Pedro Manuel Baptista Névoa

Pose Estimation using a C-Arm Fluoroscope

Instituto Superior de Engenharia do Porto
Universität zu Lübeck

2012

Pedro Manuel Baptista Névoa

BSc. degree in Computing Engineering and Medical Instrumentation
School of Engineering, Polytechnic Institute of Porto
Porto, Portugal

Pose Estimation using a C-Arm Fluoroscope

MSc. Thesis

submitted to the

University of Lübeck
Lübeck, Germany

Supervised by:

Prof. Dr. Alexander Schlaefer
Prof. Dr. Carlos Vinhais

under the ERASMUS Student Exchange Programme.

Instituto Superior de Engenharia do Porto
Universität zu Lübeck

2012

Acknowledgments

First of all, I acknowledge the support of my supervisors – Prof. Dr. Alexander Schlaefer (Institute for Robotics and Cognitive Systems, University of Lübeck) and Prof. Dr. Carlos Vinhais (Department of Physics, School of Engineering, Polytechnic School of Porto). Their continuous help and guidance motivated me to overcome challenges and problems, and to accomplish this thesis.

I highly appreciate the support from the scientific staff of the Institute for Robotics and Cognitive Systems of the University of Lübeck, the place where I developed this work. Among them, I would like to specially acknowledge Maximillian Heinig, Lars Richter, Fernando Gasca, Hamid Hajdar, Tobias Neumann and Bo Wang, who were always available to help me with different aspects related to the present project.

I am also grateful to Nuno Fernandes for his availability and concern with the success of this work; to Helge Schlegel for helping me with the writing of the german version of the abstract; to Lia Silva for sharing her wide knowledge on the field of radiology; and to Margarida Dinis for her huge support during the writing process and for helping me in the graphic component of the thesis.

Throughout the last year, I met different persons that inspired and helped me to be successful in this new experience. I wish to express my sincere gratitude to all the good friends I made during the last year, as well as to the older ones, for being such a positive influence in my life.

But above all, I would like to thank my parents and my sister for their unconditional support and encouragement throughout all my life.

Abstract

Clinical methods that are performed using image-related technologies, image-guided procedures, have been registering an increase of popularity in the last two decades. Traditional surgical methods have been replaced by minimally invasive methods in order to achieve a decrease in the overall costs and enhance productivity related aspects. Modern medical procedures such as bronchoscopy and cardiology are characterized by a focus on the minimization of invasive actions, with C-arm X-ray imaging devices having an important role in the field.

Although C-arm fluoroscope systems are a widely used technology to aid navigation in minimally invasive interventional events, they often lack on the information quality provided to the surgeon. The two-dimensional information that is obtained is not enough to provide full knowledge on the three-dimensional location of the region of interest, being the establishment of a method that is able to offer three- -dimensional information an essential task to be done.

A first step towards the accomplishment of this goal was took by defining a method that allows the estimation of the position and orientation (pose) of one object regarding the C-arm system. In order to run this tests with a C-arm system, its geometry had to be initially defined and the calibration of the system done.

The work developed and presented on this thesis focus on a method that proved itself to be sustainable and efficient enough to provide a solid basis to reach the main goal: the achievement of a technique that allows the improvement of the acquired information with a C-arm system during an intervention.

Zusammenfassung

Klinische Methoden, die Bild-Technologien verwenden, haben in den letzten zwei Jahrzehnten einen Anstieg an Popularität erfahren. Traditionelle chirurgische Methoden wurden von minimal-invasiven Methoden ersetzt, um einen Rückgang der Gesamtkosten zu erzielen und um die Aspekte der Produktivität zu steigern. Moderne medizinische Verfahren wie Bronchoskopie und Kardiologie werden durch einen Fokus auf die Minimierung von invasiven Aktionen charakterisiert. Hierbei spielt das C-Bogen-Röntgen-System eine entscheidende Rolle.

Obwohl das C-Bogen-Röntgensystem eine weit verbreitete Technologie zur Unterstützung der Navigation in den minimal-invasiven interventionellen Veranstaltungen ist, ist die Qualität der Informationen die dem Chirurgen bereitgestellt wird oft nicht ausreichend. Die erhaltene zweidimensionale Information ist nicht genug um vollständiges Wissen über die dreidimensionale Position des Interessenbereichs zu erlangen. Eine wesentliche Aufgabe ist die Einrichtung einer Methode, die in der Lage ist eine dreidimensionale Informationen zu bieten.

Ein erster Schritt zur Erreichung dieses Zieles wurde durch definieren einer Methode ermöglicht, die die Position und Orientierung des Objekts in Bezug auf das C-Bogen-Bildgebungssystem schätzt. Um diese Tests mit einem C-Bogen auszuführen, muss vorher die Geometrie definiert und das System kalibriert werden.

Die Arbeit in dieser These konzentriert sich auf eine Methode, die sich als tragfähig und effizient genug erweist, um eine solide Grundlage zu schaffen und das Hauptziel zu erreichen: die Verwirklichung einer Technik, die eine Verbesserung der erfassten Angaben mit einem C-Bogen während eines Eingriffs ermöglicht.

Resumo

Os métodos clínicos que são realizados com recurso a tecnologias de imagiologia têm registado um aumento de popularidade nas últimas duas décadas. Os procedimentos tradicionais usados em cirurgia têm sido substituídos por métodos minimamente invasivos de forma a conseguir diminuir os custos associados e aperfeiçoar factores relacionados com a produtividade. Procedimentos clínicos modernos como a broncoscopia e a cardiologia são caracterizados por se focarem na minimização de acções invasivas, com os arcos em 'C' a adoptarem um papel relevante nesta área.

Apesar de o arco em 'C' ser uma tecnologia amplamente utilizada no auxílio da navegação em intervenções minimamente invasivas, este falha na qualidade da informação fornecida ao cirurgião. A informação obtida em duas dimensões não é suficiente para proporcionar uma compreensão total da localização tridimensional da região de interesse, revelando-se como uma tarefa essencial o estabelecimento de um método que permita a aquisição de informação tridimensional.

O primeiro passo para alcançar este objectivo foi dado ao definir um método que permite a estimativa da posição e orientação de um objecto em relação ao arco em 'C'. De forma a realizar os testes com o arco em 'C', a geometria deste teve que ser inicialmente definida e a calibração do sistema feita.

O trabalho desenvolvido e apresentado nesta tese foca-se num método que provou ser suficientemente sustentável e eficiente para se estabelecer como um ponto de partida no caminho para alcançar o objectivo principal: o desenvolvimento de uma técnica que permita o aperfeiçoamento da qualidade da informação adquirida com o arco em 'C' durante uma intervenção clínica.

Contents

List of Figures

List of Tables

List of Abbreviations

2D	Two Dimensional
3D	Three Dimensional
CT	Computed Tomography
DICOM	Digital Imaging and Communications in Medicine
DOF	Degree of Freedom
FOV	Field of View
IGP	Image-Guided Procedure
ITK	Insight Segmentation and Registration Toolkit
LAO	Left Anterior Oblique
OID	Object to Image Distance
OR	Operating Room
MRI	Magnetic Resonance Imaging
PET	Positron Emission Tomography
POS	Pose from Orthography and Scaling
POSIT	POS with Iterations
RAO	Right Anterior Oblique
ROI	Region of Interest
SID	Source to Image Distance
SOD	Source to Object Distance
SPECT	Single-Photon Emission Computed Tomography

1 Introduction

Clinical procedures that are performed using image-related technologies have been manifesting an increase of popularity over the last years, with traditional surgical procedures being replaced by minimally invasive methods. Modern medical procedures such as bronchoscopy and cardiology are characterized by a focus on the minimization of invasive actions, with C-arm X-ray imaging devices being commonly used in this kind of interventions. However, the two-dimensional information obtained with this kind of method is insufficient to get a full understanding on the three-dimensional location of the region of interest. This problem can be solved if the position and orientation (pose) of one C-arm position is known regarding a second one. A first step towards the achievement of this goal was took by performing pose estimation tests with a simple and regular shaped object. In order to perform this estimation with a C-arm system, its geometry had to be initially defined and the calibration of the system performed.

1.1 Motivation

Throughout a clinical procedure, the information on the location of surgical instruments and anatomic parts can be obtained with C-arm X-ray imaging systems. Frequently, uncertainties occur in this tracking task due to the fact that the two-dimensional information offered by this kind of imaging system is not enough. This can occur as a result of overlap, vessel bifurcations and foreshortening [1, 2].

The surgeon's knowledge on the region of interest can be jeopardized, being thus desirable to have more information relative to the depth of the anatomic area where the intervention is happening. These improvements of a C-arm imaging system capabilities can be achieved by knowing the relation between different positions and orientations of the C-arm system. In order to perform the pose estimation, the system calibration arises as an inevitable initial step. The development of a method that could solve this problem would be of considerable interest for the field of IGPs.

1.2 Contributions

The goal of the presented work is to develop a method that allows the establishment of a connection between different C-arm positions and orientations, so that the location of the object of interest can be correctly determined. This thesis presents a first step given towards the accomplishment of the referred goal, being its main contributions following described:

- The test and establishment of a calibration procedure for X-ray images that defines the internal parameters of a C-arm system;
- The definition of a pose estimation method that relies in the information given by the geometry of the object of interest and its correspondent projections obtained by fluoroscopy.

Through the developed work, a sustainable and efficient method was proved to be valid; a method that can represent a solid basis to the achievement of the already mentioned main goal.

