographs and less

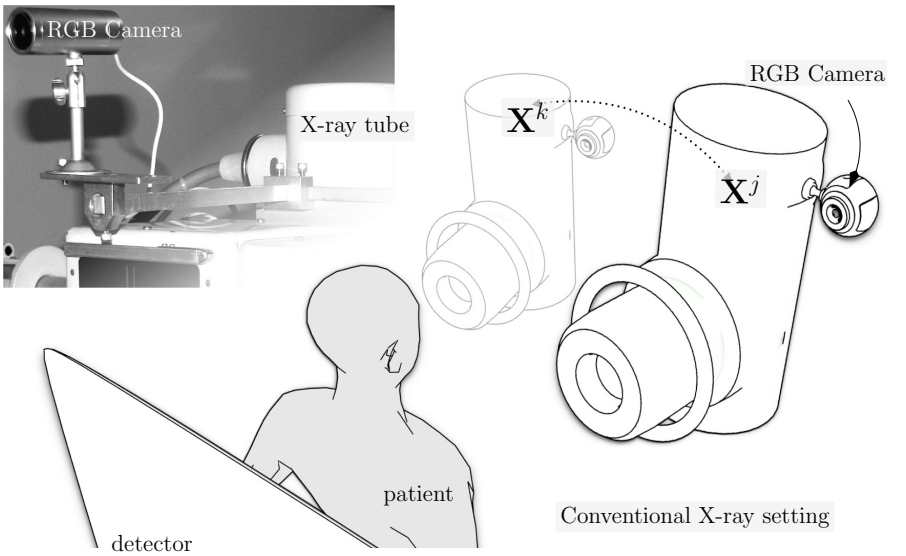
invasive frameworks is presented.

In computer vision, the process of retrieving 3D information from 2D bitmaps is usually referred as *image reconstruction* and can be achieved after a camera calibration phase. The geometric calibration of X-ray modalities starts with the inference of perspective projection matrices that map 3D scene points with their *projected* counterparts. Unfortunately, this process entails overcoming some obstacles, which will be addressed in this chapter. Perhaps, the most important one has to do with the huge level of spatial variability present in typical X-ray imaging environments, as highlighted and summarized in Section 3.2.4.

In these basic radiological settings, both the radiographed object and the X-ray imaging system can move with almost complete freedom around the room. These spatial changes are then usually coded as rigid transformations or translation vectors. As a collateral benefit, keeping track of the distances, positions, and orientations (*geometrical settings*) that dominate the radiographic scene can also play a role in beam equalization, and therefore, in dose control.

In this chapter, we account for these scene alterations in basic X-ray environments with the help of a visible light camera that is *rigidly tied* to the X-ray source (Fig. 4.1).

Figure 4.1 Proposed X-ray acquisition system with an attached digital camera that allows the tracking of the position/orientation of the X-ray source.



Alongside this goal, we also elude the use of X-ray-opaque markers (Fig. 4.4) for the reasons later elaborated in Section 4.3. These X-ray devices can be found in most healthcare centers worldwide, private and public, large and small. In the reconstruction process, only the aforementioned camera and a set of visual fiducial markers are used so that X-ray images do not get tainted with traces/projections from foreign objects. Our proposed method initially needs a calibration phase to retrieve the geometrical setting of both cameras. Afterwards, subsequent movements and/or rotations of the X-ray system or the examined object can be precisely tracked with the help of the RGB device. With this spatial information, X-ray projection matrices are then calculated and multiple radiographic image reconstruction can take place.

We also study two different scenarios (Fig. 4.2 and Fig. 4.3) that might be appropriate for X-ray examinations. In the first scenario, the camera system (X-ray + video) moves around the examined subject/object. Under this scheme, we successfully manage to *recreate* X-ray camera intrinsic parameters. The second scenario entails moving and/or rotating the radiographed patient (or item) while keeping both the camera system and

the X-ray detector fixed. Both settings can be relevant in clinical diagnosis and/or object monitoring/scanning.

Figure 4.2 First application setting: camera system moving around still patient. In this case we highlight other type of commercial markers for computer vision engineered by Intersense® from Thales Visionix Inc.

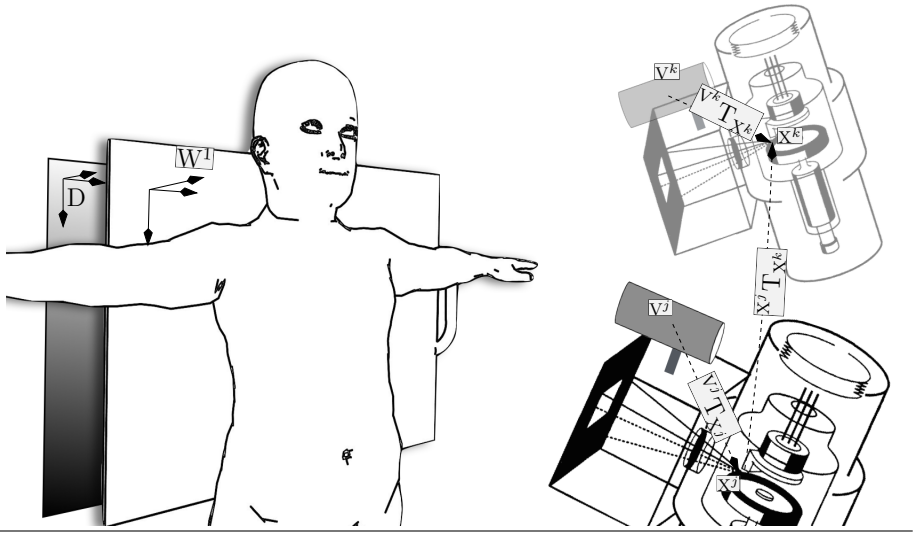
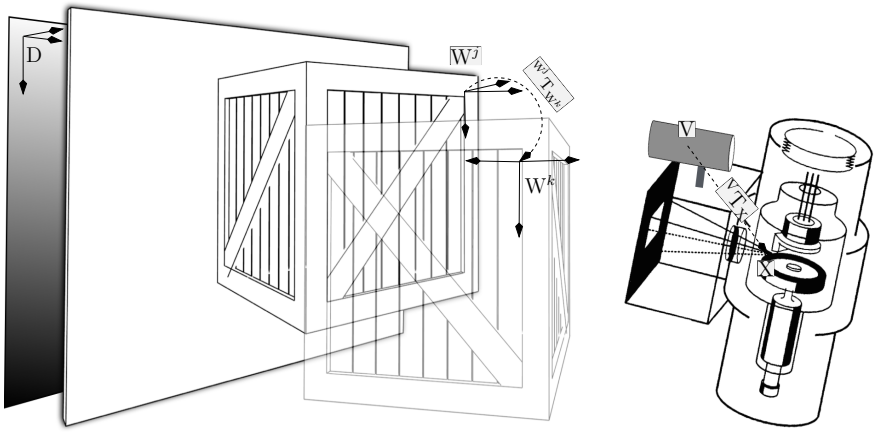
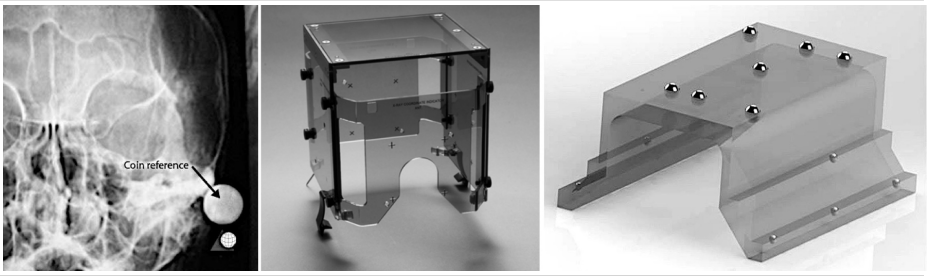


Figure 4.3 Second application setting: fixed camera system with rotating/shifting object. As an example, visible fiducials from the ARToolkit [118] project are shown.



Subsequently, we discuss some practical epipolar geometry tools and how they can be useful for diagnosis in plain X-ray imaging. Finally, a few tests carried out with a geometrical phantom are presented and their results are discussed.

Figure 4.4 X-ray opaque markers commonly used for 3D referencing in X-ray imaging.



4.2 Related work

X-ray calibration techniques and X-ray image registration procedures commonly use calibrated C-arms and CT scanners [64], which involve, of course, having access to such radiological devices. Tomosynthesis [87] devices are more affordable and are already widely used in digital screening mammography. However, their application in everyday X-ray examinations, where legacy hardware is the common denominator, may require the complete renewal of the imaging set.

In order to achieve the same goal with less expensive, more accessible/widespread radiological tools and lower dose levels, several approaches have been (or are currently being) explored. All these efforts involve the use of an external and adjacent device that interplays with the X-ray imaging apparatus. For instance, there is a trend in research focused on using depth and time-of-flight cameras in order to reconstruct 3D data and 3D models of objects being radiographed or scanned. This 3D data is then combined with X-ray images to obtain different and meaningful information. As an example, the approach followed by [236] combines

3D modeling with X-ray images in order to three-dimensionally locate and define the shape and silhouette of hidden objects inside boxes, which can have very interesting applications in surveillance and QA processes. Another precedent can be found in the study carried out by [57], who estimates patient's size, volume and appropriate dose with the help of a Microsoft Kinect™ device. Other attempts like the one described by [164] try to rebuild the X-ray system extrinsic parameters with the help of a distance meter device (in this case, a laser rangefinder located close to the X-ray emitter). The problem is further simplified by the authors of [159], who require each X-ray source and sensor to be placed at known locations. Conventional approaches use special X-ray calibration structures that accommodate fiducials that leave visible *shadows* in the radiographic image. These foreign frames remain present during each snapshot (either attached to the patient [218] or to the X-ray system) *contaminating* the acquired radiographs with their own projected traces, introducing artifacts and invading the patient's space.

As stated in Section 4.1, there also exist recent and laudable research efforts around interplaying video information with X-ray imaging. In [170, 256] and [47] researchers calibrate, model and study the clinical and surgical applications of camera augmented mobile C-arms, which also involve the precise registration of visible and X-ray images. Close to this work, we find that carried out by [175], which highlights the contribution of external cameras to radiation exposure reduction and surgery planning improvement. Radiograph mosaicking is also a subject of interest. In this direction, we find interesting citing the work performed by researchers in [255] who focus on accurate X-ray image stitching (also in C-arm modalities). Their goal requires a pure rotation around the X-ray source center which is accomplished with the help of visual fiducials and a video camera that contribute to estimating the translational part of the motion so that it can be later compensated. Similarly to this approach, we also derive the translational motion of the X-ray source with the help of visual markers and a RGB camera. However, two key differences arise. First, one of the main goals of [255] is the registration of the X-ray and RGB modalities. For this reason a set of mirrors is used so that both optical centers are made coincident. In our case this is not necessary because we do not perform image registration. Second, the X-ray source in [255] undergoes a pure rotational motion so that it is possible obtain a parallax-free mosaic of both imaging systems. Regarding our goal, we are also interested in the

translational part of the beam origin in order to derive 3D information from plain radiographs, as later tackled in Section 4.6.2.

Beyond image-to-image registration, we find increasing interest in relating bitmap content (radiographs, CT/MR slices and video) to volumes. The authors of [135] and [144] present a concise review of state-of-the-art around the topic applied to minimally invasive therapy and image guided interventions. The reconstruction of 3D structures from sets of 2D X-ray projections is studied in [208] and [52] with the help of custom designed phantoms. However, none of these research efforts explore the event of an alteration of the geometrical setting. Researchers in [192] and [21] try to tackle this problem by using image similarity measures (entropy, intensity, gradient, patterns, etc.) which can be used (with some difficulty) without the need of a phantom or an *ad hoc* calibration phase.

In addition, the literature on the combination of visual fiducials with radiology, medicine, and surgical environments is steadily growing. The works in [70, 80, 92] and [141] are just a few examples of applications of the area of augmented reality (AR) in hospital and clinical environments.

4.3 Derivation of X-ray projection matrices from visual information

Unlike other setups that can retrieve 3D information (like C-arm), ours relies on standard clinical X-ray imaging systems, like those comprised of a 4 degrees of freedom (DOF) *movable* X-ray source and a 1 or 2-DOF vertical/horizontal sliding imaging plate or IP (which receives and integrates the emitter radiation). During examinations, the patient is placed vertically next to this detector (*standing erect*) or horizontally over it (*supine anteroposterior projection*). In these radiological settings, there is no way to know with enough precision, the beam source spatial position relative to the detector or, more generally, to a coordinate reference frame (later referred to as *world*). Therefore it is not aprioristically possible to find projection matrices and achieve geometric reconstruction.

However, X-ray imaging devices have been commonly modeled as pinhole cameras, which enables the application of projective geometry. The research efforts carried out by [221, 143, 159] and [207] represent a few examples, and a more succinct introduction on the subject applied to the context of medical imaging can be found in Section 2.2.1 of [152] and chapter 20 of [42]. As stated in our introduction, in order to combine several X-ray images from the same object/patient at each different geometrical setting j, we need associated camera projection matrices P^j, P^k, etc. Each P^j relates 3D points ${}^{\mathrm{W}}\mathbf{Q}_i$ in a coordinate reference frame called *world* (W) with their 2D *observed* projections $\mathbf{q}_i^j$ on each radiographic image j. Using projective geometry, we can write:

$$\hat{\mathbf{q}}_i^j = \mathrm{P}^j \cdot {}^{\mathrm{W}}\hat{\mathbf{Q}} \tag{4.1}$$

where ${}^{\mathrm{W}}\hat{\mathbf{Q}}_i$ and $\hat{\mathbf{q}}_i$ are the homogeneous coordinates of ${}^{\mathrm{W}}\mathbf{Q}_i$ and $\mathbf{q}_i^j$, respectively. With a RQ decomposition, P^j can be expressed as:

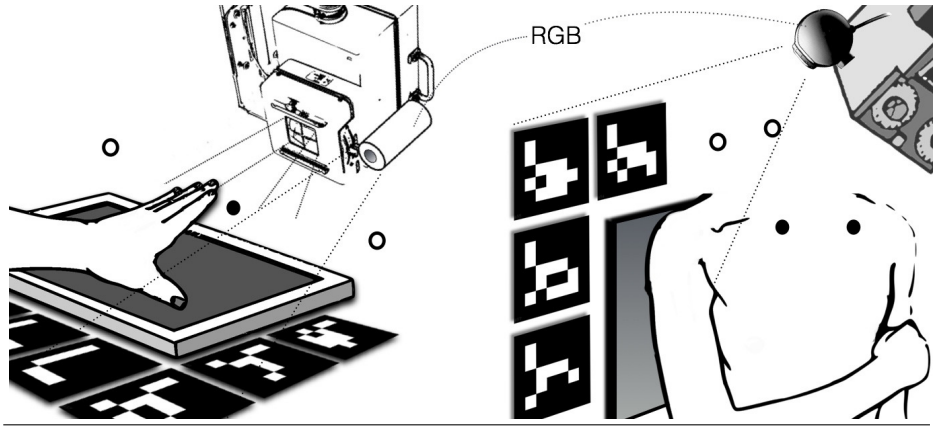
$$\mathrm{P}^j = \mathrm{K} \cdot {}^{\mathrm{X}^j}\mathrm{T}_{\mathrm{W}} \tag{4.2}$$

where K is a 3×3 upper triangular matrix that contains the intrinsic parameters of the X-ray system (for a given geometrical setting j) and ${}^{\mathrm{X}^j}\mathrm{T}_{\mathrm{W}}$ is a rigid transformation that translates 3D homogeneous points relative to W to coordinates of the X-ray camera (X^j), whose reference frame is centered at the radiation emitting anode (focal point) and one of its axes is orthogonal to the X-ray detector plane. In this document, rigid transformations are expressed with the nomenclature ${}^{\mathrm{dest}}\mathrm{T}_{\mathrm{orig}}$, that is, how points in the *origin* reference frame are translated to the *destination* coordinate system.

As already highlighted in Section 4.2, conventionally in X-ray imaging, projection matrices are obtained with the help of a calibration frame equipped with fiducials $\mathbf{Q}_i$ that are opaque to the Roentgen radiation which are then projected to $\mathbf{q}_i$ spots in the image. This frame is placed around or over the examined object/patient. Combinations of $\mathbf{Q}^j{}_i$, $\mathbf{q}_i^j$ pairs are then fed into a calibration algorithm such as Direct Linear Transform or DLT (introduced by [27] and very succinctly described in chapter 4.1 of [98]) and projection matrices can be then derived.

The problem with such radiation-opaque fiducials is that they usually leave visible traces, artifacts and extra Compton contribution in the radiographic image that can seriously obstruct the analysis and/or diagnosis process. They also can be easily projected outside the imaging sensor area when using oblique protocols, like those exemplified in [1] and in Fig. 4.5, which can harden their application. Last but not least, essential supporting frames in which they are usually accommodated, can be perceived as invasive by the patient.

Figure 4.5 Examples of lateral and oblique X-ray protocols whose projections (i.e., imaged fiducial set of black and white circles) and principal point may lie outside the final image (i.e., white circles). However the system can still be geometrically tracked through visible information and RGB cameras.



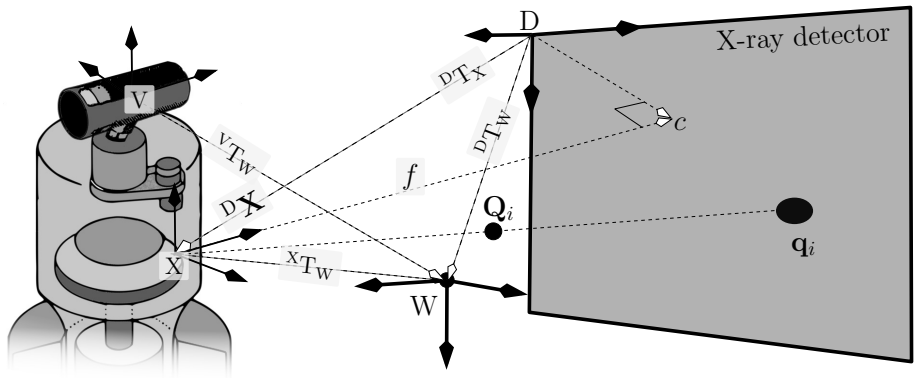
In this context, the following sections describe in detail how each P^j, for each different position j of the X-ray emitter (or patient), can be obtained using just the help of an interplaying visible light camera that has a broader field of view and whose fiducials are transparent to the radiograph production process.

4.3.1 Coordinate systems

Although some coordinate systems have already been briefly mentioned, here we review in depth the ones that are used in our proposal. These are graphically summarized in Fig. 5.2 where, for clarity, we assume the

camera system is located at a given position j (i.e., X^j and V^j).

Figure 4.6 Coordinate frames available in an X-ray setting. The point ${}^{\mathrm{D}}\mathbf{X} = (c_x, c_y, f)$ represents the anode location in detector coordinates. ${}^{\mathrm{D}}\mathrm{T}_{\mathrm{X}}$ is the rigid transformation that connects the detector and anode reference frames.



- W, or the world coordinate system, whose origin is a 3D spot in the scene.
- D, the reference system of the X-ray detector, whose origin is usually situated at the upper-leftmost pixel with one of its axes (Z) being orthogonal to the detector plane.
- X, the X-ray coordinate system, whose origin is the beam source and which also has one axis (i.e., z) orthogonal to the detector by definition. Note that there is a different X^j coordinate system for each spatial position of the X-ray emitter but all of them have their z-axis orthogonal to the detector plane.
- V, the video camera coordinate system. Since it is rigidly attached to the X-ray emitter, there will also be a different V^j for each location of the anode.

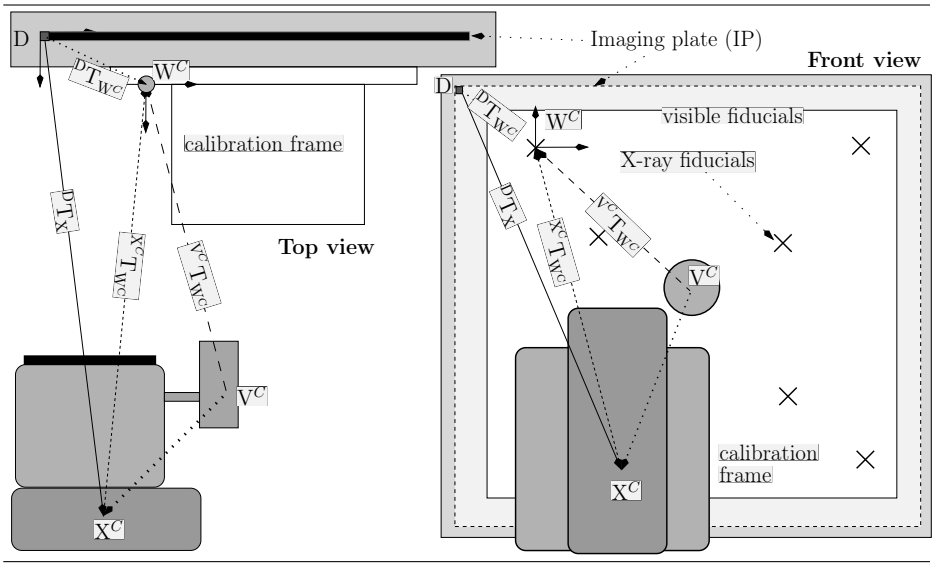
Fig. 5.2 shows that the relation between reference frames D and X is always a π rotation around the horizontal axis (x) of D and a translation in the case of X-ray cameras. This relation can be coded as a rigid transformation

expressing how to translate points (in homogeneous coordinates) from X to D as follows:

$$ {}^{\mathrm{D}}\hat{\mathrm{T}}_{\mathrm{X}} = \begin{bmatrix} \mathrm{R}_x(\pi) & {}^{\mathrm{D}}\mathbf{X} \\ \mathbf{0} & 1 \end{bmatrix} = \begin{pmatrix} 1 & 0 & 0 & c_x \\ 0 & -1 & 0 & c_y \\ 0 & 0 & -1 & f \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{4.3} $$

where ${}^{\mathrm{D}}\mathbf{X}$ is the beam origin of the X-ray expressed in D coordinates. The mark ($\hat{}$) over ${}^{\mathrm{D}}\hat{\mathrm{T}}_{\mathrm{X}}$ indicates that an extra row $(0, 0, 0, 1)$ has been appended to the transformation matrix, which allows us to work in homogeneous coordinates. Fig. 4.7 shows a front and top views of the dual imaging setup.

Figure 4.7 Top/front views of our setup during calibration.



4.3.2 Intrinsic and extrinsic parameters in X-ray cameras

In a radiological device that is modeled as a pinhole camera, the beam origin is represented by the tube anode, which also plays the role of the optical center. If there are no lenses present, we can ignore spherical aber-

rations, radial distortions, and skew ($s = 0$), w.l.o.g. Using perspective projection equations, it is possible to transform 3D points in X coordinates to 2D homogeneous radiographic image dots:

$$\hat{\mathbf{p}}_i = \mathrm{K} \cdot {}^{\mathrm{X}}\mathbf{Q}_i \tag{4.4}$$

The intrinsic parameters matrix K can be decomposed as:

$$\mathrm{K} = \begin{pmatrix} \alpha_x & 0 & x_0 \\ 0 & \alpha_y & y_0 \\ 0 & 0 & 1 \end{pmatrix} = \underbrace{\begin{pmatrix} \lambda_x & 0 & 0 \\ 0 & \lambda_y & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{\lambda} \cdot \begin{pmatrix} f & 0 & c_x \\ 0 & f & c_y \\ 0 & 0 & 1 \end{pmatrix} \tag{4.5}$$

where c_x, c_y are the coordinates of the so-called *principal point* and f is the focal length, which perpendicularly connects the anode and the detector plane. These components are expressed in *pixels*, but they can be transcribed to spatial dimensions (meters) if multiplied by a known resolution λ, which is provided by the detector manufacturer. It is conventionally assumed that λ is axis-independent, i.e., $\lambda_x = \lambda_y$. Both Fig. 5.2 and Eq. (4.5) show that the components of the optical center relative to D are coincident with the intrinsics:

$${}^{\mathrm{D}}\mathbf{X} = (c_x, c_y, f) \tag{4.6}$$

Together with Eq. (4.5), Eq. (4.6) reveals that the intrinsic parameters of any X-ray system will change if either the sensor or the emitter are moved, as already highlighted in Section 4.1. This variability lies in the fact that the detector surface and the radiation source are detached from each other. This peculiarity vividly contrasts with video camera systems, where the sensor (typically a CCD/CMOS array) stays fixed relative to the optical center.

The extrinsic parameters can be written as a rigid transformation that relates points in word coordinates (W) and X-ray camera coordinates (X) and can be derived from ${}^{\mathrm{D}}\mathbf{X}$ with the following expression:

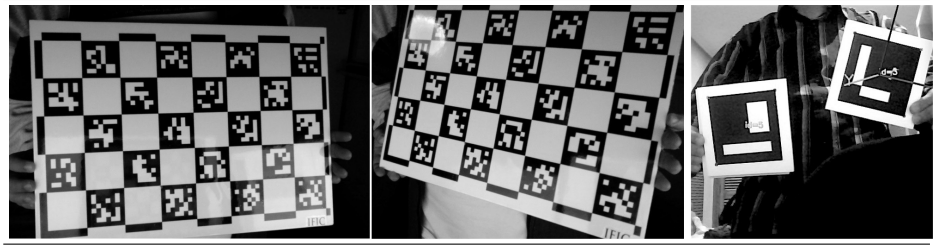
$${}^{\mathrm{X}}\hat{\mathrm{T}}_{\mathrm{W}} = \left({}^{\mathrm{D}}\hat{\mathrm{T}}_{\mathrm{X}}\right)^{-1} \cdot {}^{\mathrm{D}}\hat{\mathrm{T}}_{\mathrm{W}} \tag{4.7}$$

where ${}^{D}T_{X}$ can be obtained using Eq. (4.3) and ${}^{D}T_{W}$ is one of the results of the initial calibration process (described in Section 4.5) that maps points from W to D.

4.3.3 Scene tracking with visible fiducials

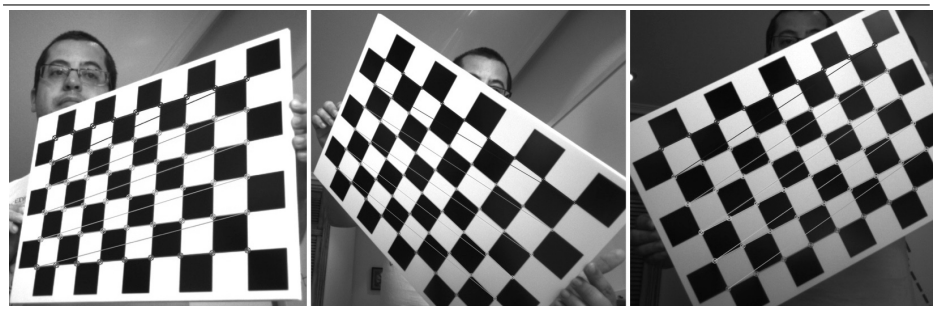
From Eq. (4.5) and Eq. (4.7), we conclude that the spatial location of the X-ray optical center ${}^{D}\mathbf{X}$ is essential for the calculation of the intrinsic and extrinsic parameters of the X-ray system.

Figure 4.8 Visual marker detection for both calibration and real-time positioning.



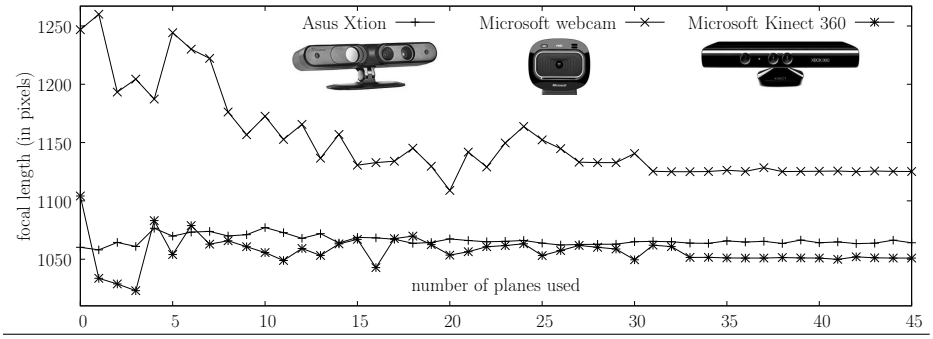
In our proposal, after a single calibration (itemized in Section 4.5), ${}^{D}\mathbf{X}$ is tracked with the help of visible fiducial markers and an ordinary RGB camera. An initial optical calibration of this device is necessary, which provides its invariant intrinsics (K_V) and contributes to minimizing distortions produced by lenses. This process is usually achievable with the help of modern computer vision toolkits, as shown in Fig. 4.9 and Fig. 4.8.

Figure 4.9 RGB camera intrinsics calibration with a computer vision toolkit such as OpenCV.



This intrinsics derivation process is grounded on the Zhang method discussed in Section 3.3.3. As also done in Section 3.8, we have performed a quality assurance phase to evaluate which planes are the correct ones to be integrated into the minimization process. In Fig. 4.10 we show the convergence of the focal length for some of the optical cameras used in this research.

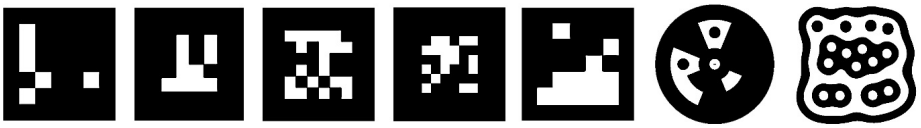
Figure 4.10 Evolution of the focal length as more images of the chessboard-like pattern (Fig. 4.9) are added to the minimization process.



The extrinsic parameters ${}^{V}T_{W}$, which connect points in W coordinates to the V reference frame, can be determined with the help of visible fiducials with known 3D coordinates. Some examples of other commonly used visible fiducial markers (Fig. 4.11) are ReacTIVision [117], Intersense [169], BinARyID [85], ARTag [79] and Cantag [204].

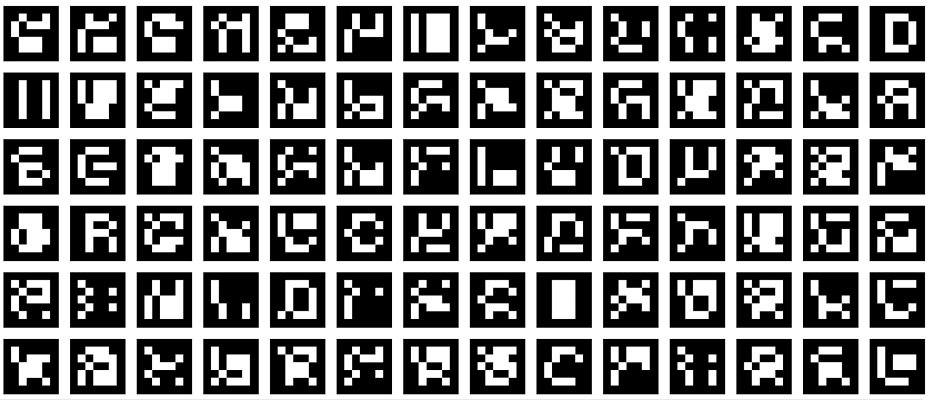
In this work, we use the Aruco markers described in [86]. The Aruco framework was chosen over other alternatives because of its robustness

Figure 4.11 Some industry-standard visual makers. From left to right: Aruco, BinARyID, ARTag, ARToolkitPlus, Cantag, Intersense and ractTIVsion.



against noise and vertex jitter. There are 1024 different instances of this fiducial system that can be easily detected in real time (Fig. 4.8). In Fig. 4.12 we show a few examples of these instances (also in Fig. 4.12). Aruco also comes with a companion C++ library and utilities that allow a straight and fast integration in our workflow. An interesting aspect from this fiducial system is that the *camera pose* (location and orientation) can be determined with just one Aruco marker with enough fidelity, if necessary. This feature makes the system very robust to occlusions made by scene objects. However, in a normal setting many Aruco markers are detected and the camera pose is retrieved with high precision.

Figure 4.12 Some of the 1024 Aruco markers.



Since the coordinates of the anode relative to the external camera (${}^{\mathrm{V}}\mathbf{X}$) are constant (obtained in the calibration stage discussed in Section 4.5) and given that both imaging systems are rigidly tied, we are allowed to write:

$$ {}^{\mathrm{D}}\hat{\mathbf{X}} = {}^{\mathrm{D}}\hat{\mathrm{T}}_{\mathrm{W}} \cdot \left({}^{\mathrm{V}}\hat{\mathrm{T}}_{\mathrm{W}} \right)^{-1} . {}^{\mathrm{V}}\hat{\mathbf{X}} \tag{4.8} $$

Together with Eq. (4.7), we can compute the X-ray intrinsic and extrinsic parameters of the X-ray imaging system.

4.4 Application scenarios

We envision two possible operation scenarios (shown in Fig. 4.2 and Fig. 4.3, respectively).

4.4.1 Moving camera scenario

This scenario is the most intuitive from the point of view of stereo and computer vision and is undoubtedly more suitable if the injury being examined prevents the patient from moving safely. In this setting, the X-ray source is placed at different locations and orientations relative to W, and an X-ray image of the patient (who stands still) is generated at each position j. Using visual fiducials and the process described in Section 4.3.3, we can determine the location of the beam source ${}^{\mathrm{D}}\mathbf{X}^{j}$ for each setting j.

Note that, in this scenario, the intrinsic part (K^{j}) of P^{j} changes when the X-ray emitter is moved. Nevertheless, K^{j} can be easily updated using Eq. (4.5) and Eq. (4.6) together with the relation ${}^{\mathrm{D}}\mathbf{X}^{j}$ for each setting j. Finally, the projection matrix for each geometrical setting can be obtained as:

$$\mathrm{P}^{j} = \mathrm{K}^{j} \cdot {}^{\mathrm{X}^{j}}\mathrm{T}_{\mathrm{W}} \tag{4.9}$$

where ${}^{\mathrm{X}^{j}}\mathrm{T}_{\mathrm{W}}$ is the outcome of Eq. (4.7).

4.4.2 Moving patient scenario

In this setting, the X-ray emitter remains fixed and the object/patient rotates and/or moves. A different radiograph is produced at each patient/object position j. If the patient is able to move him/herself and/or his/her examined anatomic part, this scenario might turn out interesting because it allows larger geometrical changes (it also depends on the examination protocol being applied). Visual markers have to be rigidly fixed somehow to the patient as in the example shown in Fig. 4.13-right and Fig. 4.3. This scenario might also be interesting for object inspection, i.e., boxes, packages, etc., that can easily be rotated and where markers can be easily stuck over flat surfaces.

The first key difference when compared with the previous setup is that, in this case, intrinsic parameters remain constant ($\mathrm{K}^j = \mathrm{K}$) because the position of the X-ray emitter (obtained during calibration) also remains fixed (i.e., ${}^{\mathrm{D}}\mathbf{X}^j = {}^{\mathrm{D}}\mathbf{X}$).

The second key difference is that what is tracked is not the absolute position of the X-ray emitter relative to W but the relative location between a new world reference system W′ (attached to the patient) and the X-ray emitter. The 3D points $\mathbf{Q}_i$ of Eq.Eq. (4.1) are then expressed in W′ coordinates (i.e., ${}^{\mathrm{W}'}\mathbf{Q}_i$).

Using the visual tracking described in Section 4.3.3, we measure ${}^{\mathrm{V}^j}\mathrm{T}_{\mathrm{W}'}$, which connects points in the W′ reference system and visible light camera coordinates for each patient location/orientation j. With this information, the projection matrix in this scenario can be expressed as:

$$\mathrm{P}^j = \mathrm{K} \cdot {}^{\mathrm{X}}\mathrm{T}_{\mathrm{W}} \cdot \left({}^{\mathrm{V}^j}\mathrm{T}_{\mathrm{W}}\right)^{-1} \cdot {}^{\mathrm{V}^j}\mathrm{T}_{\mathrm{W}'} \tag{4.10}$$

where K, ${}^{\mathrm{X}}\mathrm{T}_{\mathrm{W}}$ and ${}^{\mathrm{V}^j}\mathrm{T}_{\mathrm{W}}$ are just results of the initial calibration (presented in Section 4.5). Note that ${}^{\mathrm{V}^j}\mathrm{T}_{\mathrm{W}}$ remains constant for each patient position because the visible light camera remains fixed relative to the original W.

4.5 Calibration phase

The goal of the calibration step is to obtain the necessary (and invariant) geometrical relations connecting scene elements, imaging systems and reference frames to each other and to W. To that end, the Teflon structure shown in Fig. 3.7 of Chapter 3, was designed. Once this information is obtained, the calibration frame can be removed from the scene.

As stated in Section 3.4 of the same Chapter, the calibration frame accommodates 13 tin/copper, cross-shaped markers which are opaque to X-rays, at two different planes. It also contains 12 visible Aruco fiducials that can be easily detected using the visible light camera. The coordinate frame W is centered in one of the fiducials, and the 3D coordinates of all of the fiducials (visible and opaque to radiation) are known relative to it after a careful construction process. The calibration phase can be summarized in the following steps:

1. The calibration frame is introduced in the scene, and a radiograph and a RGB or grayscale image are generated from it (similar to those shown in Fig. 4.18).
2. ${}^{\mathrm{V}}\mathrm{T}_{\mathrm{W}}$ is calculated with the acquired photograph and the methodology in Section 4.3.3. This transformation is used in both application scenarios in Eq. (4.8) and Eq. (4.10).
3. An initial X-ray projection matrix P is computed with the DLT algorithm and combinations of 3D locations of X-ray-opaque fiducials and their corresponding 2D projections on a calibration X-ray instance.
4. Intrinsic K and extrinsic ${}^{\mathrm{X}}\mathrm{T}_{\mathrm{W}}$ parts are extracted from the RQ decomposition of P. Matrix K (unaltered) is used in the moving patient scenario, specifically in Eq. (4.10).
5. Vector ${}^{\mathrm{D}}\mathbf{X}$ and matrix ${}^{\mathrm{D}}\mathrm{T}_{\mathrm{X}}$ are rebuilt from K using Eq. (4.5) and Eq. (4.3) together with the detector resolution provided by the manufacturer. Finally, the rigid transformation ${}^{\mathrm{D}}\mathrm{T}_{\mathrm{W}}$ used in Eq. (4.8) is computed.
6. The relative position between the X-ray emitter and the camera ${}^{\mathrm{V}}\mathbf{X}$ is also obtained. This relation remains invariant and is later applied

in Eq. (4.8).

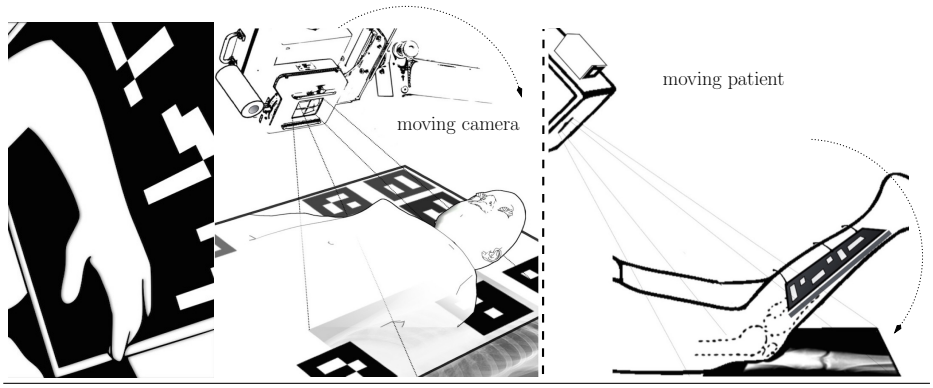
$$^{V}\mathbf{X} = {}^{V}\mathrm{T}_{W} \cdot \left({}^{D}\mathrm{T}_{W}\right)^{-1} \cdot {}^{D}\mathbf{X} \tag{4.11}$$

7. When all the aforementioned relations have been collected the calibration frame is no longer necessary.

4.6 3D information from projection matrices

In this section, we examine how our technique can be used in real application scenarios and how 3D information from correlated X-ray images can be extracted. From this point onwards, projection matrices are obtained with the help of visible fiducials as explained in Section 4.3.

Figure 4.13 Applications of the procedures detailed in this text. Depending on the scenario of application (moving patient or moving camera system), visible fiducials can be placed over the detector, a wall, or the examination table (left), or they can be tied to the patient or even stuck over scanned items (right).