1.3 Outline of the Thesis

The work presented on this thesis was carried out in the Institute for Robotics and Cognitive Systems of the University of Lübeck (Universität zu Lübeck) in the time from October 2011 to July 2012 under the supervision of Prof. Dr. Alexander Schlaefer and supported by the ERASMUS student exchange programme. This thesis is organized as follows:

Chapter 2 defines Image-Guided Procedures (IGPs) based on the point of view of different authors, the methods and techniques that are used within the scope of this field, and the imaging modalities that allow the clinical use of image technology.

Chapter 3 reviews the essentials behind C-arm fluoroscope systems, such as the nuclear physics principles related with X-ray production and interactions. This chapter also analyses the technical specifications of the C-arm system that was used.

Chapter 4 focus on the calibration procedures of the C-arm system that were done in order to be possible to perform the pose estimation method (presented in Chapter 5). This chapter starts by presenting the adopted geometry for modeling the C-arm and then moves to the calibration parameters definition and results.

Finally, Chapter 6 presents the general conclusions about the developed work.

2 Image-Guided Procedures

Clinical procedures that are performed using image-related technologies have been registering an increase of usage over the last two decades, with minimally invasive methods replacing the traditional surgical procedures. Minimally invasive procedures have several benefits when comparing to the traditional ones, such as a decrease in overall costs and the achievement of better results [3, 4]. Image-guided procedures (IGP) use medical image processing methods (segmentation, registration, visualization and calibration, among others) in order to improve the understanding of anatomical structures in spatial and temporal contexts [5]. More on this field can be found in [6–12].

2.1 Overview

The traditional used methods for performance of surgery have been substituted by image-aided methods [14, 15]. For many procedures, there has been a striking reduction of the intrusiveness level, with modifications on the performance of specific operations and on the strategic approach to surgeries.

Galloway [16] considers that IGPs have four basic components: image acquisition, image-to-physical-space registration, three-dimension tracking, and display of imaging data and location. There are five types of images with which IGPs usually deal – projective images (plane films and angiograms, computed tomography, magnetic resonance imaging, single-photon emission computed tomography and positron emission tomography); intraoperative two-dimensional imaging (ultrasound); laparoscopic, endoscopic, and microscopic images; intraoperative tomograms; and processed images. Table 2.1 classifies some imaging devices respecting their intra-operative availability, accessibility, and data dimensionality.

Yaniv and Cleary [13] suggest the description of image-guided procedures using a time-line based view and divide it in three different phases: pre-operative planning, intra-operative plan execution, and post-operative assessment. The goal

Tab. 2.1: Classification of imaging devices respecting their availability for surgical use, their accessibility to physicians, the dimensionality of the data they acquire and the type of information offered by the images. *Reproduced from* [13].

modality	**availability**	**accessibility**	**dimensionality**
optical imaging	available	high	2D projection
X-ray fluoroscopy	available	high	2D projection
X-ray	available	high	2D projection
MRI	available	high	3D
CT	available	high	3D
US	available	high	2D
fMRI	not available	moderate	3D
PET	not available	moderate	3D
SPECT	not available	moderate	3D
C-arm CT	available	low	3D

of the first phase of image-guided procedures is to create a plan for the surgery that is to be performed. The surgical task is simulated based on patient data that can be obtained from different imaging modalities: computed tomography (CT), magnetic resonance imaging (MRI), or positron emission tomography (PET). The pre-operative planning leads to a dataset that is defined in a single coordinate system. The intra-operative plan execution consists in the alignment of the coordinate system in which the plan was defined to a coordinate system established in the operating room. The tools used on the surgery and the target anatomical structures are tracked and spatially related by the image-guided system. The initial procedure plan can also be modified during this stage. The last phase, post-operative assessment, compares the post-operative images to the pre-operative ones.

For Gering *et al.* [17], image-guided systems accomplish an enhancement of the surgeon's capability to use image technology to perform a surgery. These systems allow the following functions:

- Data analysis;
- Surgical planning;
- Surgical guidance;
- Surgical guidance with intra-operative updates.

Surgical guidance systems deal with pre-operative data, tracking of surgical instruments and correlation of the tracked devices along with the data. There could also

exist the need for morphological, cortical function or metabolic activity function information. These data sets are obtained in different coordinate systems and need to be registered to a common model, being this model also registered to the patient to perform surgical guidance. These methods allow the establishment of a relationship between the patient's and the surgical data's coordinate systems.

2.2 Methods and Techniques

The main methods and techniques used in IGPs can be defined as:

- Image acquisition and processing;
- Data visualization;
- Image registration;
- Tracking systems.

The following sections will briefly describe each one of these topics.

2.2.1 Image Acquisition and Processing

An image is a collection of measurements in two-dimensional (2-D) or three-dimensional (3-D) space [18, 19]. Medical images are the main source of information for image-guided procedures and their source is radiation absorption in X-ray imaging, acoustic pressure in ultrasound, or RF signal amplitude in MRI. The image type can be scalar (a single measurement is made at each location in the image) or multi-channel (more than one measurement is done at each location in the image).

For Galloway [16], IGPs usually deal with five different kinds of images:

- Projective images: plane films, tomographic sets (CT, MRI, SPECT, PET);
- Intraoperative two-dimension imaging;
- Endoscopic/microscopic images, intraoperative plane film imaging;
- Intraoperative tomograms: CT, MRI;
- Processed images: rendered images, tomographic angiograms.

The first medical IGP is reported by John Cox, Professor of Physics at McGill University in Montreal, who used X-rays as an aid to surgery performance [20], shortly after their discovery by Roentgen in 1895.

2.2.2 Data Visualization

Images used in IGPs usually have to be handled together with data coming from other types of source, *e.g.* information about tool locations or superimposed anatomical structures. Tomographic images can be visualized in one of three different ways: surface rendering, volume re-slicing, or direct volume rendering (DVR).

When dealing with images, the definition of pixel arises as an important concept. A pixel [21] is the rectangular region surrounding the pixel center that holds the data. The position of a pixel inside the image is identified by a unique index (array of integers that define the position of the pixel along each coordinate dimension of the image) and the size of a pixel (pixel spacing) is measured between the pixel centers. These centers are represented as circles and can differ along each dimension. The origin of the image is related to the coordinates of the first pixel in the image.

2.2.3 Image Registration

Registration is the process that finds a transformation from the coordinate system of one dataset to another, so that all the characteristics that appear in both data sets are aligned [22–24]. It is important in many aspects of functional image analysis, by realizing the spatial transformation between preoperatively taken images space and robot space [25, 26]. Table 2.2 shows a classification of medical registration methods. In many medical procedures, several images are acquired from subjects at different times, often with different imaging modalities, and there is the need of joining the asset of each image modality in order to gather more precise information on the matter to be studied.

In image-guided surgery, it is required to match the preoperative images and plans to the intraoperative situation, and to determine the relative positions of surgical tools and anatomical structures. By performing space registration, it is possible to transform surgical plan or preoperatively taken images to robot commands which move the robot to guide surgical tools. The position of the surgical tools can be converted to the preoperative space and displayed in real time [25, 27].

Tab. 2.2: Classification of medical image registration methods. *Based on* [23, 28].

dimensionality	2D/2D	2D/3D	3D/3D
nature of registration basis	extrinsic	intrinsic	non-image based
nature of transformation	rigid	affine	projective
domain of transformation	local	global	•
optimization procedure	computed	searched for	•
modalities involved	mono-modal	multi-modal	•

2.2.4 Tracking Systems

One of the challenges related to IGPs is the targeting of the position and orientation of tools and anatomical structures on a preoperatively or intraoperatively acquired image of the patient [29, 30]. This can be achieved using tracking systems. They allow the performance of this task for one single object at a time or for multiple objects simultaneously. Tracking devices can be based in ultrasonic, optical or electromagnetic properties. Ultrasonic, electromagnetic and optical methods manage to perform together with a mechanical system, while ultrasound and optical systems need an unobstructed line of sight between them and the target to be tracked. The signal detected by electromagnetic devices can suffer modifications from the presence of metallic devices in the neighborhood of the sensors. Most of the tracking operations are performed by optical systems that track fiducial markers attached to the patient's body [31]. For active optical tracking, the fiducials are often light-emitting diodes, *e.g.* infrared-emitting diodes (IREDs). Retroreflective spheres (RRSs) or disks are usually the fiducials used for passive optical tracking.

2.3 Imaging Modalities

Different imaging modalities are used to locate a target in the anatomy or to plan an intervention. The following sections give some detail about three important medical imaging modalities: diagnostic ultrasound imaging; magnetic resonance imaging (MRI); and X-ray imaging systems.

2.3.1 Diagnostic Ultrasound Imaging

An ultrasound is a mechanical vibration with a frequency above the range of human hearing (approximately 20 kHz) [33, 34]. For medical application, frequencies in

the range between 500 kHz and 100 Mhz are usually employed to make the display of soft tissues such as muscles and internal organs possible [35, 36].

Signals at these frequencies are aimed into body tissues, where they travel through a medium by means of the motions of the particles in the medium [37]. These waves cross various tissues and undergo different phenomena: absorption, refraction, reflection, and scattering. The phenomena occurring depends on the acoustic properties of the tissues, such as density and absorption coefficient, being the most important the phase velocity. Table 2.3 shows the acoustic properties of different tissues. After travelling through the different tissues, the reflected waves are received by the transducer of the ultrasound device and images are produced and analyzed. The differences between acoustic impedance values at an interface determines the amount of the reflected waves at the interface. A portion of gel is used on the surface of the sounder to minimize the amount of reflected waves at the air/tissue interface. The achieved spatial resolution depends on the ultrasound wavelength, while the resolution changes with the frequency (higher frequencies give superior resolution). The wavelength λ of a sinusoidal waveform is given by:

$$c = \lambda f \tag{2.1}$$

where f is the frequency of the wave and c is the velocity of propagation (phase velocity). The phase velocity depends on the propagation medium.

Ultrasound Transducers

The ultrasounds used for medical purposes are produced and detected with a transducer, composed of one or more ceramic elements with electromechanical (piezo-

Tab. 2.3: Acoustic properties of different materials. *Reproduced from* [32].

material	**c (m/s)**	α **(dB/MHz^y-cm)**	γ	ρ **(kg/m^3)**	Z
air	343	•	•	1.21	0.0004
bone	3360	3.54	0.9	1789	6.00
blood	1550	0.14	1.31	1039	1.61
fat	1450	0.6	1.0	952	1.38
liver	1570	0.45	1.05	1051	1.65
muscle	1580	0.57	1.0	1041	1.6545
water	1482.3	2.17×10^{-3}	2.0	1000	1.482

electric) properties [38]. The piezoelectric element has the function of converting electrical energy into mechanical energy in order to produce waves, or mechanical energy into electrical energy for ultrasonic wave detection. Piezoelectric crystals have a well-defined molecular arrangement of electrical dipoles. When mechanical pressure is applied to the crystal, it creates an electric potential difference across the crystal faces. Reciprocally, the crystal is also able to convert electrical to mechanical energy. The transducer is then used as a receiver or a transmitter of ultrasounds.

2.3.2 Magnetic Resonance Imaging

Magnetic resonance imaging (MRI) is a non-invasive, non-radioactive imaging technique that uses a magnetic field and pulses of radio wave energy to produce high quality images of structures of the human body [39]. MRI can be used to scan directly in axial, sagittal, coronal, or oblique planes [40]. The hardware components of a MRI scanner are a magnet; gradient coils/amplifier; and RF coils/transmission and reception system [41].

In Table 2.4 it is possible to see the characteristics of MRI–relevant isotopes. MRI takes advantage of the high amount of hydrogen atoms present in the human body. If hydrogen is placed in a static magnetic field, the magnetic moment of the atom spins at Larmor frequency:

$$\omega_0 = -\gamma B_0 \tag{2.2}$$

where ω_0 is the Larmor or resonant frequency and is proportional to the field; γ is the gyromagnetic ratio (constant specific to the nucleus, $\gamma/2\pi = 42.57$ MHz/T for the H^1 proton); and B_0 is the magnetic field strength. An hydrogen atom has one proton in its nucleus and as the nucleus spins a weak nuclear magnetic moment can be

Tab. 2.4: Characteristics of isotopes relevant to MRI. *Reproduced from* [32].

isotope	**spin state**	γ' **(MHz/T)**	**natural isotropic abundance** (%)	**sensitivity relative to** ^{1}**H**
^{1}H	1/2	42.58	99.99	100
^{13}C	1/2	10.71	1.11	1.6
^{19}F	1/2	40.05	100	3.4
^{23}Na	3/2	11.26	100	9.3
^{31}P	1/2	17.24	100	6.6

observed [42]. The nuclear spin of the tissue constituent atoms is used as the origin of the signal for MR imaging. Some of the advantages of MRI are its high resolution performance, the low risk and non-invasiveness; while the main disadvantages are the long duration of the exam, the high costs and the existence of high magnetic fields. MR images have high application in the field of neurosurgery, where they can serve as a basis for multi-subject analyses of brain structure and function; pre-processing surface modeling; surgical planning; and brain registration [43].

2.3.3 X-Ray Imaging Systems

X-rays were discovered by Roentgen in 1895 while studying cathode rays (stream of electrons) in a gas discharge tube [40]. Roentgen observed that this kind of radiation had the ability to get through opaque matter, produce fluorescence and ionize a gas. An X-ray imaging system has an X-ray source, a collimator, and an X-ray detector. The energy of the generated X-rays is between 20 and 150 keV [32]. An X-photon production system consists in an electronic vacuum tube with a cathode (negative electrode) and an anode (positive electrode). Figure 2.1 shows a schematic representation of an X-ray tube.

The cathode is a tungsten filament that ejects electrons when heated (thermionic emission). The atomic electrons are released from their atoms where their energy is high enough due to heating. The anode is made of tungsten, molybdenum, or copper. A high voltage is applied between the anode and the cathode, with the electrons generated at the cathode being accelerated toward the anode. When the electrons hit the atoms of the anode, X-ray energy is produced by one of two possible interactions: bremsstrahlung or characteristic X-rays.

In the first, bremsstrahlung (also known as braking radiation), the accelerated electron passes close to the nucleus of the atom and is deflected by the action of

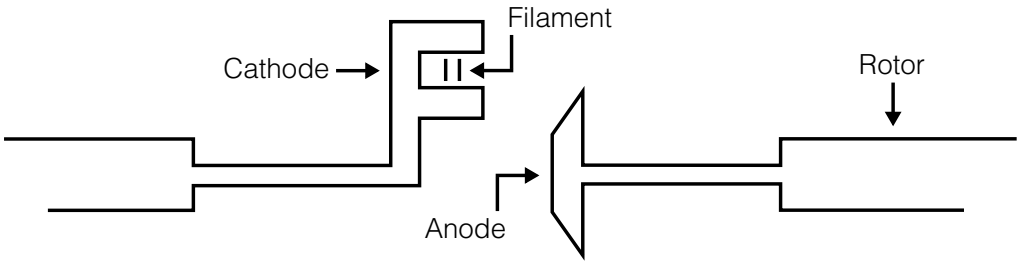


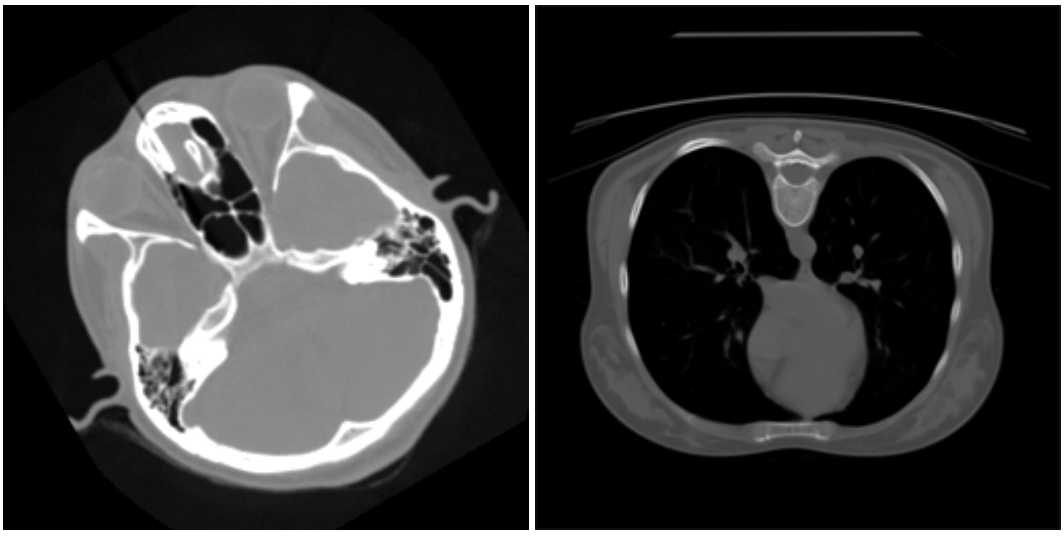
Fig. 2.1: Schematic representation of an X-ray tube. *Redrawn from* [44].

Coulomb forces of attraction, resulting in a change of direction and loss of energy by the electron. Characteristic X-rays are created when an electron interacts with the atoms of the target. An electron from an orbital shell is ejected, leaving the atom ionized. An electron from a more distant orbital fills the vacancy left open by the original electron, with energy being radiated in the form of electromagnetic radiation. This radiation is known as characteristic radiation (X-rays) because it is characteristic of the atoms in the target and of the shells between which the transitions took place.

Computed Tomography

Computed tomography allows the generation of a patient-specific attenuation coefficients map; it was developed in 1971 by Dr. G. N. Hounsfield in England and used to image the brain [46]. CT systems consist in three main components: a gantry (that includes the X-ray source/detector and the data-acquisition system), a data handling unit, and a storage facility.

A beam of X-rays goes across a patient in synchrony with a radiation detector that stands on the opposite side of the patient. When crossing the different tissues present in the human body, the X-rays are attenuated and a number of measurements for different orientations is done by the detector. A reconstruction algorithm assigns different levels (CT numbers) to different attenuation coefficients and an



(a) (b)

Fig. 2.2: Computed tomography images: (a) head; and (b) thorax. *Reproduced from* [45].

image is formed (Figure 2.2). The CT numbers can be established as Hounsfield numbers (H) – between -1000 (for air) and +1000 (assigned to bone), with water being defined with value 0. The computation of a Hounsfield number is performed by the following equation:

$$H = \frac{\mu_{tissue} - \mu_{water}}{\mu_{water}} \times 1000 \tag{2.3}$$

where μ is the linear attenuation coefficient. From this, the following conclusion arises: one Hounsfield unit is a change of 0.1% in water's attenuation coefficient.