4.6.1 Epipolar lines between radiographic images

A key step in stereo imaging (including diagnostic X-ray imaging) involves finding point correspondences in two stereo images. Using epipolar geometry, the search for a corresponding point $\mathbf{q}_i^j$, which is initially observable in image j, can be reduced to a *search through a line* on the second radiograph k. The line on this second image is called the *epipolar line*. This simple technique can dramatically contribute to resolving ambiguities when two points of interest lie very close to each other in one image. This could create some difficulty in distinguishing between them. However, they are easily recognizable in another projection (i.e., a *paired* radiograph), where they can be more efficiently matched by a corresponding epipolar line.

Given two paired X-ray images along with their corresponding projection matrices P^j and P^k, we can compute the Fundamental Matrix F presented in chapter 9 of [98] which enables the mapping of any observed point $\mathbf{q}_i^j$ on the first X-ray image j to an *artificial* infinite epipolar l_i^k in image k.

This epipolar line can be *bounded* or shortened if the rough dimensions of the object under examination (e.g., chest/patient thickness) are taken into account. This restriction simplifies the search on the second image to a much shorter segment ($\updownarrow_i^k$) instead of the infinite l_i^k. In order to obtain the limits of the segment for a particular point $\mathbf{q}_i^j$, we begin by *back-projecting* it into a 3D infinite ray $\overline{\mathsf{Q}_i^j}(\zeta)$:

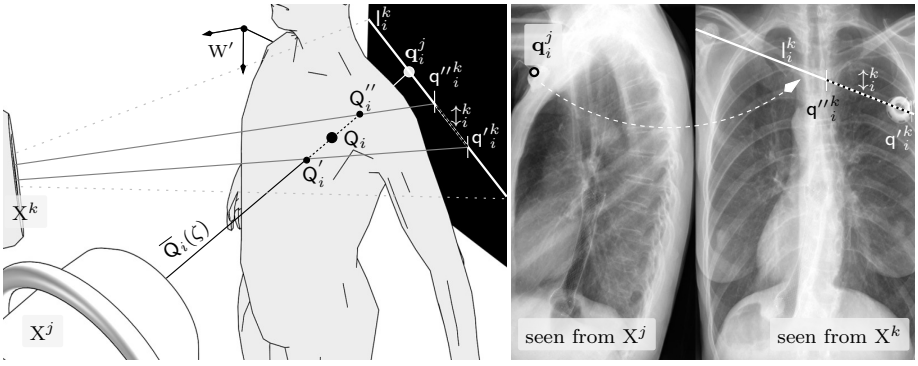
$$\overline{\mathsf{Q}}_i^j(\zeta) = \left(\mathrm{P}^j\right)^+ \cdot \hat{\mathbf{q}}_i^j + \zeta \mathbf{X}^j \tag{4.12}$$

where $\mathbf{X}^j$ is the anode location in W coordinates and ζ is a scalar that parametrizes the ray that passes over the point $\mathbf{q}_i^j$, continues towards $\mathbf{Q}_i$, and finally reaches the anode $\mathbf{X}^j$ (see Section 6.2.2 of [98] for details). Next, the boundaries Q'_i and Q''_i of the segment $\overline{\mathsf{Q}}_i$ that are coherent with the rough thickness of the studied object/person are iteratively matched. Finally, these two confining 3D points are *reprojected* on the second X-ray image k:

$$\left.\begin{array}{l} \hat{\mathsf{q}}'^k_i = \mathrm{P}^k \cdot \hat{\mathsf{Q}}'_i \\ \hat{\mathsf{q}}''^k_i = \mathrm{P}^k \cdot \hat{\mathsf{Q}}''_i \end{array}\right\} \updownarrow_i^k = \left(\mathsf{q}'^k_i, \mathsf{q}''^k_i\right) \tag{4.13}$$

and a *bounded* $\updownarrow_i^k$ between them can be plotted (Fig. 4.14).

Figure 4.14 Point reconstruction using two X-ray images with epipolars.



Real examples of bounded epipolars are shown in Fig. 4.20 for the moving camera scenario.

4.6.2 3D reconstruction from two image pairs

This process is also known as *projection-to-volume* reconstruction or *projective reconstruction* and is described in Section 12.2 of [98] and in [97]. It enables the determination of the 3D location of an observed point $\mathbf{q}_i$ in two images (j and k). Given two projection matrices P^j and P^k and using Eq. (4.1), we can write:

$$\hat{\mathrm{q}}_i^j = \mathrm{P}^j \cdot \hat{\mathbf{Q}}_i \qquad \hat{\mathrm{q}}_i^k = \mathrm{P}^k \cdot \hat{\mathbf{Q}}_i \tag{4.14}$$

Since we are working in homogeneous coordinates, the equivalence between two points has to be expressed using the cross product:

$$\begin{aligned} \hat{\mathbf{q}}_i^j \times \hat{\mathrm{q}}_i^j &= \hat{\mathbf{q}}_i^j \times \mathrm{P}^j \cdot \hat{\mathbf{Q}}_i = \mathbf{0} \\ \hat{\mathbf{q}}_i^k \times \hat{\mathrm{q}}_i^k &= \hat{\mathbf{q}}_i^k \times \mathrm{P}^k \cdot \hat{\mathbf{Q}}_i = \mathbf{0} \end{aligned} \tag{4.15}$$

Each of these expressions determines two linearly independent equations that can be written in the form of a linear system. When solved through a single value decomposition method (SVD), we can derive the 3D location of a specific point $\hat{\mathbf{Q}}_i$ observed in the two images.

4.6.3 Registration of visible and X-ray images

If pinhole camera models are assumed, any two images obtained with different camera systems (X-ray and RGB) can be related up to a scale factor (ζ). In order to geometrically register a pixel ($\mathbf{q}_i^{\mathrm{V}}$) in the RGB image with a $\mathbf{q}_i^{\mathrm{X}}$ point in a matching radiograph, we have first to back-project $\mathbf{q}_i^{\mathrm{V}}$ to its corresponding 3D coordinates ($\mathbf{Q}_i$) in the world reference frame with the inverse of Eq. (4.5):

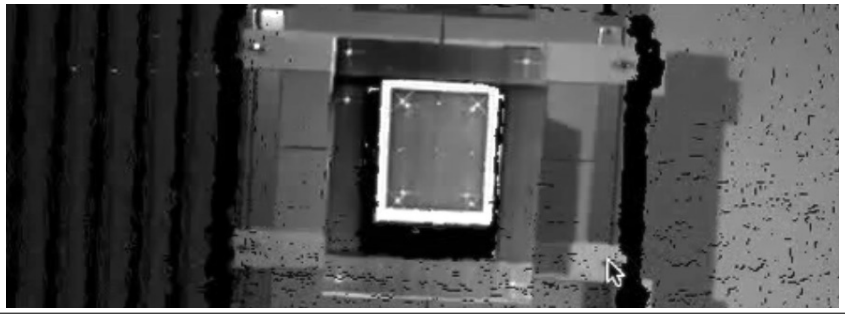
$$\hat{\mathbf{Q}}_i = \left({}^{\mathrm{V}}\mathrm{T}_{\mathrm{W}}\right)^{-1} \cdot \left(\mathrm{K}_{\mathrm{V}}\right)^{-1} \cdot \hat{\mathbf{q}}_i^{\mathrm{V}} \tag{4.16}$$

and then find the ray equation (as shown in Fig. 3.18) that connects the center of the RGB camera (also expressed in the W reference frame) and $\mathbf{Q}_i$, given by:

$$\overline{\mathbf{Q}}_i(\zeta) = \left({}^{\mathrm{V}}\mathrm{T}_{\mathrm{W}}\right)^{-1} \cdot (0,\, 0,\, 0,\, 1)^{\intercal} + \zeta \cdot \hat{\mathbf{Q}}_i \tag{4.17}$$

where ζ is a scalar (manually chosen) representing the coordinate along the ray. Then, a specific 3D point in $\overline{\mathbf{Q}}_i(\zeta_m)$ can be re-projected onto the X-ray image with Eq. (4.1), obtaining a registered 2D location $\mathbf{q}_i^{\mathrm{X}}(\zeta_m)$, which also depends on the selected value for ζ.

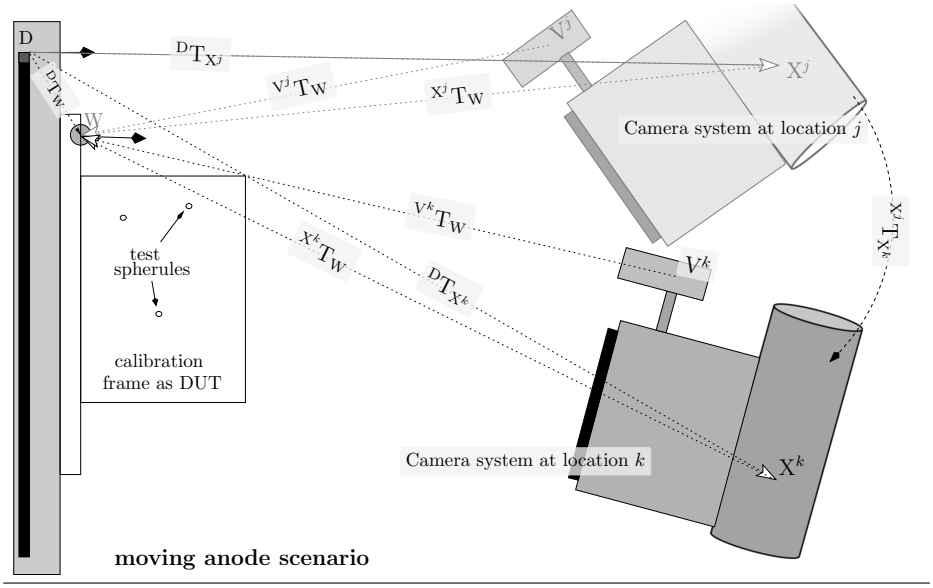
Figure 4.15 Radiograph of the calibration frame registered with the corresponding RGB picture as captured by the video camera (for a specific value of the ζ parameter).



4.7 Tests on the calibration frame acting as a device under test

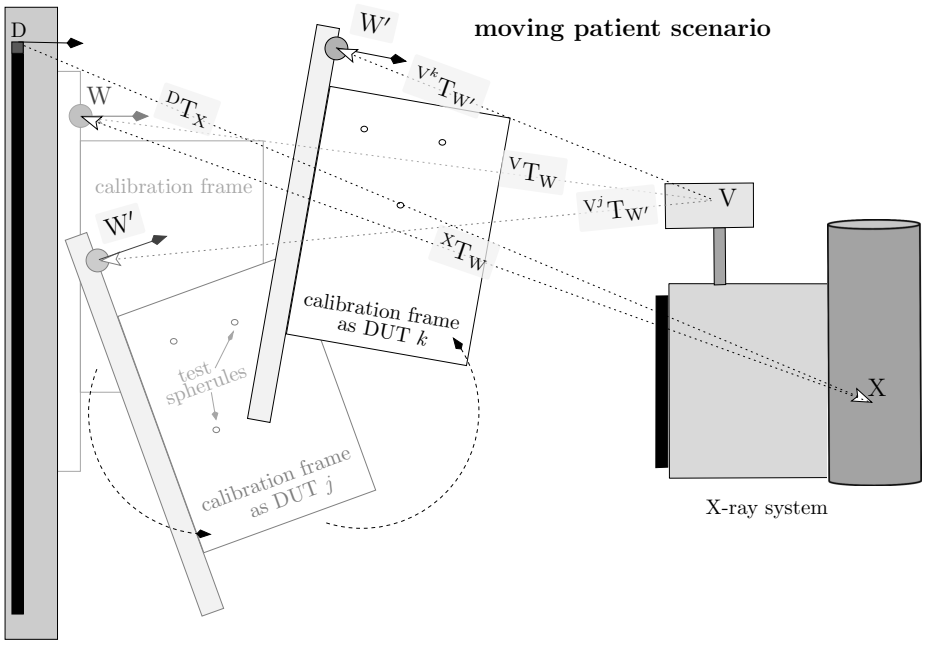
In this section, the set of techniques described from Section 4.3 to Section 4.6 are tested. Now the calibration frame will play the role of a *device under test* or DUT. Graphical representations of the two application scenarios are shown in Fig. 4.16 and Fig. 4.17. Results are provided for both the DUT and the phantom, and in each case, for the moving camera/patient scenarios. Within this test, the X-ray fiducials used

Figure 4.16 Moving camera system scenario consisting of a moving anode. D, X and V stand for the coordinate frames described in Section 4.3.1. The rigid transformations ${}^{\mathrm{D}}\mathrm{T}_{\mathrm{X}^j}$ and ${}^{\mathrm{D}}\mathrm{T}_{\mathrm{X}^k}$ are estimated from visual fiducials only detected at each camera location V_j and V_k as explained in Section 4.4.1. X-ray projection matrices P^j, P^k, etc., are derived with Eq. (4.9).



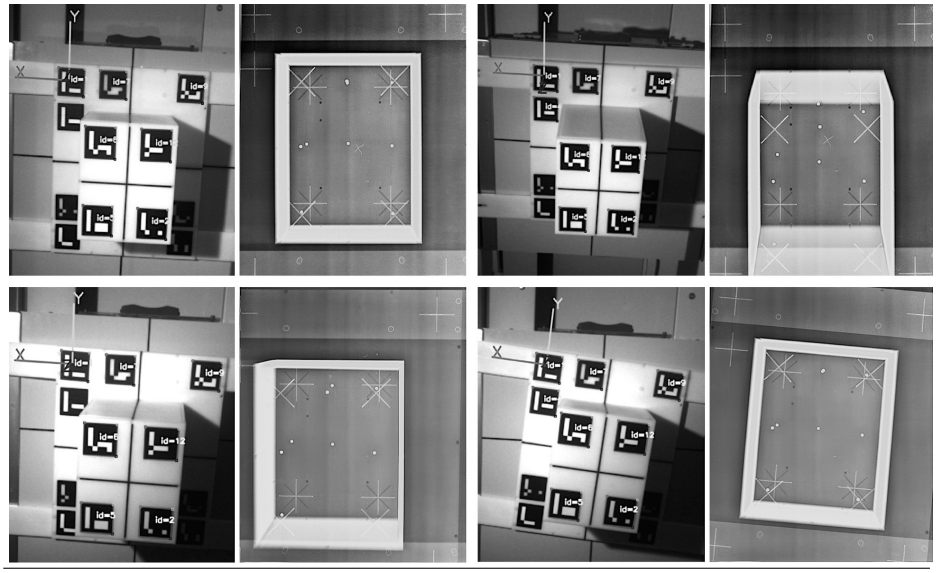
during calibration are completely ignored. This means that each P^j is now derived using only visual information. In order to evaluate the algorithm, we placed nine 2 mm radius lead spherules or bearings at different well-known positions inside the calibration frame (see Fig. 4.19-left).

Figure 4.17 Moving patient application scenario with the calibration frame now acting as a device under test (DUT). The DUT is shifted to other positions while the anode remains fixed. Intrinsics remain constant (i.e., $\mathrm{K}^j = \mathrm{K}^k = \mathrm{K}$) and each DUT position is tracked using visual information. X-ray projection matrices P^j, P^k, etc., are obtained with Eq. (4.10) and as discussed in Section 4.4.1.



The DUT was radiographed ten times for each of the application scenarios. In Fig. 4.18 we show a few examples of the captured images. We established several evaluation metrics in order to check the goodness of the results. All metrics make a distance comparison between the *real location* of the lead spherules and the *estimated location* predicted by using the generated projection matrices.

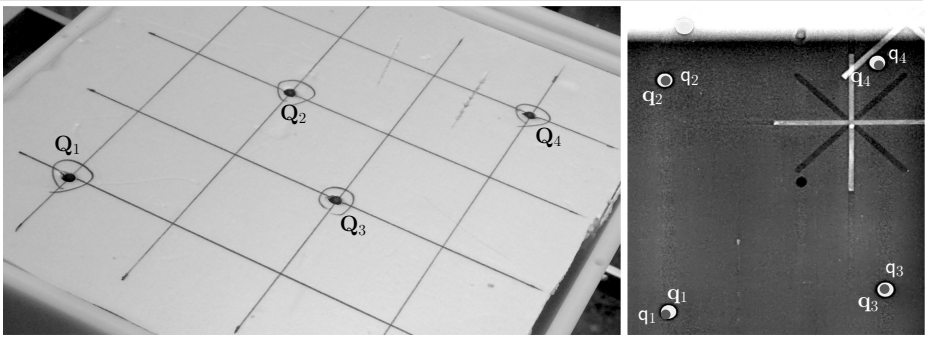
Figure 4.18 Moving camera (top) and moving patient (bottom) scenarios.



4.7.1 The mean distance between *generated* projections of known landmark locations and their observed image coordinates

This is the most basic quality test for the projection matrices. It consists of measuring the mean 2D image distance between the *observed projection* $\mathbf{q}_i$ (on the image i) of each of the 9 spherules ($\mathbf{Q}_i$) and their *computed projection* q_i, estimated with Eq. (4.1). The results (highlighted in Fig. 4.19-right) show this mean distance is equal to 12 px (with a dispersion of 8 px) for the moving camera scenario and 8 px (with a dispersion of 6 px) for the patient moving scenario. As expected, the difference in accuracy between the two scenarios is not significant since the geometry involved in both problems is very similar. The detector resolution is the same specified in Table 3.1.

Figure 4.19 (Left) Lead spherules inside the calibration frame (acting now as a DUT) placed at 3D known locations $\mathbf{Q}_i$. (Right) An example of the level of agreement between some lead spherule projection centers $\mathbf{q}_i$ (white dots) and their predicted image location $\mathfrak{q}_i$ (gray dots) by using the estimated projection matrix P^j. The cross-shaped mark at the top-right corner corresponds to one of the fiducial set that previously played a role during calibration (Fig. 3.7) but is now ignored during the whole testing phase.



4.7.2 The distance between epipolars and projections of spherules

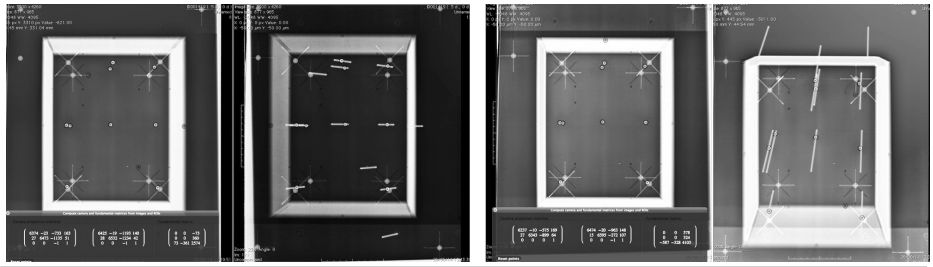
The proposed metric consists of computing the mean 2D distance (for all possible radiograph combinations) between each lead spherule, whose observed projection is $\mathbf{q}_i^k$, and its corresponding (and calculated) bounded epipolar line $\updownarrow_i^k$. Again this mean distance for both moving camera and moving patient scenarios were obtained.

Epipolars were bounded assuming an average DUT thickness of 24 cm. In Fig. 4.20 and Fig. 4.21 we illustrate this test and some of its results, which confirm a mean distance of 13 px with a dispersion of 8 px for the camera moving scenario and a mean distance of 10 px with a dispersion of 6 px for the patient moving scenario. Again the differences between scenarios is not significant.

4.7.3 The distance between back-projections and real 3D locations

Since the 3D location of the spherules is very well known (Fig. 4.19-left), we test the accuracy of the algorithm described in Section 4.6.2 by calculating the mean value of the distances between the known 3D locations $\mathbf{Q}_i$ and predicted ones $\mathbf{Q}_i$ using all possible radiograph combinations and all spherules. Outcomes indicate a mean offset of 2 mm (deviation of 2 mm) when the spherules are located 2 meters away from the anode (approximately) and a mean baseline of 45 cm (displacement of X-ray anode between radiographs).

Figure 4.21 More bounded epipolars (right images) between different stereoscopic X-ray images.



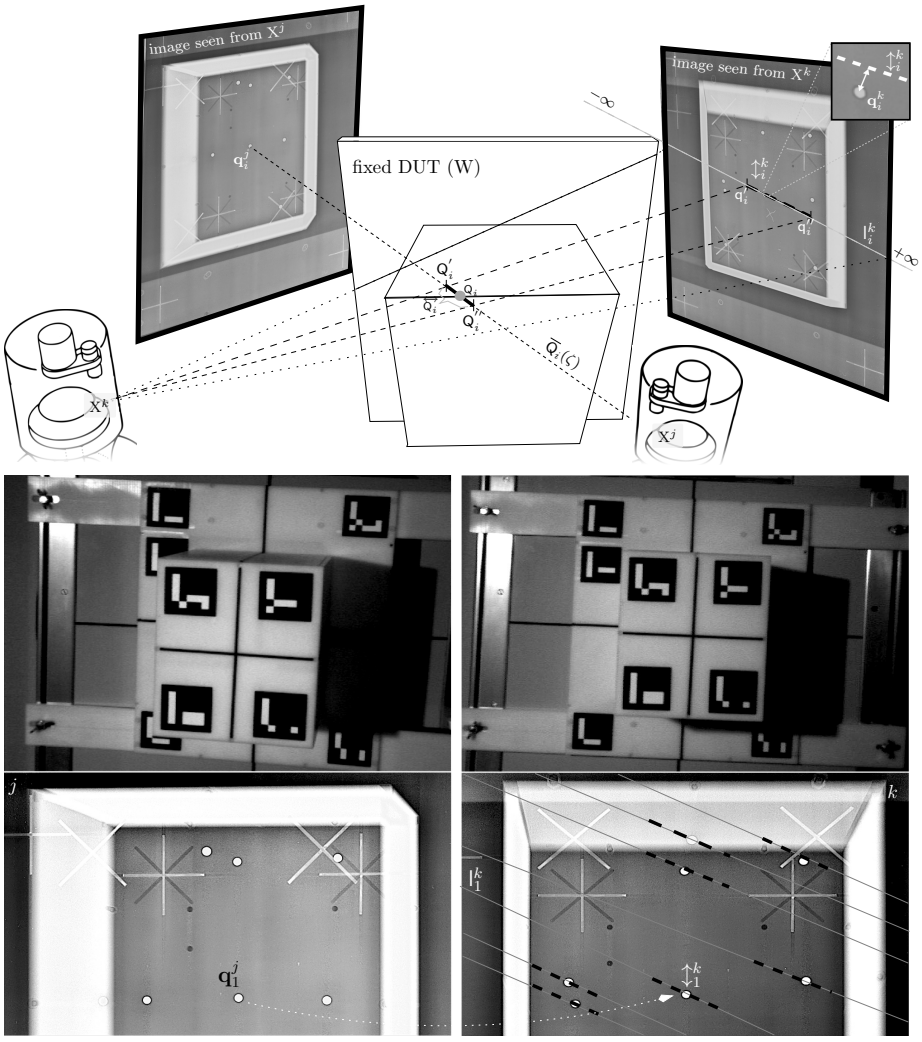
4.8 Conclusions

We have introduced an innovative methodology for determining the geometrical setting in standard X-ray imaging systems. This information is essential for 3D reconstruction using several X-ray images, which could be a relevant tool for diagnosis and object inspection. Contrary to the usual approach, ours is not based on the projection of external opaque fiducials on each radiograph (as we did in Chapter 3). Instead, we account for all the necessary spatial information and modifications thanks to an external camera and visible markers present in the scene. Our technique enables the removal of foreign reference marks and interposed frames, contribut-

ing to reducing the complexity and invasiveness of the X-ray diagnostic process. The same technique also allows the combination of X-ray images produced at acuter angles/protocols, where conventional fiducials would be projected outside the radiographic plane. Two application scenarios have been discussed, each one involving different geometrical, optical, and mathematical challenges and application suitability. The accuracy attained in both scenarios is equivalent and the choice of one or the other will depend on the final application. The *moving patient* approach may contribute to enhance X-ray based object scanning frameworks by for instance, providing the exact 3D location and size of items masked in boxes, packages, etc.

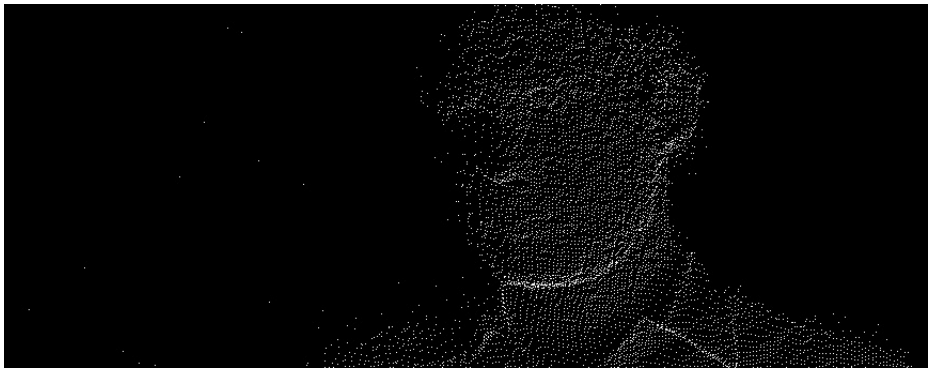
The experiments on a boxlike phantom show that it is possible to estimate epipolar lines in a second image from given points in a first one with a very good level of precision. This simple but effective subtlety can help radiologists correlate points/areas of interest between different radiographs. Finally, 3D reconstruction using ordinary X-ray images represents an elegant alternative to invasive techniques like CT and similar radiological equipment. In Section 7.1 we apply the described techniques to a arm-shaped anthropomorphic phantom in a more realistic X-ray environment.

Figure 4.20 Top: schematic representation of the geometrical relations described in Section 4.6.1 for a moving camera scenario (detector appears twice and in two positions for the sake of clarity). The segment $\overline{\mathsf{Q}_i}$ roughly spans the mean length of the calibration frame ($|\mathsf{Q}"_i - \mathsf{Q}'_i| \approx 0.24\,\mathrm{m}$). The projection of this segment in the X-ray image is $\updownarrow_i^k$. Bottom: example of resolved bounded epipolars in a moving patient scenario.



5

Geometrical X-ray setting definition from depth data



In this chapter we present another method for deriving 3D internal information in conventional X-ray settings. In contrast with Chapter 4, this method now augments an ordinary X-ray device with a consumer RGB-D camera. The patient's rotation around the craniocaudal axis is tracked relative to this camera thanks to the depth information provided and the application of a modern surface-mapping algorithm. The measured spatial information is then translated to the reference frame of the X-ray imaging system. Using the intrinsic parameters of the diagnostic equipment, epipolar geometry, and X-ray images of the patient at different

angles, 3D internal positions can be obtained. Both the RGB-D and X-ray instruments are first geometrically calibrated to find their joint spatial transformation. With the proposed approach, internal 3D reconstructed coordinates and distances can be provided to the physician. It also contributes to reducing the invasiveness of ordinary X-ray environments and can replace other types of clinical explorations mainly aimed at measuring or geometrically relating elements that are present inside the patient's body.

The concepts and techniques here summarized extend the research carried out in [9] and [12].

5.1 Relevance of projection-to-volume techniques

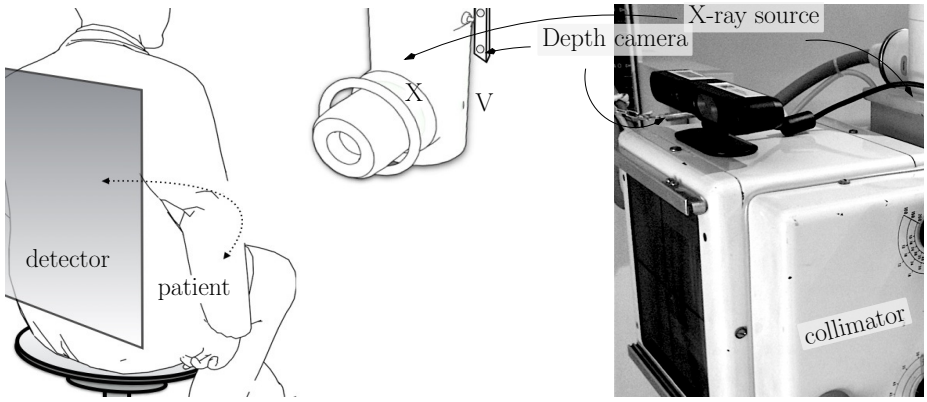
As explained in Section 4.6.1, the common procedure for obtaining 3D information (locations, distances, angles, etc.) from two images j and k, also known as *projection-to-volume registration* or P2VR, requires two camera projection matrices P^j and P^k. In conventional radiography, projection matrices can be obtained by using calibration frames with N radio-opaque fiducials placed at known locations $\mathbf{Q}_i$ (with $i = 1 \cdots N$) that are projected onto each radiograph at 2D coordinates $\mathbf{q}_i^j$ and $\mathbf{q}_i^k$ for radiographs j and k, respectively. The 3D locations of these fiducials are usually expressed relative to a common *world* coordinate frame (W).

P2VR in radiology has been a subject of interest for a long time. It is worth citing one of the pioneering research works carried out by Caponetti et al. [45], where a *shape-from-contour* algorithm and the back-lighting from two perpendicular projections is used, under the assumption of a parallel beam, to approximate the 3D surface of a bone.

For our contribution to P2VR, we present an alternative method with special a focus on its deployment in ordinary and primary diagnostic X-ray settings and the elimination of the need for calibration frames. Our approach makes use of a rigidly attached RGB-D sensor or *depth camera* that can resolve the rigid transformation that represents the patient's

movement between two instants t_j and t_k and relative to a fixed X-ray system. The specific movement studied in this work is the rotation around the patient's own craniocaudal axis, from which the corresponding P^j and P^k matrices can be derived, thereby enabling the accurate distillation of 3D locations and lengths. In this context, it is worth mentioning that modern depth cameras, though often seen as mere consumer products, provide high performance and good spatial resolution as Khoshelham [120, 119] demonstrates (a mean value of 1 mm in the depth direction for calibrated devices). Accuracy can be further improved with the methods described in Section 5.5.

Figure 5.1 Depiction of the presented setup. The patient rotates in supine anteroposterior (AP) or posteroanterior (PA) position while motion is continuously tracked with the RGB-D sensor (V). This device is rigidly attached to the housing of the anode (X). The patient is radiographed at two angles relative to the vertical axis.



5.2 Combination of clinical and depth information

The augmentation of general purpose radiological modalities with other interplaying sensors has been proposed for other applications. The research carried out by Aoki et al. [20] highlights how modern consumer depth cameras have huge potential in many clinical fields, such as radiotherapy and radiography. Badal et al. [28] have engineered a dose-monitoring system based on the tracking of the location of patients and staff during interventional fluoroscopy sessions using depth sensors. The authors in Cook at al. [57] estimate the patient's size with the help of a Microsoft Kinect device in order to normalize the dose received, whereas Kozono K. [126] uses a similar approach to monitor the location of the patient and better assess the X-ray entrance dose. In the analysis performed by Tahavori et al. [237], a Kinect device is again used to capture the surface of the patient with the goal of detecting possible misalignments during beam radiotherapy. Three-dimensional imaging of the breast is studied in Wheat et al. [259] using a Kinect-based system, and the researchers in Bauer et al. [30] perform a similar task related to patient alignment in computerized tomographies (CT). Respiratory variations during positron emission tomographies (PET) are also determined in Noonan et al. [176].

Besides the use of depth data in association with radiological modalities, we observe a growing trend in the use of RGB-D cameras in many other fields of medicine. Digital biometry, for instance, is one of these. The tools developed by Reyes et al. [202] and Sandau et al. [215] analyze body posture and estimate the range of movement of skeletal joints. Researchers in [167] have created an experimental auto-learning system that enables patients to self-position during radiotherapy sessions. Three-dimensional imaging of the breast is studied in Wheat et al. [259] and achieved using a Kinect-based system. The project developed in Hsu et al. [103] consists of a depth camera-based system for improving spinal condition. In other projects carried out in several hospitals by Banerjee et al. [29] and Li et al. [134], patients are monitored using unobtrusive depth sensors while resting in their beds. The rehabilitation of patients using these devices is also a subject of interest tackled in Da Gama et al. [63].

Depth sensors are also occupying a niche in the operating room (OR). The work presented in [102] concludes that gesture detection provides an interesting human-computer interface for radiological image handling and selection in both ORs and other sterile environments. Another area of interest for the application of this technology is organ movement tracking. The research detailed in [129] measures a beating heart with the help of a Kinect sensor. Real-time respiratory motion monitoring has been tackled in [136] and [15], among other studies.

Finally, the use of augmented reality in medical education is becoming a trend too. In [154], the researchers demonstrate the usefulness of the Microsoft Kinect equipment in tracking students and then generating an on-the-fly overlay animation of their anatomy. The methodology tackled in [35] learns the most common user gestures in clinical environments without the need for specific training.

This research shares some of the goals of the work carried out by Albiol et al. [11] who perform P2VR using videocamera-augmented X-ray equipment and visual markers. The key difference is that the methodology described in [11] is more appropriate for X-ray protocols in which the patient remains still and the imaging setup is the one that moves from location j to k.

However, despite all of these research examples on the subject of *external sensors and depth cameras in medicine and radiology*, it is very difficult to find citations about the use of RGB-D data for the determination of the patient's location/orientation relative to a conventional X-ray imaging setup (such as the one shown in Fig. 5.1). As the following sections demonstrate, this information can be non-invasively measured with enough precision, making a significant impact on the elaboration of the diagnosis.

5.3 System overview

The proposed system consists of a conventional X-ray setup and a depth sensor that is rigidly attached to the housing of the X-ray anode. During examinations, the X-ray imaging system (source/detector) and the RGB-D device both remain fixed. A RGB-D camera is comprised of depth and video component and both data streams are registered by default. The RGB-D device used measures distances by analyzing the speckle pattern of a projected infrared light.

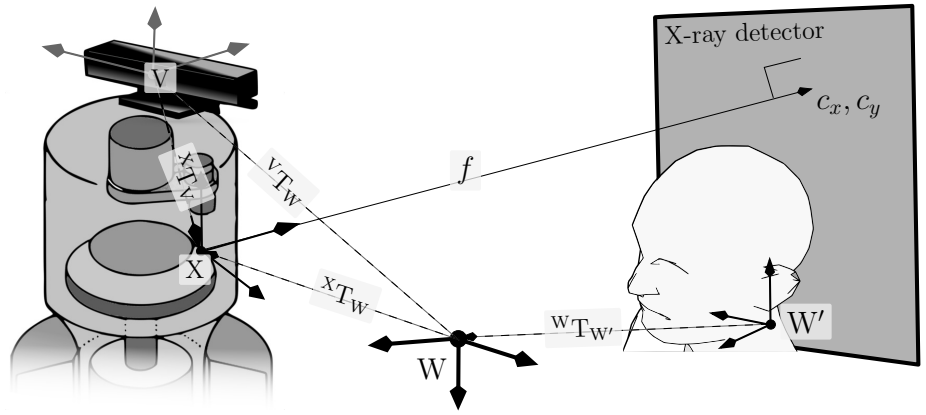
Initially, the patient is asked to stand erect in front of the depth sensor and is rigidly rotated with the help of a swivel stool or a spinning platform. This motion is continuously tracked by the depth sensor and two radiographs are taken at different orientations j and k (instants t_j, t_k). By using the patient's displacement relative to the depth sensor in combination with epipolar geometry, P^j and P^kcan be updated. However, before proceeding, the dual imaging system must first be geometrically calibrated.

5.4 System calibration

As explained in Section 4.5, the goal of the calibration phase is to obtain both the intrinsic parameters of the X-ray setting and the geometrical relation between the two imaging systems which remain invariant during the examination. The same calibration frame shown in Fig. 3.7 was designed to do this. As a quick reminder: it accommodates 13 copper, cross-shaped markers $\mathbf{Q}_i^{\mathrm{cal}}$ (with $i = 1 \cdots 13$) that are opaque to the Roentgen radiation. It also contains a matching number of visible markers that can be easily detected using the incorporated video camera of the depth sensor. Both fiducial types are made coincident to ease calculations. Besides, the frame defines the origin and orientation of the W coordinate frame. The 3D coordinates of all of the calibration markers are known by construction and referenced relative to W.

Three other coordinate systems are defined, which are shown in Fig. 5.2. The first one is the X coordinate system, whose origin is at the X-ray anode and has one axis that is orthogonal to the detector plate. The second one is the camera coordinate system (V), whose origin is the optical center of the built-in video camera. V also defines the origin of the depth data given that both streams (video and depth) are registered. Finally, an object-dependent coordinate frame W′ is used for testing purposes, as discussed in Section 7.2. All of the coordinate systems (X, V, W, and W′) can be related using rigid transformations, namely: ${}^{\mathrm{V}}\mathrm{T}_{\mathrm{W}}$, ${}^{\mathrm{X}}\mathrm{T}_{\mathrm{W}}$, ${}^{\mathrm{X}}\mathrm{T}_{\mathrm{V}}$, and ${}^{\mathrm{W}}\mathrm{T}_{\mathrm{W}'}$.

Figure 5.2 Coordinate frames and geometrical relations used in the proposed imaging system. W is a common coordinate frame whose origin is somewhere known in the room. W′ has its origin inside the object under study and is linked to it. V and X are the proper coordinate frames of the anode and depth camera, respectively.

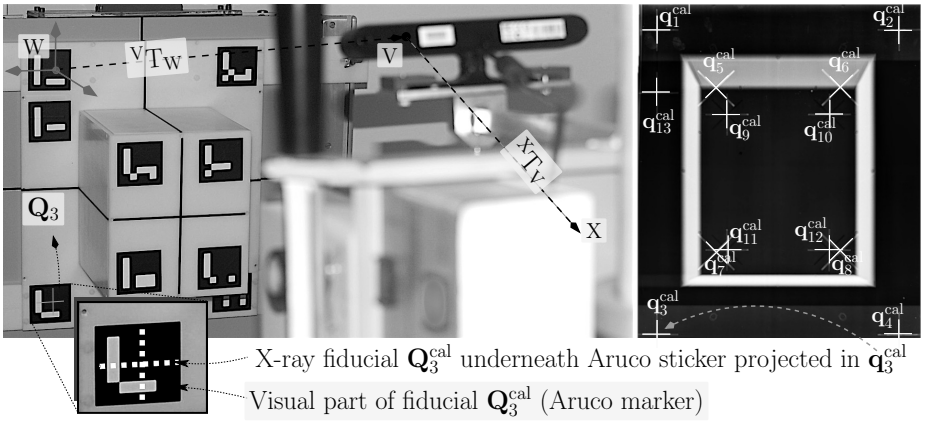


The calibration process can be summarized as follows:

1. The calibration frame is introduced in the scene and a photograph (with the RGB-D camera) and a radiograph are generated from it (Fig. 5.3-right). The structure is then removed and is no longer necessary during the examination of the patient.

2. Using the Direct Linear Transform (DLT) algorithm [98], the 3D location $\mathbf{Q}_i^{\mathrm{cal}}$ of each radio-opaque marker (relative to W) and the coordinates (manually outlined) of the corresponding 2D projection

Figure 5.3 Left: information gathered during calibration: a photograph and a radiograph of the frame. Both images enable the derivation of the transformation connecting both imaging systems (${}^{\mathrm{X}}\mathrm{T}_{\mathrm{V}}$). Visible fiducials help find the relation between the world and the depth camera (${}^{\mathrm{V}}\mathrm{T}_{\mathrm{W}}$). Right: projections of the cross-shaped X-ray markers (hidden beneath the visual fiducials) in the radiograph.



$\mathbf{q}_i^{\mathrm{cal}}$ in the radiograph obtained in the previous step, we build an initial X-ray projection matrix $\mathrm{P}^{\mathrm{cal}}$.

3. The intrinsic (K) and extrinsic (${}^{\mathrm{X}}\mathrm{T}_{\mathrm{W}}$) parameters of the X-ray system are extracted from $\mathrm{P}^{\mathrm{cal}}$ using the RQ decomposition: $\mathrm{P}^{\mathrm{cal}} = \mathrm{K} \cdot {}^{\mathrm{X}}\mathrm{T}_{\mathrm{W}}$. In ordinary cameras, intrinsic parameters are fixed. However, in X-ray imaging, detector and source are decoupled, which entails that K will be altered if either the imaging plate or anode are shifted. Conventionally, the elements of K contain the principal point c_x, c_y and the focal length f, which perpendicularly connects the anode and the detector. The extrinsic part relates 3D coordinates between W and X frames.