2.3.4 Comparison of Imaging Modalities

The previously described modalities – diagnostic ultrasound imaging, magnetic resonance imaging (MRI), and X-ray imaging systems (CT included) – can be compared in order to choose the best option for the given case. Table 2.5 [32] reflects some of the aspects that can be compared in terms of what is imaged, access, spatial resolution, penetration, safety, operational speed, associated costs, and portability.

Tab. 2.5: Comparison of imaging modalities. *Reproduced from* [32].

	ultrasound	**x-ray**	**CT**	**MRI**
what is imaged	mechanical properties	mean tissue absorption	tissue absorption	biochemistry (T1 & T2, & PD)
access	small windows adequate	2 sides needed	circumferential around body	circumferential around body
spatial resolution	3 to 0.3 mm	≈1mm	≈1mm	≈1mm
penetration	3 to 25 cm	excellent	excellent	excellent
safety	very good	ionizing radiation	ionizing radiation	very good
speed	100 fps	minutes	0.5 min – 1 min	minutes
cost	$	$	$$$$	$$$$$$$$
portability	excellent	good	poor	poor

3 C-Arm Fluoroscope Systems

C-arm systems are a widely used technology to aid navigation in minimally invasive interventional events such as bronchoscopy, making use of X-ray principles to provide images from structures of the human body. After a more general introduction to image-guided procedures done in the previous chapter, this chapter describes in a more detailed way C-arm fluoroscope systems and its operating principles (such as X-ray production and interactions). In the last section, an analysis of the technical specifications of the model that was used (Philips BV Endura) is done.

3.1 Overview

Fluoroscopic C-arms are devices that allow the acquisition of images during a surgery [47]. They are commonly used to provide real-time feedback and intervention guidance in minimally invasive interventional procedures related with bronchoscopy, cardiology, and orthopedics.

C-arm systems offer real-time X-ray images and are the most popular technology in the field of interventional X-ray imaging, due to their high versatility, flexibility and mobility [48]. A C-arm fluoroscope consists in [49]:

- an X-ray tube/generator/control;
- a ‘C’ shaped arc;
- an image intensifier;
- a control panel with a footswitch;
- a computerized image display system.

The tube voltage, tube current, and irradiation time define the contrast perceptibility and dose, and are regulated by the X-ray tube, the generator, and the control

Tab. 3.1: Radiologic units. *Reproduced from* [52].

quantity	**description**	**conventional unit**
fluence	number of photons per unit area	m^{-2}
flux (fluence rate)	fluence per unit time	$m^{-2}s$
intensity (energy fluence)	number of photons times photon energy per unit area	$keVJm^{-2}$
exposure (X)	charge produced per unit mass of air from X and gamma -rays	roentgen (CKg^{-1})
kerma (K)	kinetic energy released in matter per unit mass	rad (JKg^{-1} *or* gray)
dose (D)	energy absorbed per unit mass	rad (JKg^{-1} *or* gray)

system [50]. An X-ray beam crosses the patient and is detected by the image intensifier, being the image displayed on a TV screen. There are some negative aspects associated with the use of fluoroscopy [51], such as: image distortion; registration to other imaging modalities; calibration of specific model parameters; and need for pose estimation and correlation when using more than one image. The calibration problem is addressed in Chapter 4, while Chapter 5 focus on the estimation of the position and orientation for an image taken with a C-arm system. The following sections cover the operational principles of a C-arm system and its specifications.

3.2 X-Ray Basic Physics

As other medical imaging systems, C-arm fluoroscopes quantify the transmission of X-rays through the human body. This section covers the fundamentals of X-ray production, its interactions with matter, and the attenuations that occur while they are crossing the human body. The application of X-ray physics principles to the clinical use of C-arm system is also discussed.

3.2.1 Quantities and Units

Some definitions and units are used in radiologic physics to express the properties of ionizing radiation (X and gamma -rays). In Table 3.1 it is possible to see some radiologic definitions, that are important for the understanding of this sec-

tion. The following definitions of flux, energy flux, fluence, energy fluence, fluence rate, energy-fluence rate, particle radiance, and energy radiance were taken from the report *Fundamental Quantities and Units for Ionizing Radiation*, ICRU[1] [53].

Flux and Energy Flux

The flux, N, is the quotient of dN by dt, where dN is the increment of the particle number in the time interval dt:

$$N = \frac{dN}{dt}, \tag{3.1}$$

being N expressed in s^{-1}. The energy flux, R, is the quotient of dR by dt, where dR is the increment of radiant energy in the time interval dt:

$$R = \frac{dR}{dt}, \tag{3.2}$$

being R expressed in W. These definitions are normally used for limited spatial regions (the flux of particles emerging from a collimator). For source emission, the flux of particles emitted in all directions is generally considered.

For visible light and related electromagnetic radiations, the energy flux is defined as power emitted, transmitted, or received in the form of radiation and termed radiant flux/power.

Fluence and Energy Fluence

The fluence, ϕ, is the quotient of dN by da, where dN is the number of particles incident on a sphere of cross-sectional area da:

$$\phi = \frac{dN}{da}, \tag{3.3}$$

being ϕ expressed in m^{-2}. The energy fluence, ψ, is the quotient of dR by da, where dR is the radiant energy incident on a sphere of cross-sectional area da:

$$\psi = \frac{dR}{da}, \tag{3.4}$$

being ψ expressed in Jm^{-2}. The use of cross-sectional area da comes from the fact that the area da is considered to be orthogonal to the direction of each particle.

[1] ICRU: International Commission on Radiation Units and Measurements (`http://icru.org/`)

In cases where the radiation interactions are independent of the direction of the incoming particles, fluence and energy fluence are used. In dosimetric calculations, fluence is frequently expressed in terms of the lenghts of the particle trajectories. The fluence, ϕ, can be given by:

$$\phi = \frac{dl}{dV}, \tag{3.5}$$

where dl is the sum of lengths of particle trajectories in the volume dV. For a radiation field that is constant over a time interval, t, and composed of particles with velocity v, the fluence, ϕ, is given by:

$$\phi = nvt, \tag{3.6}$$

where n is the particle number density. The distributions, ϕ and ψ, of the fluence and energy fluence with respect to energy are given by:

$$\phi_E = \frac{d\phi}{dE}, \quad \psi_E = \frac{d\psi}{dE}, \tag{3.7}$$

where $d\phi$ is the fluence of particles of energy between E and $E + dE$, and $d\psi$ is their energy fluence. The relationship between the two distributions is given by

$$\psi_E = E\phi_E. \tag{3.8}$$

The energy fluence is related to the quantity radiant exposure defined, for fields of visible light, as the quotient of the radiant energy incident on a surface element by the area of that element. When a parallel beam is incident at an angle θ with the normal direction to a given surface element, the radiant exposure is equal to $\psi\cos\theta$.

Fluence Rate and Energy-Fluence Rate

The fluence rate, ϕ, is the quotient of $d\phi$ by dt:

$$\phi = \frac{d\phi}{dt}, \tag{3.9}$$

where $d\phi$ is the increment of the fluence in the time interval dt. ϕ is expressed in $\mathrm{m}^{-2}\mathrm{s}^{-1}$. The energy-fluence rate, ψ, is the quotient of ψ by dt, where $d\psi$ is the

increment of the energy fluence in the time interval dt:

$$\psi = \frac{d\psi}{dt}, \tag{3.10}$$

where ψ is expressed in Wm^{-2}. For a radiation field composed of particles of velocity v, the fluence rate, ϕ, is given by:

$$\phi = nv, \tag{3.11}$$

where n is the particle number density.

Particle Radiance and Energy Radiance

The particle radiance, ϕ_Ω, is the quotient of $d\phi$ by $d\Omega$, where $d\phi$ is the fluence rate of particles propagating within a solid angle $d\Omega$ around a specified direction:

$$\phi_\Omega = \frac{d\phi}{d\Omega}, \tag{3.12}$$

where ϕ_Ω is expressed in $\mathrm{m}^{-2}\mathrm{s}^{-1}\mathrm{sr}^{-1}$. The energy radiance, ψ_Ω, is the energy fluence rate of particles propagating within a solid angle $d\Omega$ around a specified direction:

$$\psi_\Omega = \frac{d\psi}{d\Omega}, \tag{3.13}$$

where ψ_Ω is expressed in $\mathrm{Wm}^{-2}\mathrm{sr}^{-1}$. In a spherical coordinate system with polar angle θ and azimuthal angle φ, $d\Omega$ is equal to $sin\theta d\theta d\varphi$. For visible light and related electromagnetic radiations, the particle radiance and energy radiance are termed photon radiance and radiance, respectively. The distributions of particle radiance and energy radiance respect to energy are given by:

$$\phi_{\Omega,E} = \frac{d\phi_\Omega}{dE}, \quad \psi_{\Omega,E} = \frac{d\psi_\Omega}{dE}, \tag{3.14}$$

where $\mathrm{d}\phi_\Omega$ is the particle radiance for particles of energy between E and $E + dE$, and $\mathrm{d}\psi_\Omega$ is their energy radiance. The two distributions are related by:

$$\psi_{\Omega,E} = E\phi_{\Omega,E}. \tag{3.15}$$

The quantity $\phi_{\Omega,E}$ is sometimes termed angular flux or phase flux in radiation-transport theory. Apart from aspects that are of minor importance in the present

context (*e.g.* polarization), the field of any radiation of a given particle type is completely specified by the distribution, $\phi_{\Omega,E}$, of the particle radiance with respect to particle energy, as this defines number, energy, local density, and arrival rate of particles propagating in a given direction. This quantity, as well as the distribution of the energy radiance with respect to energy, can be considered as basic in radiometry.

3.2.2 X-Ray Production

The process of X-ray production happens when highly energetic electrons interact with matter and convert their kinetic energy into electromagnetic radiation [52]. This type of radiation is composed by bundles of energy called photons, whose energy is defined by:

$$E = h\upsilon = h\frac{c}{\lambda}, \tag{3.16}$$

and its speed by:

$$c = \lambda\upsilon, \tag{3.17}$$

where h is the Planck's constant and c, λ, and υ are the speed, wavelength, and frequency of the radiation. The energy of a photon is expressed in electron volt. The electron volt is defined as the energy required to move one electron against a potential difference of one volt [54]. It is equivalent to 1.6×10^{-19} J. Electromagnetic radiation has the properties of a wave where the electric and magnetic field components move perpendicularly to the direction of the wave propagation. Figure 3.1 shows a schematic representation of the electric and magnetic field components of electromagnetic radiation.

Electromagnetic radiation has the ability of travelling in vacuum; it does not need a medium to conduct it. The speed of electromagnetic radiation has a constant value (c = 3.8 $\times$ 10^8 ms^{-1}). In vacuum, no matter what is its wavelength or frequency, electromagnetic radiation always travels at this speed. The production of X-rays happens in an X-ray tube, which consists in a high vacuum glass envelope. This tube has an anode and a cathode, with a power source attached to it. The cathode is negatively charged respecting the anode. The cathode is responsible for thermionic emission, *i.e.*, the raise of atomic electrons energy to a level that is sufficient for them to release themselves from the atom. After they are released from

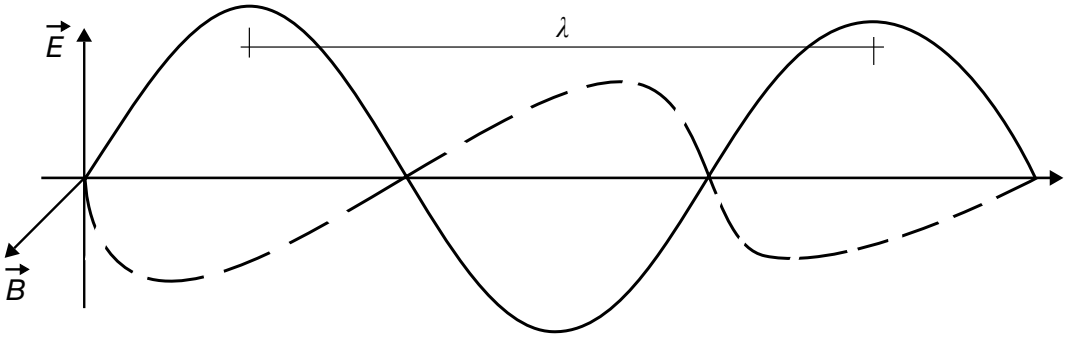


Fig. 3.1: Schematic representation of the electric ($\vec{E}$) and magnetic ($\vec{B}$) field components of electromagnetic radiation. λ: Wavelength. *Redrawn from* [52].

the cathode, the electrons are attracted by the positive pole of the tube (the anode). Due to the potential difference between the cathode and the anode, the electrons are accelerated toward the anode and gain kinetic energy. When the electrons arrive to the anode (the target), their kinetic energy is lost by three different phenomena: excitation, ionization, and radiation. Excitation is the process by which electrons move to more external orbital shells (higher energy states) by receiving energy from a charged particle. In ionization, a neutral atom acquires a positive or a negative charge. If the energy received by the incident particle is high enough, an electron from the atom is ejected and the atom becomes positively charged. In radiation, the energy lost from the charged particle is directly used to create a photon. X-rays are produced by two different mechanisms: bremsstrahlung and characteristic X-rays.

Bremsstrahlung

X-rays produced by radiation are known as bremsstrahlung (or braking radiation). Bremsstrahlung is the mechanism that happens when the accelerated electron passes near an atom of the target material and is deflected from its path by the action of Coulomb forces. After this deflection, the electron loses part or all of its energy, being it propagated as electromagnetic radiation. The direction of emission of bremsstrahlung photons depends on the energy of the incident photons. The bremsstrahlung is responsible for the continuous component of the X-ray spectra.

Characteristic X-rays

The accelerated electron can lose its energy through ionization, giving origin to a characteristic X-ray as a subsequent emission of a photon when the electronic va-

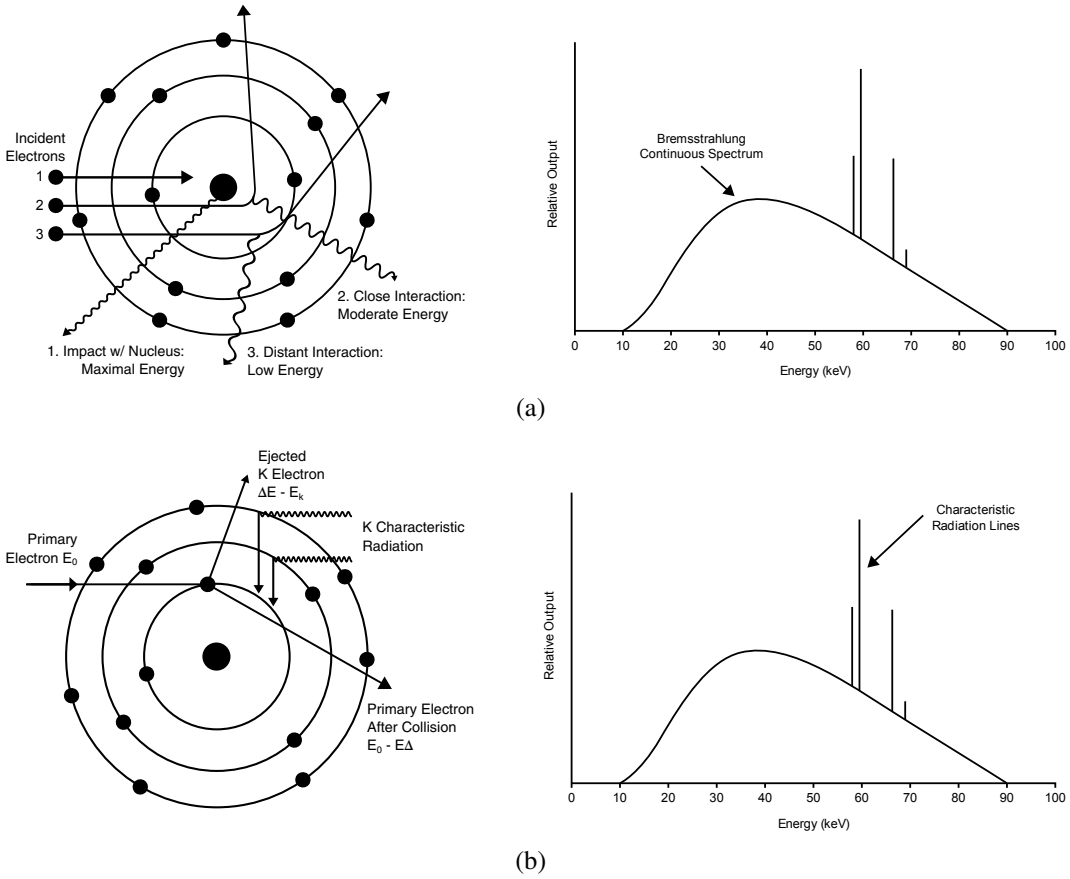


Fig. 3.2: X-ray production mechanisms: (a) bremsstrahlung; and (b) characteristic X-rays. The X-ray spectrum with the identified range of energies for each mechanism can be seen (for a tube potential of 90 kVp). *Redrawn from* [40, 52].

cancy is filled. In this situation, an incident electron causes the ejection of one orbital electron (K, L, or M electron). After this interaction, the atom is ionized and a vacancy is created in one of the orbits. An outer electron will fall to this new vacancy, causing the emission of electromagnetic radiation. This is called characteristic radiation due to the fact that each element has electronic orbitals of characteristic energy. Its wavelength is given by Planck's Law:

$$\lambda = h\frac{c}{E}, \tag{3.18}$$

where h is the Planck's constant; c is the speed of electromagnetic radiation; and E is the characteristic energy of the photon. Table 3.2 refers to the major characteristic transitions, respective transitions and X-ray energies produced in a tungsten target.

Tab. 3.2: Characteristic X-ray energies for Tungsten. *Reproduced from* [40].

series	lines	transition	energy (keV)
K Series	$K\beta_2$	$N_{III} - K$	69.09
	$K\beta_1$	$M_{III} - K$	67.23
	$K\alpha_1$	$L_{III} - K$	59.31
	$K\alpha_2$	$L_{II} - K$	57.97
L Series	$L\gamma_1$	$N_{IV} - L_{II}$	11.28
	$L\beta_2$	$N_{IV} - L_{III}$	9.96
	$L\beta_1$	$M_{III} - L_{II}$	9.67
	$L\alpha_1$	$M_{II} - L_{III}$	8.40
	$L\alpha_2$	$M_{IV} - L_{III}$	8.33

3.2.3 Interactions of Photons with Matter

As photons pass through matter, they interact with atoms. The type of interaction that occurs depends on the energy of the photons and the chemical composition of matter (atomic number of elements present in matter). Three different interactions can occur: photoelectric effect; Compton scattering; and pair production (not relevant for nuclear medicine, it only occurs for photon energies above 1020 keV).