4. Using the photograph of the calibration frame (and visual markers) taken in the first step, it is possible to obtain the position and orientation of the video camera (and depth sensor) relative to W, i.e., the sought rigid transformation ${}^{\mathrm{V}}\mathrm{T}_{\mathrm{W}}$ that links W and V (i.e., how points in the world are transformed to the depth camera *point-of-view*). In this work, we use an automatic process for visual marker detection based on the Aruco library, described in Section 4.3.3.

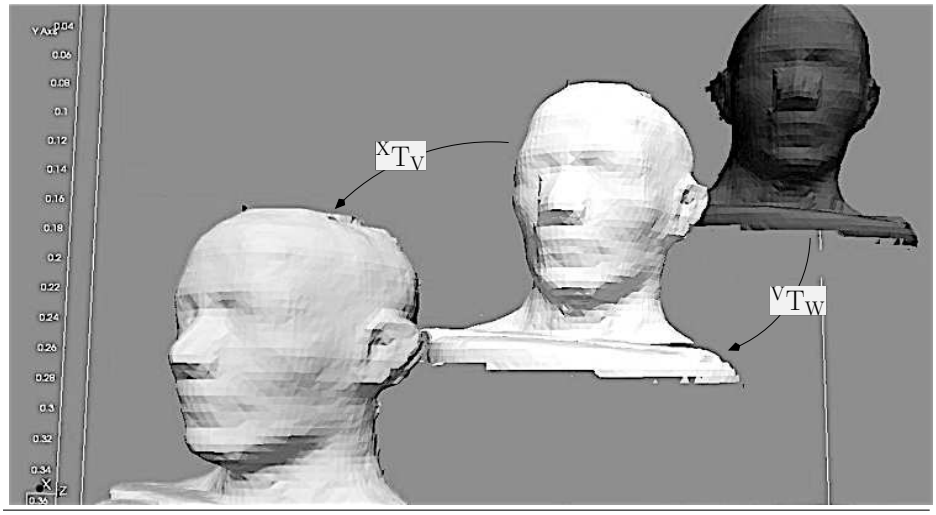
5. Finally, the relation between the X-ray and RGB-D coordinate system can be obtained as:

$$^{X}T_{V} = {}^{X}T_{W} \cdot \left({}^{V}T_{W}\right)^{-1} \tag{5.1}$$

This transformation remains constant as long as the relative position between devices remains fixed. With Eq. (5.1), we can translate the derived spatial information to the reference frame of the X-ray imaging system. The part $\left({}^{V}T_{W}\right)^{-1}$ is the *pose* of the depth camera (i.e., its position/orientation relative to W).

The derived transformations allow to go from one coordinate system to the other, as Fig. 5.4 shows.

Figure 5.4 Reconstructed surfaces and their transformations between different coordinate systems. Of course, all three meshes are viewed from an external point of view for didactic reasons.

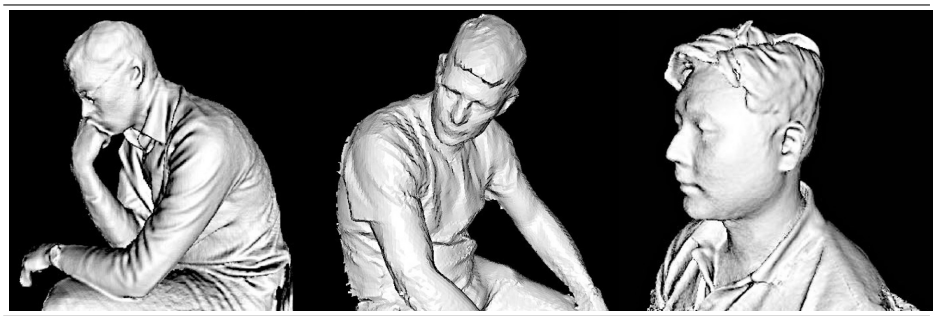


5.5 Estimation of motion and projection matrices

Once the system is calibrated, the patient enters the scene and comfortably sits or stands on a rotating platform. At this initial instant t_j, the first radiograph is captured and then the platform starts spinning gently while depth images are continuously generated. During this process, the patient should remain still to ensure that a motion be as close to rigid as possible. The captured depth data (also known as *point cloud*) is continuously analyzed until the patient reaches a second orientation at t_k. At this instant, the second radiograph is taken and the process ends.

In this work, we use the KinectFusion technology developed by Newcombe et al. [172] to estimate the rigid motion of a patient ${}^{V^k}T_{V^j}$ between consecutive X-ray snapshots obtained at t_j and t_k. This algorithm was originally designed to reconstruct 3D scenes robustly by moving the Microsoft Kinect sensor around an object or person and performing a Dense Surface Mapping (DSM) described by Tong et al. [241] and shown in Fig Fig. 5.5. KinectFusion also adds extra accuracy to the 3D derived geometry. For instance, Meister et al. [153] report having achieved a 2 mm precision for the Euclidean error of a scanned 40 cm high human-like statue, the camera/object distance being ∼1 m. This error decreases if the tracked object is bigger because more points contribute to the estimation of the motion. Apart from the mentioned surface, KinectFusion also refines the location/orientation of the sensor relative to W (i.e., its pose). Volume and surface reconstruction are further tackled in Section 6.1.

Figure 5.5 Examples of surface mappings obtained from depth data.

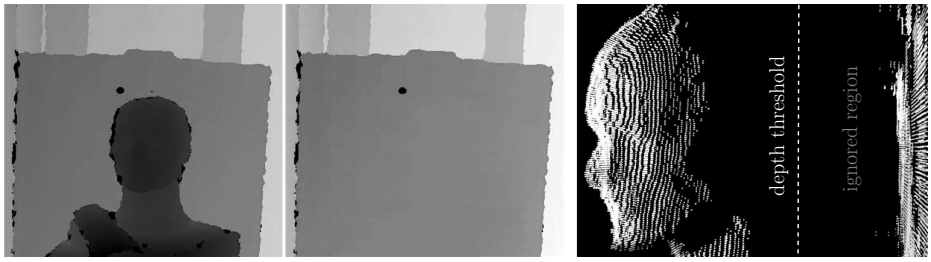


KinectFusion mandatorily needs computing devices with GPU computing capabilities, however, recent computers are usually equipped with

sufficiently powerful graphics cards.

Alternatively, in our devised setup, the patient rotates relative to a fixed Kinect-like device. This is not a problem for the KinectFusion algorithm because it only needs the relative motion between the patient and the depth sensor. This situation can be alternatively understood as a *moving virtual camera system* around the fixed W reference but with the key advantage of preserving the intrinsic parameters (K) of the X-ray equipment and the transformation (${}^{X}T_{V}$) between the X-ray equipment and the depth camera, which are obtained during calibration. The only significant difference with respect to the default KinectFusion application scenario is that, in our *virtual camera approach*, the background information has to be subtracted from the point dataset frame by frame. To achieve this, it is possible to use any of the methods already applied by many Kinect-based applications, such as the gesture recognition studied by Biswas and Basu [37], i.e., obtaining a depth snapshot of the room and keeping it for background removal or establishing a length threshold beyond which point clouds are no longer considered (Fig. 5.6). This last approach is also known as a *passthrough* filter.

Figure 5.6 Left: background depth capture and removal for one of the head-like phantoms used in this work. Right: application of a depth threshold to the acquired point cloud. In this case, the elements that lay beyond an established distance (dotted line) from the RGB-D sensor (i.e., chest walls) are ignored and not taken into account by KinectFusion.



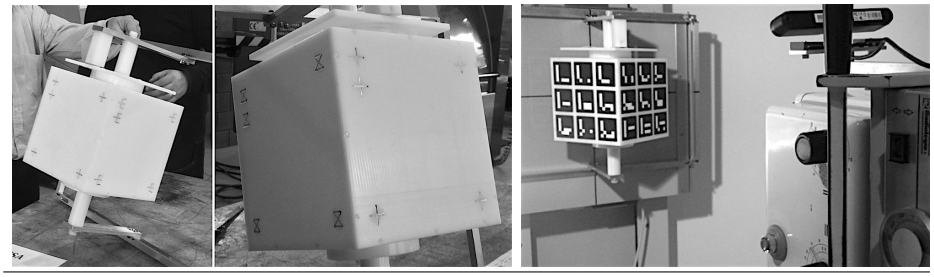
The X-ray projection matrices P^j and P^k for instants t_j and t_k can be calculated as:

$$\begin{aligned} P^j &= K \cdot \left\{ {}^{X^j}T_{V^j} \cdot {}^{V^j}T_W \right\} \\ P^k &= K \cdot \left\{ {}^{X^k}T_{V^k} \cdot {}^{V^k}T_{V^j} \cdot {}^{V^j}T_W \right\} \end{aligned} \tag{5.2}$$

where ${}^{X^j}T_{V^j}$ and ${}^{X^j}T_{V^j}$ are the result of Eq. (5.1) and constant and equal to ${}^{X}T_{V}$. The parts inside the brackets in Eq. (5.2) correspond to the composition of the involved rigid transformations and represent the extrinsic parameters of each projection matrix. The patient's motion and the corresponding angular span around the vertical axis (included in the term ${}^{V^k}T_{V^j}$) between t_j and t_k should be chosen following medical criteria. Besides, it is recommended that each radiograph has a diagnostic meaning on its own. An important practical question is how the patient's breathing may affect the estimation of the rigid motion. In a real scenario, these motions are unnoticeable under a hospital gown. It is however convenient that the X-ray images are taken with a similar state of the lungs to minimize organ displacements and provide more precise 3D information. Additionally, X-ray images and motion information should be in sync, but this is normally the default case given that radiographs are produced by DICOM compliant hardware and depth cameras also append time information, i.e., t_j and t_k. Altogether, the process summarized here should not take more than a few seconds in a fully-digital X-ray setting or about a minute in a CR-equipped setup, representing minor annoyances for the patient.

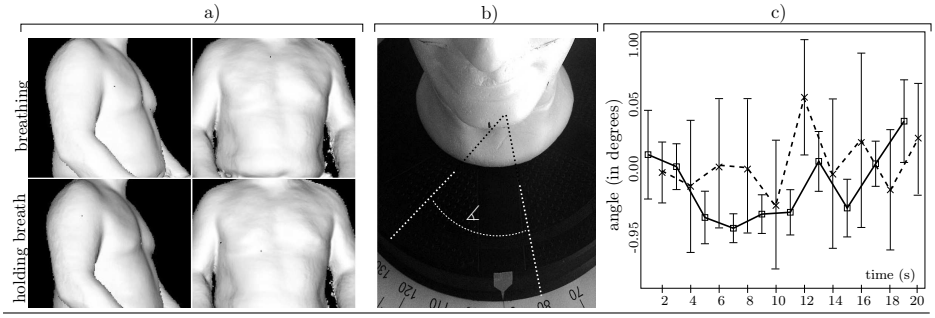
The precision of the rotation around the vertical axis was also tested with a special rotating device, shown in Fig. 5.7, a head-like phantom and a real patient (Fig. 5.8), with very accurate results (even when the patient breaths and compared against ground-truth).

Figure 5.7 Rotating phantom designed to check the precision of the Kinfu algorithm when tracking movements around the vertical axis. The top part contains an angle meter.



Finally, the evaluation of the motion performed by KinectFusion is more appropriate for relatively large parts, such as the torso, but it can also work with smaller ones such as the head.

Figure 5.8 Evaluation of the accuracy of KinectFusion. a) LAT and AP patient surface model (changes produced by breathing are unnoticeable). b) Head phantom used to measure the precision of the derived angular span. c) Evolution of the estimated angle (around vertical axis) for the LAT patient orientation while breathing (dashed lines) and holding breath (solid lines) for ∼20 s. Points represent the mean value for each 2 s. window (∼44 depth frames), and error bars show the dispersion for that time segment.



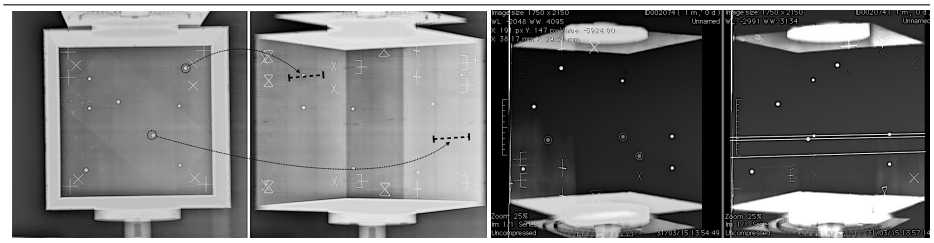
5.6 3D point reconstruction

In a similar way as explained in Section 4.6.2, given the two matrices P^j and P^k from Eq. (5.2) and two observed landmarks $\mathbf{q}_a^j$ and $\mathbf{q}_a^k$ in images j and k, respectively, it is possible to determine the 3D location Q_a (derived) of the imaged point $\mathbf{Q}_a$ (ground-truth). In our implementation, the radiologist should be the one that manually locates $\mathbf{q}_a^j$ and $\mathbf{q}_a^k$, assisted by the automatic drawing of epipolar lines (like l_a^k and l_b^k shown in Fig. 7.5). It is therefore important for these traces to be visible in both radiographs. As Hartley and Zisserman [98] recall, projective geometry establishes that 3D points and the corresponding projections are related using the cross products defined in Eq. (4.15), where all points are expressed in homogeneous coordinates. Each of these relations determines two linearly independent equations that can be solved by using a Singular Value Decomposition (SVD). If this is repeated for a second point $\mathbf{Q}_b$ projected on the same pair of X-ray images j and k, it is also possible to derive the 3D length of the segment $\overleftrightarrow{\mathsf{Q}}_{a,b}$ defined by the two back-projected ends Q_a and Q_b, i.e., $\left|\overleftrightarrow{\mathsf{Q}}_{a,b}\right|$.

5.7 X-ray tests on a rotating frame

Similar tests to those carried out in Section 4.7 have been performed with the rotating cube-shaped frame made of polytetrafluoroethylene and shown in Fig. 5.7. The results show (Fig. 5.9) we can draw similar (bounded) epipolar lines (introduced in Section 4.6.1) with millimetric precision. The rigid transformation was also compared to that obtained with the visual markers, obtaining a very similar result.

Figure 5.9 Tests on a rotating phantom.



5.8 Conclusions

In this chapter we have presented a method for deriving 3D locations from plain radiographs in fiducial-less ordinary diagnostic settings. With this methodology based on depth cameras and the tracking of the patient's motion, projection matrices can be derived in ordinary X-ray settings and 3D inner locations and lengths can be determined through epipolar geometry.

One of the key advantages of the presented techniques is that they require no dedicated fiducial system during the examination, contributing to the patient's calmness and the reduction of the sense of invasiveness. In other words, the patient is his/her own fiducial system, in contrast with the use of an external one, as we did in Chapter 4 using visible fiducials.

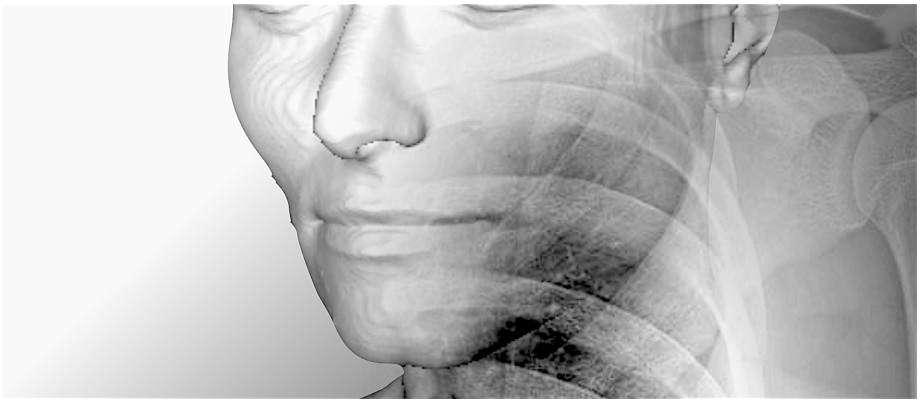
The patient just feels he/she is undergoing an ordinary standing erect X-ray examination that produces two radiographs, as many clinical protocols and trauma evaluations already require. The only difference is that he/she has to smoothly rotate between X-ray snapshots. This movement can be further eased with the help of a simple rotating platform.

The 3D information derived can be also used to connect points and areas in different radiographs through epipolar geometry, as shown in Section 5.7 with a cuboid-like rotating phantom. Epipolar lines (tackled in Chapter 3) can be drawn between two instances, as shown in Fig. 7.7, obtaining remarkable results which can turn out very useful in real clinical environments as the one demonstrated in Fig. 7.9.

Tests with X-ray phantoms will show in Section 7.2 that a millimetric level of precision can be achieved. The proposed techniques can be used in healthcare scenarios where 3D measurements are relevant and as an alternative to other modalities which may be felt as more invasive by or involve higher doses for the patient. In Chapter 7 we will also show some examples of the presented techniques with real patients.

6

Densitometric X-ray images



Conventional X-ray imaging does not take into account the patient's physical volume. This fact undoubtedly represents a lose of valuable information given the fact that X-rays may travel different distances through the patient's body until they reach the detector. Among other things, knowing the patient's volume enables the generation of *density images* instead of *absorption ones*. Until now, the most common solution to this problem was the *dual-energy X-ray absorptiometry* (discussed below in Section 6.6), which involves the comparison of two X-ray images produced with different voltages (150 and 80 kVp, for instance).

The goal of this chapter is to delimit the patient's body geometry and apply it to the original X-ray image, which entails deriving *a new transfer function for X-ray imaging*. A depth data-based methodology to account for the patient's real external volume is also discussed, which will allow us to derive a *mask* with the traversed lengths for all the Roentgen rays reaching each detector pixel. This mask is finally applied to the radiograph as a pixel-wise operation. The fact of managing the patient's volume and obtaining this *traversed length mask* or *L-buffer* heavily relies on the techniques described in previous chapters, notably Chapter 5. The result of this *merge* is a new type of X-ray image that can reveal new and interesting diagnostic details about the patient's health situation by enhancing soft tissues over the skeletal structure and other high density elements. These *density enhanced* X-ray images represent a new type of radiological image, easy to generate and to work with and very simply integrable as part of the diagnostic toolbox. The concepts and methodologies here tackled further develop those originally presented in [8].

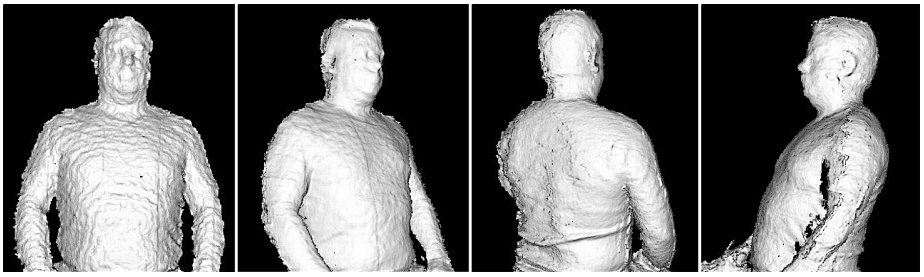
6.1 Patient volume estimation and surface reconstruction with KinectFusion

Surface reconstruction has always been a subject of interest since the inception of virtual and augmented reality. Many methods and technologies have been devised to achieve it. In 1996, Curless and Levoy [61] described a method for volumetric integration of complex models from range images (VRIP). The volumetric integration basically consists of a cumulative weighted signed distance function (SDF). This method is able to integrate high-detail models, in the order of a million triangles. However, the execution time can be in the order of hours and it is not suitable for AR applications. The range images used in this work were captured by laser scanners. Laser scanners provide range images with high accuracy, but the drawback of them is the high cost of the hardware. In 2002, Rusinkiewicz et al. [210] described a method for real-time 3D model acquisition. Using a real-time low-quality structured-light 3D scanner, they aligned the range

images from different viewpoints to produce complete 3D rigid objects. Different from the method proposed by Curless and Levoy, it operated at $\approx$ 10 Hz with lower cost hardware but did not reconstruct high-quality models. It was the first system to reconstruct and display the 3D models in real-time and it increased the possibility to do *markerless* AR with surface reconstruction. In 2010, Cui et al. [60] described a method for 3D object scanning using a time-of-flight (ToF) camera. In this work, Cui et al. showed a super-resolution method that improves significantly the quality of the depth maps acquired from a ToF camera. One drawback of this method is that it does not run in real-time. Compared to the other scanners presented, time-of-flight cameras have the lowest cost and provide range images with the lowest accuracy.

As mentioned in Chapter 5, in order to obtain the patient volume in an X-ray setting we use the KinectFusion algorithm, which operates on point clouds and depth data obtained from the patient. KinectFusion [111, 171] integrates depth maps from the depth camera into a *truncated signed distance formula* (TSDF) representation [62]. The TSDF is discretized into a voxel grid, typically 512×512×512, that represents a physical volume of space (normally a 3×3×3 m cube, which faithfully represents the dimensions of an X-ray room). Each voxel $\mathbf{v}$ contains two numbers: a signed distance d indicating how far that cell is from a surface and an integer weight w representing confidence in the accuracy of the distance. If $d < 0$ then $\mathbf{v}$ is *inside* a surface; if $d > 0$ then it is *outside*. Only depth values within a truncation band $-T < d < T$ are stored (a typical value is $T = 0.03$ m). The remaining voxels are sentinels with either $w = d = 0$ (uninitialized) or $d = T$ (empty space). The actual *world surfaces* are encoded as the zero crossings of the distance field and can be extracted by ray casting or marching cubes. The computational expense of this approach is mitigated by a highly parallelized implementation on newly available GPUs with up to 512 or more floating point cores and several GB of memory. The original algorithm can typically process each new frame in well under the 30 ms available before the next frame arrives. As the Kinect moves each new depth frame is used to incrementally localize its pose within previously observed geometry using the generalized iterative closest point algorithm (GICP) [219] with projective data association. The new readings are then integrated by sweeping through the TSDF: every cell which would appear in the camera is updated based on the previously stored values and the new depth map using a projective distance metric.

Figure 6.1 Surface reconstruction example with KinectFusion.



The result is the *dense surface reconstruction* (Fig. 6.1), already presented and discussed in Section 5.5.

In our implementation, we use the same technique explained in the aforementioned section, that is: the *virtual camera* approach, which is similar to the *moving patient scenario* described in Section 4.4.2.

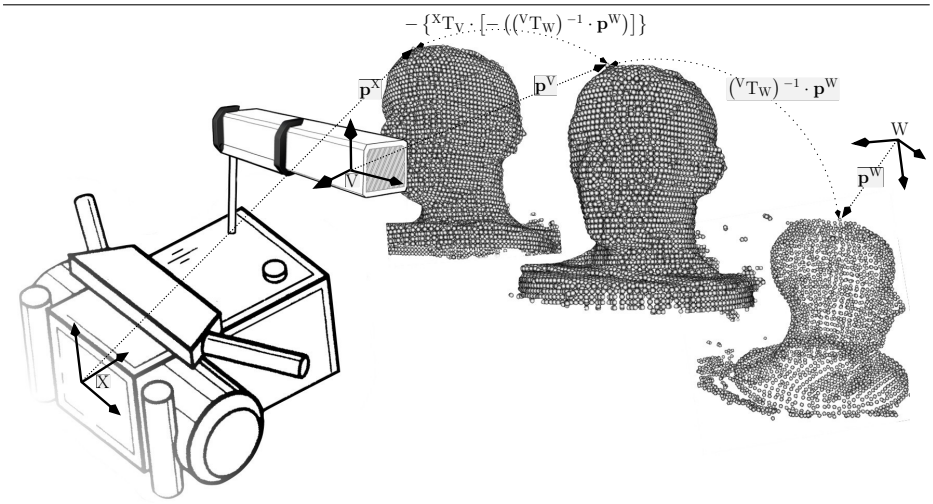
6.2 Volume translation to the X-ray reference frame

Once the patient's volume has been computed, it has to be translated to the X-ray reference frame, i.e., to the point of view of the anode. A 3D point belonging to the volume in world coordinates ($\mathbf{p}^{\mathrm{W}}$), can be translated to the anode's reference frame X with the following expression (schematically represented in Fig. 6.2):

$$\mathbf{p}^{\mathrm{X}} = -\left\{ {}^{\mathrm{X}}\mathrm{T}_{\mathrm{V}} \cdot \left[-\left(\left({}^{\mathrm{V}}\mathrm{T}_{\mathrm{W}} \right)^{-1} \cdot \mathbf{p}^{\mathrm{W}} \right) \right] \right\} \tag{6.1}$$

where ${}^{\mathrm{X}}\mathrm{T}_{\mathrm{V}}$ is the transformation that links the anode and depth camera reference frames. The term $\left({}^{\mathrm{V}}\mathrm{T}_{\mathrm{W}} \right)^{-1} \cdot \mathbf{p}^{\mathrm{W}}$ is the translation vector of a point in the volume in the depth camera's reference frame ($\mathbf{p}^{\mathrm{V}}$). Finally, $\left({}^{\mathrm{V}}\mathrm{T}_{\mathrm{W}} \right)^{-1}$ is the *pose* of the depth camera, as discussed in Section 5.4.

Figure 6.2 Translation of volume points between reference frames.



With this translated set of points $\left[\mathbf{p}^{\mathrm{X}}\right]$, it is now possible to accurately follow the path of each ray abandoning the anode and reaching the detector, while traversing the patient's (or examined object's) volume during its flight.

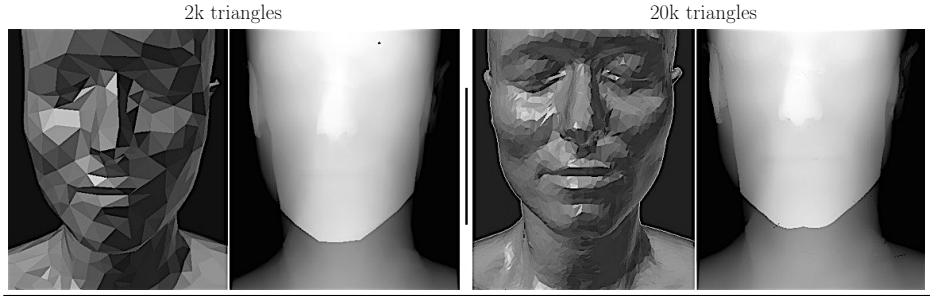
6.3 Derivation of the traversed length

Once the patient's volume has been obtained and expressed in the coordinates of the X-ray system, we can compute the so-called *L-buffer*. For each pixel $p(x, y)$ in the detector (with size $w \times h$), we can associate a traversed length l_p. A complete *X-ray beam traversed length mask* $w \times h \times 1$ matrix L or L-buffer can be then linked to the radiograph. If this mask is taken into account, it is possible to enhance each tissue type density in a more efficient and realistic way (specially, soft-tissue).

An initial method for calculating the traversed length path is by calculating the intersection of a ray and a 3D mesh of triangles, as described in [161]. The main advantage of this *meshing* approach is that the quality

of the generated L-buffer varies very little with the triangle density, as Fig. 6.3 shows.

Figure 6.3 Different mesh resolutions generate similar L. A total of 2k triangles were use to produce the first one (shown on the left), while approximately 20k triangles were used to compute the mask on the right.



Specifically, in order to efficiently compute the path length of a ray through an object, we use the algorithm described in [84]. For each scanned object, an image is generated. The intensity of each pixel corresponds to the distance covered by the ray within the object, from the source to the pixel centre in the detector. An example of a L-buffer image is shown in Fig. 6.4.

In Fig. 6.5 we show that the ray penetrates into the object when the dot product between direction vector of the ray and the object surface normal is negative. This dot product is positive if the ray leaves an object. Thus the path length of the ray into an object can be written as follows:

Figure 6.4 Real L-buffer computation from scanned volume of a head-like anthropomorphic phantom.

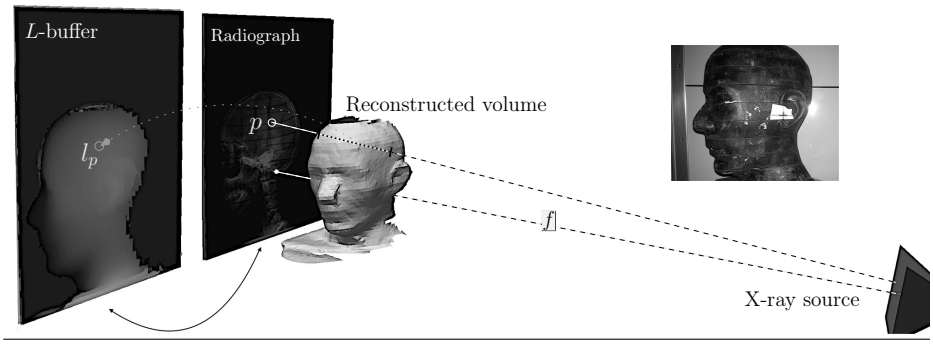
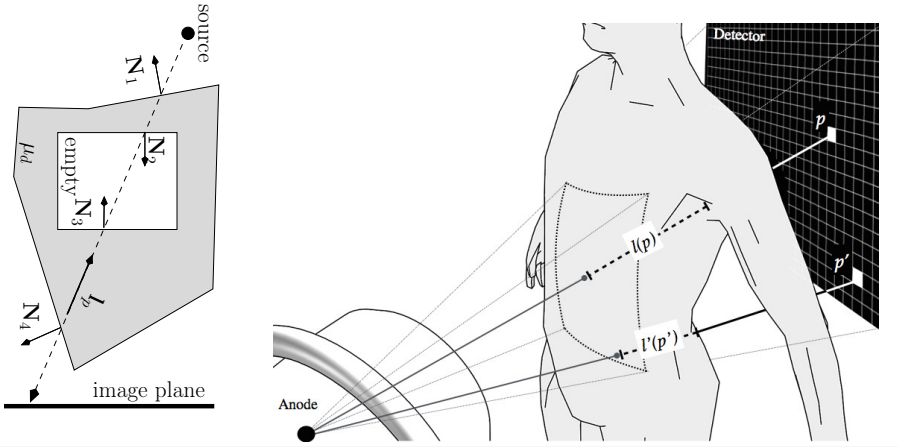


Figure 6.5 Left: computation of l_p (left). Right: different paths followed by X-rays when traversing the patient's body: each pixel (p, p') can be paired with a specific patient's traversed length (l, l').



$$l_p = \sum_i \text{sgn}\left(\mathbf{l}_p \cdot \mathbf{N}_i\right) \cdot d_i \tag{6.2}$$

where $\mathbf{l}_p$ is the *viewing vector* (i.e. the unit vector from the emission point to the detector's pixel p). The parameter i refers to the i^{th} intersection between the ray and the object surface, d_i is the distance from the X-ray source to the i^{th} intersection point and $\mathbf{N}_i$ is the vector normal to the object surface at the i^{th} intersection. The part $\text{sgn}\left(\mathbf{l}_p \cdot \mathbf{N}_i\right)$ is the sign of the dot product between $\mathbf{l}_p$ and $\mathbf{N}_i$. The intersections are found in an arbitrary order, which avoids sorting through them. This implementation of L-buffer (also described in [252]) very efficient and GPU-friendly.

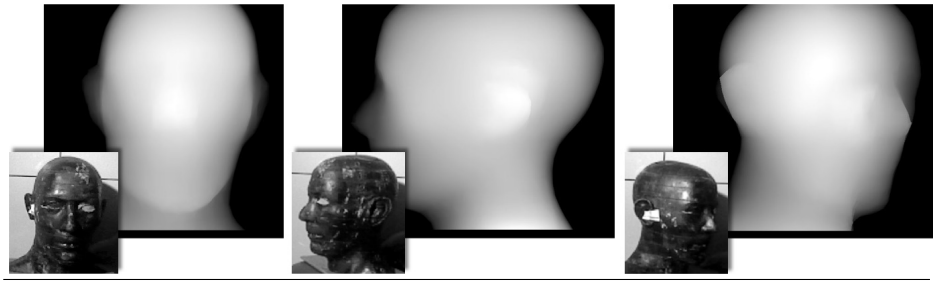
In Fig. 6.5 it is represented a 2D schematic representation of X-ray scene and the application of Eq. (6.2). Let μ_d be the attenuation coefficient of the outside mesh. In this case, the path length is given by:

$$l_p = (d_2 - d_1) + (d_4 - d_3) \tag{6.3}$$

where d_1, d_2, d_3 and d_4 are the distances from the X-ray source to the successive intersection points of the ray with the volume mesh.

The set integrated by all path lengths is the L-buffer. With this buffer it is now possible to generate densitometric images, as discussed below. Some more examples of L-buffers are shown in Fig. 6.6, where the amount of traversed length (by the X-ray beam) has been re-expressed using a gray level scale.

Figure 6.6 Some L-Buffers (expressed as gray levels, where black means $l_p = 0$ and white $l_p = \max$) at some positions of a head-like anthropomorphic phantom just before being radiographed.



6.4 Densitometric imaging

Once the L-buffer has been calculated we can obtain the mean density traversed by each X-ray arriving at a pixel x, y, according to the equation:

$$\mathcal{P}(x,y) = -\frac{\ln\left(\mathcal{I}(x,y)/\mathcal{I}_0(x,y)\right)}{\mu \cdot l(x,y)} \tag{6.4}$$

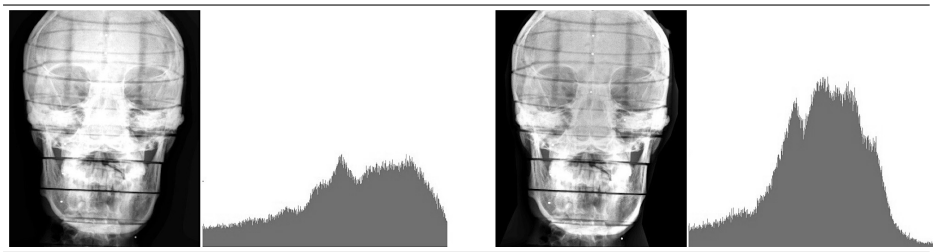
where $\mathcal{I}$ is the original absorption image obtained with Eq. (2.6) and $\mathcal{I}_0$ is the image corresponding to the same geometrical, energy and exposure configuration when no object/patient is interposed in the beam's path. With more detail, a pixel x, y in $\mathcal{I}_0$ corresponds to:

$$\mathcal{I}_0(x,y) = \int_0^{\mathrm{kVp}} N_0(E) \cdot e^{-\mu_a(E)\cdot l(x,y)} dE \tag{6.5}$$

where $\mu_a(E)$ is the linear attenuation coefficient for air and for energy E. The term $l(x, y)$ is the distance to pixel x, y, which will roughly be similar to the SID parameter (tackled in Section 2.7.2).

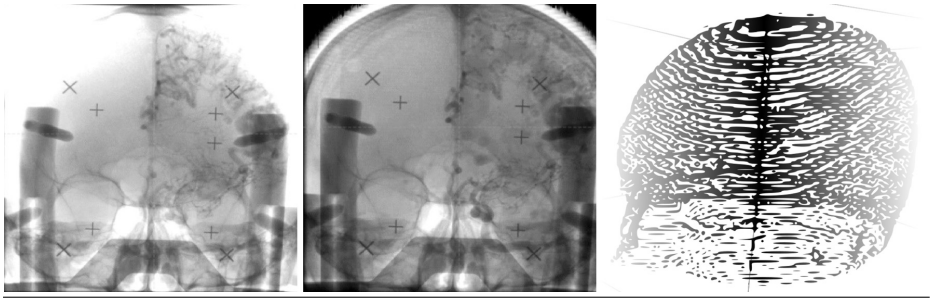
As stated in the introduction, this enables the obtention of densitometric images instead of absorption ones, the latter being the ones normally produced by conventional X-ray equipment. Theses images objectively have a higher quality, following the methodology described in [6] and in Appendix A. More specifically, densitometric images have a more compressed dynamic range (Fig. 6.7), which contributes to highlight soft-tissues over bony structures.

Figure 6.7 Compression of image histogram for densitometric images (right), compared against plain radiographs (left).



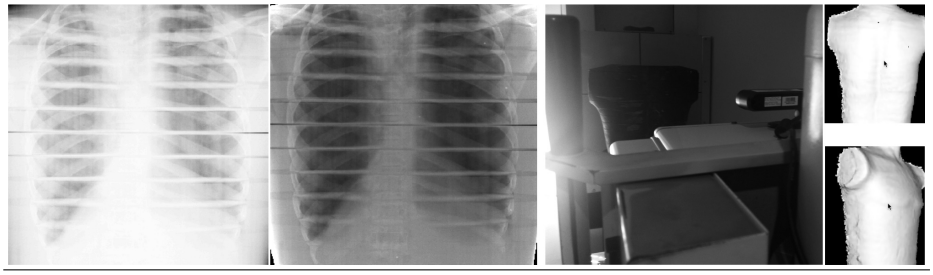
In Eq. (6.4) it is represented a *new transfer function* for X-ray imaging that accounts for the traversed length of each X-ray and the mean density encountered by it until reaching the detector.

Figure 6.8 Example of a densitometric image (middle image). The left radiograph represents the original X-ray instance and the rightmost image represents the traversed volume.



In some cases, this compression can dramatically enhance the resulting X-ray image, as shown in Fig. 6.9. As stated above, this enhancement can be objectively quantified with the metric described in Appendix A.

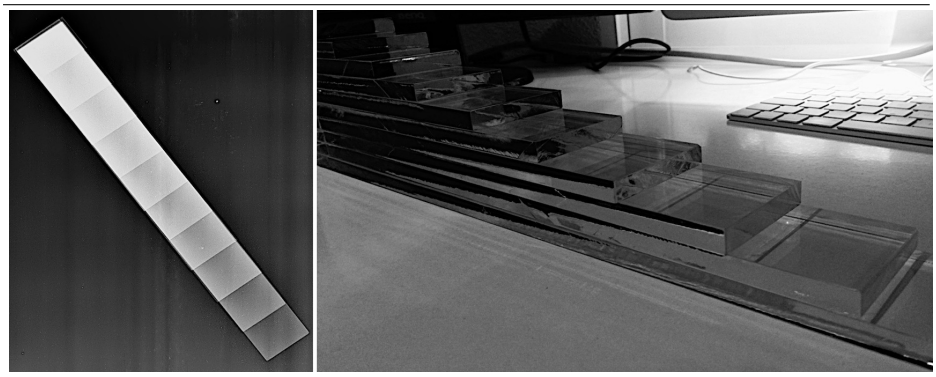
Figure 6.9 Another example of a densitometric image (middle image). The left radiograph represents the original X-ray instance.



Clearly, results obtained with Eq. (6.4) depend on the chosen value for μ. However, this parameter loses its relevance if the pixel values in Eq. (6.4) are resample/rescaled to a given grayscale range. By using Eq. (A.1), we can obtain a new contrast-stretched image $\tilde{\mathcal{P}}$. If $a = i_{\min}$ and $b = i_{\max}$, we would be performing a Global Contrast Stretching (GCS), which is a contrast operation commonly performed in radiology screening [212]. in GCS, the values $i_{\min}$ and $i_{\max}$ are the minimum and the maximum pixel intensities of the 12-bit original image (which will very likely correspond to a number close to 0 and 4095, respectively).

As a quick test, with Eq. (6.4) and the known geometry of the methacrylate staircase-like phantom, we have obtained a density of 1.16 g/cm^2, which is very close to the ground-truth.

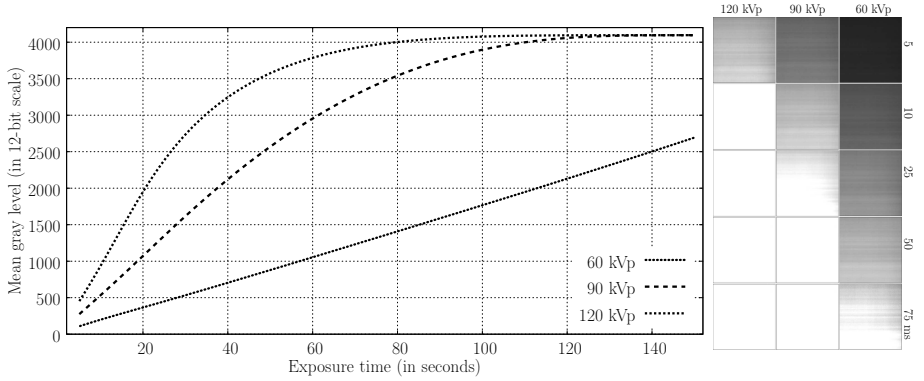
Figure 6.10 Methacrylate (Poly(methyl 2-methylpropenoate)) staircase-like phantom and a sample X-ray image of it.



6.5 Detector response

A final operation to be performed is taking into account the detector response. In this spirit, we have also tested (and measured) the behavior of the used CR detector behavior as shown in Fig. 6.11. For each pixel and beam energy, we have obtained a coefficients tuple m, n. These coefficients allows us to modify each image taking into account the detector response (which as expected and in our case, it happens to be linear) following Eq. (6.6).

Figure 6.11 Detector response: mean grey level (12-bit) depending on the chosen tube energy and exposure time. Common X-ray techniques remain in the linear region.



$$\tilde{\mathcal{P}}_c = C_m \cdot \tilde{\mathcal{P}} + C_n \tag{6.6}$$

where C_m and C_n are the matrices containing the linear coefficients obtained after interpolating the curves in Fig. 6.11. The linearity of the detector response can also be easily seen in Fig. 6.10.

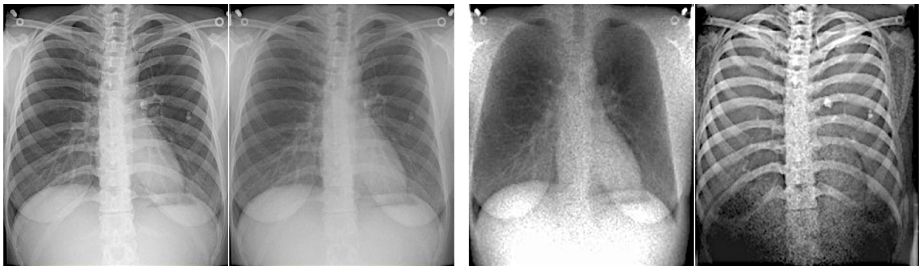
6.6 Bone-only images

The main goal of dual X-ray absorptiometry is enhancing the visibility of soft-tissue or the bone structures. Dual X-ray imaging involves obtaining two images at two different energies (kVp). Each image is weighted accordingly in order to null the signal due to bone for a *soft-tissue only* image, or to null the signal due to soft tissue for a *bone-only* image.

Soft-tissue only images can be achieved, for instance, by multiplying the high energy image ($\mathcal{H}$) by 2 and the low energy image ($\mathcal{L}$) by 1, subtracting the weighted high from the weighted low image, and scaling the residual tissue signal over a range. A similar operation can be done

to generate a bone-only representation. This last representation is the one which we are able to mimic very efficiently by using the densitometric imaging approach.

Figure 6.12 Example of dual X-ray imaging to highlight soft-tissue ($\mathcal{D}_S$) or bone ($\mathcal{D}_B$) from a low energy image ($\mathcal{L}$) generated at 60 kVp and a high energy radiography ($\mathcal{H}$) produced at 120 kVp.



$\mathcal{L}$ $\mathcal{H}$ $\mathcal{D}_S$ $\mathcal{D}_B$

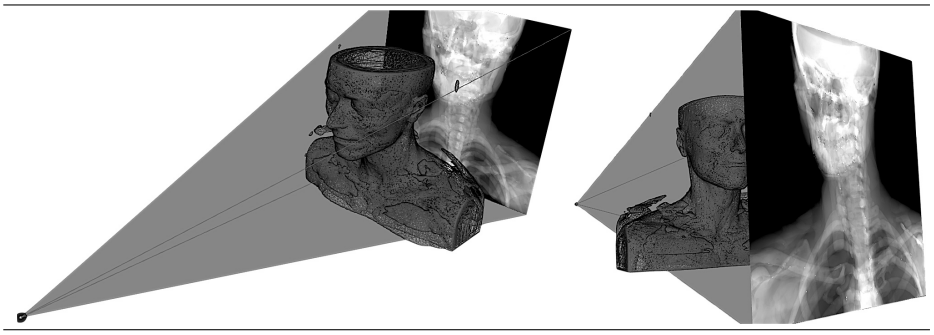
The simplest method to achieve a dual-energy image ($\mathcal{D}_S$ for soft-tissue and $\mathcal{D}_B$ for bone-only) is to perform a simple logarithmic subtraction:

$$\begin{aligned} \ln(\mathcal{D}_S) &= \ln(\mathcal{H}) - w \cdot \ln(\mathcal{L}) \\ \ln(\mathcal{D}_B) &= w \cdot \ln(\mathcal{L}) - \ln(\mathcal{H}) \end{aligned} \tag{6.7}$$

where w is a weighting factor different for bone (w_B) and soft-tissue (w_S), respectively.

As commented above, the dual energy X-ray technique, by definition, needs a second low-energy exposure in addition to the ordinary high-energy projection, increasing exposure. However, we can achieve similar results by building a virtual water phantom (as shown in Fig. 6.14-right) and generating a *virtual X-ray image* from it ($\mathcal{H}'$). A more realistic example of the virtual X-ray methodology is shown in Fig. 6.13.

Figure 6.13 Example of the application of the deterministic algorithm to simulate and generate X-ray images.



The aforementioned made-up water-filled volume (Fig. 6.14) is just the result of assigning the density of water to the patient's volume obtained with the methods discussed in Section 6.1.

A new a $\mathcal{D}'_B$ image can be generated with the following equation:

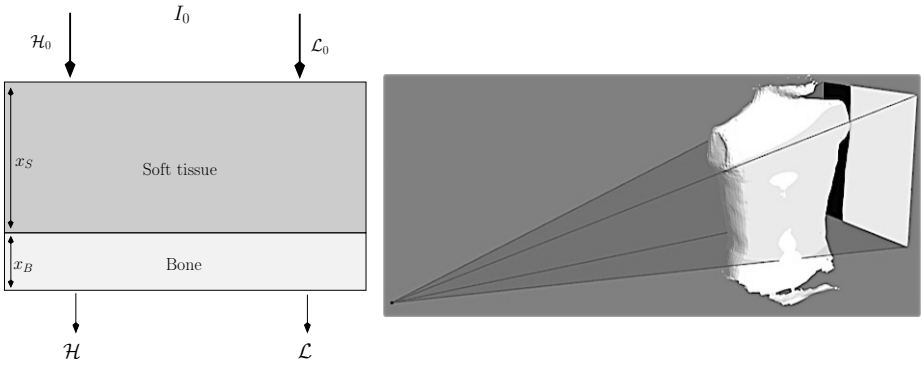
$$\ln(\mathcal{D}'_B) = w_B \cdot \ln(\mathcal{L}) - \ln(\mathcal{H}') \tag{6.8}$$

This process requires just one X-ray image from the patient (which in turn involves less dose) and measuring the patient's volume with the previously discussed methods. The virtual soft-tissue radiograph is computed with the ray-tracing principle [84] over a triangle-meshed version of the resolved volume with the methods described in Section 6.3. In turn, radiation attenuation is obtained by considering the thickness penetrated by each ray going through the object characterized by its density (water, in this case), its attenuation coefficient, beam energy and Eq. (2.6).

For each ray, the total path length through the volume is determined using the geometrical computations. Finally, the attenuation of X-rays for a given pixel is computed using the recorded path lengths and X-ray attenuation coefficients in Eq. (2.6). With more detail, for each object i, we can find l_i, which is the path length of a ray in the ith object. It can be decomposed to illustrated the different rendering passes:

1. compute the path length of each traversed object l_i, then
2. make use of the first pass to compute $\sum_i \mu_i l_i$, and finally,

Figure 6.14 Left: Beer's Law for two different incident energies. Right: virtually radiographed water phantom built to obtain bone-only images.



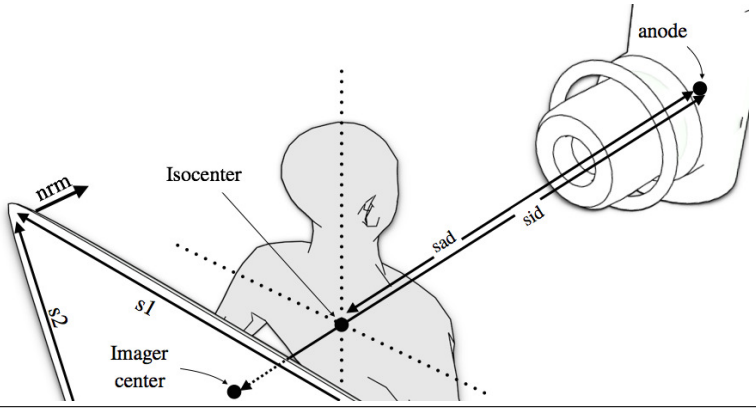
3. compute, in a second pass, the number of transmitted photons using the attenuation law with Eq. (2.6).

In Section 6.6 we will show a practical example in which we compute the $\mathcal{D}'_B$ instance of the Rando phantom.

6.7 Enhancement of digitally reconstructed radiographs

Computer or Digitally Reconstructed Radiographs (CRR/DRRs) from CT scans represent nowadays an essential tool in the planning and verification of radiotherapy treatments sessions. However, they can also provide valuable diagnostic information by themselves and can be, for some type of assessments, as useful as plain radiographs and a diagnosis complement to the CT studies they are derived from. For instance, the research carried out by [38] shows how acetabular fractures can be identified as efficiently as with plain X-ray images and authors in [195] arrive to a similar conclusion for intracranial lesions. Besides, they are commonly used in the operating room for direct comparisons with fluoroscopic images [109] and

Figure 6.15 Geometrical scenario defined by Plastimatch. The patient's volume (*isocenter* or the origin of the reference frame defined by Plastimatch) is virtually positioned in a *synthetic* conventional X-ray setting with a fictional detector area and anode. In Plastimatch's nomenclature *sad* stands for the distance between the isocenter and anode and *SID* is the orthogonal distance between the detectors' center and the anode. The *nrm* vector defines the orientation of the imaging system and hence the angle in which the patient would be imaged.



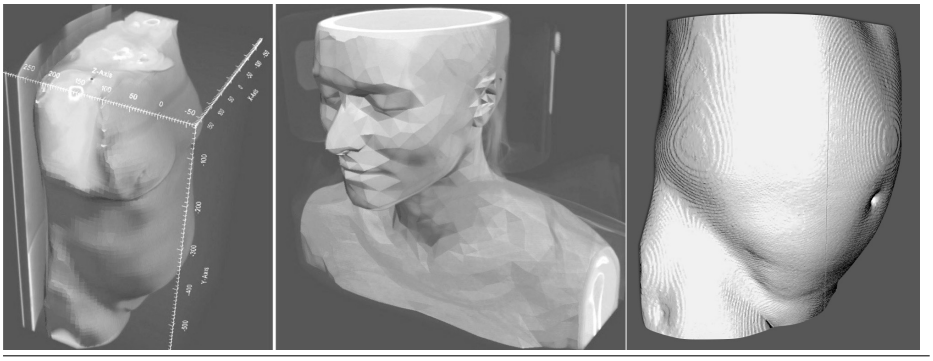
after surgery they can also be suitable for verifications with immediate postoperative images [67].

As with plain (and physical) radiographs, DRR algorithms do not take into account the patient's physical volume (inherently present in the CT data) which is, undoubtedly, a lose of valuable information. This is understandable, since they try to mimic as faithfully as possible a real X-ray scenario.

Many vendors ship and/or provide DRR software and tools. For our research we have chosen the Plastimatch package [223] mainly developed at Harvard university. Plastimatch implements a fork of the Siddon ray tracing method [228]. This algorithm uses the original exact path length based on the intersection of rays with the image voxels. From here, voxel interpolation is also applied which contributes to increase the apparent resolution of the final DRR. Both multi-core and GPU versions are available. The geometrical framework defined by Plastimatch is shown in Fig. 6.15, which is very similar to that described in Section 2.7.2.

Once the initial DRR has been generated, the next step is isolating the patient's volume from the available CT data. Many techniques do exist to this effect (3D edge-detection, graph-searching algorithms, deformable models, stochastic techniques, region growing and neural networks, etc.) and are very succinctly summarized in [130]. In our implementation we have just computed isosurfaces [185]. These are defined by connecting voxels whose intensity is equal to a value, i.e., Hounsfield units. A isosurface with a Hounsfield value of -200 to -500 will normally arise patient's skin (Fig. 6.16).

Figure 6.16 Some isosurfaces obtained from CT scans.

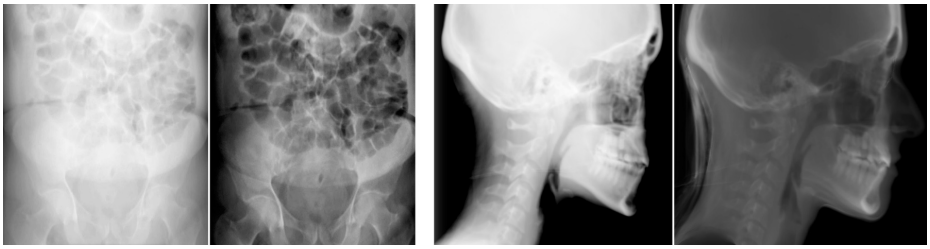


With the computed volume, we just have to apply the same methods discussed in Section 6.4 to produce *volume-enhance DRRs* like the ones shown in Fig. 6.17. As with densitometric images, these enhanced DRRs better highlight soft-tissues when compared against vendor-generated ones (with Plastimatch, in our specific case). This fact may significantly contribute to the diagnostic process.

6.8 Conclusions

In this chapter we have presented the foundations of densitometric imaging. Density enhanced images highlight soft-tissue areas over high density

Figure 6.17 Some vendor-generated DRRs (left) and their *volume-enhanced version* (right).



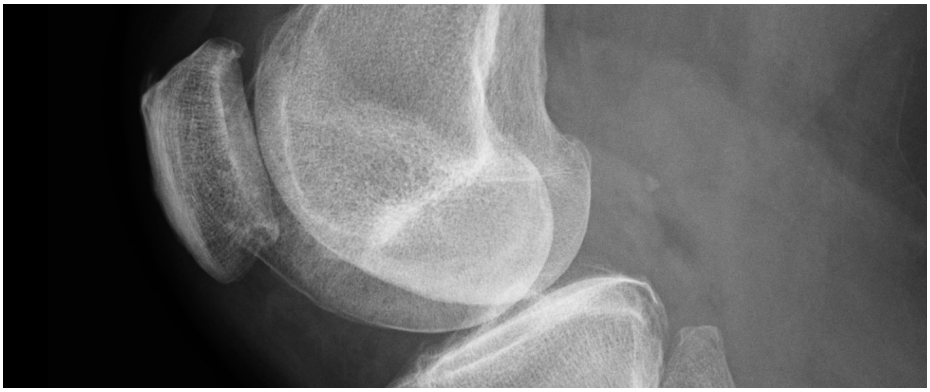
ones by compressing the histogram of intensity levels. A densitometric X-ray image can reveal more information about the density of the radiographed element/patient and can have a significant impact on the final diagnosis. In order to produce these images, it is necessary to compute the volume of the X-rayed object/person and rigidly translate it to anode's point of view. Finally, a new transfer function that incorporates the traversed X-ray paths in the interfering volume is derived.

This process relies on the techniques summarized in previous chapters and can also be used to derive the so-called *bone-only* images from which a value for the bone density can be assessed. Reconstructed bone-only images only require a low dose X-ray exposure, thus further contributing to the patient's health. Finally, we have shown that the same techniques can also be applied to enhance the quality of digitally reconstructed radiographs, which were previously generated from CT scans.

In Chapter 7 we will show some examples of densitometric images, their capabilities and we will also test their improved quality.

7

Tests and experiments



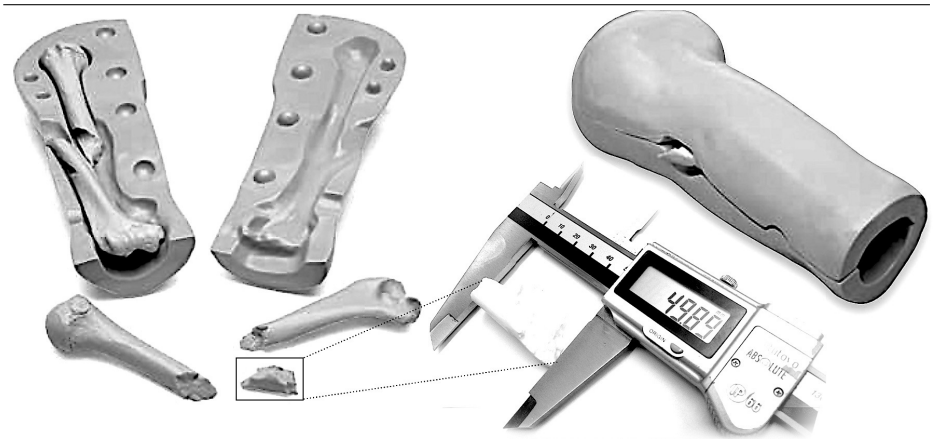
In this chapter we apply in a more realistic environment the methodologies previously summarized regarding the augmentation of X-ray equipment, generated radiographs and their graphical information. This is achieved in two ways. The first one involves the geometrical conjugation of several X-ray images taken at different patient or imaging system positions. With these techniques (originally presented in [53]) we can derive precise 3D information formerly available only in more complex (and more radiative) modalities such as CT scans. The second way focuses on the X-ray image itself, which can be applied physical filters involving the patient's volume in order to obtain densitometric/density-enhanced im-

ages, following the recipes introduced in Chapter 6. All these experiments were carried out in a real clinical scenarios with both research phantoms and real patients.

7.1 3D data from radiographs and visual information

In this section we apply the techniques introduced in Chapter 4 on a more realistic target (other than the tests performed with the calibration DUT in Section 4.7). This time we use an arm-shaped anthropomorphic phantom from Life/form® that includes a splinter fracture (Fig. 7.1).

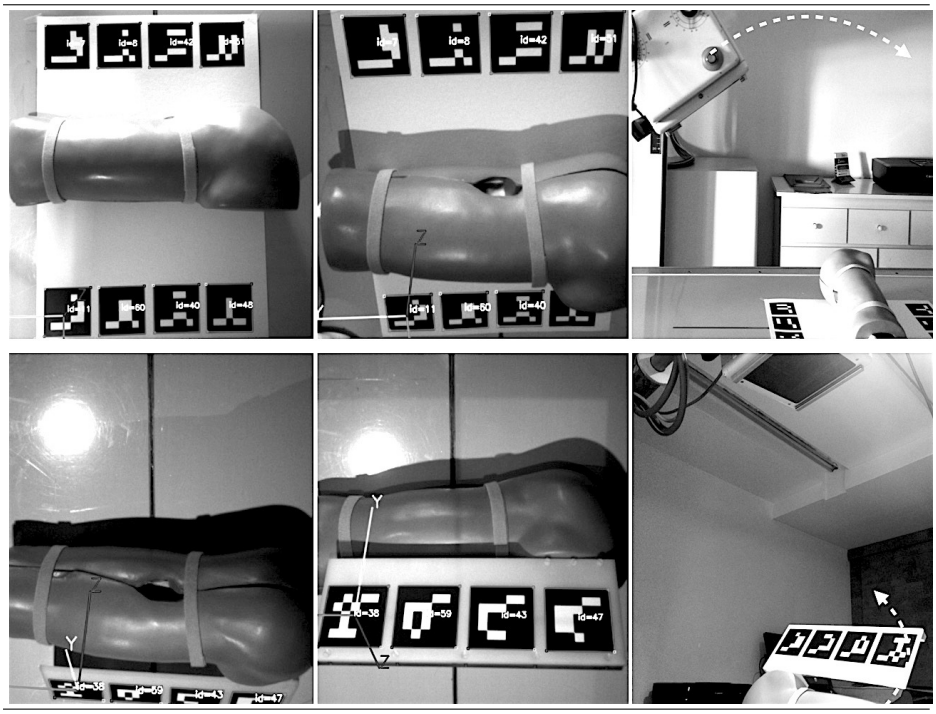
Figure 7.1 Anthropomorphic arm-shaped phantom with a 5 cm splinter fracture.



More specifically, we apply the techniques described in Section 4.7.3 in order to compute the splinter length from X-ray image pairs and to compare it with the real length (~ 5 cm). Obviously, in cases with severely injured patients, the moving camera scenario (immobilized patient tackled in Section 4.4.1) might be more appropriate. However, we also analyzed the moving patient scenario for the sake of completeness. Both of them

are graphically summarized in Fig. 7.2.

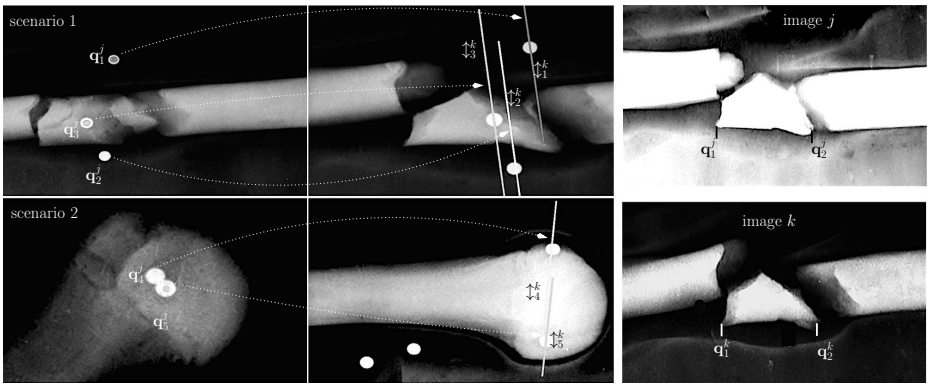
Figure 7.2 Anthropomorphic phantom tested for the camera moving scenario (top) and patient moving scenario (bottom). In the camera moving scenario, a board with visible fiducials at known 3D coordinates relative to W is used and several radiographs are generated at different angles/locations. In the patient moving scenario, the board of visible fiducials is rigidly tied to the object under examination, which arbitrarily moves and rotates while the X-ray imaging system remains fixed.



Combining all available image pairs (similar to those in Fig. 7.3-bottom) we obtained a mean length of 4.9 cm with a deviation of 0.1 cm for the fixed arm setting. In the case of the moving phantom, we retrieved a mean distance of 5.0 cm with a deviation of 0.2 cm.

Bounded epipolars are drawn in Fig. 7.3-top using the algorithm described in Section 4.7.2. For this experiment, several lead spherules were added as we did with the DUT in Section 4.7. The distances between computed epipolars and real observed projections was under 15 px for the

Figure 7.3 Left: some examples of bounded epipolar lines between stereo X-ray snapshots. Points 4 and 5 lie very close to each other and are difficult to discern in the left image. However, they can be clearly differentiated in the stereo pair on the right. Right: two radiographs (produced in a moving camera setting) of our anthropomorphic phantom. Two points have been manually selected on each image at both ends of the splinter. With this information, we can estimate $|\mathsf{Q}_2 - \mathsf{Q}_1|$ with the mathematical tools described in Section 4.6.2.

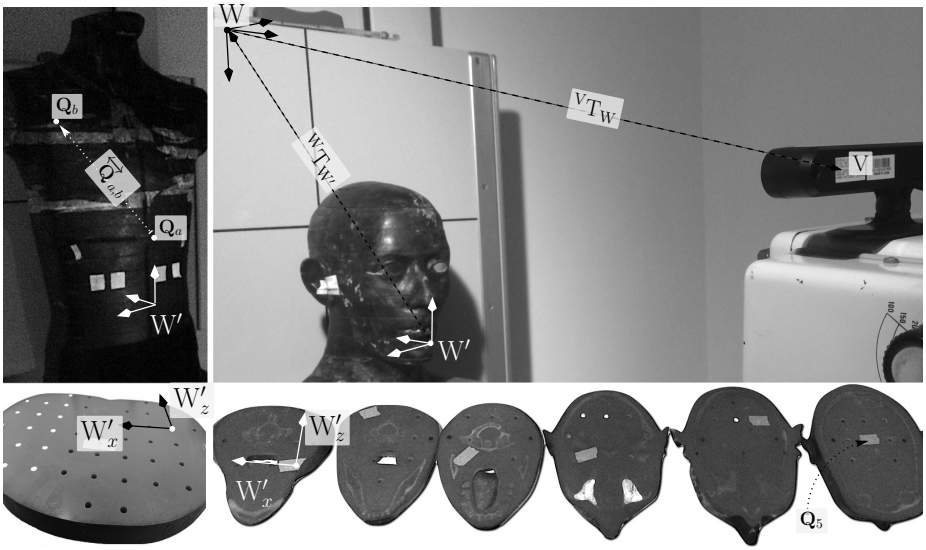


camera moving scenario and under 12 px for the phantom moving scenario. This result eases a good identification of hidden facets which are missing or indistinguishable in one image, but that are quite evident in a second radiograph.

7.2 3D data from radiographs and depth information

To test the proposed methodology in Chapter 5, we have initially used three phantoms. The first two consist of two anthropomorphic phantoms (head and torso, studied separately) shown in Fig. 7.4.

Figure 7.4 Tested anthropomorphic phantoms (torso and head). Several slices are shown and the local reference frame W′ is defined in two of them (one for each phantom). Spherical lead bearings are placed in each slice. As an example, two of them ($\mathbf{Q}_a$ and $\mathbf{Q}_b$) are represented, with $\overleftrightarrow{\mathbf{Q}}_{a,b}$ being the segment that connects them.



A set of spherical lead bearings (1 mm size) was manually placed in the phantoms at well-known locations relative to a local coordinate system W′ (shown in Fig. 7.4) bound to each phantom. The reason for choosing W′ is that we can easily know the location of the lead bearings relative to W′ (by construction), but we do not know the exact position of each phantom relative to W.

Both phantoms were placed ∼1.40 m away from the anode and their rotation was tracked using the techniques described in Section 5.5. Radiographs were produced at arbitrary angles relative to the X-ray imaging system and their projection matrices were calculated with Eq. (5.2). In Fig. 7.5 we show two of these X-ray snapshots and some steps of the dense surface mapping process. X-ray images are 35×43 cm with a resolution of 10^4 px/m.

Figure 7.5 A pair of X-ray images generated from the anthropomorphic phantom (torso) at two positions j and k with an angular amplitude of 85° (around the craniocaudal axis) between them. The projection of the chosen local reference frame W′ is also represented. Notice the white dots corresponding to the projections of the lead bearings. One of these, belonging to the torso ($\mathbf{q}_b$), is highlighted in the stereo image in the right. The dashed line tagged with and l_b^k is the corresponding epipolar line in image k associated to its stereo counterpart (projection $\mathbf{q}_b^j$). Some stages of the DSM algorithm representing the continuous tracking of the patient are shown on the right.

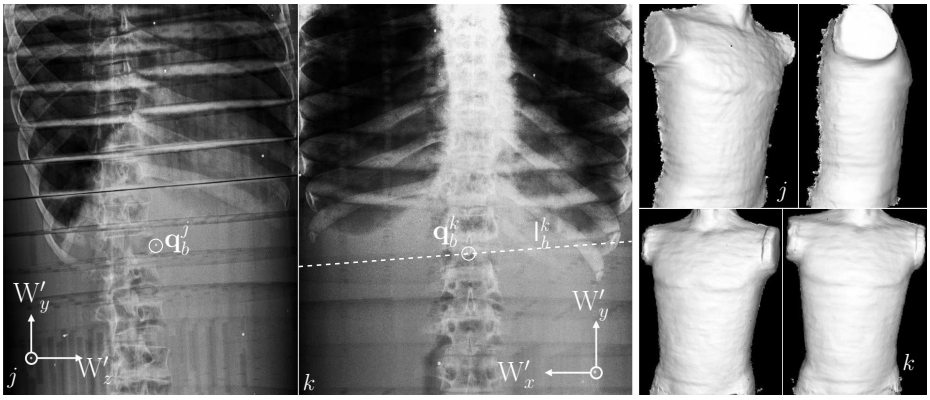


Figure 7.6 Tested head-like polystyrene phantom (also partially shown in Fig. 5.8-b). The location of some horizontal and vertical (dotted) fixed length pins $\overleftrightarrow{\mathbf{Q}}_i$ are shown. Right: one of the 30 mm length metal pins.

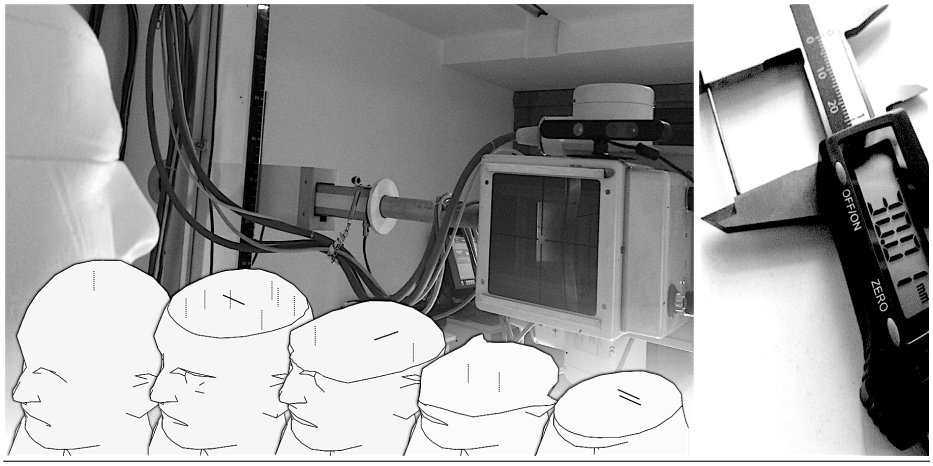
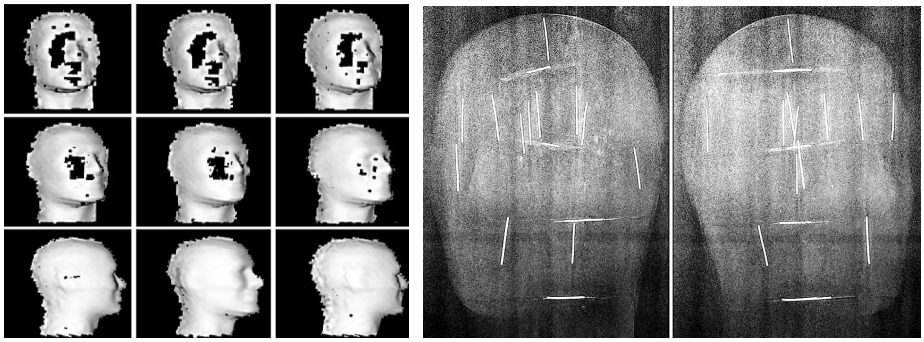


Figure 7.7 Surface mappings from the polystyrene phantom and two radiographs showing implanted needles.



These anthropomorphic phantoms (together with the radiograph pairs generated from them) will be used to estimate the 3D locations of the spherical lead bearings. These derived locations will then be compared against the ground-truth positions expressed in the W′ coordinate system. It is then mandatory to also find the optimal transformation (i.e., ${}^{\mathrm{W}}\mathrm{T}_{\mathrm{W}'}$) that best aligns the two sets: Q_i, which is derived in coordinates of W; and $\mathbf{Q}_i$, which is known relative to W′. This entails performing a Euclidean transform, which preserves shape and size, as tackled by Besl and McKay [34]. In this work, ${}^{\mathrm{W}}\mathrm{T}_{\mathrm{W}'}$ is found with a Procrustes analysis described by Dryden and Mardia [74] but with scale invariance. We will apply this transformation later when comparing the mean Euclidean distance between the 3D coordinates of both sets.

The third phantom is made of polystyrene (Fig. 7.6). On this occasion, 15 medical needles of known and precise lengths 36.1, 18.4, and 9.2 mm (5 of each type) were introduced at random locations, in both vertical and horizontal orientations. The phantom was placed ∼1.57 m away from the X-ray anode and its motion was tracked using the same techniques described in Section 5.5. The angular span between t_j and t_k was 70 degrees. In this case, we compared the estimated needle lengths $\left|\overleftrightarrow{\mathrm{Q}}_i\right|$ against the ground-truth. To derive these lengths, we follow the procedure described at the end of Section 5.6. The only role of this second phantom is to house the aforementioned needles at different and varied positions. This phantom could have had any shape; however, we thought the experiment would be more realistic if the phantom resembled a human part (a head, in this case).

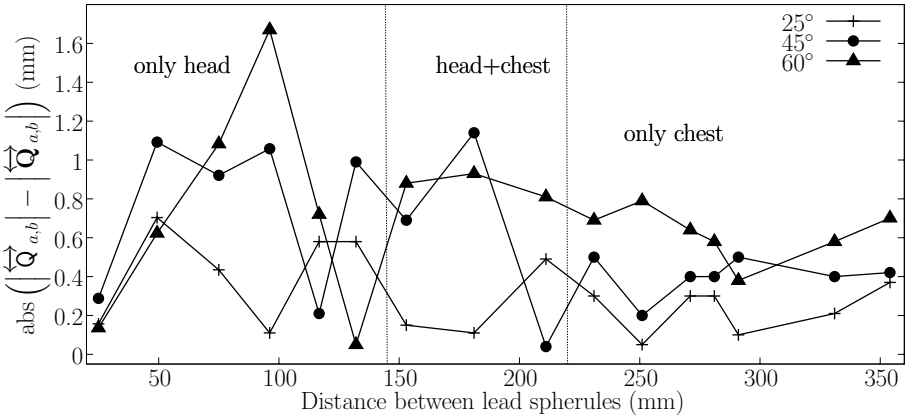
Using Eq. (4.15), we obtained the 3D location Q_i of each lead fiducial from combinations of radiograph pairs of the anthropomorphic phantoms at different orientations and the 2D pixel coordinates of each spherule projection in each image. The alignment operation ${}^{\mathrm{W}}\mathrm{T}_{\mathrm{W}'}$ was applied to express their components relative to the W reference frame. Table 7.1 shows the mean Euclidean difference ($\triangle$) between all Q_i and $\mathbf{Q}_i$. Each row also specifies the angular amplitude ($\measuredangle$) around the vertical axis for each orientation/radiograph combination. The experiment was carried out for the head and torso parts separately to test the tracking algorithm with a relatively small (head) and a relatively large body part (chest).

Table 7.1 Mean Euclidean span ($\triangle$) between ground-truth positions of spherical lead bearings inside the anthropomorphic phantom (head and torso) and the derived ones for some angular extents ($\measuredangle$).

	Head phantom		Torso phantom	
$\measuredangle$ (°)	$\triangle$ (mm)	σ (mm)	$\triangle$ (mm)	σ (mm)
25	1.4	0.7	0.8	0.5
45	0.9	0.5	0.9	0.5
60	1.3	0.6	1.0	0.6
85	1.6	0.9	1.2	0.7

We also compared the differences in the distances between some spherical bearings in the anthropomorphic phantom (head and torso). Given that we know where each spherule is located (in the W′ system), we can also determine the ground-truth Euclidean distance $\left|\overleftrightarrow{\mathbf{Q}}_{a,b}\right|$ between any $\mathbf{Q}_a$, $\mathbf{Q}_b$ pair. This ground-truth distance is then compared against the estimated one $\left|\overleftrightarrow{\mathsf{Q}}_{a,b}\right|$ using the previously derived locations Q_i of each spherical lead bearing. In Fig. 7.8 we represent these differences graphically for three angles and for several known distances (ground-truth). These distances range from the closest gap between two near spherical lead bearings (24.6 mm) to the widest gap (355.1 mm). Again, the experiment was executed separately for the head and torso anthropomorphic phantoms, but for the sake of simplicity the results are shown combined in Fig. 7.8. The same plot also distinguishes between which distances are measured in the head and torso phantoms.

Figure 7.8 Difference (in absolute value) between computed distances between pairs of spherical lead bearings and ground-truth. The results for three angular amplitudes (25°, 45°, and 60°) are shown.



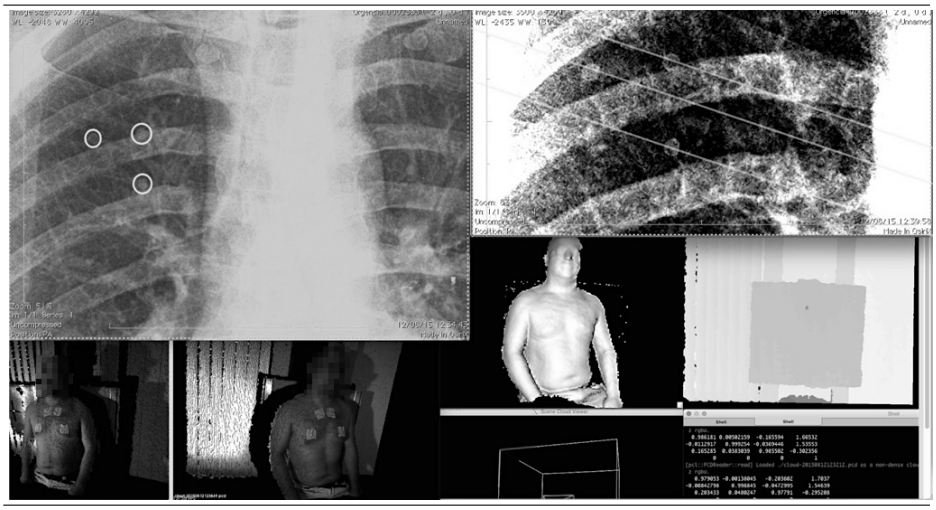
In the case of the polystyrene phantom, we obtain the mean needle lengths $\left|\overleftrightarrow{\mathbf{Q}}_i\right|$ shown in Table 7.2 and Fig. 7.7, and we compare them against the ground-truth $\left|\overleftrightarrow{\mathbf{Q}}_i\right|$.

Table 7.2 Derived lengths $\left|\overleftrightarrow{\mathbf{Q}}_i\right|$ and ground-truth $\left|\overleftrightarrow{\mathbf{Q}}_i\right|$ (in mm).

$\left\|\overleftrightarrow{\mathbf{Q}}_i\right\|$	$\left\|\overleftrightarrow{\mathbf{Q}}_i\right\|$	$\left\|\overleftrightarrow{\mathbf{Q}}_i\right\|$	$\left\|\overleftrightarrow{\mathbf{Q}}_i\right\|$	$\left\|\overleftrightarrow{\mathbf{Q}}_i\right\|$	$\left\|\overleftrightarrow{\mathbf{Q}}_i\right\|$
36.1	36.1±0.3	18.4	18.5±0.1	9.2	9.1±0.2

These techniques involving 3D reconstruction have also been applied with real patients. In Fig. 7.9 we show an example of a patient to whom several X-ray opaque fiducials were sticked and how their location can be easily traceable between radiographs through the drawing of epipolar lines. This patient was asked to rotate between two posterior oblique positions (summarized in Section 2.7.1).

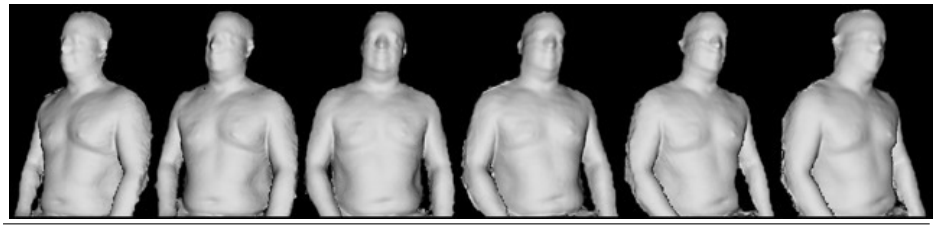
Figure 7.9 Epipolar geometry used within a real clinical scenario and a patient, allowing the identification of simple fiducial markers between two radiographs.



7.3 Densitometric images

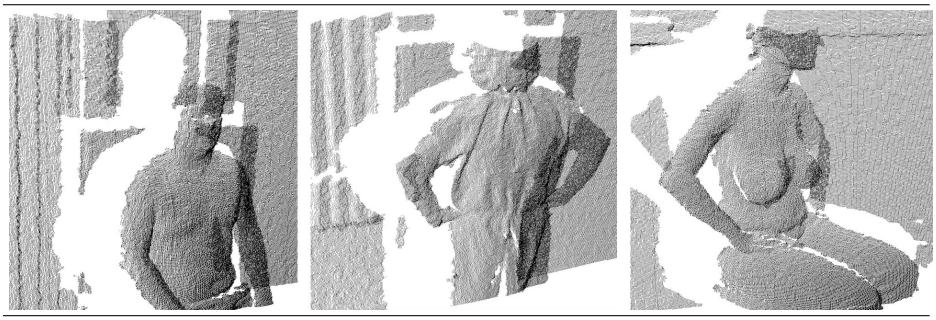
This section presents the results of several experiments about the obtention of densitometric images, following the procedures described in Chapter 6. In order to verify the amount in quality of densitometric version over the plain X-ray image, we have followed the methods presented in Appendix A. In all cases, the patient's volume was reconstructed following the techniques described in Section 5.5 and translated to the X-ray anode's point-of-view (Fig. 7.12).