Compton Scattering

In Compton scattering, an incident photon transfers part of its energy to valence shell electrons (which binding energy value is much less than the energy of the incident photon) [55]. The following equations express the collision between the two particles (electron and photon) that happens in Compton scattering:

$$E = h\nu_0 \frac{\alpha(1-\cos\phi)}{1+\alpha(1-\cos\phi)}, \tag{3.19}$$

$$h\nu' = h\nu_0 \frac{1}{1+\alpha(1-\cos\phi)}, \tag{3.20}$$

$$(1+\alpha)\tan\frac{\phi}{2}, \tag{3.21}$$

where $h\nu_0$, $h\nu'$, and E are the energies of the incident photon, scattered photon, and electron; and $\alpha = h\nu_0/m_0c^2$, where m_0c^2 is the rest energy of the electron (0.511 MeV). If $h\nu_0$ is expressed in MeV, then $\alpha = h\nu_0/0.511$. The electron receives energy from the photon and, according to the energy that is transferred from the

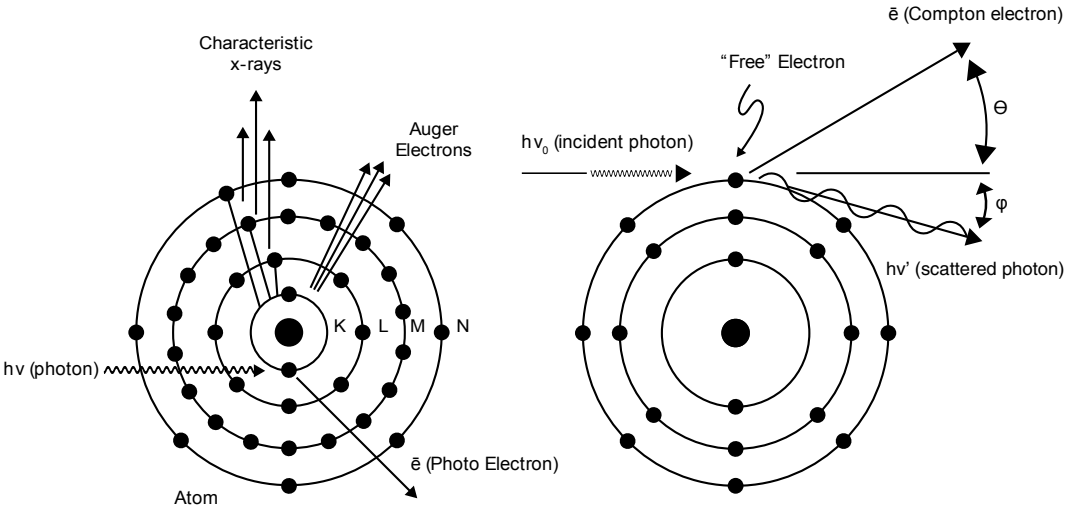


Fig. 3.3: Schematic representation of the phenomena that happens when photons interact with matter: photoelectric effect and Compton scattering. *Redrawn from* [40].

photon to the electron, it is scattered at a determined angle that can range from nearly 0° to 180°. It is the dominant process for low atomic number materials, like soft tissue ($Z = 7.5$) and bone ($Z = 13$).

After being ejected, the electron loses its kinetic energy by excitation and ionization caused by the interactions with the atoms of surrounding material. The Compton scattered photon may traverse through the medium and go through new interactions: Compton scattering, photoelectric absorption, or Rayleigh scattering.

Photoelectric Effect

The photoelectric effect is characterized by the interaction between a γ particle (low energy) and an orbital electron, to which the γ particle transfers its energy (usually, to an orbital electron located in an inner-shell of the atom). The vacancy left open by the ejected electron is filled by an outer-shell electron, leaving the atom in an excited state. The excess energy, *i.e.*, the difference between the incident gamma ray energy and the binding energy of the inner-shell electron,

$$E_{e^-} = E_\gamma - E_b, \tag{3.22}$$

is released as X-ray radiation (photoelectron). E_{e^-}, E_γ, and E_b stand for the energy of the photoelectron, the incident gamma ray, and the inner-shell electron binding. The photoelectric effect is the dominant process for high atomic number materials.

3.2.4 Beam Attenuation

In its way through the human body from the source to the detector, the X-ray beam (or stream of photons) suffers attenuation – modifications in the intensity of the beam caused by absorption (the energy of the incident photons is transferred to the atoms of the target material) or scattering (photons deflection) [56]. Depending on the X-ray beam and on the characteristics of the human tissues, the photons bundle gets attenuated. This attenuation is considered as an exponential function given by:

$$\frac{I_{out}}{I_{in}} = e^{-(\nu x)}, \tag{3.23}$$

where x is the thickness of the attenuator (in cm) and ν is the linear attenuation coefficient (Table 3.3), expressed in cm^{-1}. If N is considered as the photon flux of the incident beam, the number of photons interacting with particles of matter and removed from beam, in a layer of thickness ds, is given by:

$$dN = -\nu N ds. \tag{3.24}$$

The linear attenuation coefficient is higher for dense tissue (*e.g.* bone) than for soft tissue. It is given by:

$$\nu = \frac{\Delta N}{N \Delta x}, \tag{3.25}$$

where ΔN is the number of photons removed from the X-ray beam in thickness Δx.

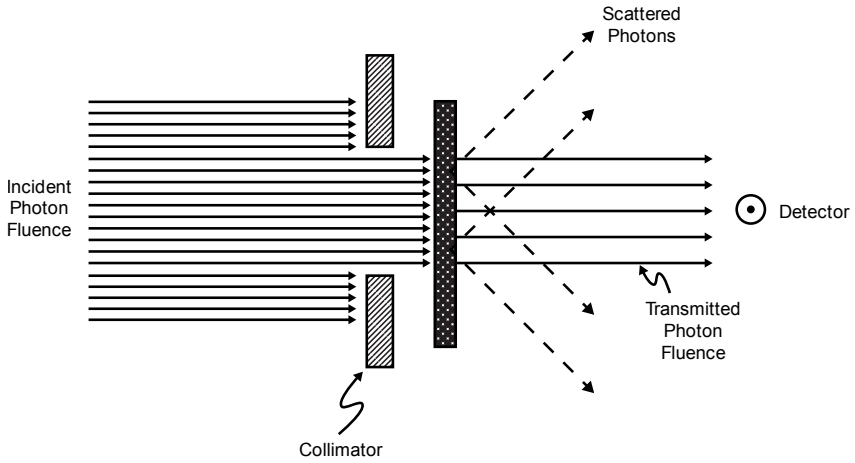


Fig. 3.4: Schematic representation of absorption and scattering mechanisms happening to an X-ray beam. *Redrawn from* [40].

Tab. 3.3: Linear Attenuation Coefficients. *Reproduced from* [56].

material	**atomic number** (Z)	**density** ($\mathbf{gcm^{-3}}$)	**linear attenuation coefficient** ($\mathbf{cm^{-1}}$)
water	7.4	1.0	0.214
ice	7.4	0.917	0.196
water vapor	7.4	0.000598	0.000128
compact bone	13.8	1.85	0.573
air	7.64	0.00129	0.00029
fat	5.92	0.91	0.193

The mass attenuation coefficient[2] is the rate of photon interactions per unit area mass and is independent of the physical state of the material. It is given by the linear attenuation coefficient divided by the attenuator density and usually expressed in gram per centimeter squared (cm^2g^{-1}).

3.3 C-Arm System Description

This section describes the properties, applications, and components of the C-arm system that was used to do the image acquisition. The X-rays images used to perform the pose estimation were obtained using a Philips BV Endura.

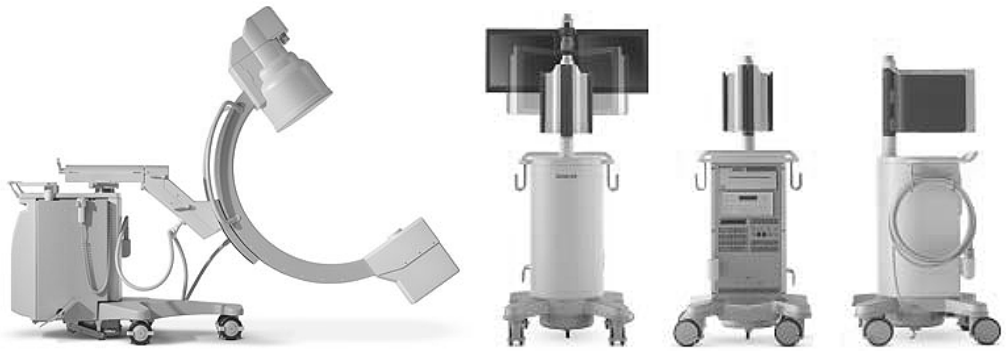
3.3.1 Applications and Components

The BV Endura is a mobile X-ray system for the acquisition and visualization of diagnostic X-ray images. It is intended for medical use in operating rooms. The system is designed to perform in the following medical areas:

- Neurosurgery;
- Orthopedic, abdominal and thoracic surgery;
- Vascular procedures.