Figure 7.10 Reconstructed patient's volume from several DSM snapshots.



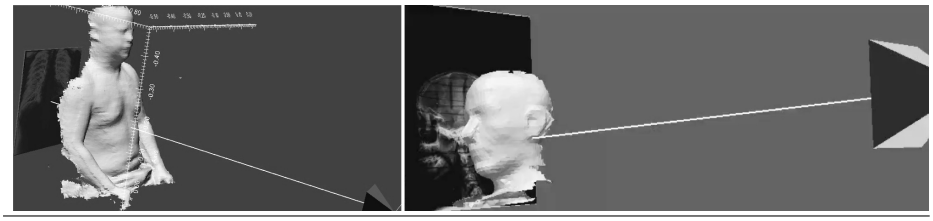
Some examples of this volume assessment process is shown in Fig. 7.10 and Fig. 7.11.

Figure 7.11 Point clouds from patients acquired during X-ray examinations.



These volumes were later translated to the anode's reference frame (Fig. 7.12) in order to obtain the L-buffer described in Section 6.2.

Figure 7.12 3D representation of the X-ray examination scene where the patient and phantom volume have been aligned accordingly to the anode's reference frame.



In Fig. 7.13, Fig. 7.14, Fig. 7.15, Fig. 7.16 and Fig. 7.17 we show some examples of densitometric images and the increase in the amount of

information (in %) relative to the normal absorption radiograph.

Figure 7.13 Densitometric image (right) from a plain radiograph (left) of the chest of the Rando phantom. The increase in the mutual information is 4%.

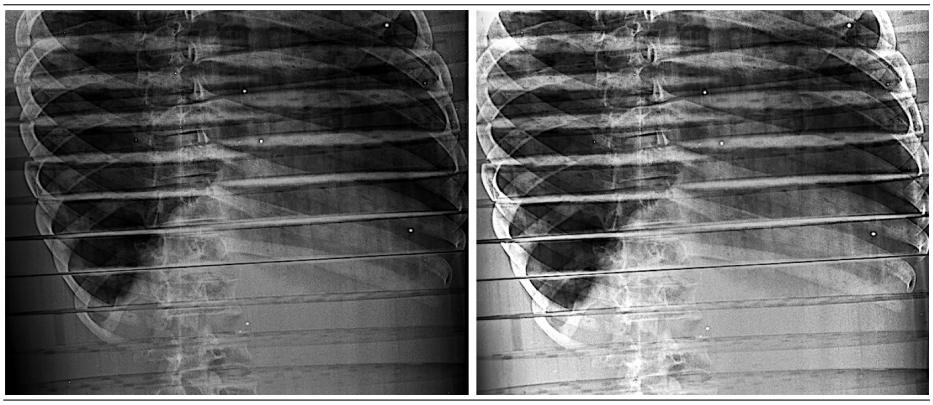


Figure 7.14 Densitometric image (right) from a plain radiograph (left). In this case, the increase in the mutual information is 3%.

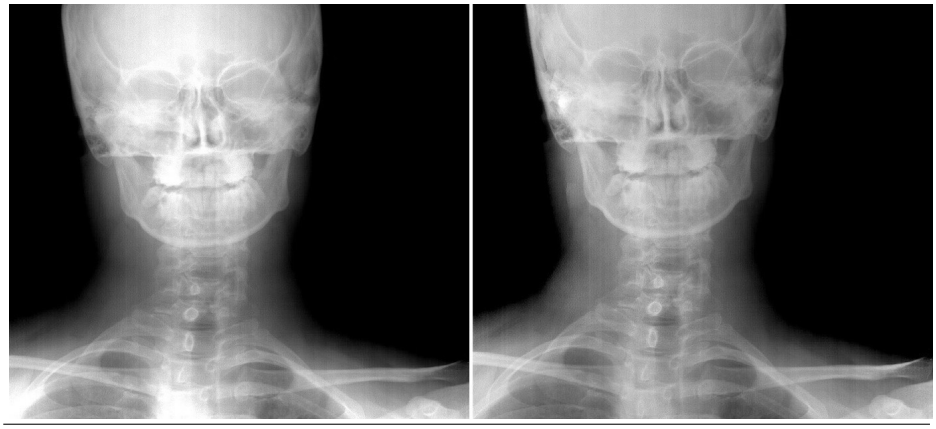


Figure 7.15 Densitometric image (right) from a plain radiograph (left). The increase in the mutual information is 2%.

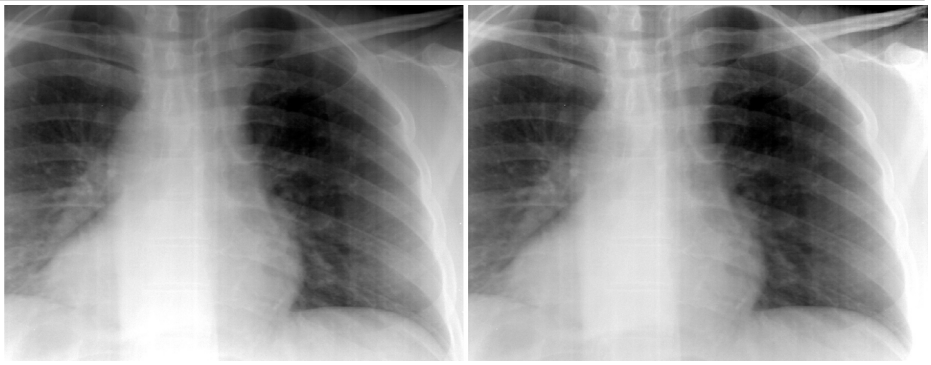


Figure 7.16 Densitometric image (right) from a plain radiograph (left). The increase in the mutual information is 2%.

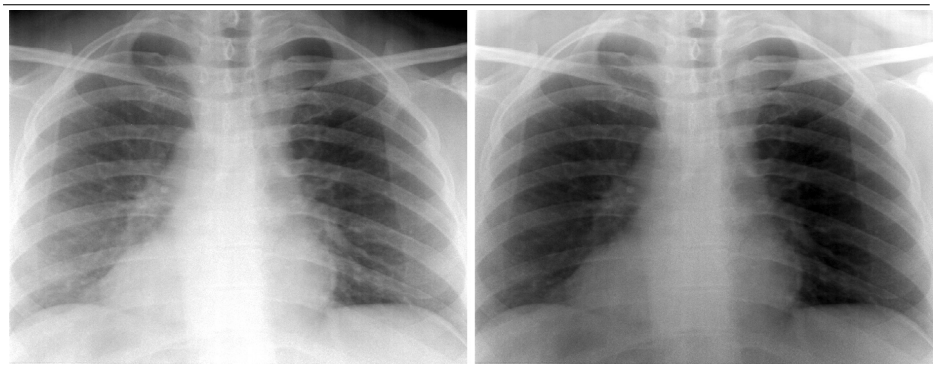


Figure 7.17 Densitometric image (right) from a plain radiograph (left). The increase in the mutual information is 1%.

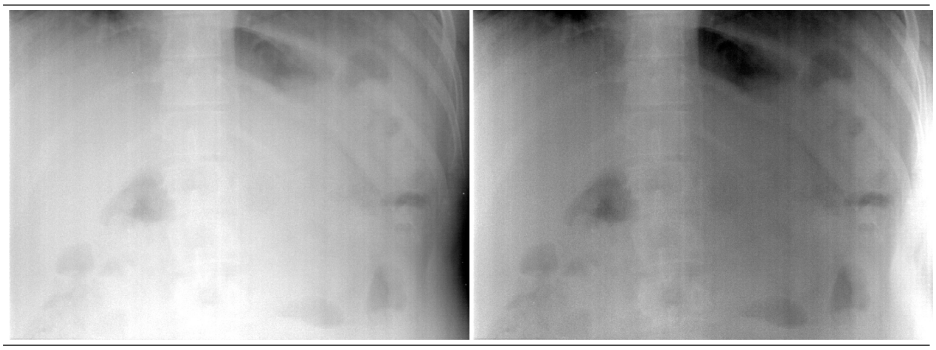


Figure 7.18 Two CT scans from the same Rando Phantom in a Discovery CT750 HD from GE Medical Systems and in a Phillips Brilliance 64 equipment.

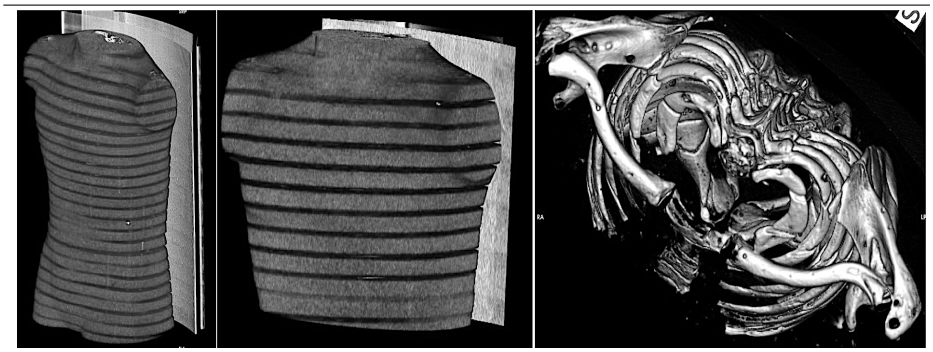


Figure 7.19 $\mathcal{D}_B$ image obtained with the methodology described in this section (right). Left image represents the original radiograph and the central image is a grayscale representation of the L-buffer.

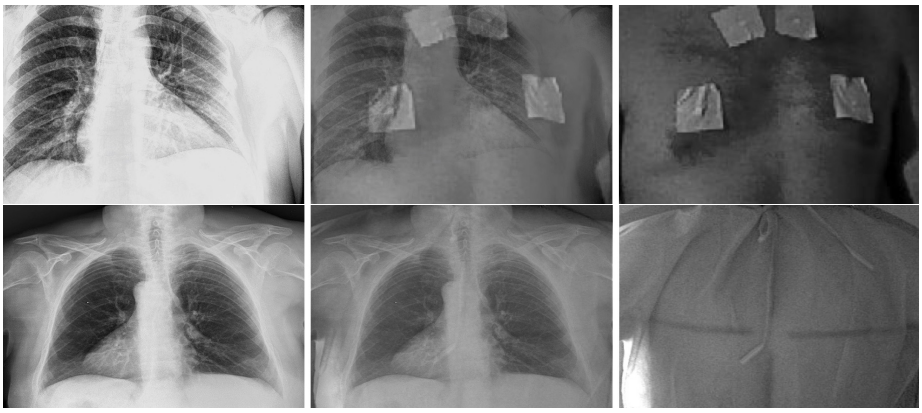
7.4 Bone density assessment

In order to test the techniques described in Section 6.6 we have compare the bone density obtained with the methodology there explained with the one derived from two different CT scans from a Rando® phantom shown in Fig. 7.4. The bone density assessed by both scanners is almost the same, however, a mean value was used for the sake of completeness (Fig. 7.18). According to Eq. (6.4) and Eq. (6.8), we obtain a very similar density (1.01 g/cm^2) for the clavicle.

7.5 Multimodality image registration

Finally, and as minor test, Fig. 7.20 shows how it is possible to perform multi-sensor registration (visible image and radiograph, in this case). In order to obtain these compositions, we have applied an epipolar constraint between both sensors, as explained in Section 4.6.3.

Figure 7.20 Two RGB + X-ray image registration examples (for a given value of the ζ parameter). From left to right, 0% of the RGB image is shown, 50% and finally, 100%.



7.6 Conclusions

In this chapter we have shown the results of several experimental tests carried out in real clinical environments.

To test our 3D reconstruction approach, we have run and described several experiences with phantoms that resemble human body parts, such as the head and torso. In these experiments, we have successfully derived 3D data from some embedded lead bearings and metal needles. The results have demonstrated that we can achieve a millimetric level of precision. As Fig. 7.8 and Table 7.2 highlight, millimetric precision is achievable when

comparing the derived lengths against ground-truth. We also obtain a nice agreement between the derived 3D locations and their ground-truth position, as shown in Table 7.1. These results leave the door open for the application of this technique as a complementary tool for other radiological or clinical exams whose main goal is measuring distances or geometrically relating 3D spots whose projections are observable in two radiographs. In a real scenario, these spots can be splinters, cysts, tumors, benign corpuscles, etc., and basically anything worth measuring or locating inside the human body. Our method leaves the identification of these common points to the health professional in both radiographs. The proposed methodology can also have a niche of application as a substitute of CT scans whose principal objective is the extraction of 3D information or as complementary or preliminary tool for monitoring the evolution of a patient's disease over time.

Regarding the generation of densitometric images, we have shown some examples and compared them to the original X-ray images. All density images show a superior quality according the metric presented in Appendix A. We have also shown an extension of previous application related to the assessment of the bone mineral density with fairly good results.

8

Augmented γ radiation detectors

Portable particle detectors lack the methods and resources needed to accurately track the 3D location of radioactive sources. In this chapter we present a new application of the idea of merging particle detectors and environment recognition tools in order to provide this missing information.

More concretely, we summarize the collaboration with the Gamma Spectroscopy group (IFIC) in the project GUALI. In the context of this project, a high-energy photon detector device has been built which consists of a pinhole gamma camera with a monolithic CeBr3 scintillating crystal coupled to a 2D position-sensitive photomultiplier. The system is able to spatially locate radioactive sources thanks to the interplay with

a rigidly attached video camera. This coupled camera system exploits the technology reviewed in Chapter 4 and extends it to a new useful and interesting scenario beyond the realm X-ray imaging.

8.1 Gamma detectors and scene recognition

As commented in the introduction, in this section we examine a new dual-camera system composed of a gamma ray detection unit and an attached RGB sensor. When the information of both cameras is combined, this apparatus enables the determination of the spatial location of radioactive sources. This setup is currently being deployed in the Jose Cabrera nuclear power station (currently being dismantled) as part of the available nuclear waste control toolbox.

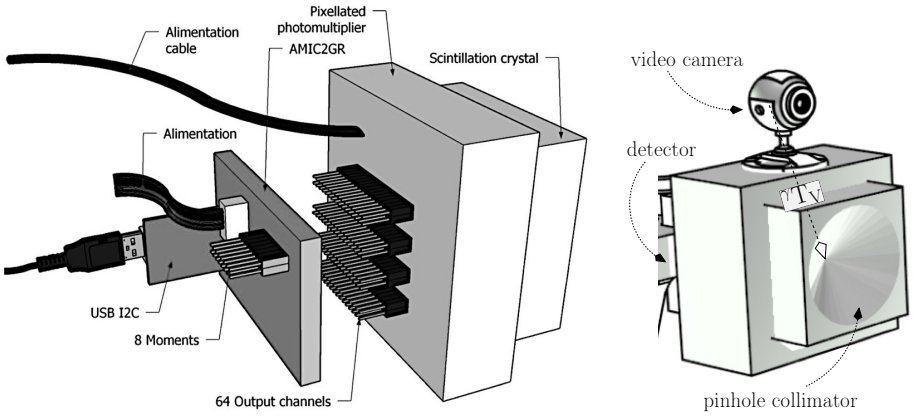
A gamma camera finds itself in the mid path between X-ray equipment and visible devices. On one hand, the intrinsic parameters of a gamma camera do not change over time as the device is moved, just what happens with visible cameras (Fig. 3.2-left). On the other hand, a gamma camera does not collect reflected light, but the directly emitted photons from radioactive sources.

8.1.1 GUALI γ detector overview

In the GUALI project (*Gamma Unit Advanced Locator Imager*), researchers explore for the first time the applicability of using γ-ray imaging in neutron capture measurements to identify and suppress spatially localized background. For this aim, a pinhole gamma camera is assembled, tested and characterized in terms of energy and spatial performance. It consists of a monolithic CeBr3 scintillating crystal coupled to a position-sensitive photomultiplier and readout through an integrated circuit. The pinhole collimator is a massive carven block of lead. In contrast with traditional scintigraphy with γ-cameras and Compton cameras are based on an in-

vasive method: a radioactive source needs to be present inside the object (also other 3D techniques like CT, PET, etc). The original description of this γ detector and its technical/physical foundations are discussed in [73, 140].

Figure 8.1 Left: schematic view of the γ detection system. The pixelated photomultiplier is optically coupled to a monolithic CeBr3 crystal and readout using an integrated circuit AMIC2GR. Right: schematic view of the pinhole lead collimator, the detection system and the geometrical relation between the camera systems.



This γ-ray imager is based on a bulky lead collimator and therefore, its performance for neutron capture measurements is relatively good. However, the rather simple device we use still allows to implement and explore the applicability of gamma imaging under the severe background conditions, such as nuclear power plants (subject tackled in Section 8.2). The GUALI scintillator has been specifically calibrated for detecting photoemissions from Cs-137 and Co-60.

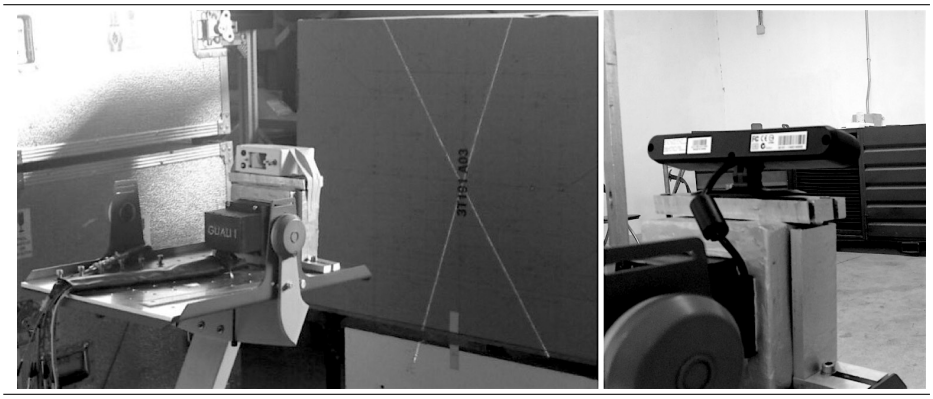
Caesium-137 is a radioactive isotope which is formed from the nuclear fission of Uranium-235 and other fissionable isotopes in nuclear reactors and nuclear weapons. It is among the most problematic of the short-to-medium-lifetime fission products because it easily moves and spreads in nature due to the high water solubility of Caesium's most common chemical compounds. Caesium-137 has a half-life of about 30.17 years. About 95% decays by β emission to a metastable Barium-137m. The main photon peak of Ba-137m is 662 keV, which is the one detected by

GUALI.

Cobalt-60, is a synthetic radioactive isotope of Cobalt with a half-life of ∼5 years. It is produced artificially in nuclear reactors after neutron capture ($^{59}_{27}\mathrm{Co}+\mathrm{n} \longrightarrow {}^{60}_{27}\mathrm{Co} \longrightarrow {}^{60}_{28}\mathrm{Ni}+\mathrm{e}^{-}+\mathrm{v_e}+\gamma$). The main advantage of this isotope is that it is a high intensity γ-ray emitter with a relatively long half-life with energies around 1.3 MeV.

The GUALI detector initially made use of other positioning methods such as a laser range finder, like the one shown in Fig. 8.2-left. In the context of the collaboration with the present thesis project, these devices have been replaced with more modern scene tracking sensors and methods.

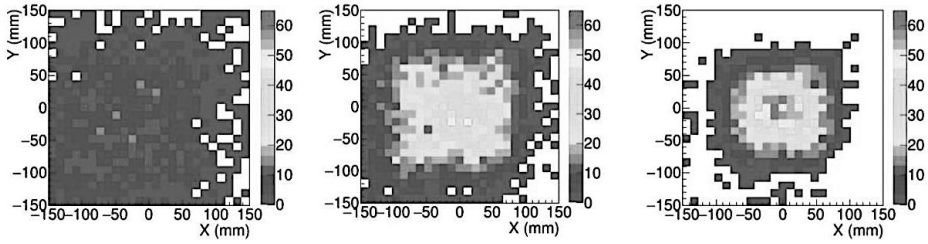
Figure 8.2 GUALI, before (using a laser range finder) and after (using a coupled RGB camera).



8.1.2 Scene identification of γ sources

Given the fact that the presented gamma-camera has a bulky lead collimator acting as a pinhole, projective geometry related tools can be also applied in order to derive the 3D location of radioactive sources from their 2D location in gammagraphic images Υ (Fig. 8.3).

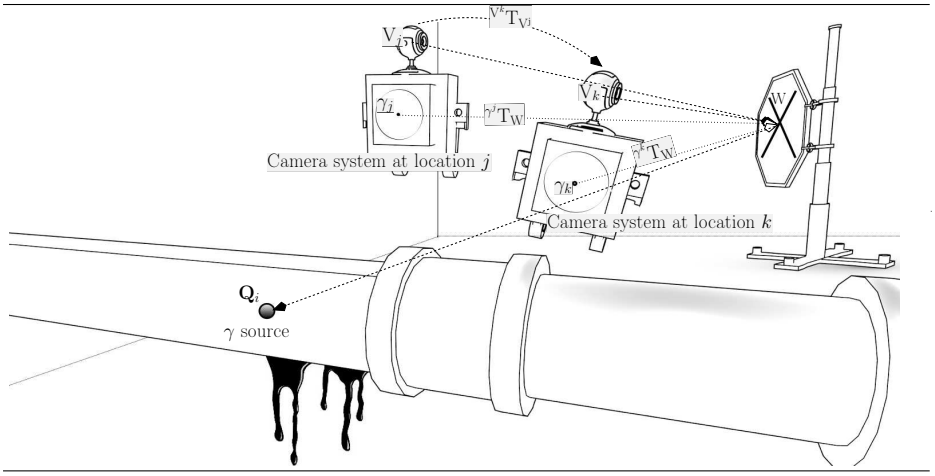
Figure 8.3 Three gammagraphic images (Υ) obtained with GUALI.



Besides, the addition of a RGB camera allows, as already studied in Chapter 4, the 3D identification of each γ source in the scene.

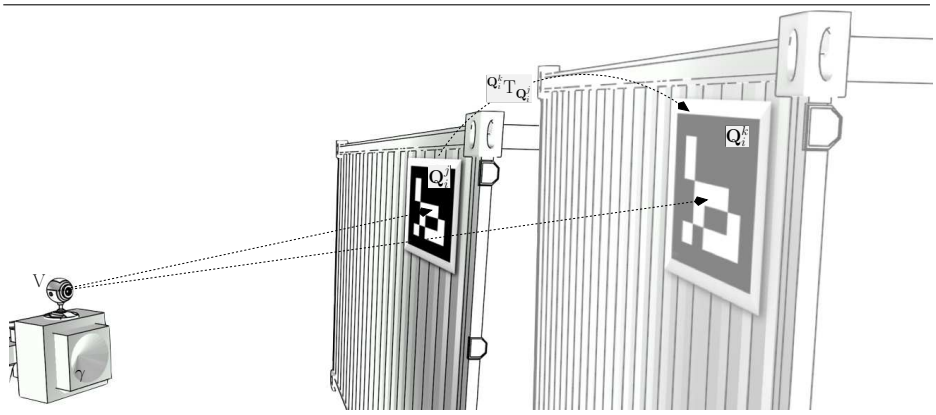
This identification can happen if the camera system (GUALI+RGB) is moving (Fig. 8.4) or if the γ source moves (Fig. 8.5).

Figure 8.4 Example of a *moving detector scenario* (involving a rigid translation T), analog to Fig. 4.16.



In the case of a moving detector (which is the planned scenario for GUALI), it is recommended to establish a fixed (and external) visual reference frame (world).

Figure 8.5 Example of a *moving* γ *source scenario*, analog to Fig. 4.17.



As just stated, the interplay of the proposed γ-camera with and external positioning/reference device enables the derivation of the 3D location of radioactive sources. The adopted device used in this research work is an ordinary RGB or video camera, which, as in the case of X-ray imaging is rigidly attached to the gamma one and whose role in 3D reconstruction is very similar to that performed in the aforementioned field. Basically, this type of cameras, if properly calibrated, can contribute to retrieve the spatial location of a radioactive source $\mathbf{Q}_i$ from the 2D coordinates $\mathbf{q}_j$, $\mathbf{q}_k$, etc., of the detected emission in two different gammagraphic images, following the same mathematical background studied in Section 4.6.2. The main challenge in order to obtain these 3D positions consists in the derivation of the transformation that connects both camera systems ${}^{\gamma}\mathrm{T}_{\mathrm{V}}$ (registration) which is tackled in Section 8.1.3.

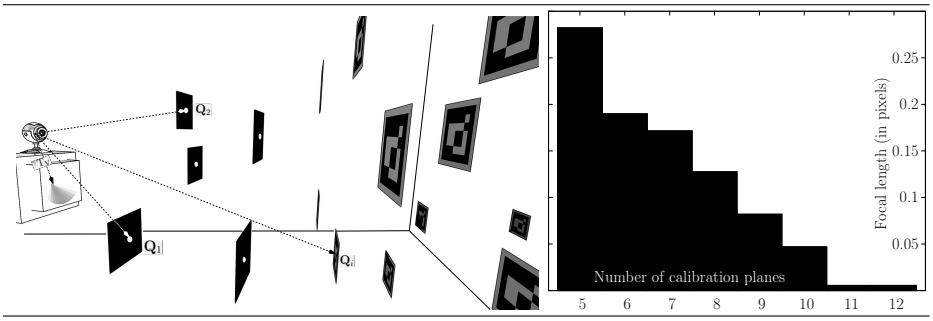
8.1.3 Geometrical calibration of the γ-RGB camera system

As stated above, calibrating the described system entails finding ${}^{\gamma}\mathrm{T}_{\mathrm{V}}$ and the intrinsic parameters of the GUALI γ-camera (K_{γ}). This calibration process follows similar steps to those described in Section 4.5 but with one big difference. The new calibration frame should be much bigger and house both visual and γ-emitter fiducials. The visual part is again configured with the help of the Aruco framework (discussed in Section 4.3.3).

For practical reasons, instead of making use of several and simultaneous γ-emitter fiducials, we add them one-by-one to the calibration space (Fig. 8.6). Again and for easiness, their real location ($\mathbf{Q}_i$) is made coincident with the location of each visual maker, which is in turn easily known thanks to the Aruco framework.

The GUALI detector outputs γ-images (*intensity maps*) like those shown in Fig. 8.3. These intensity maps represent the 3D projection of γ-sources onto each 2D canvas. The highest intensity centers in these maps represent each $\mathbf{q}_i$ location.

Figure 8.6 Graphical summary of the calibration phase (left) and evolution of the computed focal length of the GUALI camera as more 3D $\longleftrightarrow$2D correspondences are used (right).



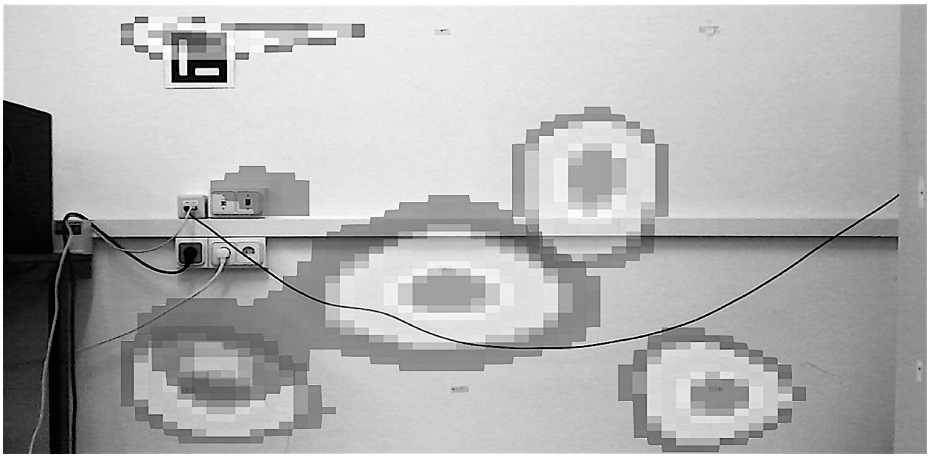
With the collected information, an initial γ-ray projection matrix P_γ is computed. This matrix can in turn be SVD-decomposed as described in Section 8.1.3 in order to obtain the sought geometrical relation:

$$\mathrm{P}_\gamma = \overbrace{\mathrm{K}_\gamma}^{\text{intrinsic matrix}} \cdot \overbrace{{}^{\gamma}\mathrm{T}_\mathrm{V}}^{\text{extrinsic matrix}} \tag{8.1}$$

where the elements of K_γ are similar to those contained in Eq. (3.4).

In Fig. 8.7 we show some photographs taken with the RGB device during the calibration phase.

Figure 8.7 γ calibration points taken during a calibration session and a snapshot of the visible reference frame (Aruco marker) obtained with the attached RGB camera.



The reason for choosing such a different board (compared to the one used in Chapter 4 and shown in Fig. 3.7) is that, in the case of the spatial identification of radioactive sources we do not need a great resolution (a few centimeters is still fine), but on the contrary, a good application range is important (2 to 6 m). Finally, it is not convenient to use several γ sources at the same time, given the high risk they may represent to the operator's health.

8.1.4 3D reconstruction of gamma sources positions

In the case of the 3D reconstruction of γ sources it is important to analyze the emission from as many points of view as possible in order to achieve something similar to a *tomography scan*. The main reason for this is the own radiation shielding coming from the objects being analyzed (mainly, nuclear waste containers). This is in contrast with the methodology followed during X-ray 3D reconstruction explained in previous chapters, where not only two stereoscopic X-ray views are more than enough, but also a reasonable limit to the radiation inflicted on the patient.

In this spirit, there exist a set of techniques known as *multi-view*

stereovision, *bundle adjustment, triangulation* and *structure from motion* (SfM). The 3D information obtained with these techniques is initially applied to the RGB camera of the GUALI device, but can later be translated to the γ camera if both imaging systems are calibrated (as explained in Section 8.1.3).

In our case, we use the *gold standard* reconstruction algorithm, which minimizes the sum of squared errors between the measured and predicted image positions of the 3D point in all views in which they are visible:

$$\mathbf{Q}_i = \arg\min_{\mathbf{Q}_j} \sum_i \left\| \mathbf{q}_i^j - \mathfrak{q}_i^j(\mathrm{P}_i^j, \mathbf{Q}_i) \right\| \tag{8.2}$$

where $\mathbf{q}_i^j$ and $\mathfrak{q}_i^j(\mathrm{P}_i^j, \mathbf{Q}_i)$ are the measured and predicted image positions for point $\mathbf{Q}_i$ in view j. With Eq. (8.2) we can derive the extrinsic matrices for the linked RGB camera at each location where a radiation measurement is to be performed. This camera has been already calibrated (its intrinsic and distortion parameters are known) with the methods discussed in Section 4.3.3. Each of the extrinsic parts can be distilled with the computer vision techniques described in Section 3.2.2.

In our practical implementation of Eq. (8.2), we have used the computer library described in [168]. This framework allows us to perform the mapping and localization from a large set of planar markers (a *marker map*). Internally, the framework runs on a video sequence by first collecting all available observations to create a quiver with the relative poses of the observed markers. Then, an initial pose graph is created that is later refined by distributing the rotation and translational errors around the cycles. Using the initial marker poses from the refined graph, an initial estimation of the frame poses are obtained. Eventually, all poses are refined using a Levenberg-Marquardt optimization to reduce the reprojection error of the marker corners in all observed frames and thus solve Eq. (8.2). The proposed optimization function ensures that the markers geometry is kept during the optimization process. Each marker in the marker map can be, for instance, an Aruco pattern.

The next step consists in translating the images obtained with the gamma camera to the *space* or point of view of the visible camera, so that we can apply the results of the gold standard method discussed above and derive the 3D location of γ sources. The movement from one position to

another is identified by computing changes in the pose relative to a specific maker manually tagged as the world (W).

8.1.5 Registration of RGB and γ images

In contrast with what we discussed in Section 4.6.3 and Section 7.5, the visual registration of gammagraphic and video images is very important because it neatly shows the rough position of radioactive sources inside nuclear waste containers or within the inspected the scene as a whole. In order to geometrically register the information of the two stereo cameras (RGB and γ), we have to apply again what we learnt in Section 3.5.3 and specifically in Eq. (4.16), but re-expressed for a different paired imaging system composed of a γ and a RGB (V) device:

$$\hat{\mathbf{Q}}_i = ({}^{\gamma}\mathrm{T}_{\mathrm{V}})^{-1} \cdot (\mathrm{K}_{\gamma})^{-1} \cdot \hat{\mathbf{q}}_i^{\gamma} \tag{8.3}$$

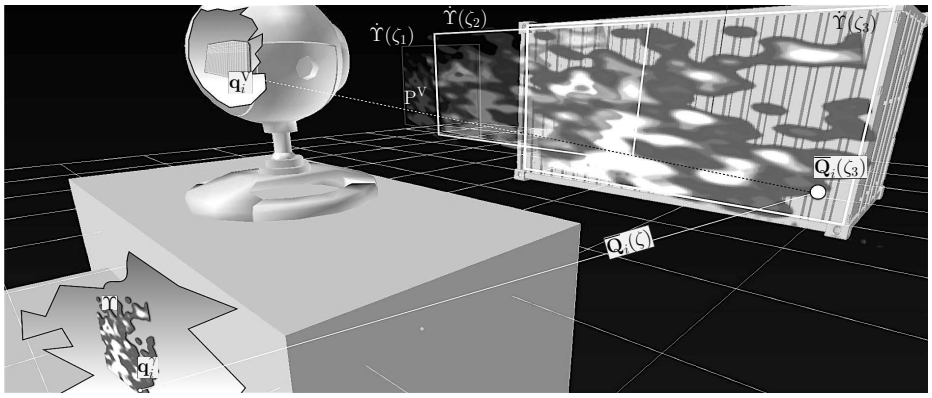
where $\mathbf{q}_i^{\gamma}$ is a point i in the original gammagraphic image Υ and $\mathbf{Q}_i$ in world (W) coordinates. We then have to find the ray equation that connects the center of the γ camera (also expressed in the W reference frame) and $\mathbf{Q}_i$, given by:

$$\overline{\mathbf{Q}}_i(\zeta) = ({}^{\gamma}\mathrm{T}_{\mathrm{W}})^{-1} \cdot (0,\, 0,\, 0,\, 1)^{\intercal} + \zeta \cdot \hat{\mathbf{Q}}_i \tag{8.4}$$

where ζ is a scalar (manually chosen) representing the coordinate along the ray. A portrayal of this ray is more or less shown in Fig. 3.18 for the case of X-ray imaging system, but can be easily transferred to the present gamma measurement scenario.

Next, a specific 3D point in $\overline{\mathbf{Q}}_i(\zeta_m)$ can be re-projected onto the visible image with Eq. (4.1), obtaining a registered 2D location $\mathbf{q}_i^{\mathrm{V}}(\zeta_m)$, which also depends on the selected value for ζ. In order to obtain the final registered image ($\dot{\Upsilon}(\zeta_m)$), we just have to iterate over all the i points in the $\mathbf{q}_i^{\mathrm{V}}(\zeta_m)$ set and apply Eq. (8.3) and Eq. (8.4) for a shared (and manually chosen) depth ζ. This scenario is schematically represented in Fig. 8.8 and an example of RGB-γ registration is shown in Fig. 8.14.

Figure 8.8 Registration of γ and RGB images.



8.2 Field tests

Several experiences have been carried out in the José Cabrera (Zorita) nuclear power station (Fig. 8.9). The tests and experiences consisted on measuring several containers and wall pieces which were contaminated with traces of radioisotopes during the time of operation of the aforementioned facility.

8.2.1 The Zorita nuclear power station

The Zorita plant is located in the municipality of Almonacid de Zorita (Guadalajara). It started operating in 1968 and ended 38 years later, on 30th April 2006. Enresa (*Empresa Nacional de Residuos Radiactivos S.A.*) became the license of the facility the nuclear power station in 2010 and is in charge of the safe management of the radwaste and of the restoration of the sites where nuclear and uranium mining activities have taken place. The Spanish Parliament created Enresa in 1984 as a public, non-profit organization responsible for the management of radioactive waste.

The dismantling operation yields more than 100.000 tons [69]. The work performed also includes the segmentation of the reactor internals. The site will be completely cleared only after the low and medium level waste will be transported to the disposal center in El Cabril (Córdoba,

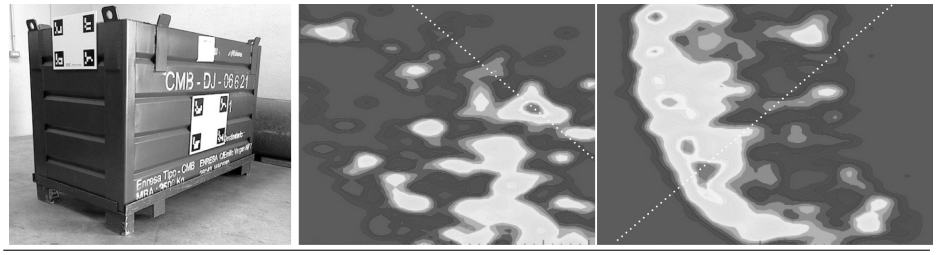
Spain). The dismantling operation also requires all the components of the primary circuit, including the plant reactor, to be segmented and a large part of the resulting materials to be classified and placed in special containers (like the one shown in Fig. 8.10-left).

Figure 8.9 The José Cabrera Nuclear Power Station.

8.2.2 Examined nuclear waste in Zorita

The examined elements in Jose Cabrera consisted in radwaste containers and concrete walls. In the case of the examined containers, two γ/RGB snapshots were obtained and the same stereo/epipolar geometry tools tackled in Section 3.5.3 was used. In Fig. 8.10 we represent the application of these techniques to locate the 3D position of a radioactive element inside the container.

Figure 8.10 3D reconstruction of radioactive source in a waste container with epipolar geometry.

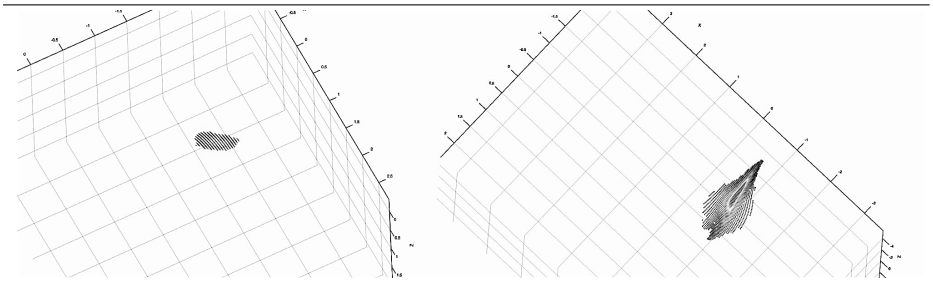


In later visits to José Cabrera, several concrete blocks were examined, like the ones shown in Fig. 8.11 and Fig. 8.13.

Figure 8.11 Contaminated concrete blocks being inspected.

In this occasion, we used the framework presented in Section 8.1.4 to obtain the extrinsic parameters, i.e., the relative pose between the GUALI instrument and each block, as introduced in Section 3.2.2. We have been able to reconstruct the 3D locations of Cs-137 and Co-60 sources and match the position already roughly diagnosed by staff at Enresa. In Fig. 8.12 we show point clouds of the situation of two γ sources and Fig. 8.13-right shows a picturesque 3D reconstruction with the rough location and intensity shape of a Cs-137 source inside an irregular piece.

Figure 8.12 Reconstruction of the 3D positions of a Cs-137 (left) and C0-60 sources.

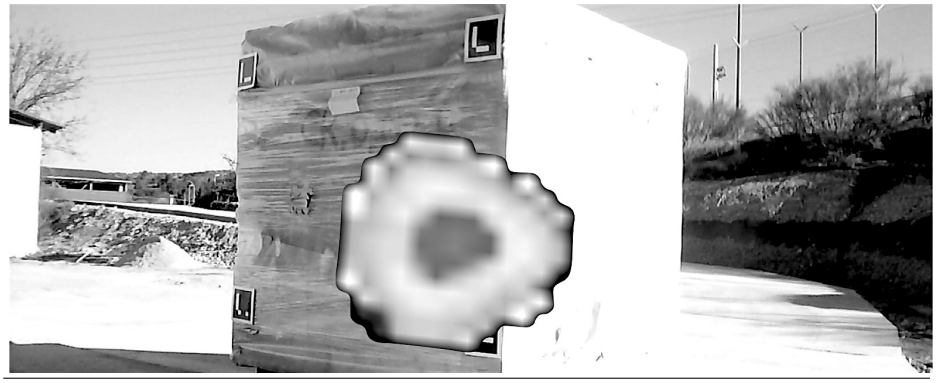


In Fig. 8.13 we show some elements whose geometry is being successfully tracked in real-time with a map of markers, as also described in Section 8.1.4. These markers and their location are identified with the RGB camera.

Figure 8.13 Irregular concrete piece and a waste container tracked with a marker map composed of visible fiducials.

Finally, Fig. 8.14 shows a γ+RGB registered pair of images using the mathematical background described in Section 8.1.5.

Figure 8.14 Registration of a gammagraphic image $\dot{\Upsilon}(\lambda)$ (obtained from a Cs-137 contaminated concrete block) and a visible/RGB one, for a given value of the λ parameter (depth).

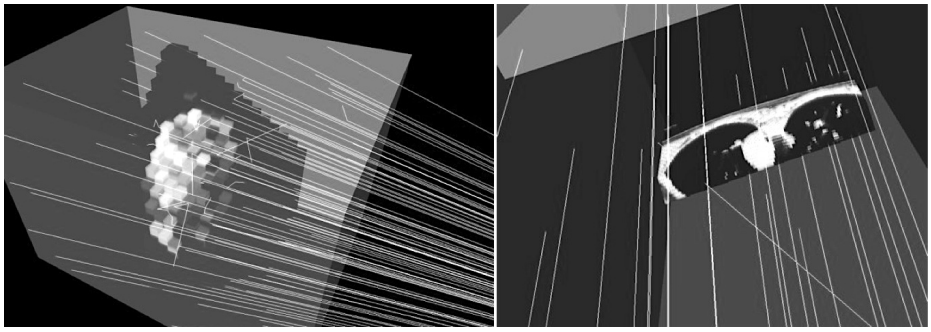


8.3 Conclusions

In this Chapter we have transferred the explored technologies in previous chapters to a new area of application: nuclear waste management. In contrast with X-rays and external positioning devices, in this field we interplay a similar video camera and a gamma ray detector optimized for the tracking of photons originated after the decay of the radioisotopes Co-60 and Cs-137. The results show it is possible to track and assign 3D information to the location of γ sources with a very good level of precision for this type of scenarios. It is also possible to seamlessly overlay gammagraphic and visible images, which turns out very useful during control inspections.

9

Conclusions and outlook



In this research work we have introduced a number of methods and techniques around the interplay of external scene determination tools and radiation detection setups, with a special focus on conventional X-ray diagnostic equipment, given their widespread use and presence. These tools have allowed us to derive accurate 3D information from plain absorption and/or intensity images (radiographs, gammagraphies, etc.) and perform precise 3D anatomical/inner reconstructions. The same techniques have also been applied to the definition of a new transfer function and the generation of X-ray densitometric images from absorption ones. These *augmented* images offer greater quality over plain radiographs (as demonstrated by a new image quality metric also discussed later in the Appendix A of this work). A similar approach has been applied to gammagraphic images generated with a γ camera in the process of measuring

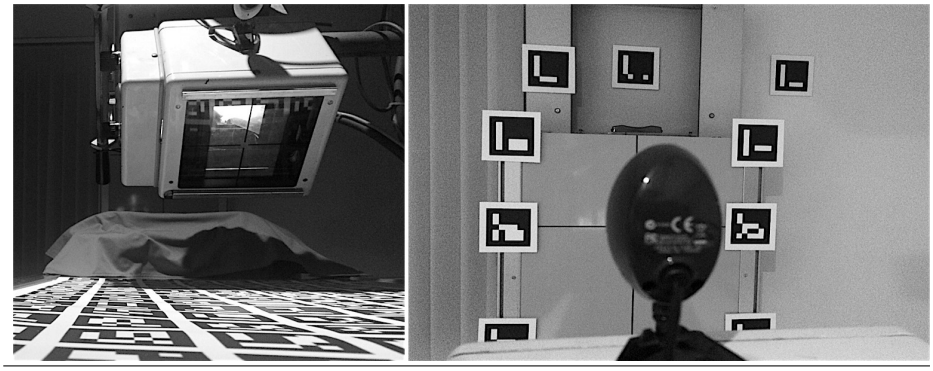
radioactive sources. In general, all the presented techniques can be applied in other scenarios that comprise radiation detectors and the lack of spatial information.

In spite of all these proven advantages and their suitability in many fields of application, there is still work to do in some areas, which are summarized below.

Enhancement of patient and X-ray imaging system tracking

Regarding the combination of RGB cameras, visible fiducials and X-ray environments (as discussed in Chapter 4), we can envision many other possibilities of their application in real clinical environments (Fig. 9.1).

Figure 9.1 Other possibilities of combining visual fiducials in X-ray rooms for both horizontal and vertical positioning protocols.



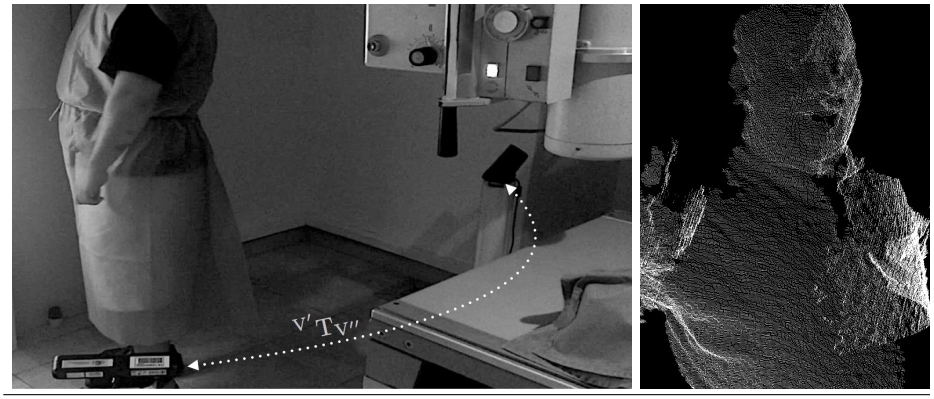
These combinations would specifically contribute to better track the position of the X-ray imaging system during the generation of radiographs.

Regarding the interplay of depth sensors and X-rays, the main current drawback is the need for patient's rotation around his/her craniocaudal axis (as described in Section 5.1 and Section 6.1), which inherently goes

against the patient's comfort. We foresee new methodologies involving the fusion of several calibrated Kinect-like devices (*a priori* just three:V, V′ and V″) that together would fully map the X-ray room, including the patient, who would now remain comfortable still in one of the endorsed protocols discussed in Section 2.7.1. The transformations connecting each depth camera (${}^{V}T_{V'}$, ${}^{V'}T_{V''}$ and ${}^{V}T_{V''}$,) can be found with a visible marker board similar to those already used in this research (Fig. 3.7) or with an extra geometrical calibration object whose 3D configuration is partially visible from the point of view of all three cameras (like the rotating cube described in Fig. 5.7). This 3D frame would enable the application of the iterative closest point (ICP) algorithm in its simplest version in order to find the aforementioned transformations. In the ICP algorithm, one point cloud is kept fixed while the other one is transformed to best match the former. The algorithm iteratively revises the transformation needed to minimize an error metric. Researchers in [71, 31] propose more refined methods for the calibration and registration of depth sensors into a joint coordinate system.

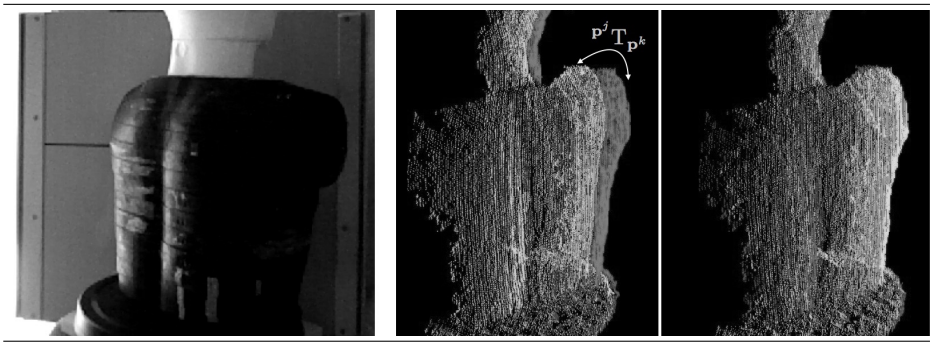
The composition of all point clouds (captured with each device) would be achieved through the synchronous application of Eq. (6.1). Background subtraction would be done using the methods described in Section 5.5 or through a spatial tree-based data structure search on *unorganized* point cloud data [263]. This last case is more appropriate when the background and patient's point clouds are *blended*, which represents a more realistic scenario. Unorganized point clouds are characterized by non-existing point references between points from different point clouds due to varying size, resolution, density and/or point ordering. Once the patient's point cloud has been resolved (and isolated from the background), a surface mesh can be found with legacy methods like marching cubes, Poisson, etc., or the ones summarized at the beginning of Section 6.1, as well as many other modern reconstruction algorithms [44, 101, 151, 17]. An example of these ideas is represented in Fig. 9.2.

Figure 9.2 Combination of several RGB-D sensors in a radiographic scenario and geometrical stitching of two point clouds acquired with two calibrated depth sensors. Three registered sensors would be necessary to capture the whole patient's volume.



In scenarios where patient rotation is applicable, but no GPU hardware is available (and thus, the KinectFusion method described in Section 5.5 is not usable), we are considering the use of an offline pairwise incremental registration process. This technique adopts again the iterative closest point algorithm in order to incrementally register a series of point clouds two by two. It can be time consuming, but can run on very humble computer equipment. This registration entails finding the rigid transformation $({}_{\mathbf{p}^j}\mathrm{T}_{\mathbf{p}^k})$ for the same point $\mathbf{p}$ between each two patient positions j and k, k and l, l and m, iteratively and as shown in Fig. 9.3.

Figure 9.3 Pairwise iterative closest point phase between two point clouds of the Rando phantom. The image on the right shows the successful registration of the two point clouds represented by $\mathbf{p}^j$ and $\mathbf{p}^k$.



Volume-enhanced X-ray images

Although the superiority in quality of densitometric images has been proven through the several examples shown in Chapter 7 and the metric defined in Appendix A, it is necessary to further extend the set of images, patients, examined areas and applied X-ray protocols (different kVp, mAs, patient positioning, etc.). In addition, the methods described in Section 6.6 should be further explored both clinically and theoretically in order to fully present the concept of densitometric imaging as a complete alternate framework to dual X-ray imaging.

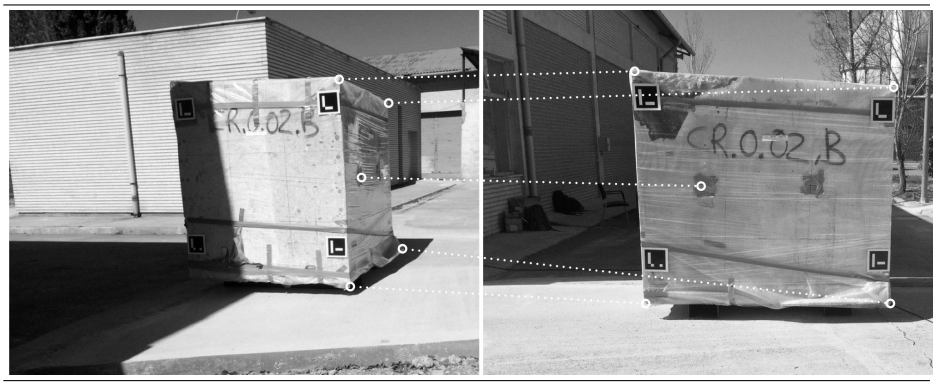
Regarding the generation bone-only images, we think we can achieve similar results to those discussed in [33], where authors present great quality versions of this type of images by applying a synthetic background subtraction representing soft tissues and air, which is very similar to our approach described in Section 6.6. However, their research still needs two radiographs (high and low energy) to be generated.

Enhancement of setting identification with gamma cameras

Finally, the methods tackled in Section 8.1.4 could be further simplified with the application of other modern computer vision techniques, such as RANSAC. The *random sample consensus* is an iterative feature-based image matching algorithm that finds the transformation that best transforms the first image to the second. By finding this matrix, we would be accounting for the same mathematical relations needed in Eq. (4.8), Eq. (4.10) and Eq. (5.2). This would entail the expendability of added reference markers such as the Aruco ones described in Section 4.3.3.

In Fig. 9.4 we show and example of application of this algorithm with the same contaminated concrete block studied in Section 8.2. In this context, this proposed method behaves very similarly to the fact expressed

Figure 9.4 The RANSAC method applied to the reconstruction of a 3D scene in a nuclear waste management scenario.



in Section 5.8: *the scene is its own fiducial system.*

Volume and dose assessment

Although an active field of research and practice in medicine and radiology, there is still room for improvement in the realm of dose assessment in ordinary X-ray imaging. In the following paragraphs we review its current status and discuss possible enhancements based on the methods and materials presented in this research work. Radiation protection and dose assessment in conventional X-ray is a subject of active worldwide research and concern. In Fig. 9.5 we show, for instance, some geolocalized publications on the subject.

The common final goal of these researchers, healthcare institutions, interest groups and radiation protection professionals is the same: keep dose *as low as reasonably achievable* (ALARA criterion) and measure this dose as accurately as possible. It is well known that patient care units tend to abuse this type of preventive imaging procedures, even when they rarely reveal clinically unsuspected findings [251].

Patient doses in radiography examinations are mainly derived from

the *entrance surface dose* (ESD) or *dose-area product* (DAP). ESD is computed either by direct measurements, normally sticking *thermoluminescent dosimeters* (TLDs) to patients [224] or physical phantoms [54]. ESD can also be computed after its relation to the *backscatter* factor along with known tube parameters [78, 188]. Together with ESD, mathematical models (including phantoms like MIRD5 [147], NCAT [220], and Cristy [59]), voxelized CT scans [148, 261, 65, 75, 95]), and tabulated recommendations [46, 191] are also commonly used. This models usually take X-ray tube type, anode material and available filtrations as input parameters. Measurements/simulations are made at the center of the X-ray beam with a fixed field size.

The final patient *effective dose* is proportional to the estimated ESD, depending also on the X-ray penetrating power and the region being examined. There exist international regulations that tackle this physical derivation, like ICRP-103 [108] and ICRP-60 [107]. *Global effective* dose is, in turn, calculated as a weighted sum of the effective doses on each organ. In [184] and [191] is included recent tabulated calculations of organ doses from a variety of diagnostic radiology procedures. Effective dose needs knowledge of the dose of the body and then the weighted sum of the equivalent doses to 22 organs and to each of the tissues exposed. This is complicated when parts of the body are exposed heterogeneously as in the case of diagnostic radiology and is therefore more difficult to apply routinely.

By simulation of a limited number of photons in an X-ray field, a mean value of the absorbed energy in a specific organ and the global effective dose to a patient can be estimated. The well-known relation is represented by the equation:

$$E_{\text{eff}} = \sum_R W_R \sum_T W_T \cdot D_{T,R} \tag{9.1}$$

where $D_{T,R}$ is the absorbed dose averaged over the organ or tissue T, due to radiation of type R and W_R and W_T are the radiation and tissue *weighting factors* respectively.

In order to automate these parameter-filling and calculations, specific software packages are available. They are almost permanently used and represent a *de facto* standard regarding dose assessment. A few examples are: Caldose [127], PCXMC [238], Orgdose [183], BEAMnrcMP [206],

DoseCal [66], WinDose [116], MCNPX [258] and CALDose_X [128].

The same concern about correct dose assessment is present in the CT modality scenario, where radiation protection professionals face a similar dilemma. Initially, the *CT dose index* or CTDI [2] was developed to provide a standardized method to compare radiation levels between different CT scanners using a reference phantom. It thus provides information regarding only the scanner output: it does not address patient size, and hence does not quite estimate patient dose. This is the reason why the *volume computed tomography dose index* or CTDI_{vol} [150] was defined. Almost all current CT scanners display CTDI_{vol}.

Our proposal for the improvement of the procedure of assessing the patient's dose in conventional X-ray settings is based on the *dose-per-volume* concept. Thanks to the methods devised in this thesis work, we can now accurately measure the patient's volume and assign a precise position in the radiographic scene. Once this geometry has been determined, the dose can be assessed by comparing it with tabulated values. These tabulated values have to be previously obtained by radiating a phantom filled with TLDs with different configurations (energy, exposure time, distance from source, materials, etc.).

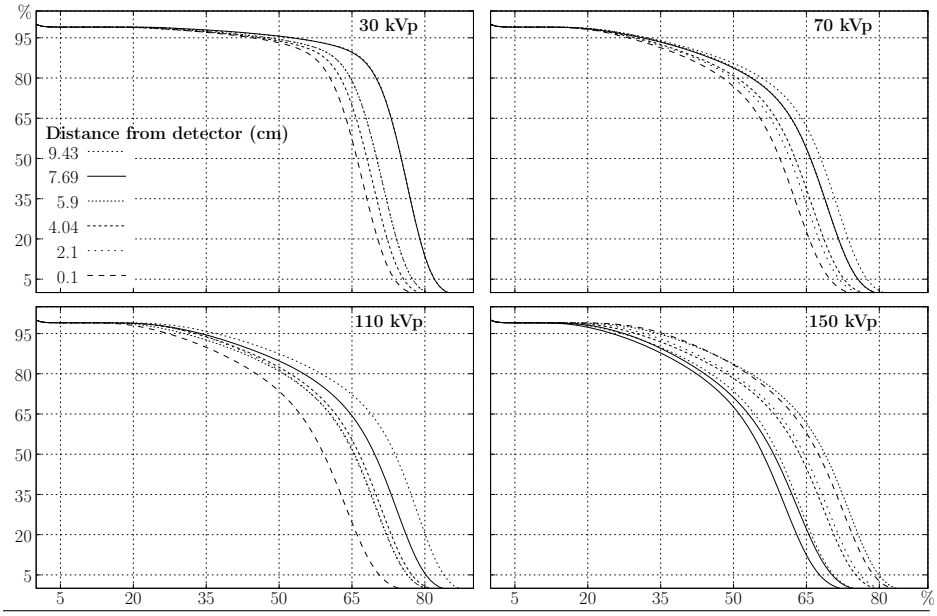
The dose received by a patient during an X-ray test can be very geometry sensitive. In the two following computer simulations, we demonstrate how the absorbed energy can be very position-dependent, even for small spatial changes.

The first simulation is based on the Monte Carlo approach [18], while the second is *determinist* (or analytic) and its based on ray-tracing. This Monte Carlo simulation has been carried Geant4 (GEometry ANd Tracking) is a platform for simulating the passage of particles through matter using Monte Carlo methods. The Monte Carlo method consists of individually tracking each photon during its different interactions with matter at each step of the simulation. This method can produce very accurate images, but they are computationally expensive. We have simulated a cuboid and a beam of X-rays very similar to the previous Section. However, this time the cuboid is translated along the transversal axis, modifying the OID factor. In Fig. 9.6 we show the dose volume histogram (DVH) of the dose received by the virtual water phantom. A DVH graphically summarizes the simulated radiation distribution within the volume of a

patient or phantom, originally as a result of a radiation treatment plan. In other words, the goal of DVHs is to summarize 3D dose distributions in a graphical 2D chart format.

This time, a polychromatic beam has been used, previously generated with the Spektr [229] package (other software packages such as SpekCalc [193], are also available). Spektr was designed to provide a flexible, extensible tool for calculation of X-ray spectra, application of X-ray filters, and analysis of spectral characteristics. Four different kVp configurations have been emulated (30, 70, 110 and 150 kVp) in order to reflect as closely as possible common radiographic techniques.

Figure 9.6 Dose-volume histograms derived with the Monte Carlo simulation.

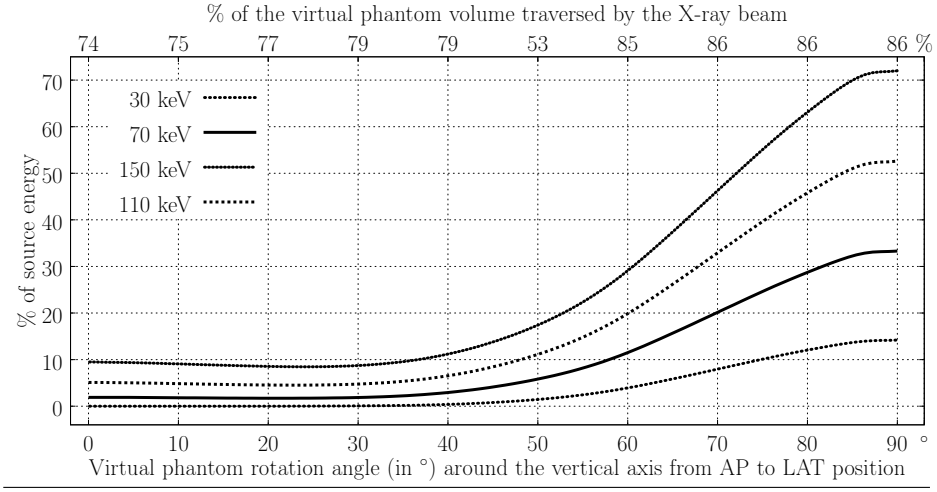


The Monte Carlo simulation shows that little variations in the OID parameter entails non-negligible changes in the computed DVHs in the water phantom. This is again another example of how the spatial environment can severely affect the dose and deposited energy in the patient being radiographed.

We have also carried out a deterministic simulation of the energy

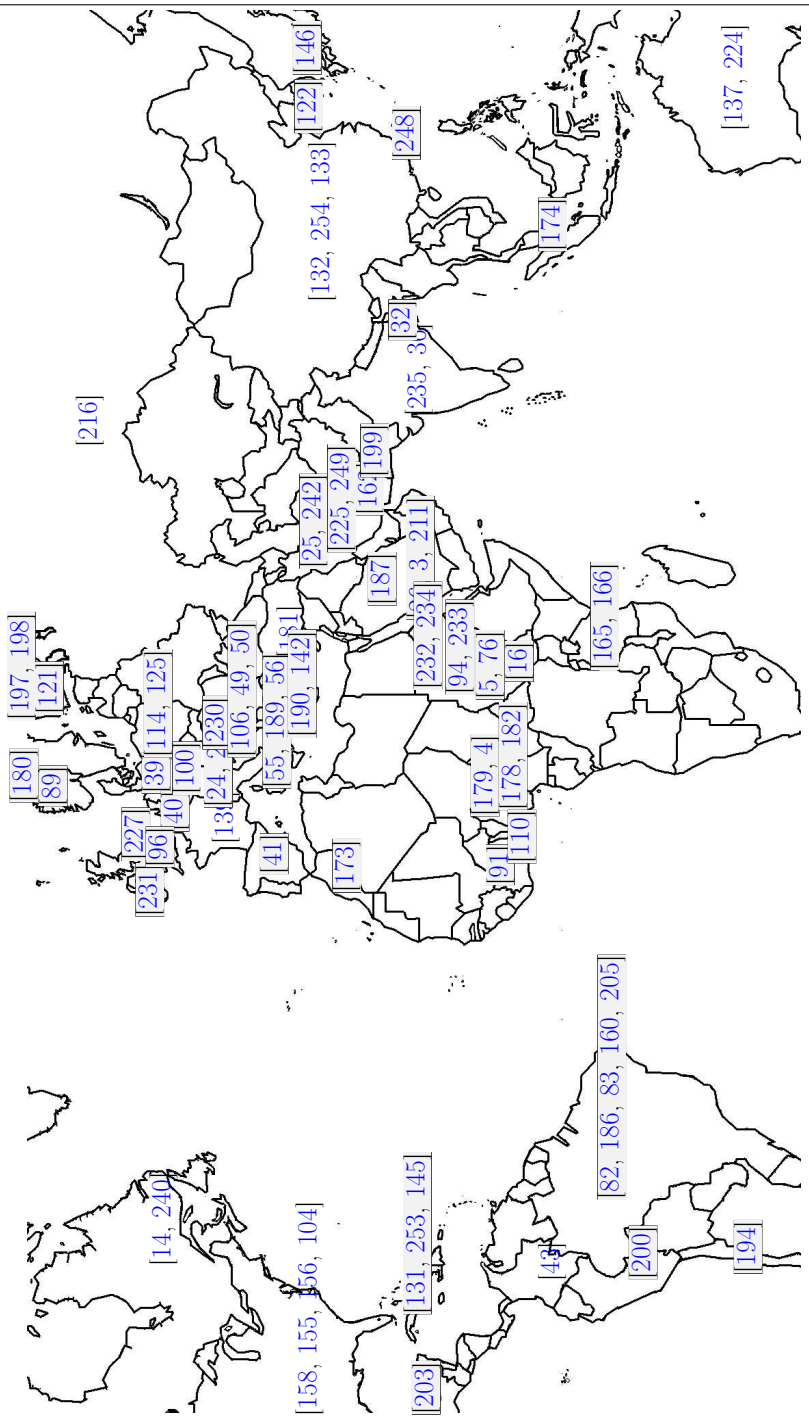
received by a rotating cuboid water phantom in a virtual X-ray system. The deterministic simulation uses a virtual rotating water cuboid. This geometrical phantom is *beamed* at several angles by several monochromatic X-ray sources (30, 70, 150 and 110 keV) and the % of the source energy is computed for each angle of rotation. The results are shown in Fig. 9.7.

Figure 9.7 Results of the deterministic simulation carried out with a rotating water phantom. The chart on the top shows the % of the source energy that is deposited in the virtual cuboid. The geometry and distances are very similar to those shown in Fig. 9.6.



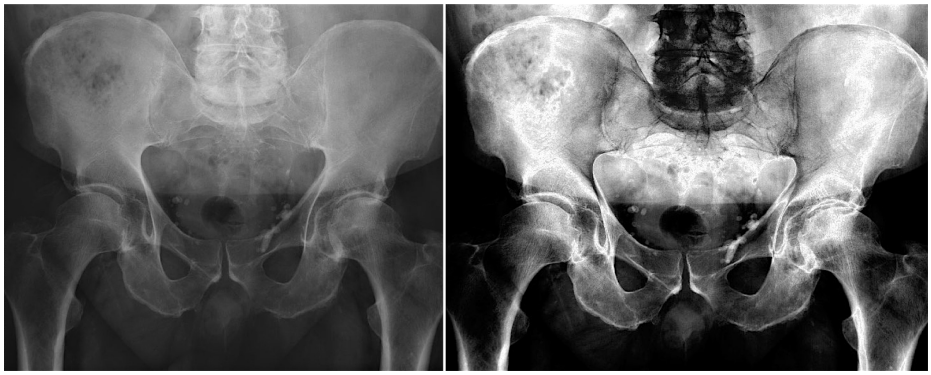
Clearly, as the results of the simulation show, above 40°, small angle variations may have important repercussions in the deposited energy and hence, even small changes in the patient positioning may have a direct impact in the final received dose.

Figure 9.5 Research works related to dose assessment in X-rays.



A

Image quality assessment in X-ray imaging



In this appendix we introduce a genuine methodology to assess the quality of radiological images which will help us to objectively compare *volume enhanced radiographies* (presented in Chapter 6) against conventional ones. This method performs an intensity windowing operation, which is in turn based on the maximization of the mutual information (MI) between a perceptual decomposition of the original 12-bit sources and their screen displayed 8-bit version. This methodology is based on a human visual system algorithm. Its primary goal is to measure the quality of a radiological image (i.e., a radiograph in our specific case) based on the MI between the original instance and its Gabor-filtered derivations. In order to do so, the

algorithm performs an automatic intensity windowing process that seeks the WL/WW that best displays each image for examination. It starts with the default, high contrast, wide dynamic range 12-bit data, and then maximizes the graphical information presented in ordinary 8-bit displays. This quality assessment method is an important tool used in Chapter 7. Defining an objective quality assessment method is essential in order to attain experiment reproducibility and diagnostic repeatability. The proposed quality assessment method consists of an interplay of the concepts of Mutual Information (MI), Human Visual System (HVS), Gabor filtering and linear contrast manipulation. The method begins by computing the MI between the Gabor-filtered representations of both the original and the displayed images. Then, the WL/WW combination that makes this set of MI values maximum is iteratively sought. The suitability of this methodology and its advantages has been tested in [7] which describes a set of experiments involving the participation of a panel of radiologists.

A.1 Basic contrast manipulation in radiological images

Linear contrast adjustment (which is equivalent to establishing a window level and width) represents a basic operation that is performed in everyday screening and in almost all radiological disciplines. Its foundations are graphically represented in Fig. A.1 and in the following equation:

$$j = j(i,a,b) = \begin{cases} 0 & \forall \quad i < a \\ 255 \times \frac{i-a}{b-a} & \forall \quad a \leq i \leq b \\ 255 & \forall \quad b < i \end{cases} \tag{A.1}$$

where i and j account for the intensities of the original 12-bit ($\mathcal{I}$) and the contrast-stretched 8-bit images ($\tilde{\mathcal{I}}$), respectively. From Eq. (A.1), it is easy to derive that a basic IW operation renders the lowest intensity pixels of $\mathcal{I}$ equal to black ($i = 0$) and the highest intensity ones equal to white ($i = 255$). These intensity thresholds are determined by the a and b parameters,

respectively, which are usually manually established and modified by the health professional. However, they can also be automatically determined or predefined by a group of presets (i.e., to highlight certain types of tissues or densities).

Figure A.1 Top: histogram and functional representation of a basic image contrast operation. As an image histogram, intensities below and above the a and b thresholds are cumulatively transferred from the original image $\mathcal{I}$ to the 0 and 255 bins in the 8-bit windowed version $\tilde{\mathcal{I}}$, respectively. As an intensity transform function, an intensity value of j in $\tilde{\mathcal{I}}$ is the result of applying Eq. (A.1) to an intensity level i in $\mathcal{I}$. Bottom: an example of a raw radiographic image with low contrast and narrow dynamic range and next to it, the result of applying an appropriate IW operation.

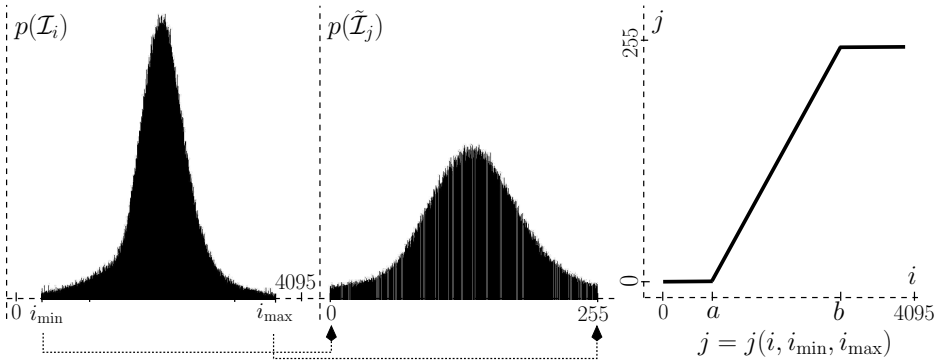
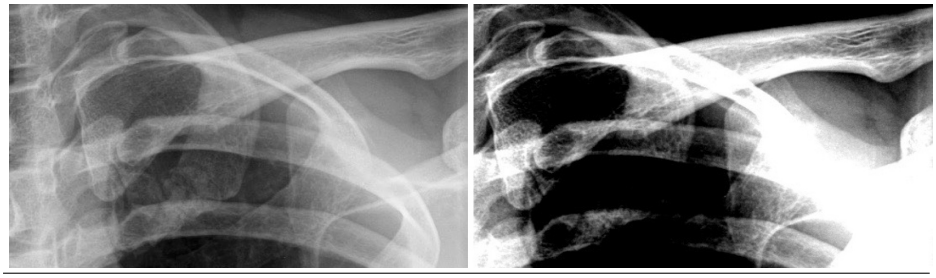


Figure A.2 Quality of an X-ray image as a function of the contrast settings (WL/WW = -2000/3000 for the image on the left and WL/WW = -2000/800 for the image on the right).



Depending on the chosen contrast settings, undoubtedly the quality

of an image will vary accordingly, as shown in Fig. A.2.

A.2 Traditional information theory-based image quality assessment

Some of the earliest methods applied to image quality determination are the Peak Signal to Noise Ratio (PSNR) and Mean Square Error (MSE) which have been widely used in the context of image coding due to their easy implementation. Another commonly used quality metrics related to information theory are entropy and MI.

Figure A.3 Example of the application of PSNR. From left to right: original image and three instances with the same of PSNR equal to 225, which are a proof of the fact that this method can be easily fooled in the case of image quality assessment.

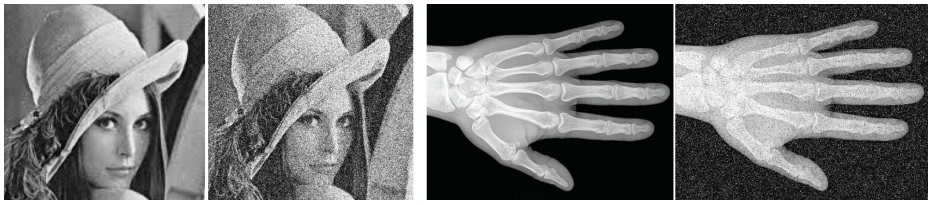
Intuitively, we can consider that a displayed image has the *highest quality* when the maximum amount of relevant information from the source image is preserved. Thus, the concept of entropy [124] naturally emerges. Shannon borrowed this concept from physics as a measure of the amount of *uncertainty* of an information source. Entropy has been used in many image processing applications such as spatial registration [77]. In the case of digital images, entropy is computed as:

$$H(\mathcal{I}) = -\sum_{i=0}^{4095} p(\mathcal{I}_i) \cdot \log_2(p(\mathcal{I}_i)) \tag{A.2}$$

where $p(\mathcal{I}_i)$ is the probability distribution of the pixel intensities of image

$\mathcal{I}$. With Eq. (A.2) we measure the sharpness and shape of the image histogram, which is indirectly related to the image texture. For instance,

Figure A.4 Example of added white noise to induce false entropy.



images with large homogeneous areas have a very peaked histogram (low entropy). On the other hand, images which rich textures have a lot of contrast and a flatter histogram, and thus a higher entropy. Although image entropy has been regularly used for bitmap quality quantification, it has also been proven to be easily fooled. For instance, the entropy of an image can be easily manipulated by simply adding white noise to it, as Fig. A.4 shows.

Mutual information [246] can also play a role in image quality assessment [226]. MI measures (in bits) the reciprocal dependence between two variables. In other words, it quantifies the information obtained from the first variable through the second one. In mammography screening, MI has traditionally been used for image registration [260] and diagnosis through template matching [244, 243]. MI can be mathematically expressed as:

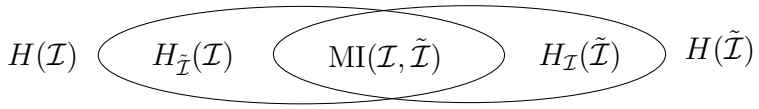
$$\mathrm{MI}(\mathcal{I},\tilde{\mathcal{I}}) = H(\mathcal{I}) + H(\tilde{\mathcal{I}}) - H(\mathcal{I},\tilde{\mathcal{I}}) \tag{A.3}$$

where $H(\mathcal{I})$ and $H(\tilde{\mathcal{I}})$ are the correspondent proper entropies of each image. The term $H(\mathcal{I},\tilde{\mathcal{I}})$ is defined as the joint entropy, which can be estimated with the 4096×256 bi-dimensional histogram of the intensities of $\mathcal{I}$ and $\tilde{\mathcal{I}}$:

$$H(\mathcal{I},\tilde{\mathcal{I}}) = -\sum_{i=0}^{4095}\sum_{j=0}^{255} p(\mathcal{I}_i \cap \tilde{\mathcal{I}}_j) \cdot \log_2(p(\mathcal{I}_i \cap \tilde{\mathcal{I}}_j)) \tag{A.4}$$

where $p(\mathcal{I}_i \cap \tilde{\mathcal{I}}_j) = p(\mathcal{I}_i) \cdot p(\tilde{\mathcal{I}}_j \mid \mathcal{I}_i)$ is the probability that corresponding pixels in $\mathcal{I}$ and $\tilde{\mathcal{I}}$ have intensities i and j, respectively. An easy and

Figure A.5 Venn diagram showing the relation among each image entropy, conditional entropies $H_{\tilde{\mathcal{I}}}(\mathcal{I})$, $H_{\mathcal{I}}(\tilde{\mathcal{I}})$, and MI.



very common way of understanding the relation between these information entities is by a Venn diagram, which is shown in Fig. A.5.

The diagram makes use of the term *conditional entropies* $H_{\tilde{\mathcal{I}}}(\mathcal{I})$ and $H_{\mathcal{I}}(\tilde{\mathcal{I}})$. The conditional entropies reflect the part of information in one image that cannot be explained when the other image is known. Following the previous example, if we add white noise to an image, its entropy grows because the conditional entropy is higher, but the MI between the original and the *corrupted* version remains the same.

Unfortunately, none of the aforementioned methods can measure by themselves the quality perceived by a human observer [213]. In contrast, techniques based on HVS have shown a better performance in image quality estimation [257, 48]. Even information-based quality assessment methods like entropy and MI assume that the image pixels are statistically independent (which is obviously a wrong principle) and they do not take into account how the visual cortex and human perception work.

In order to address this limitation, a new image quality metric that complements MI with a HVS approach is proposed in Appendix A.3.1. This metric, based on Gabor filters and linear contrast manipulation, is significantly more faithful and consistent with the quality perceived by the human brain.

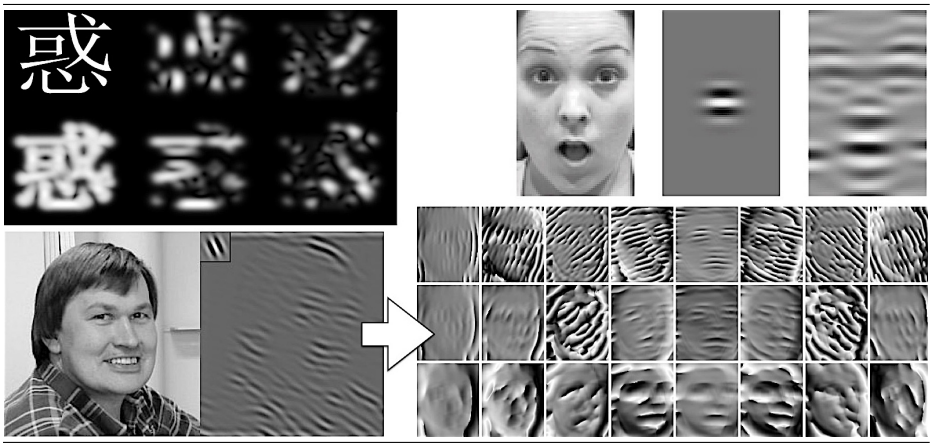
A.3 A human visual system-based quality assessment approach

In this section, we present the theoretical background related to the proposed HVS-based image quality assessment method.

A.3.1 An image quality metric based on mutual information and Gabor filtering

Gabor filters are an excellent tool for texture analysis of images. In short, the responses of Gabor filters [81] correspond to those of single cells in the visual cortex. These cells extract contours and directional patterns. Gabor filters are commonly grouped in *banks* where each filter *captures* the image information in the vicinity of a frequency and at a specific direction (Fig. A.6).

Figure A.6 Examples of the application of Gabor filtering to common images and use cases (i.e., character and face recognition).



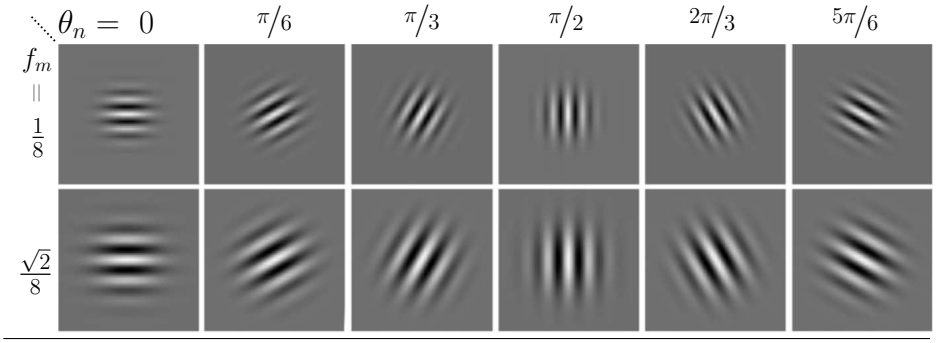
The output of each filter is related to the contours of the image at a given scale and orientation. These filters have been extensively used in texture analysis and object classification [113] and have recently been proposed for image quality assessment [250]. They are also conquering a niche [196, 112] in computer-aided diagnosis (CAD).