The system consists in two main components: the C-arm stand and the image viewing station. Figure 3.5 shows the C-arm stand (a) and the image viewing station (b)

[2]The Physical Measurement Laboratory of the National Institute of Standards and Technology (NIST) offers a detailed lecture in X-ray mass attenuation coefficients. It can be accessed at `http://nist.gov/pml/data/xraycoef/index.cfm`



(a) (b)

Fig. 3.5: Philips BV Endura: (a) 'C' shaped arc; and (b) computerized image display system. *Reproduced from* [57].

of the Philips BV Endura. The X-ray source is equipped with a durable solid anode. The X-ray collimator restricts to the actual field of view of the image intensifier. A shutter made of solid lead can be independently moved and rotated in order to prevent direct exposure of the image intensifier and to reduce scattered radiation. An additional radiation filter reduces the surface dose to the patient by 40% when compared to conventional filters. The image intensifier has a compact CCD television camera and a carbon fiber X-ray grid to reduce the scattered radiation. Two image intensifier modes are available: 23/17/14 cm (9/7/5 inches) and 31/23/17 cm (12/9/7 inches). A handset and a foot set connected to the C-arm allow the activation of the operation mode control. The mechanical movements of the C-arm support are fully compensated. The height adjustment is motorized. The control handles are coupled to each other and control the rear wheels. The front wheel can move freely. The removable cartridge fits on the entrance window of the image intensifier. The size of the X-ray beam can be selected on the operator console. The cassette holder allows a rotation of 360° and the following dimensions: 24×24 cm; 4×30 cm.

3.3.2 Technical Data and Calibration

The technical data of Philips BV Endura can be found in Appendix A. In order to perform the method described in Chapter 5, the system had to be calibrated. The next chapter covers the procedures associated with the calibration of the system.

4 C-Arm Geometry and Calibration

In order to use a C-arm system to estimate the position and orientation of an object, its geometry has to be initially defined and the calibration of the system performed. A C-arm system can be regarded as a pinhole camera [58, 59]. The camera calibration has to be done in order to characterize the internal geometry of the camera and its optical properties [60, 61]. For a more detailed treatment on camera geometry and calibration, the reader is referred to [33, 62, 63].

4.1 C-Arm Systems Geometry

The definitions that describe the orientation of the C-arm and the distances between its main components have to be taken into account [58, 64, 65]:

- RAO/LAO (right anterior oblique/left anterior oblique) angle: rotation of the C-arm regarding the patient's right and left side;
- CRAN/CAUD (cranial/caudal) angle: rotation of the C-arm regarding the patient's head (cranial) or feet (caudal) direction;
- SID (source-to-image distance): distance between the X-ray emitting source and the radiation detector;
- SOD (source-to-object distance): distance between the X-ray emitting source and the object that is being scanned;
- OID (object-to-image distance): distance between the object that is being scanned and the radiation detector.

In Figure 4.1 it is possible to see a schematic representation of a C-arm system with its geometric definitions. Since X-ray techniques have been used as a complement to IGPs, research in the field of X-ray imaging has been done to solve the problem

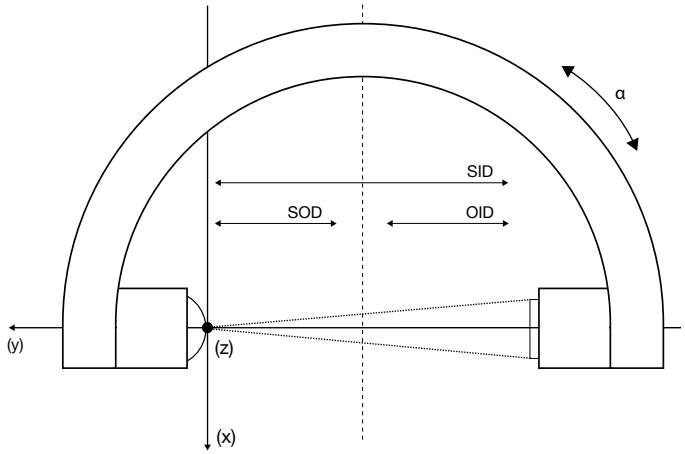


Fig. 4.1: Schematic representation of a C-arm with its geometric definitions. SID: source-to-image distance; SOD: source-to-object distance; OID: object-to-image distance; α: RAO/LAO angle.

of obtaining the X-ray source position and image plane orientation [66]. The pinhole model camera can be used as a geometric model for the C-arm and projection matrices can be computed in order to find the 3D point–projection image map.

4.1.1 Projective Geometry

Euclidean geometry describes the properties of objects that are unchanged by rigid motions: lengths of objects, angles of intersecting lines, parallelism of two lines that lie in the same plane and never meet [67, 68]. In camera-related operations, Euclidean geometry is not suitable because points at infinity cannot be considered, the definitions of lengths and angles are not valid, and the intersection of parallel lines can occur. For these reasons, when it comes to camera geometric modeling, another subset is used: projective geometry. Projective geometry is the study of geometric properties that are invariant under projective transformations, allowing a much larger class of transformations than Euclidean geometry. Table 4.1 shows the relation between four different geometries (projective, Euclidean, similarity and affine) by describing the transformations allowed in each, and the measures that remain invariant under those transformations. Projective geometry deals with points, lines and their relations in 2-D and 3-D [62].

In Euclidean space [68], a point can be represented by $(X,Y,Z)^T$. Moreover, projective space uses homogeneous coordinates to define a point. The homogenous coordinates associated with a point in a projective space of dimension n are typically

Tab. 4.1: The four referred geometries, the transformations allowed in each, and the measures that remain invariant under those transformations. *Reproduced from* [67].

	Euclidean	**similarity**	**affine**	**projective**
transformations				
rotation	•	•	•	•
translation	•	•	•	•
uniform scaling		•	•	•
nonuniform scaling			•	•
shear			•	•
perspective projection				•
composition of projections				•
invariants				
length	•			
angle	•	•		
ratio of lengths	•	•	•	
parallelism	•	•	•	
incidence	•	•	•	•
cross ratio	•	•	•	•

expressed as an (n+1)–dimensional vector [69]. The mapping from a point in the n–dimensional Euclidean space to an (n+1)–dimensional projective space can be written as:

$$(X_1, X_2, ..., X_n)^T \rightarrow (\lambda X_1, \lambda X_2, ..., \lambda X_n, \lambda)^T, \tag{4.1}$$

where λ is a non-zero value and describes a scaling factor usually denominated as the homogeneous scaling factor. The relationship between the coordinates of each point, for each space, is given by:

$$X = \frac{X_1}{X_4}, \;\; Y = \frac{X_2}{X_4}, \;\; Z = \frac{X_3}{X_4}, \tag{4.2}$$

where X_4 is different than zero. The coordinates (X_1, X_2, X_3, X_4) are called homogeneous coordinates. A projective transformation is the correspondence between a point P in the physical world with coordinates (X, Y, Z) and a point p on the image plane with coordinates (x, y):

$$P(X, Y, Z) \rightarrow q(x, y) \tag{4.3}$$

The projection screen is considered as being a projective space, therefore homo-

geneous coordinates are used to represent a projected point q that, in this case, is defined by a three-dimensional vector:

$$q = (q_1, q_2, q_3). \tag{4.4}$$

The following section 4.1.2 focus on the pinhole camera model, based on the projective geometry basis that was shown in the present section.

4.1.2 Pinhole Camera Model

The pinhole camera model is a simple imaging model that reproduces accurately the geometry of perspective projection. Perspective projection can be defined as the mapping between a 3D point and its corresponding projections onto a projection plane. According to the pinhole camera model, rays of light enter a camera through an infinitesimally small aperture and an image of an object is formed by the intersection of the light rays with an image plane.

Figure 4.2 shows a schematic representation of a pinhole camera. The optical center of the camera is identified by O and it is the origin of the optical axis, the line that goes from the optical center and is perpendicular to the image plane. The point where the optical axis meets the image plane is denoted as the principal point. The

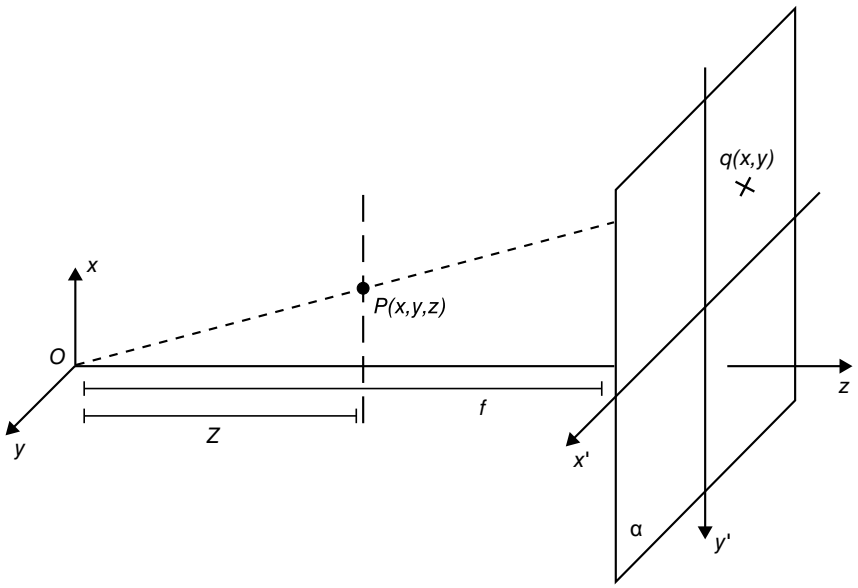


Fig. 4.2: Pinhole camera model. A point $P(X,Y,Z)$ is projected onto an image plane α by the ray passing in the center of projection O. The image point is identified by the point $q(x,y)$. f: focal length of the camera; Z: optical center/object distance.

focal length f expresses the distance between the optical center of the camera and the image plane. It is the only factor relevant to the size of the image that is formed in the projection surface and it can be defined as the same as the SID of a C-arm.