Mathematically, a Gabor filter consists of a sinusoidal wave modulated by a Gaussian envelope. The impulse response of a complex Gabor filter for an image pixel x, y is defined as:

$$G_{f_m,\theta_n}(x,y) \quad = \quad e^{\left(\frac{-x'^2}{2\sigma^2}\right)} \cdot e^{(2\pi f_m x'\sqrt{-1})} \tag{A.5}$$

where $x' = x\sin\theta + y\cos\theta$, f_m is the spatial frequency, θ_n is its orientation relative to the x-axis In this work, we have initially used six different values for θ_n $(0, \pi/6, \pi/3, \pi/2, 2\pi/3$ and $5\pi/6)$ and three different values for f_m $(1/8, \sqrt{2}/8$ and $1/4)$, although other combinations of all these parameters have been tested for the sake of completeness in [6]. The term $\sigma = 1/(2f_m)$ represents the spatial deviation for each filter. We generate a total of 18 complex responses, some of which are shown in Fig. A.7. Each image

Figure A.7 Real response of a sample Gabor filter bank $\mathrm{Re}(G_{f_m,\theta_n})$ generated with three frequencies $(M = 3)$ and six orientations $(N = 6)$.



pixel $\mathcal{I}(x, y)$ is then linearly convolved to obtain a complex Gabor response version $\mathcal{R}_{f_i,\theta_j}$ (for a frequency f_m and angle θ_n) with the expression:

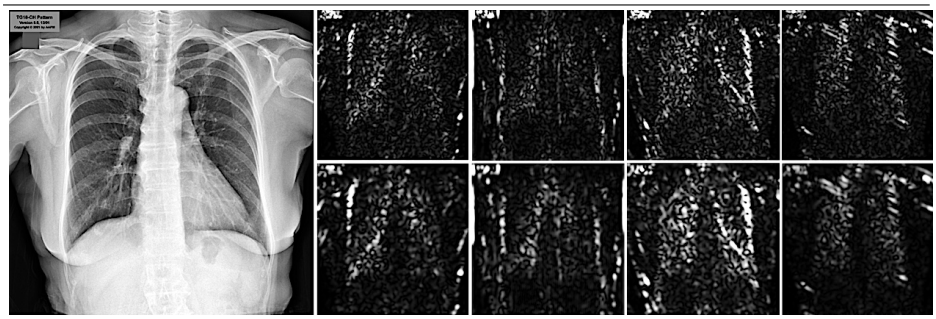
$$\mathcal{R}_{f_m,\theta_n}(x,y) = \mathcal{I}(x,y) * G_{f_m,\theta_n}(x,y) \tag{A.6}$$

After an image has been filtered, we define our Gabor response for a pixel x, y as:

$$R_{f_m,\theta_n}(x,y) = |\mathcal{R}_{f_m,\theta_n}(x,y)| \tag{A.7}$$

which corresponds to obtaining the amplitude of Eq. (A.6). Two examples of these Gabor responses are shown in Fig. A.8 and Fig. A.9.

Figure A.8 Gabor response for a chest radiography and for several filter configurations.

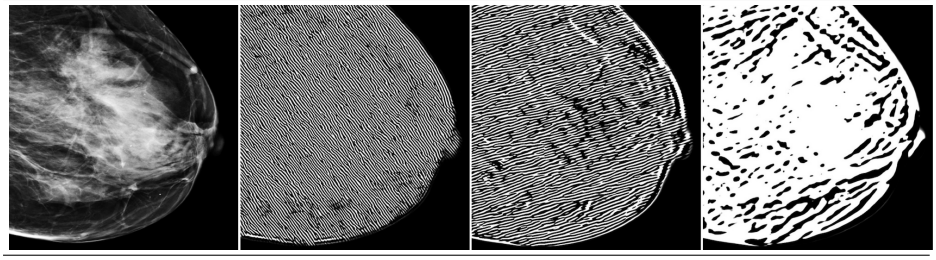


To assess the quality of a displayed image, we use the MI between the Gabor decompositions of $\mathcal{I}$ and $\tilde{\mathcal{I}}$. From here, if we let R_{f_m,θ_n} and $\tilde{R}_{f_m,\theta_n}$ be the Gabor responses of $\mathcal{I}$ and $\tilde{\mathcal{I}}$ obtained with Eq. (A.7), respectively, we can measure the perceived quality as:

$$\mathcal{Q}(\mathcal{I},\tilde{\mathcal{I}}) = \sum_{n=1}^{N}\sum_{m=1}^{M} \mathrm{MI}(R_{f_m,\theta_n},\, \tilde{R}_{f_m,\theta_n}) \tag{A.8}$$

The value of $\mathcal{Q}$ is upper bounded by the entropy of the input image. Moreover, the pixels are not assumed to be independent (in contrast to conventional information-based methodologies, such as those reviewed in Appendix A.2) because the statistical dependencies between pixels are taken into account by the Gabor filters. In Eq. (A.8) we propose a HVS function that will be maximized (as described in Appendix A.3.2) to find the IW limits that assure the best quality when presenting 12-bit mammograms in 8-bit screening software and hardware.

Figure A.9 Gabor response for a mammogram and for several filter configurations.

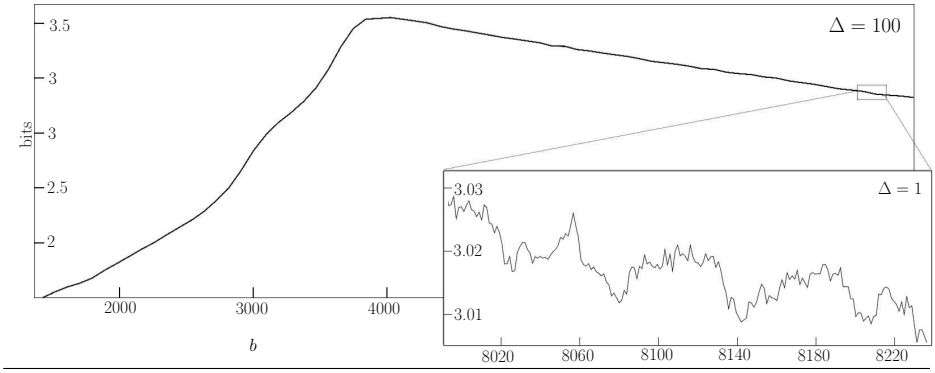


A.3.2 Auto-adjustment of image intensity levels

Internally, the proposed algorithm iteratively needs to seek the best contrast settings for the image whose quality is being assessed. It does so by using the perceptual image quality metric related to Gabor filtering proposed in Appendix A.3.1 from which we can define an objective function $\mathcal{F}_{\mathcal{I}}(a,b)$ based on Eq. (A.8) that can be maximized via a and b. The optimization of $\mathcal{F}_{\mathcal{I}}(a,b)$ will in turn assure the highest MI possible between $\mathcal{I}$ and $\tilde{\mathcal{I}}$.

However, our preliminary experiments showed that this function has plenty of local maxima, and, for this reason, it is difficult to find the optimum range using a gradient-based approach. On the other hand, a thorough search for the best parameters is computationally very expensive because it depends quadratically on the image grayscale depth. Our method hierarchically and iteratively (Fig. A.11) maximizes $\mathcal{F}_{\mathcal{I}}(a,b)$, optimizing the intensity threshold values a and b until convergence.

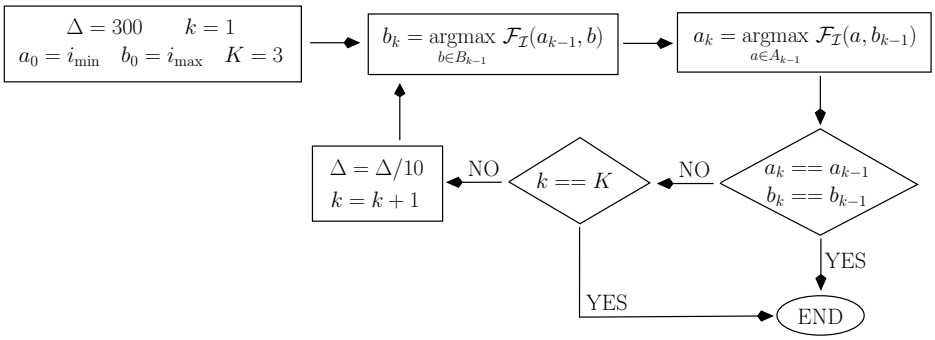
Figure A.10 Values of $\mathcal{F}_{\mathcal{I}}(a,b)$ for different grids Δ.



For each loop iteration (tagged with the parameter k), we define a grid of low (A_k) and high search values (B_k). The spacing of this grid is defined by Δ, which starts with a predefined and relatively large custom size (300, 200, 100, etc.) and is reduced, for instance, by a tenth in each k iteration ($\Delta/10$). The range of search intensities for each k is determined by the previous values a_{k-1} and b_{k-1}, which at the beginning ($k = 0$) are set to $\mathcal{I}_{\min}$ and $\mathcal{I}_{\max}$, respectively. The algorithm stops when the intensities found in an iteration are equal to those in the previous one or when it reaches a predefined iteration limit (K). From the obtained a and

b, we can easily derive $\mathrm{WL} = \frac{1}{2}(b-a)$ and $\mathrm{WW} = b - a$.

Figure A.11 Operational diagram of the proposed algorithm that seeks the optimal a and b intensity levels that maximize the MI between a source 12-bit image ($\mathcal{I}$) and its 8-bit displayed version ($\tilde{\mathcal{I}}$). It does so by iteratively optimizing an objective function $\mathcal{F}_{\mathcal{I}}(a,b)$ based on Eq. (A.8).



A.4 Validation and open implementation

A complete opensource implementation (including helper subroutines [93]) of the metric discussed in this appendix is available at https://github.com/TheAnswerIsFortyTwo/GRAIL (Fig. A.13). The fact of freely offering the here described algorithm contributes to deepening the idea of open and *reproducible research*. In turn, this development (and its validation) could not have been possible without the support of research projects committed to host open imaging databases. Specifically, we have made use of sets of mammograms compiled by the following public research institutions:

UPMC breast tomography collection from the University of Pittsburgh Medical Center (UPMC). All of the images contain hamartomas, subtle cancers, lobular carcinomas, cysts, papillomas, invasive ductal carcinomas, atherosclerotic calcifications, radial scars, vascular calcifications, benign ducts, oil cysts and fat necrosis.

Society of Breast Imaging collection or SBI. Its database contains

diagnostic images used during workshops and annual meetings. They mainly include calcifications and surgical clips.

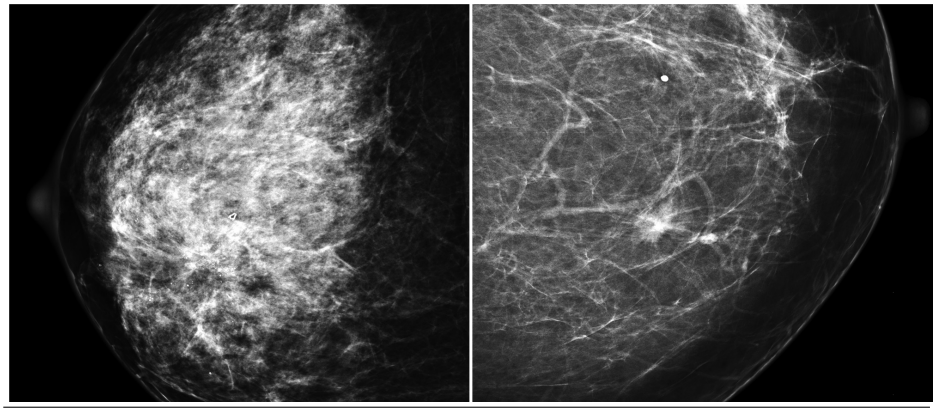
Cancer Genome Atlas collection or TCGA Research Network, part of the National Cancer Institute (part of the National Institute of Health). Images mainly contain invasive carcinoma.

Integrating the Healthcare Enterprise or IHE. It is an initiative developed by healthcare professionals to improve the way computer systems in hospitals and clinics share information. The image collection is part of their MESA software package, that was engineered at the Mallinckrodt Institute of Radiology together with the Healthcare Information and Management Systems Society (HIMSS) and the Radiological Society of North America (RSNA).

Cancer Imaging Archive which holds an important breast diagnosis collection [51]. This compilation contains cases with high-risk normals, ductal carcinoma *in situ*, fibroids, and lobular carcinomas.

Task Group 18 or TG18 from the AAPM. This task force evaluates the performance of electronic display devices [239]. For this purpose, they have efficiently gathered a set of high quality images, including not only geometrical and grayscale patterns but also anatomical ones, such as the two wide dynamic range mammograms shown in Fig. A.12.

Figure A.12 Mammograms TG18MM2 and TG18MM1 from the AAPM used for tests and validation of the presented quality quantification algorithm.



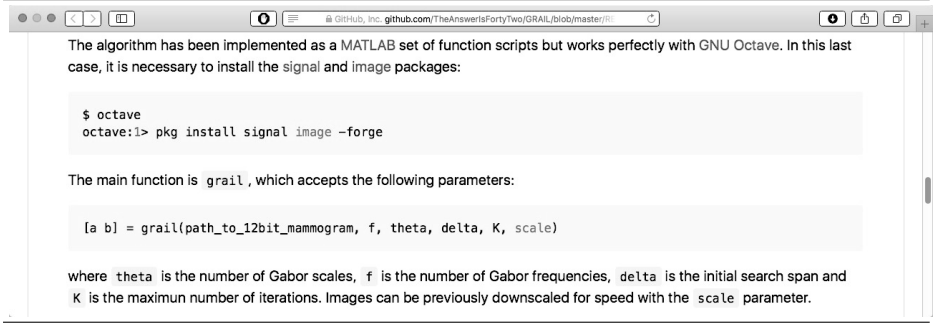
We were interested in collections that hosted fully-digital 12-bit images acquired with relative modern equipment. We were also concerned about incorporating and reflecting a broad spectrum of features, densities, associated health statuses, presence or not of foreign elements, and different qualities in order to account for as many clinical scenarios as possible. All these sets of images include the typical examination views: mediolateral oblique, craniocaudal, mediolateral and lateromedial.

According to the Breast Imaging Reporting and Data System (BI-RADS) from the ACR [177] and its four categories of breast density, the image set comprises:

- 33 mammograms associated to almost *entirely fatty* breasts (BI-RADS A),
- 36 representing *scattered fibroglandular* densities (BI-RADS B), 55 *heterogeneously dense* cases (BI-RADS C),
- and 35 images identifiable as *extremely dense* breasts (BI-RADS D).

Also, 29 instances contain foreign X-ray-opaque elements (i.e., surgical staples, fiducial markers, etc.) which reveal some sort of surgical intervention. Implants appear in 7 images. Around half of the mammograms were generated from a craniocaudal angle, the rest were obtained with lateral protocols such as mediolateral, oblique or lateromedial. Finally, 59 images contain high-density elements, including the aforementioned foreign items, makers, calcifications and other abnormalities.

Figure A.13 Website of GRAIL (Gabor-relying adjustment of image levels), that is, the implementation of the quality metric introduced in this appendix.



A.5 Conclusions

In this appendix, we have presented an innovative method to objectively assess the quality of an image by seeking the intensity window that maximizes the visual information when displaying a radiological image. The proposed technique is in turn based on Gabor filters, the human visual system and linear contrast manipulation. This methodology can be used to determine the quality of plain and densitometric X-ray images.

Bibliography

[1] *X-ray Patient Positioning Manual*. Aunt Minnie, 2008. 92

[2] Computer Tomography Subcommittee Task Group 23. The measurement, reporting and management of radiation dose in ct. Technical report, AAPM, 2008. 186

[3] Mohamad Anwar K. Abdelhalim, Reuqaya O. Al-Darwish, and Mohammad S Al-Ayed. Assessment of patient doses levels during x-ray diagnostic imaging using tl dosimeters and comparison with local and international levels. *Trends in Medical Research*, 2008. 189

[4] IR Ajayi and A Akinwumiju. Measurement of entrance skin doses to patients in four common diagnostic examinations by thermoluminescence dosimetry in nigeria. *Radiation protection dosimetry*, 87(3):217–220, 2000. 189

[5] N. Alatts and A. Abukhiar. Radiation doses from chest x-ray examinations for pediatrics in some hospitals of khartoum state, 2013. 189

[6] Alberto Albiol, Alberto Corbi, and Francisco Albiol. Automatic intensity windowing of mammographic images based on a perceptual metric. *Medical Physics*, 2017. 137, 198

[7] Alberto Albiol, Alberto Corbi, and Francisco Albiol. Automatic intensity windowing of mammographic images based on a perceptual metric. *Medical Physics*, 44(4):1369–1378, 2017. 192

[8] Antonio José Albiol, Alberto Albiol, Avelina Fernandez, Germán Rodrigo, and Francisco Albiol. Dispositivo y procedimiento de obtención de imá-

genes densitométricas de objetos mediante combinación de sistemas radiológicos y cámaras de profundidad. *Consejo Superior de Investigaciones Científicas*, 2012. 130

[9] F. Albiol, A. Corbi, and A. Albiol. 3d measurements from x-ray images and dense surface mappings. In *2016 Global Medical Engineering Physics Exchanges/Pan American Health Care Exchanges (GMEPE/PAHCE)*, pages 1–5, April 2016. 114

[10] F. Albiol, A. Corbi, and A. Albiol. Geometrical calibration of x-ray imaging with rgb cameras for 3d reconstruction. *IEEE Transactions on Medical Imaging*, PP(99):1–1, 2016. 84

[11] Francisco Albiol, Alberto Corbi, and Alberto Albiol. Geometrical calibration of X-ray imaging with RGB cameras for 3D reconstruction. *IEEE Transaction in Medical Imaging*, 35:1952–1961, 2016. 117

[12] Francisco Albiol, Alberto Corbi, and Alberto Albiol. 3d measurements in conventional x-ray imaging with rgb-d sensors. *Medical Engineering and Physics*, 42:73 – 79, 2017. 114

[13] Francisco Albiol, Alberto Corbi, and Alberto Albiol. Evaluation of modern camera calibration techniques for conventional diagnostic x-ray imaging settings. *Radiological Physics and Technology*, 10(1):68–81, 2017. 54

[14] John E Aldrich and Jan Williams. Change in patient doses from radiological examinations at the vancouver general hospital, 1991-2002. *Canadian Association of Radiologists*, 56(2):94—99, 2005. 189

[15] M. Alnowami, B. Alnwaimi, F. Tahavori, M. Copland, and K. Wells. A quantitative assessment of using the Kinect for Xbox 360 for respiratory surface motion tracking. In SPIE Proceedings, editor, *SPIE Medical Imaging*, volume 8316, 2012. 117

[16] K. Alzimami, A. Sulieman, A. Yousif, E. Babikir, and I. Salih. Evaluation of radiation dose to neonates in a special care baby unit. *Radiation Physics and Chemistry*, 104(0):150 – 153, 2014. 1st International Conference on Dosimetry and its Applications. 189

[17] Nina Amenta, Sunghee Choi, and Ravi Krishna Kolluri. The power crust. In *Proceedings of the sixth ACM symposium on Solid modeling and applications*, pages 249–266. ACM, 2001. 181

[18] Pedro Andreo. Monte carlo techniques in medical radiation physics. *Physics in medicine and biology*, 36(7):861, 1991. 186

[19] Hafeez Anwar, Irfanud Din, and Kang Park. Projector calibration for 3d scanning using virtual target images. *International Journal of Precision Engineering and Manufacturing*, 13(1):125–131, 2012. 55

[20] Michirou Aoki, Masaya Ono, Youhei Kamikawa, Kenta Kozono, Hidetaka Arimura, and Fukai Toyofuku. Development of a real-time patient monitoring system using Microsoft Kinect. In *World Congress on Medical Physics and Biomedical Engineering*, pages 1456–1459, 2013. 116

[21] Souha Aouadi and Laurent Sarry. Accurate and precise 2d-3d registration based on X-ray intensity. *Computer Vision and Image Understanding*, 2008. 90

[22] Xavier Armangué and Joaquim Salvi. Overall view regarding fundamental matrix estimation. *Image and Vision Computing*, 2003. 70

[23] A Aroua, R Bize, I Buchillier-Decka, J-P Vader, J-F Valley, and P Schnyder. X-ray imaging of the chest in switzerland in 1998: a nationwide survey. *European radiology*, 13(6):1250–1259, 2003. 189

[24] A Aroua, B Burnand, I Decka, J-P Vader, and J-F Valley. Nation-wide survey on radiation doses in diagnostic and interventional radiology in switzerland in 1998. *Health physics*, 83(1):46–55, 2002. 189

[25] Mohsen Asadinezhad and Mohammad Taghi Bahreyni Toossi. Doses to patients in some routine diagnostic x-ray examinations in iran: proposed the first iranian diagnostic reference levels. *Radiation protection dosimetry*, 132(4):409–414, 2008. 189

[26] M. Ayad. Risk assessment of an ionizing-radiation energy in diagnostic radiology. *Applied Energy*, 65(1–4):321 – 328, 2000. 189

[27] Abdel Y. I. Aziz and H. M. Karara. Direct linear transformation into object space coordinates in close-range photogrammetry. In *Proc. of the Symposium on Close-Range Photogrammetry*, pages 1–18, Urbana, Illinois, 1971. 64, 91

[28] Andreu Badal, Fahad Zafar, Han Dong, and Aldo Badano. A real-time radiation dose monitoring system for patients and staff during interventional fluoroscopy using a GPU-accelerated Monte Carlo simulator and an automatic 3D localization system based on a depth camera. In SPIE, editor, *Medical Imaging 2013: Physics of Medical Imaging*, volume 8668, 2013. 116

[29] T. Banerjee, M. Enayati, J.M. Keller, M. Skubic, M. Popescu, and M. Rantz. Monitoring patients in hospital beds using unobtrusive depth sensors. In *36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 5904–5907, 2014. 116

[30] S. Bauer, J. Wasza, S. Haase, N. Marosi, and J. Hornegger. Multi-modal surface registration for markerless initial patient setup in radiation therapy using Microsoft's Kinect sensor. In *IEEE International Conference on Computer Vision Workshops (ICCV Workshops)*, pages 1175–1181, 2011. 116

[31] S. Beck and B. Froehlich. Volumetric calibration and registration of rgbd-sensors. In *2015 IEEE Virtual Reality (VR)*, pages 151–152, March 2015. 181

[32] Z Begum. Entrance surface, organ and effective doses for some of the patients undergoing different types of x ray procedures in bangladesh. *Radiat Prot Dosimetry*, 95(3):257–262, 2001. 189

[33] François Berthiaume, Pierre Gravel, and Jacques a de Guise. Bone enhancement in digital dual energy radiographs from normalization with a synthetic background image. *Physics in medicine and biology*, 53(5):1259–75, 2008. 183

[34] Paul J Besl and Neil D McKay. Method for registration of 3-D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):239–256, 1992. 153

[35] A. Bigdelou, Tobias Benz, Loren Arthur Schwarz, and Nassir Navab. Simultaneous categorical and spatio-temporal 3D gestures using Kinect. In *IEEE Symposium on 3D User Interfaces (3DUI)*, pages 53–60, 2012. 117

[36] K Biju and PS Nagarajan. Computed normalised effective doses to an indian adult in conventional diagnostic x ray chest examinations. 2000. 189

[37] K. K. Biswas and S. K. Basu. Gesture recognition using microsoft kinect. In *International Conference on Automation, Robotics and Applications*, pages 100–103, 2011. 123

[38] Joseph Jr Borrelli, Michael Peelle, Elizabeth McFarland, Bradley Evanoff, and William M Ricci. Computer-reconstructed radiographs are as good as plain radiographs for assessment of acetabular fractures. *Am J Orthop (Belle Mead NJ)*, 37(9):455–459, Sep 2008. 143

[39] G. Brix, E. Nekolla, and J. Griebel. Strahlenexposition von patienten durch diagnostische und interventionelle röntgenanwendungen. *Der Radiologe*, 45(4):340–349, 2005. 189

[40] Marco JP Brugmans, Wilhelmina CAM Buijs, Jacob Geleijns, and Johan Lembrechts. Population exposure to diagnostic use of ionizing radiation in the netherlands. *Health physics*, 82(4):500–509, 2002. 189

[41] Jaume Buscà, Alicia Vigil, and Raúl Medina. Radiología digital en los servicios de radiodiagnóstico. parámetros dosimétricos. *Imagen Diagnóstica*, 1(2):70 – 72, 2010. 189

[42] S.C. Bushong. *Radiologic Science for Technologists: Physics, Biology, and Protection*. Elsevier - Health Sciences Division, 2012. 91

[43] Rafael Caicedo, Gloria Argüelles, and Alberto Alzate. Exposición a dosis bajas de radiación ionizante en el hospital universitario del valle, cali, 1980-1992. *Colombia Med*, 1996. 189

[44] Fatih Calakli and Gabriel Taubin. Ssd-c: Smooth signed distance colored surface reconstruction. In *Expanding the Frontiers of Visual Analytics and Visualization*, pages 323–338. Springer, 2012. 181

[45] L Caponetti and AM Fanelli. 3d bone reconstruction from two x-ray views. In *Engineering in Medicine and Biology Society, 1990., Proceedings of the Twelfth Annual International Conference of the IEEE*, pages 208–210. IEEE, 1990. 114

[46] J. H. E. Carmichael, C. Maccia, B. M. Moores, J. W. Oestmann, H. Schibilla, D. Teunen, R. Van Tiggelen, and B. Wall. European guidelines on quality criteria for diagnostic radiographic images (eur 16260). Technical report, European Commission, 1996. 185

[47] Xin Chen, Hemal Naik, Lejing Wang, Nassir Navab, and Pascal Fallavollita. Video guided calibration of an augmented reality mobile C-arm. In *IJCARS*, 2014. 89

[48] Lark Kwon Choi, Todd Goodall, and Alan C Bovik. Perceptual Image Enhancement Keywords. pages 1–37. 196

[49] Olivera Ciraj, Srpko Markovic, and Dusko Kosutic. First results on patient dose measurements from conventional diagnostic radiology procedures in serbia and montenegro. *Radiat Prot Dosimetry*, 113(3):330–335, 2005. 189

[50] Olivera F Ciraj-Bjelac, Srpko Marković, and Duško D Košutić. Patient dosimetry in diagnostic radiology. *Nuclear Technology and Radiation Protection*, 18(1):36–41, 2003. 189

[51] Kenneth Clark, Bruce Vendt, Kirk Smith, John Freymann, Justin Kirby, Paul Koppel, Stephen Moore, Stanley Phillips, David Maffitt, Michael Pringle, et al. The Cancer Imaging Archive: maintaining and operating a public information repository. *Journal of Digital Imaging*, 26(6):1045–1057, 2013. 202

[52] Bernhard EH Claus. Geometry calibration phantom design for 3D imaging. In *Medical Imaging*, 2006. 90

[53] Francisco Albiol Colomer, Alberto Corbí Bellot, Celso Bellot Romero, and Alberto Albiol Colomer. Dispositivo para extraer información tridimensional de imágenes radiográficas de un objeto; procedimiento de calibración de dicho dispositivo; y procedimiento para generar dichas imágenes radiográficas, 2015. 147

[54] Gaetano Compagnone, Laura Pagan, and Carlo Bergamini. Comparison of six phantoms for entrance skin dose evaluation in 11 standard x-ray examinations. *J Appl Clin Med Phys*, 6(1):101–113, 2005. 185

[55] Gaetano Compagnone, Laura Pagan, and Carlo Bergamini. Effective dose calculations in conventional diagnostic x-ray examinations for adult and paediatric patients in a large italian hospital. *Radiation Protection Dosimetry*, 114(1-3):164–167, 2005. 189

[56] Gaetano Compagnone, Laura Pagan, and Carlo Bergamini. Local diagnostic reference levels in standard x-ray examinations. *Radiat Prot Dosimetry*, 113(1):54–63, 2005. 189

[57] T. S Cook, G. Couch, T. J Couch, W. Kim, and W. Boonn. Using the Microsoft Kinect for patient size estimation and radiation dose normalization. *J Digit Imaging*, 2013. 89, 116

[58] A. Corbi, F. Albiol, and A. Albiol. Joint calibration of rgb and x-ray cameras. In *2016 Global Medical Engineering Physics Exchanges/Pan American Health Care Exchanges (GMEPE/PAHCE)*, pages 1–5, April 2016. 84

[59] M. Cristy. Mathematical phantoms representing children of various ages for use in estimates of internal dose. Technical report, NUREG/CR-1159 ORNL/NUREG/TM-367 (Oak Ridge National Laboratory), 1980. 185

[60] Yan Cui, Sebastian Schuon, Derek Chan, Sebastian Thrun, and Christian Theobalt. 3d shape scanning with a time-of-flight camera. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pages 1173–1180. IEEE, 2010. 131

[61] Brian Curless and Marc Levoy. A volumetric method for building complex models from range images. In *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, pages 303–312. ACM, 1996. 130

[62] Brian Curless and Marc Levoy. A volumetric method for building complex models from range images. In *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, pages 303–312. ACM, 1996. 131

[63] Alana Da Gama, Pascal Fallavollita, Veronica Teichrieb, and Nassir Navab. Motor rehabilitation using Kinect: A systematic review. *Games for Health Journal*, 4(2):123–135, 2015. 116

[64] M J Daly, J H Siewerdsen, Y B Cho, D A Jaffray, and J C Irish. Geometric calibration of a mobile C-arm for intraoperative cone-beam ct. *Med Phys*, 2008. 88

[65] G David and S C Price. Performing jcaho dose calculations with the aid of a microcomputer spreadsheet program. *AJR Am J Roentgenol*, 154(4):863–866, 1990. 185

[66] Martin Davies, Hazel McCallum, Gillian Whiter, Julie Brown, and Mike Helm. Patient dose audit in diagnostic radiography using custom designed software. *Radiography*, 3(1):17–25, 1997. 186

[67] John R Van Sörnsen de Koste, Hans C.J de Boer, Regine H Schuchhard-Schipper, Suresh Senan, and Ben J.M Heijmen. Procedures for high precision setup verification and correction of lung cancer patients using ct-simulation and digitally reconstructed radiographs (drr). *International Journal of Radiation Oncology*Biology*Physics*, 55(3):804 – 810, 2003. 144

[68] Daniel Cardoso de Moura, Jorge Manuel Gomes Barbosa, João Manuel Ribeiro da Silva Tavares, and Ana Reis. Calibration of bi-planar radiography with minimal phantoms. In *Symposium on Informatics Engineering*, 2008. 56

[69] Empresa Nacional de Residuos Radiactivos S.A. Report on the decommissioning of the josé cabrera nuclear power plant. Technical report, ENRESA, 2015. 173

[70] J. de Siebenthal and F. Langlotz. Use of a new tracking system based on artoolkit for a surgical simulator: accuracy test and overall evaluation. 2002. 90

[71] Teng Deng, Jean-Charles Bazin, Tobias Martin, Claudia Kuster, Jianfei Cai, Tiberiu Popa, and Markus Gross. Registration of multiple rgbd cameras via local rigid transformations. In *Multimedia and Expo (ICME), 2014 IEEE International Conference on*, pages 1–6. IEEE, 2014. 181

[72] James T. Dobbins, David L. Ergun, Lois Rutz, Dean A. Hinshaw, Hartwig Blume, and Dwayne C. Clark. Dqe(f) of four generations of computed radiography acquisition devices. *Medical Physics*, 22(10), 1995. 59

[73] César Domingo-Pardo, Namita Goel, Tobias Engert, Juergen Gerl, Masahiro Isaka, Ivan Kojouharov, and Henning Schaffner. A position sensitive gamma-ray scintillator detector with enhanced spatial resolution, linearity, and field of view. *IEEE transactions on medical imaging*, 28(12):2007–2014, 2009. 165

[74] Ian L Dryden. *Statistical shape analysis*. Wiley, 1998. 153

[75] J. Dutreix. Evaluation of radiation dose during radiodiagnostic examinations. *J Radiol*, 1991. 185

[76] E. H. A. Elshiekh. *Radiation doses from paediatric x-ray examinations in some hospitals in Khartoum Area*. PhD thesis, Sudan Academy of Sciences, 2007. 189

[77] Shu Kai S. Fan and Yu Chiang Chuang. Entropy-based image registration method using image intensity difference on overlapped region. *Machine Vision Applications*, 2011. 194

[78] K Faulkner, DA Broadhead, and RM Harrison. Patient dosimetry measurement methods. *Applied Radiation and Isotopes*, 50(1):113–123, 1999. 185

[79] Mark Fiala. Artag, a fiducial marker system using digital techniques. In *IEEE Computer Society Conf. on Computer Vision*, 2005. 97

[80] J. Fischer, M. Neff, D. Freudenstein, and D. Bartz. Medical augmented reality based on commercial image guided surgery. In *10th Eurographics Conf. on Virtual Environments*, 2004. 90

[81] Itzhak Fogel and Dov Sagi. Gabor filters as texture discriminator. *Biol. cybernetics*, 61:103, 1989. 197

[82] MB Freitas and EM Yoshimura. An overview of doses to patients and irradiation conditions of diagnostic chest x ray examinations carried out in hospitals of the city of sao paulo, brazil. *Radiation protection dosimetry*, 103(2):141–148, 2003. 189

[83] MB Freitas and EM Yoshimura. Dose measurements in chest diagnostic x rays: adult and paediatric patients. *Radiation protection dosimetry*, 111(1):73–76, 2004. 189

[84] N Freud, P Duvauchelle, JM Létang, and D Babot. Fast and robust ray casting algorithms for virtual x-ray imaging. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, 248(1):175–180, 2006. 134, 142

[85] B Fröhlich, R Blach, and R van Liere. A lightweight id-based extension for marker tracking systems. 2007. 97

[86] S. Garrido-Jurado, R. Muñoz-Salinas, F.J. Madrid-Cuevas, and M.J. Marín-Jiménez. Automatic generation and detection of highly reliable fiducial markers under occlusion. *Pattern Recognition*, 2014. 97

[87] David G Grant. Tomosynthesis: a three-dimensional radiographic imaging technique. *IEEE Transactions on Biomedical Engineering*, 1972. 88

[88] M Green. The angular distribution of characteristic X radiation and its origin within a solid target. *Proceedings of the Physical Society*, 83(3):435, 1964. 35

[89] P Grøn, HM Olerud, G Einarsson, W Leitz, A Servomaa, BW Schoultz, and O Hjardemaal. A nordic survey of patient doses in diagnostic radiology. *European radiology*, 10(12):1988–1992, 2000. 189

[90] Manoj Gupta, Suchi Upadhyay, and Ashok K. Nagawat. Camera calibration technique using tsai's algorithm, 2011. 63

[91] Prince Kwabena Gyekye, Adu Simon, Emi-Reynolds Geoffrey, Yeboah Johnson, Inkoom Stephen, Cynthia Kaikor Engmann, and Wotorchi-Gordon Samuel. Radiation dose estimation of patients undergoing lumbar spine radiography. *Journal of medical physics/Association of Medical Physicists of India*, 38(4):185, 2013. 189

[92] Tomasz Hachaj and MarekR. Ogiela. Augmented reality interface for visualization of volumetric medical data. In RyszardS. Choraś, editor, *Image Processing and Communications Challenges 2*, volume 84 of *Advances in Intelligent and Soft Computing*, pages 271–277. Springer Berlin Heidelberg, 2010. 90

[93] M. Haghighat, S. Zonouz, and M. Abdel-Mottaleb. CloudID: trustworthy cloud-based and cross-enterprise biometric identification. *Expert Systems with Applications*, 42(21):7905–7916, 2015. 201

[94] MA Halato, ST Kafi, AM Ahmed, FA Sid Ahamed, Z Ibrahim, MF Suliman, and II Suliman. Dosimetry for patients undergoing radiographic examinations in sudan. In *IX Radiation Physics and Protection Conference*, 2008. 189

[95] M D Harpen. A mathematical spread sheet application for production of entrance skin exposure nomograms. *Med Phys*, 23(2):241–242, 1996. 185

[96] D Hart and BF Wall. Uk population dose from medical x-ray examinations. *European journal of radiology*, 50(3):285–291, 2004. 189

[97] R. Hartley. Projective reconstruction and invariants from multiple images. *Pattern Analysis and Machine Intelligence*, 1994. 104

[98] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2004. 63, 64, 91, 103, 104, 119, 125

[99] RichardI. Hartley. Self-calibration from multiple views with a rotating camera. In Jan-Olof Eklundh, editor, *Computer Vision — ECCV '94*, volume 800 of *Lecture Notes in Computer Science*, pages 471–478. Springer Berlin Heidelberg, 1994. 60

[100] Thomas Herrmann. Forschung mit ionisierenden strahlen in deutschland (röntgen, ct, pet). In Christian Lenk, Gunnar Duttge, and Heiner Fangerau, editors, *Handbuch Ethik und Recht der Forschung am Menschen*, pages 27–32. Springer Berlin Heidelberg, 2014. 189

[101] Dirk Holz and Sven Behnke. Fast range image segmentation and smoothing using approximate surface reconstruction and region growing. *Intelligent autonomous systems 12*, pages 61–73, 2013. 181

[102] Andreas M Hötker, Michael B Pitton, Peter Mildenberger, and Christoph Düber. Speech and motion control for interventional radiology: requirements and feasibility. *International Journal of Computer Assisted Radiology and Surgery*, 8(6):997–1002, 2013. 117

[103] Sheng-Jhe Hsu, Wen-Sheng Tseng, Fu-Chieh Hsu, and Yung-Ying Lo. iSpine: A motion-sensing edutainment system for improving children's spinal health. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems*, 2013. 116

[104] Walter Huda, Edward L Nickoloff, and John M Boone. Overview of patient dosimetry in diagnostic radiology in the usa for the past 50 years. *Medical physics*, 35(12):5713–5728, 2008. 189

[105] DC Hunt, O Tousignant, and JA Rowlands. Evaluation of the imaging properties of an amorphous selenium-based flat panel detector for digital fluoroscopy. *Medical physics*, 31(5):1166–1175, 2004. 28

[106] O Iacob and C Diaconescu. Doses to patients from diagnostic radiology in romania. In *IAEA-85*, 2001. 189

[107] ICRP. Recommendations of the international commission on radiological protection, annals of icrp 21. Technical report, International Commission on Radiological Protection, 1990. 185

[108] ICRP. the 2007 recommendations of the international commission on radiological protection, annals of icrp. Technical report, International Commission on Radiological Protection, 2007. 185

[109] Hiroshi Imamura, Noriaki Ida, Naozo Sugimoto, Shigeru Eiho, Shin-ichi Urayama, Katsuya Ueno, and Kanji Inoue. Registration of preoperative cta and intraoperative fluoroscopic images for assisting aortic stent grafting. In *Medical Image Computing and Computer-Assisted Intervention*. Springer Berlin Heidelberg, 2002. 143

[110] Stephen Inkoom, James Togobo, Geoffrey Emi-Reynolds, Adrian Oddoye, Theophilus Ofosu Ntiri, and Prince Kwabena Gyekye. Retrospective patient dose analysis of ghana's first direct digital radiography system. *Health physics*, 103(2):133–137, 2012. 189

[111] Shahram Izadi, David Kim, Otmar Hilliges, David Molyneaux, Richard Newcombe, Pushmeet Kohli, Jamie Shotton, Steve Hodges, Dustin Freeman, Andrew Davison, et al. Kinectfusion: real-time 3d reconstruction and interaction using a moving depth camera. In *Proceedings of the 24th annual ACM symposium on User interface software and technology*, pages 559–568, 2011. 131

[112] Stefan Jaeger, Alexandros Karargyris, Sameer Antani, and George Thoma. Detecting tuberculosis in radiographs using combined lung masks. In *Conference on Engineering in Medicine and Biology Society*, 2012. 197

[113] Anil K Jain, Nalini K Ratha, and Sridhar Lakshmanan. Object detection using Gabor filters. *Pattern Recognition*, 30(2):295–309, 1997. 197

[114] J Jankowski and MA Staniszewska. Methodology for the set-up of a quality control system for diagnostic x ray units in poland. *Radiation protection dosimetry*, 90(1-2):259–262, 2000. 189

[115] George C. Kagadis, Alisa Walz-Flannigan, Elizabeth A. Krupinski, Paul G. Nagy, Konstantinos Katsanos, Athanasios Diamantopoulos, and Steve G. Langer. Medical imaging displays and their use in image interpretation. *RadioGraphics*, 33(1):275–290, 2013. 47

[116] WA Kalender, B Schmidt, M Zankl, and M Schmidt. A pc program for estimating organ dose and effective dose values in computed tomography. *European radiology*, 9(3):555–562, 1999. 186

[117] Martin Kaltenbrunner and Ross Bencina. reacTIVision: a computer-vision framework for table-based tangible interaction. In *Int. Conf. on Tangible and Embedded Interaction*, 2007. 97

[118] Hirokazu Kato. Artoolkit: library for vision-based augmented reality. *IEICE, PRMU*, 2002. vi, 87

[119] K. Khoshelham. Accuracy Analysis of Kinect Depth Data. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 38:133–138, 2012. 115

[120] Kourosh Khoshelham and Sander Oude Elberink. Accuracy and resolution of Kinect depth data for indoor mapping applications. *Sensors*, 12(2):1437–1454, 2012. 115

[121] T Kiljunen, H Järvinen, and S Savolainen. Diagnostic reference levels for thorax x-ray examinations of paediatric patients. 2014. 189

[122] You-hyun Kim, Jong-hak Choi, Chang-kyun Kim, Jung-min Kim, Sung-soo Kim, Yu-whan Oh, Chang-yeap Lee, Dae-hyun Kang, Young-bae Lee, Pyong-kon Cho, et al. Patient dose measurements in diagnostic radiology procedures in korea. *Radiation protection dosimetry*, 123(4):540–545, 2007. 189

[123] Tom Kimpe and Tom Tuytschaever. Increasing the number of gray shades in medical display systems - How much is enough? *Digital Imaging*, 20:422, 2007. 44

[124] Eugenia N. Kirsanova and Michael G. Sadovsky. Entropy approach in the analysis of anisotropy of digital images. *Open Systems & Information Dynamics*, 2002. 194

[125] K Kisielewicz, A Truszkiewicz, S Wach, and M Wasilewska-Radwańska. Evaluation of dose area product vs. patient dose in diagnostic x-ray units. *Physica Medica*, 27(2):117–120, 2011. 189

[126] Kenta Kozono. A study on a real-time X-ray entrance dose monitoring system in interventional radiology using Microsoft Kinect. In *Japan Radiology Congress*, 2013. 116

[127] R Kramer, HJ Khoury, and JW Vieira. Caldose_x—a software tool for the assessment of organ and tissue absorbed doses, effective dose and cancer risks in diagnostic radiology. *Physics in medicine and biology*, 53(22):6437, 2008. 185

[128] R Kramer, HJ Khoury, and JW Vieira. Caldose_x—a software tool for the assessment of organ and tissue absorbed doses, effective dose and cancer risks in diagnostic radiology. *Physics in medicine and biology*, 53(22):6437, 2008. 186

[129] Gerhard Kurz, Peter Hegedus, Gabor Szabo, and Uwe D Hanebeck. Experimental evaluation of Kinect and inertial sensors for beating heart tracking. In *Deutsche Gesellschaft für Computer- und Roboterassistierte Chirurgie (CURAC)*, 2013. 117

[130] Sarang Lakare and Arie Kaufman. 3d segmentation techniques for medical volumes. *Center for Visual Computing, Department of Computer Science, State University of New York*, 2000. 145

[131] San Lázaro and Centro Habana. Evaluación de dosis que recibe el paciente durante el diagnóstico de hernia discal: Resultados preliminares. In *VIII Congreso Regional de Seguridad Radiológica y Nuclear*, 2010. 189

[132] LB Li, JP Wang, XR Yu, SS He, FH Yu, and CH Ding. Medical radiation usage and exposures from medical x ray diagnosis in shandong province of china. *Radiation protection dosimetry*, 93(3):261–266, 2001. 189

[133] LB Li, JP Wang, XR Yu, SS He, FH Yu, and CH Ding. Medical radiation usage and exposures from medical x ray diagnosis in shandong province of china. *Radiation protection dosimetry*, 93(3):261–266, 2001. 189

[134] Yun Li, L. Berkowitz, G. Noskin, and S. Mehrotra. Detection of patient's bed statuses in 3D using a Microsoft Kinect. In *Conference of the IEEE Engineering in Medicine and Biology Society*, pages 5900–5903, 2014. 116

[135] Rui Liao, Li Zhang, Ying Sun, Shun Miao, and C. Chefd'hotel. A review of recent advances in registration techniques applied to minimally invasive therapy. *IEEE Transactions on Multimedia*, 2013. 90

[136] S.H. Lim, E. Golkar, and A.A. Abd Rahni. Respiratory motion tracking using the Kinect camera. In *IEEE Conference on Biomedical Engineering and Sciences (IECBES)*, pages 797–800, 2014. 117

[137] Sandra A Lowe. Diagnostic radiography in pregnancy: risks and reality. *Aust N Z J Obstet Gynaecol*, 44(3):191–196, 2004. 189

[138] Quan-Tuan Luong and Olivier Faugeras. The fundamental matrix: Theory, algorithms, and stability analysis. *Inter. Journal of Computer Vision*, 1996. 70

[139] C. Maccia, M. Benedittini, C. Lefaure, and F. Fagnani. Doses to patients from diagnostic radiology in france. *Health Physics*, 54(4), 1988. 189

[140] DL Pérez Magán, L Caballero, César Domingo-Pardo, Jorge Agramunt-Ros, F Albiol, A Casanovas, A González, C Guerrero, J Lerendegui-Marco, A Tarifeño-Saldivia, et al. First tests of the applicability of gamma-ray imaging for background discrimination in time-of-flight neutron capture measurements. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 823:107–119, 2016. 165

[141] J. B. Antoine Maintz and Max A. Viergever. A survey of medical image registration. *Med. Image Analysis*, 1998. 90

[142] T Makri, E Yakoumakis, D Papadopoulou, G Gialousis, V Theodoropoulos, P Sandilos, and E Georgiou. Radiation risk assessment in neonatal radiographic examinations of the chest and abdomen: a clinical and monte carlo dosimetry study. *Phys Med Biol*, 51(19):5023–5033, 2006. 189

[143] P. Markelj and B. Likar. *Registration of Three-dimensional and Two-dimensional Medical Images: PhD Thesis*. P. Markelj, 2010. 56, 91

[144] P. Markelj, D. Tomaževič, B. Likar, and F. Pernuš. A review of 3D/2D registration methods for image-guided interventions. *Med. Image Analysis*, 2012. 90

[145] A Martinez, L Jova, J Carrazana, E Diaz, R Mora, C Guevara, and I Fleitas. Results of the study of entrance surface dose from conventional examinations in diagnostic radiology. In *IAEA-85*, 2001. 189

[146] T Maruyama, Y Kumamoto, Y Noda, K Iwai, N Mase, K Nishizawa, and Y Furuya. Determinations of organ or tissue doses and collective effective dose equivalent from diagnostic x ray examinations in japan. *Radiation Protection Dosimetry*, 43(1-4):213–216, 1992. 189

[147] Cynthia H McCollough and Beth A Schueler. Calculation of effective dose. *Medical physics*, 27(5):828–837, 2000. 185

[148] E L McGuire and P A Dickson. Exposure and organ dose estimation in diagnostic radiology. *Med Phys*, 13(6):913–916, 1986. 185

[149] Phillip J. Mckerrow and Jo Abrantes. Problems encountered in the implementation of tsai's algorithm for camera calibration. In *Australasian Conference on Robotics and Automation*, 2002. 63

[150] Michael F McNitt-Gray. Aapm/rsna physics tutorial for residents: topics in ct: radiation dose in ct 1. *Radiographics*, 22(6):1541–1553, 2002. 186

[151] Boris Mederos, Nina Amenta, Luiz Velho, and Luiz Henrique De Figueiredo. Surface reconstruction for noisy point clouds. Citeseer. 181

[152] Gerard Medioni and Sing Bing Kang. *Emerging Topics in Computer Vision*. Prentice Hall PTR, Upper Saddle River, NJ, USA, 2004. 66, 91

[153] Stephan Meister, Shahram Izadi, Pushmeet Kohli, Martin Hämmerle, Carsten Rother, and Daniel Kondermann. When can we use kinectfusion for ground truth acquisition. In *Workshop on Color-Depth Camera Fusion in Robotics, IROS*, volume 2, 2012. 122

[154] Ma Meng, Pascal Fallavollita, Troels Blum, Ulrich Eck, Christian Sandor, Simon Weidert, Jens Waschke, and Nassir Navab. Kinect for interactive AR anatomy learning. In *IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 227–278, 2013. 117

[155] Fred A. Jr Mettler, Bruce R. Thomadsen, Mythreyi Bhargavan, Debbie B. Gilley, Joel E. Gray, Jill A. Lipoti, John McCrohan, Terry T. Yoshizumi, and Mahadevappa Mahesh. Medical radiation exposure in the u.s. in 2006: Preliminary results. *Health Physics*, 95(5), 2008. 189

[156] R Michel and S C Perle. Effective dose equivalent estimates in diagnostic radiology with single dosimetry. *Health Phys*, 79(2 Suppl):S17–9, 2000. 189

[157] Peter Mildenberger, Marco Eichelberg, and Eric Martin. Introduction to the dicom standard. *European radiology*, 12(4):920–927, 2002. 50

[158] Donald L. Miller, Deukwoo Kwon, and Grant H. Bonavia. Reference levels for patient radiation doses in interventional radiology: Proposed initial values for u.s. practice. *Radiology*, 253(3):753–764, 2009. PMID: 19789226. 189

[159] TimothyJ. Miller and EnricoC. Quintana. Stereo x-ray system calibration for three-dimensional measurements. In Helena Jin, Cesar Sciammarella, Sanichiro Yoshida, and Luciano Lamberti, editors, *Advancement of Optical Methods in Experimental Mechanics, Volume 3*, Conference Proceedings of the Society for Experimental Mechanics Series, pages 201–207. Springer International Publishing, 2014. 56, 89, 91

[160] K.E.M. Mohamadain, A.C.P. Azevedo, L.A.R. da Rosa, M.R.N. Guebel, and M.C.B. Boechat. Dose measurements using thermoluminescent dosimeters and dosecal software at two paediatric hospitals in rio de janeiro. *Applied Radiation and Isotopes*, 59(1):53 – 57, 2003. 189

[161] Tomas Möller and Ben Trumbore. Fast, minimum storage ray/triangle intersection. In *ACM SIGGRAPH 2005 Courses*, page 7. ACM, 2005. 133

[162] Ehsan Moudi, Sina Haghanifar, Hoora Hadian, Ali Shabestani Monfared, Gholamreza Deilam, and Nika Bahemmat. Assessment of radiation exposure of head and neck during panoramic radiography in different age groups. *Science*, 3(1):10–14, 2014. 189

[163] Daniel C. Moura, Jorge G. Barbosa, Ana M. Reis, and João Manuel R. S. Tavares. A flexible approach for the calibration of biplanar radiography of the spine on conventional radiological systems. *Computer Modeling in Engineering and Sciences*, 2010. 56

[164] Daniel C. Moura, Jorge G. Barbosa, Manuel R.S. Tavares, João, and Ana M. Reis. Calibration of bi-planar radiography with a rangefinder and a small calibration object. 5358:572–581, 2008. 89

[165] WE Muhogora, AM Nyanda, US Lema, and JE Ngaile. Typical radiation doses to patients from some common x ray examinations in tanzania. *Radiation protection dosimetry*, 82(4):301–305, 1999. 189

[166] WE Muhogora, A Trianni, F Toso, A Devetti, R Padovani, P Msaki, and R Kazema. Comparison of image quality and patient dose for chest x-ray examinations on conventional and low cost computed radiography systems. *Radiography*, 18(4):275–278, 2012. 189

[167] Tara Mullaney, Björn Yttergren, and Erik Stolterman. Positional acts: Using a Kinect sensor to reconfigure patient roles within radiotherapy treatment. In *8th International Conference on Tangible, Embedded and Embodied Interaction*, pages 93–96. ACM, 2013. 116

[168] Rafael Munoz-Salinas, Manuel J Marín-Jimenez, Enrique Yeguas-Bolivar, and Rafael Medina-Carnicer. Mapping and localization from planar markers. *arXiv preprint 1606.00151*, 2016. 171

[169] Leonid Naimark and Eric Foxlin. Fiducial detection system, 2007. US Patent 7,231,063. 97

[170] Nassir Navab, Sandro-Michael Heining, and Joerg Traub. Camera augmented mobile C-arm (CAMC): calibration, accuracy study, and clinical applications. *IEEE Transactions on Medical Imaging*, 2010. 89

[171] Richard A Newcombe, Shahram Izadi, Otmar Hilliges, David Molyneaux, David Kim, Andrew J Davison, Pushmeet Kohi, Jamie Shotton, Steve Hodges, and Andrew Fitzgibbon. Kinectfusion: Real-time dense surface mapping and tracking. In *10th IEEE international symposium on Mixed and augmented reality (ISMAR)*, pages 127–136. IEEE, 2011. 131

[172] Richard A. Newcombe, Shahram Izadi, Otmar Hilliges, David Molyneaux, David Kim, Andrew J. Davison, Pushmeet Kohli, Jamie Shotton, Steve Hodges, and Andrew Fitzgibbon. KinectFusion: Real-time dense surface mapping and tracking. In *10th IEEE international symposium on Mixed and augmented reality (ISMAR)*, pages 127–136, 2011. 122

[173] Khadija Nfaoui, Farida Bentayeb, O El Basraoui, and Ana Cecilia Pedrosa de Azevedo. Evaluation of paediatric x-ray doses in moroccan university hospitals. *Radiation protection dosimetry*, page ncq278, 2010. 189

[174] KH Ng, P Rassiah, HB Wang, AS Hambali, P Muthuvellu, and HP Lee. Doses to patients in routine x-ray examinations in malaysia. *The British journal of radiology*, 71(846):654–660, 1998. 189

[175] S. Nicolau, P. Y. Lee, H. S. Wu, and M. H. Huang. Fusion of c-arm x-ray image on video view to reduce radiation exposure and improve orthopedic surgery planning : first in-vivo evaluation. In *CARS*, 2011. 89

[176] P.J. Noonan, J. Howard, D. Tout, I. Armstrong, H.A. Williams, T.F. Cootes, W.A. Hallett, and R. Hinz. Accurate markerless respiratory tracking for gated whole body PET using the Microsoft Kinect. In *IEEE Nuclear Science Symposium and Medical Imaging Conference*, pages 3973–3974, 2012. 116

[177] American College of Radiology Committee. Bi-rads atlas 5th edition. Technical report, ACR, 2014. 203

[178] FO Ogundare, CL Ajibola, and FA Balogun. Survey of radiological techniques and doses of children undergoing some common x-ray examinations in three hospitals in nigeria. *Medical physics*, 31(3):521–524, 2004. 189

[179] FO Ogundare, CZ Uche, and FA Balogun. Radiological parameters and radiation doses of patients undergoing abdomen, pelvis and lumbar spine x-ray examinations in three nigerian hospitals. 2014. 189

[180] HM Olerud and G Saxebøl. Diagnostic radiology in norway from 1983 to 1993-examination frequency and collective effective dose to patients. *Radiation protection dosimetry*, 74(4):247–260, 1997. 189

[181] T Olgar, D Bor, G Berkmen, and T Yazar. Patient and staff doses for some complex x-ray examinations. *Journal of Radiological Protection*, 29(3):393, 2009. 189

[182] C.J. Olowookere, R.I. Obed, I.A. Babalola, and T.O. Bello. Patient dosimetry during chest, abdomen, skull and neck radiography in {SW} nigeria. *Radiography*, 17(3):245 – 249, 2011. 189

[183] Ernest K Osei and Rob Barnett. Software for the estimation of organ equivalent and effective doses from diagnostic radiology procedures. *Journal of Radiological Protection*, 29(3):361, 2009. 185

[184] Ernest K Osei and Johnson Darko. A survey of organ equivalent and effective doses from diagnostic radiology procedures. *ISRN radiology*, 2013, 2012. 185

[185] Stanley Osher and James A Sethian. Fronts propagating with curvature-dependent speed: algorithms based on hamilton-jacobi formulations. *Journal of computational physics*, 79(1):12–49, 1988. 145

[186] OA Osibote and ACP De Azevedo. Estimation of adult patient doses for common diagnostic x-ray examinations in rio de janeiro, brazil. *Physica Medica*, 24(1):21–28, 2008. 189

[187] H Osman, A Elzaki, M Elsamani, J Alzaeidi, K Sharif, and Ahmad Elmorsy. Assessment of peadiatric radiation dose from routine x-ray examination: A hospital based study, taif pediatric hospital. *Wulfenia Journal*, 2014. 189

[188] H Osman, A Sulieman, II Suliman, and AK Sam. Radiation dose measurements in routine x ray examinations. 2011. 185

[189] R Padovani, G Contento, M Fabretto, MR Malisan, V Barbina, and G Gozzi. Patient doses and risks from diagnostic radiology in north-east italy. *The British journal of radiology*, 60(710):155–165, 1987. 189

[190] D Papadimitriou, A Perris, MG Molfetas, N Panagiotakis, A Manetou, G Tsourouflis, J Vassileva, P Chronopoulos, O Karapanagiotou, and S Kottou. Patient dose, image quality and radiographic techniques for common x ray examinations in greece and comparison with the european guidelines. *Radiation protection dosimetry*, 95(1):43–48, 2001. 189

[191] R A Parry, S A Glaze, and B R Archer. The AAPM/RSNA physics tutorial for residents. Typical patient radiation doses in diagnostic radiology. *Radiographics*, 19(5):1289–1302, 1999. 185

[192] G.P. Penney, J. Weese, J.A Little, P. Desmedt, D. L G Hill, and D.J. Hawkes. A comparison of similarity measures for use in 2-d-3-d medical image registration. *IEEE Transactions on Medical Imaging*, 1998. 90

[193] Gavin G Poludniowski. Calculation of x-ray spectra emerging from an x-ray tube. part ii. x-ray production and filtration in x-ray targets. *Medical physics*, 34(6):2175–2186, 2007. 187

[194] Alfredo Ramírez, Ana María Silva, Eric Farias Ch, and Daniella Fabri. Consideraciones sobre el uso de rayos x en la práctica médica convencional para la necesaria aplicación de normas básicas de seguridad. *Revista HCUCh*, 18:131–40, 2007. 189

[195] C R Ramsey, D Arwood, D Scaperoth, and A L Oliver. Clinical application of digitally-reconstructed radiographs generated from magnetic resonance imaging for intracranial lesions. *Int J Radiat Oncol Biol Phys*, 45(3):797–802, Oct 1999. 143

[196] Rangaraj M. Rangayyan, Fábio J. Ayres, and J. E. Leo Desautels. A review of computer-aided diagnosis of breast cancer: Toward the detection of subtle signs. *Journal of the Franklin Institute*, 344(3-4):312–348, 2007. 197

[197] S Rannikko, A Servomaa, I Ermakov, L Masarskii, L Saltukova, M Razumnaya, and V Nikitin. Calculation of the estimated collective effective dose equivalent (se) due to x-ray diagnostic examinations-estimate of the se in finland. *Health physics*, 53(1):31–36, 1987. 189

[198] Simo Rannikko, KTK Karila, and Matti Toivonen. Patient and population doses of x-ray diagnostics in finland. Technical report, Radiation and Nuclear Safety Authority, Helsinki (Finland), 1997. 189

[199] H. F. Rasool. *Assessment and control of medical exposure in Pakistan*. PhD thesis, Vienna University of Technology, 2009. 189

[200] C Regalado, A Benavente, O Celedonio, Q Ramirez, A Vasquez, et al. Radiological risk study at the hospital in trujillo-peru. In *IAEA*, 1995. 189

[201] F. Remondino and C. Fraser. Digital camera calibration methods: considerations and comparisons. In Isprs, editor, *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume Vol. XXXVI, Dresden, Germany, 2006. 63

[202] Miguel Reyes, Albert Clapés, José Ramírez, Juan R Revilla, and Sergio Escalera. Automatic digital biometry analysis based on depth maps. *Computers in Industry*, 64(9):1316–1325, 2013. 116

[203] S. Reyes, G. Gama T., M. Beristain, and H. Espino. Evaluación de la dosis de entrada en superficie, dosis en órgano y dosis efectiva e, recibidas por personal y pacientes en estudios de colangiopancreatografía retrógrada endoscópica en el hospital general de méxico. *Anales de Radiología México*, 2007. 189

[204] Andrew C Rice, Alastair R Beresford, and Robert K Harle. Cantag: an open source software toolkit for designing and deploying marker-based vision systems. In *Int. Conf. on Pervasive Computing*, 2006. 97

[205] Guilherme Oberto Rodrigues, Silvia Goldmeier, Carlos Jader Feldman, Rogério Fachel de Medeiros, and Gustavo Glotz de Lima. Radiografías de tórax de calidad con dosis menores que los límites recomendados. *Imagen Diagnóstica*, 4(1):6–12, 2013. 189

[206] D. W. O. Rogers, B. A. Faddegon, G. X. Ding, C. M. Ma, J. We, and T. R. Mackie. Beam: A monte carlo code to simulate radiotherapy treatment units. *Medical Physics*, 22(5):503–524, 1995. 185

[207] A. Rougé, C. Picard, C. Ponchut, and Y. Trousset. Geometrical calibration of X-ray imaging chains for 3D reconstruction. *Computerized Medical Imaging and Graphics*, 1993. 56, 91

[208] Anne Rougee, Catherine L Picard, Yves L Trousset, and Cyril Ponchut. Geometrical calibration for 3d x-ray imaging. In *Medical Imaging*, 1993. 90

[209] J A Rowlands. The physics of computed radiography. *Phys Med Biol*, 47(23):R123–66, Dec 2002. 59

[210] Szymon Rusinkiewicz, Olaf Hall-Holt, and Marc Levoy. Real-time 3d model acquisition. *ACM Transactions on Graphics (TOG)*, 21(3):438–446, 2002. 130

[211] M.K. Saeed and J.M. Al-Qahtani. Paediatric dose measurements for chest x-ray examinations at maternity and children hospital in najran - saudi arabia. *Australasian Physical and Engineering Sciences in Medicine*, 35(2):215–219, 2012. 189

[212] Shahiana Izma Sahidan, Mohd Yusoff Mashor, Aida Sharmila Wati Wahab, Zaleha Salleh, and Hasnan Ja'afar. *Local and Global Contrast Stretching For Color Contrast Enhancement on Ziehl-Neelsen Tissue Section Slide Images*. Springer Berlin Heidelberg, 2008. 138

[213] Tina Samajdar and Md. Iqbal Quraishi. *Information Systems Design and Intelligent Applications: 2nd International Conference INDIA 2015*, chapter Analysis and Evaluation of Image Quality Metrics. Springer India, 2015. 196

[214] Jorge A. Sánchez, Eduardo A. Destefanis, and Luis R. Canali. Plane-based camera calibration without direct optimization algorithms. In *IV Jornadas Argentinas de Robótica, Córdoba*, 2006. 77

[215] Martin Sandau, Henrik Koblauch, Thomas B. Moeslund, Henrik Aanæs, Tine Alkjær, and Erik B. Simonsen. Markerless motion capture can provide reliable 3d gait kinematics in the sagittal and frontal plane. *Medical Engineering and Physics*, 36(9):1168–1175, 2014. 116

[216] Svetlana Sarycheva, Vladislav Golikov, and Sergey Kalnicky. Studies of patient doses in interventional radiological examinations. *Radiation protection dosimetry*, page ncq021, 2010. 189

[217] Steffen Schumann, Benedikt Thelen, Steven Ballestra, Lutz-P Nolte, Philippe Buchler, and Guoyan Zheng. X-ray image calibration and its application to clinical orthopedics. *Med Eng Phys*, 36(7):968–974, Jul 2014. 56

[218] Steffen Schumann, Benedikt Thelen, Steven Ballestra, Lutz-P Nolte, Philippe Büchler, and Guoyan Zheng. X-ray image calibration and its application to clinical orthopedics. *Medical Engineering & Physics*, 2014. 89

[219] Aleksandr Segal, Dirk Haehnel, and Sebastian Thrun. Generalized-icp. In *Robotics: Science and Systems*, volume 2, 2009. 131

[220] WP Segars, G Sturgeon, S Mendonca, Jason Grimes, and Benjamin MW Tsui. 4d xcat phantom for multimodality imaging research. *Medical physics*, 37(9):4902–4915, 2010. 185

[221] Boris Peter Selby, Georgios Sakas, Wolfgang-Dieter Groch, and Uwe Stilla. Patient positioning with x-ray detector self-calibration for image guided therapy. *Australas Phys Eng Sci Med*, 34(3):391–400, Sep 2011. 91

[222] BP Selby, Georgios Sakas, Stefan Walter, and Uwe Stilla. Geometry calibration for x-ray equipment in radiation treatment devices. In *1st Thematic conference on computational vision and medical image processing, Porto, Portugal*, pages 247–252, 2007. 56

[223] JamesA Shackleford, Nadya Shusharina, Joost Verberg, Guy Warmerdam, Brian Winey, Markus Neuner, Philipp Steininger, Amelia Arbisser, Polina Golland, Yifei Lou, Chiara Paganelli, Marta Peroni, Marco Riboldi, Guido Baroni, Paolo Zaffino, MariaFrancesca Spadea, Aditya Apte, Ziad Saleh, JosephO Deasy, Shinichro Mori, Nagarajan Kandasamy, and GregoryC Sharp. Plastimatch 1.6 - current capabilities and future directions. *Int Conf Med Image Comput Comput Assist Interv*, 15(WS), 10 2012. 144

[224] G.A. Shah and A.J. Buxton. Survey of dose associated with chest radiography in the hunter valley region of australia. *Radiography*, 5(1):23 – 27, 1999. 185, 189

[225] D Shahbazi-Gahrouei. Entrance surface dose measurements for routine x-ray examinations in chaharmahal and bakhtiari hospitals. *Iran J Radiat Res*, 4:29–33, 2006. 189

[226] Hamid R Sheikh and Alan C Bovik. Image information and visual quality. *IEEE Transactions on Image Processing*, 15(2):430–444, 2006. 195

[227] P.C Shrimpton, B.F Wall, and D Hart. Diagnostic medical exposures in the u.k. *Applied Radiation and Isotopes*, 50(1):261 – 269, 1999. 189

[228] R L Siddon. Prism representation: a 3d ray-tracing algorithm for radiotherapy applications. *Physics in Medicine and Biology*, 30(8):817, 1985. 144

[229] JH Siewerdsen, AM Waese, DJ Moseley, S Richard, and DA Jaffray. Spektr: A computational tool for x-ray spectral analysis and imaging system optimization. *Medical physics*, 31(11):3057–3067, 2004. 187

[230] Damijan Škrk, Urban Zdešar, and Dejan Žontar. Diagnostic reference levels for x-ray examinations in slovenia. *Radiology and Oncology*, 40(3), 2006. 189

[231] JA Soye and A Paterson. A survey of awareness of radiation dose among health professionals in northern ireland. 2014. 189

[232] I. I. Suliman and E. H. A. Elshiekh. Radiation doses from some common paediatric x-ray examinations in sudan. *Radiation Protection Dosimetry*, 132(1):64–72, 2008. 189

[233] IbrahimIdris Suliman and F.I. Habbani. Effective dose calculations in conventional diagnostic x-ray examinations for patients in major sudanese hospitals. In R. Magjarevic and J.H. Nagel, editors, *World Congress on Medical Physics and Biomedical Engineering 2006*, volume 14 of *IFMBE Proceedings*, pages 2087–2091. Springer Berlin Heidelberg, 2007. 189

[234] II Suliman, N Abbas, and FI Habbani. Entrance surface doses to patients undergoing selected diagnostic x-ray examinations in sudan. *Radiation protection dosimetry*, 123(2):209–214, 2007. 189

[235] SJ Supe, PS Iyer, JB Sasane, SG Sawant, and VK Shirva. Estimation and significance of patient doses from diagnostic x ray practices in india. *Radiation Protection Dosimetry*, 43(1-4):209–211, 1992. 189

[236] P. Taddei and V. Sequeira. X-ray and 3D data fusion for 3D reconstruction of closed receptacle contents. In *Conf. on 3D Vision*, 2013. 88

[237] F. Tahavori, M. Alnowami, J. Jones, P. Elangovan, E. Donovan, and K. Wells. Assessment of Microsoft Kinect technology (Kinect for Xbox and Kinect for Windows) for patient monitoring during external beam radiotherapy. In *Nuclear Science Symposium and Medical Imaging Conference*, pages 1–5, 2013. 116

[238] Markku Tapiovaara, M Lakkisto, and Antti Servomaa. Pcxmc. a pc-based monte carlo program for calculating patient doses in medical x-ray examinations. Technical report, Finnish Centre for Radiation and Nuclear Safety (STUK), Helsinki (Finland), 1997. 185

[239] Task Group 18 Imaging Informatics Subcommittee. Report no. 03: Assessment of display performance for medical imaging systems. Technical report, AAPM, 2005. 202

[240] Homer C. Tien, Lorraine N. Tremblay, Sandro B. Rizoli, Jacob Gelberg, Fernando Spencer, Curtis Caldwell, and Frederick D. Brenneman. Radiation exposure from diagnostic imaging in severely injured trauma patients. *Journal of Trauma and Acute Care Surgery*, 62(1), 2007. 189

[241] Jing Tong, Jin Zhou, Ligang Liu, Zhigeng Pan, and Hao Yan. Scanning 3D full human bodies using Kinects. *IEEE Transactions on Visualization and Computer Graphics*, 18:643–650, 2012. 122

[242] Mohammad Taghi Bahreyni Toosi and Mohsen Asadinezhad. Local diagnostic reference levels for some common diagnostic x-ray examinations in tehran county of iran. *Radiation protection dosimetry*, 2007. 189

[243] Georgia D Tourassi, Brian Harrawood, Swatee Singh, Joseph Y Lo, and Carey E Floyd. Evaluation of information-theoretic similarity measures for content-based retrieval and detection of masses in mammograms. *Medical Physics*, 34(1):140–150, 2007. 195

[244] Georgia D Tourassi, Rene Vargas-Voracek, David M Catarious Jr, and Carey E Floyd Jr. Computer-assisted detection of mammographic masses: A template matching scheme based on mutual information. *Medical Physics*, 30(8):2123–2130, 2003. 195

[245] Emanuele Trucco and Alessandro Verri. *Introductory techniques for 3-D computer vision*, volume 201. Prentice Hall Englewood Cliffs, 1998. 56

[246] Du-Yih Tsai, Yongbum Lee, and Eri Matsuyama. Information entropy measure for evaluation of image quality. *Digital Imaging*, 2007. 195

[247] R.Y. Tsai. *A Versatile Camera Calibration Technique for High Accuracy 3D Machine Vision Metrology Using Off-the-shelf Tv Cameras and Lenses*. Research report. IBM T.J. Watson Research Center, 1985. 62, 63

[248] CJ Tung, HY Tsai, SH Lo, CN Guan, and YB Chen. Determination of guidance levels of dose for diagnostic radiography in taiwan. *Medical physics*, 28(5):850–857, 2001. 189

[249] Nsikan U Esen, Rachel Obed, et al. Doses received by patients during thorax x-ray examinations. *Iranian Journal of Medical Physics*, 9(4):245–251, 2013. 189

[250] Esteban Vazquez-Fernandez, Angel Dacal-Nieto, Fernando Martin, and Soledad Torres-Guijarro. Image analysis and recognition. In *ICIAR International Conference*, 2010. 197

[251] Vishal Verma, Viswanath Vasudevan, Praveen Jinnur, Sasikanth Nallagatla, Ayanava Majumdar, Farhad Arjomand, and M Scott Reminick. The utility of routine admission chest x-ray films on patient care. *Eur J Intern Med*, 22(3):286–288, 2011. 184

[252] Franck P Vidal and Pierre-Frédéric Villard. Development and validation of real-time simulation of x-ray imaging with respiratory motion. *Computerized Medical Imaging and Graphics*, 49:1–15, 2016. 135

[253] Gonzalo Walwyn-Salas, Alina Martínez-González, Niurka González-Rodríguez, Alex Vergara-Gil, and Stefan Gutiérrez-Lores. Establecimiento y diseminación en cuba de patrones dosimétricos aplicados a radiología diagnóstica. *Boletín Científico Técnico INIMET*, (1):17–28, 2011. 189

[254] Ji-Xian Wang, Peter D Inskip, John D Boic, Ben-Xiao Li, Jing-Yuan Zhang, and Joseph F Fraumeni. Cancer incidence among medical diagnostic x-ray workers in china, 1950 to 1985. *International journal of cancer*, 45(5):889–895, 1990. 189

[255] Lejing Wang, Joerg Traub, Simon Weidert, Sandro Michael Heining, Ekkehard Euler, and Nassir Navab. Parallax-free intra-operative X-ray image stitching. *Med. Image Analysis*, 2010. 89

[256] Lejing Wang, Rui Zou, Simon Weidert, Jürgen Landes, Ekkehard Euler, Darius Burschka, and Nassir Navab. Modeling kinematics of mobile C-arm and operating table as an integrated six degrees of freedom imaging system. In Technische Universität München, editor, *Computer Aided Medical Procedures*, 2010. 89

[257] Zhou 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*, 2004. 196

[258] Laurie S Waters et al. Mcnpx user's manual. Technical report, Los Alamos National Laboratory, 2005. 186

[259] J.S. Wheat, S. Choppin, and A. Goyal. Development and assessment of a Microsoft Kinect based system for imaging the breast in three dimensions. *Medical Engineering and Physics*, 36(6):732–738, 2014. 116

[260] Michael A Wirth, Jay Narhan, and Derek WS Gray. Nonrigid mammogram registration using mutual information. In *Medical Imaging 2002*, pages 562–573. International Society for Optics and Photonics, 2002. 195

[261] R G Zamenhof, S Shahabi, and H T Morgan. An improved method for estimating the entrance exposure in diagnostic radiographic examinations. *AJR Am J Roentgenol*, 149(3):631–637, 1987. 185

[262] Zhengyou Zhang. A flexible new technique for camera calibration. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 22(11):1330–1334, Nov 2000. 63, 64, 77

[263] Kun Zhou, Qiming Hou, Rui Wang, and Baining Guo. Real-time kd-tree construction on graphics hardware. *ACM Transactions on Graphics (TOG)*, 27(5):126, 2008. 181

[264] Helmut Zollner and Robert Sablatnig. *Comparison of methods for geometric camera calibration using planar calibration targets.* Digital Imaging in Media and Education, 2004. 63

Publicaciones, premios y patentes

El trabajo expuesto en este documento de tesis está originalmente basado en el estudio y desarrollo de la siguiente patente, la cual establece las bases para la generación de imágenes densitométricas (descritas en el Capítulo 6 de este trabajo):

- F. Albiol, A. Albiol, A. J. Albiol *et al.*, *Device and method for obtaining densitometric images of objects by a combination of radiological systems and depth-sensing cameras*. Número de publicación: US20150222875A1. Número PCT: PCT/ES2013/070502. Fecha de publicación: 31 de julio, 2012.

Asimismo, durante el desarrollo de esta investigación se han obtenido las siguientes patentes:

- A. Corbi, F. Albiol, A. Albiol *et al.*, *Device for extracting three-dimensional information from X-ray images of an object, method for calibrating such device, and method for generating this type of X-ray images*. Número de publicación: WO2016156646A1. Número de solicitud: PCT/ES2016/070216. Fecha de publicación: 6 octubre, 2016.
- A. Corbi, F. Albiol, A. Albiol *et al.*, *Sistema y método para la identificación volumétrica e isotópica de distribuciones de escenas radioactivas*. Número de solicitud: P201730164. Número de referencia: ES1641.1226. Fecha de recepción: 10 febrero 2017.

El trabajo científico que da soporte a estas patentes son las siguientes publicaciones:

- F. Albiol, A. Corbi, and A. Albiol, Geometrical calibration of X-ray imaging with RGB cameras for 3D reconstruction, *IEEE Transactions on Medical Imaging*, vol. 35, pp. 1952–1961 (2016).
- F. Albiol, A. Corbi, and A. Albiol, Evaluation of modern camera calibration techniques for conventional diagnostic X-ray imaging settings, *Radiological Physics and Technology*, vol. 10, pp. 68–81 (2017).

- F. Albiol, A. Corbi, and A. Albiol, 3D measurements in conventional X-ray imaging with RGB-D sensors, *Medical Engineering and Physics*, vol. 42, pp. 73–79 (2017).

Como método de verificación de la calidad de la imagen densitométrica se ha publicado el siguiente resultado de investigación (resumido en el apéndice de este trabajo):

- A. Albiol, A. Corbi, and F. Albiol, Automatic intensity windowing of mammographic images based on a perceptual metric. *Medical Physics*, vol. 44, pp. 1369–1378 (2017).

Igualmente, han tenido lugar las siguientes publicaciones e intervenciones en congresos:

- A. Corbi, A. Albiol, and F. Albiol, Joint Calibration of RGB and X-ray Cameras, *Medical Engineering Physics Exchanges/Pan American Health Care Exchanges (GMEPE/PAHCE)*, 4-9 April, 2016.
- F. Albiol, A. Albiol, and A. Corbi, 3D Measurements from X-ray Images and Dense Surface Mappings, *Medical Engineering Physics Exchanges/Pan American Health Care Exchanges (GMEPE/PAHCE)*, 4-9 April, 2016
- A. Albiol, A. Corbi, and F. Albiol, Measuring X-ray image quality using a perceptual metric, *Medical Engineering Physics Exchanges/Pan American Health Care Exchanges (GMEPE/PAHCE)*, 4-9 April, 2016
- A. Albiol and A. Corbi, Towards Densitometric Imaging, *VI Jornadas CPAN*, 20-22 octubre, 2014.
- A. Corbi, Desarrollo de nuevos modelos de mejoras de detectores basados en imagen 3D, *VIII Jornadas CPAN*, 28 noviembre, 2016.
- A. Corbi, Incorporando la información 3D a la radiografía convencional, *XIII Jornadas CIPFP Ausiàs March*, 8 de febrero, 2017.
- I Jornadas RSEF/IFIMED de Física Médica, 11 de marzo, 2016 (asistente).

- Curso de control de instalaciones radioactivas en centros de investigación, Área de prevención de riesgos laborales, CSIC, 28 de junio, 2016 (asistente).

Se han recibido también los siguientes premios y galardones:

- A. Corbi, premio Alumni Plus al talento del Consejo Social de la Universidad de Valencia, diciembre de 2015. El galardón recibido reconoce la trayectoria general durante la ejecución del presente trabajo de tesis.
- A. Corbi y F. Albiol, VI premios CPAN a la divulgación en física de partículas, astropartículas y nuclear, categoría *experimentos*, octubre, 2015. El galardón recibido está relacionado con detección de radiación al mismo tiempo que, en lo que concierne a este trabajo de tesis, ha contribuido a la realización de pruebas para el estudio de dispositivos de disparo de rayos-X.

Por último, destacar que se ha complementado la labor de investigación con varios eventos de divulgación tanto en prensa como en centros públicos de enseñanza y se ha dirigido un proyecto de fin de carrera relacionado con la materia tratada (Juan José Rubio Guillamón, *Calibración de sistema radiográfico mediante sensores externos*, presentado el 24 de junio de 2015 en la Universidad Politécnica de Valencia).

Agradecimientos

Deseo agradecer muy especialmente la implicación y confianza de la empresa IST Storage S.L. (Burjassot) y de la Empresa Nacional de Residuos Radiactivos S.A. (ENRESA, Madrid) durante la duración de este trabajo de tesis. Agradezco también la amable y siempre dispuesta ayuda del Departamento de Radiofísica del Hospital Universitari i Politècnic La Fe (Valencia) mediante el préstamo de su fantoma antropomórfico en muchas ocasiones. De igual manera, quisiera reconocer el afecto y compañía

brindada por el grupo de espectroscopia gamma del IFIC. Para finalizar y desde estas últimas lineas, expreso mi más sincero reconocimiento a los compañeros y personal del IFIC que me han apoyado durante esta trayectoria.