Camera Intrinsic Parameters

From Figure 4.2 it is possible to see that the coordinate values x' and y' of an image plane point $q(x,y)^T$, correspondent to a 3D point $P(X,Y,Z)^T$, are given by:

$$x = \frac{Xf}{Z}, \;\; y = \frac{Yf}{Z}, \tag{4.5}$$

where f is the focal length and Z is the distance between the object and the optical center. The previous equation can be written using the projective framework as:

$$(\lambda x, \lambda y, \lambda) = (Xf, Yf, Z)^T, \tag{4.6}$$

and in matrix notation by:

$$\alpha \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}, \tag{4.7}$$

where λ represents the homogeneous scaling factor.

Some changes in the above matrix have to be done in order to consider the origin of the image coordinate system. The principal point $(o_x, o_y)^T$ coincides with the origin of the coordinate system, leaving the projection matrix as:

$$\alpha \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} f & 0 & o_x & 0 \\ 0 & f & o_y & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}. \tag{4.8}$$

The mapping between a 3D point and its corresponding projection onto a projection plane is now expressed in the above matrix (intrinsic camera parameter matrix) by the inclusion of the intrinsic parameters. These parameters have to be found through calibration, which is described in the next section (4.2). The parameters that express the position and the orientation of the camera (extrinsic parameters) are covered by the object pose estimation problem, discussed in Chapter 5.

4.2 Camera Calibration

In the context of three-dimensional machine vision, camera calibration is the process of determining the internal camera geometric and optical characteristics [2, 70, 71] (the intrinsic parameters, seen in Section 4.1.2). The perform of calibration allows the mathematical adjustment for the principal divergences that occur in the pinhole camera model. This section describes the calibration body that was used to find the focal length, the optical center, and to correct the distortions present in the X-ray images; as well as the achieved results.

4.2.1 Calibration Body

The calibration body consists in a square with 25 (5×5) metallic spheres. A schematic representation of the body with its dimensions can be seen in Figure 4.3. Several X-ray images of the calibration body were acquired using the C-arm system in different positions and orientations. Some selected images can be seen in Figure 4.4. The calibration was performed using OpenCV[1], an open source computer vision library written in C and C++ (Appendix B). An image pre-processing step was done in order to define a region of interest (ROI) that consists in the isolation of the field of view (FOV) of each image (Figure 4.5 shows the FOV of some of those images). This was done to assure that the used images were coincidental with the projection plane area (defined by the radiation detector of the C-arm system).

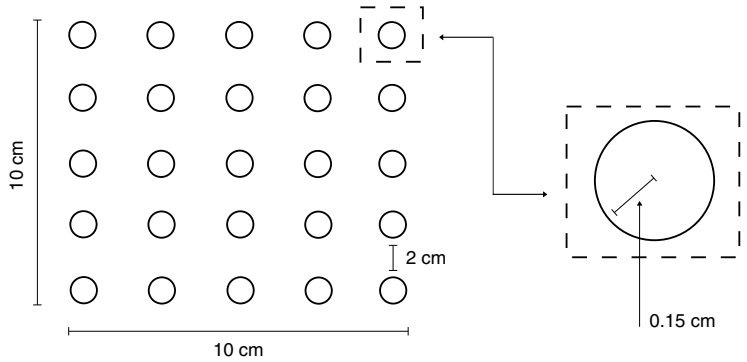


Fig. 4.3: Schematic representation of the pattern that was used to calibrate the C-arm system. The pattern consists in a symmetric group of 25 (5×5) metal spheres. The size of the pattern is 10×10 cm and each sphere has a volume of ≈ 0.015 cm^3. The distance between each sphere is 2 cm.

[1] A tutorial on camera calibration with OpenCV can be seen at the library's online documentation, accessed at `http://opencv.itseez.com`

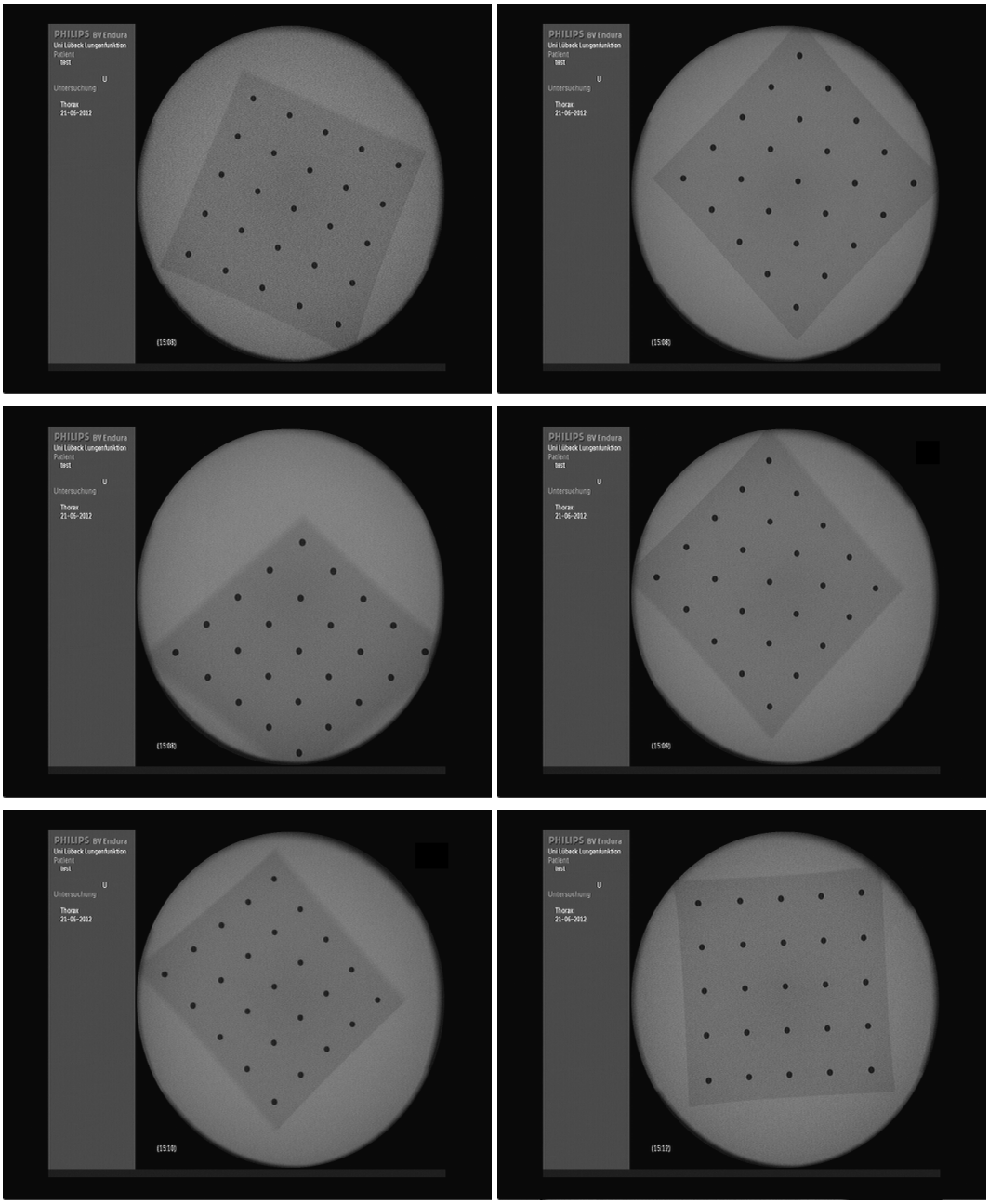


Fig. 4.4: Some of the calibration body images that were acquired with the C-arm. A total of 120 images were acquired.

4.2.2 Lens Distortions

The lens of cameras usually involve significant distortions due to manufacturing reasons or lens/imager alignment [69]. The pinhole camera model considers this distortions in the obtained images and they can be considered as constants. Lens

Tab. 4.2: Types of lens distortions, their causes and effects. The constant parameters that describe each kind of distortion are also shown.

distortion	**cause**	**effect**	**parameters**
radial	lens shape	edge pixels distortion	k_1, k_2, k_3
tangential	lens/image plane alignment	image displacement	p_1, p_2

distortions can be radial or tangential. Radial distortion manifests in the form of 'barrel' or 'fish-eye' effect and it can be mathematically corrected by:

$$x_c = x(1 + k_1 r^2 + k_2 r^4 + k_3 r^6), \tag{4.9}$$

$$y_c = y(1 + k_1 r^2 + k_2 r^4 + k_3 r^6), \tag{4.10}$$

where (x, y) stands for the coordinates of a pixel in the input image and (x_c, y_c) stands for the coordinates of a pixel in the undistorted output image. Tangential distortion happens due to the fact that the imaging lenses are not perfectly parallel to the imaging plane. The correction of tangential distortion can be achieved by:

$$x_c = x + [2p_1 y + p_2(r^2 + 2x^2)], \tag{4.11}$$

$$y_c = y + [2p_1(r^2 + 2y^2) + 2p_2 x]. \tag{4.12}$$

Thus, there are five distortion parameters that can be organized in a 1-by-5 matrix:

$$D = (k_1 \quad k_2 \quad p_1 \quad p_2 \quad k_3), \tag{4.13}$$

where k_1, k_2 and k_3 characterize the radial distortions; p_1 and p_2 describe the tangential distortions. Other type of distortions can occur in imaging systems, but they are not so important as the radial and tangential ones. Table 4.2 summarizes the properties of radial and tangential distortions.

4.2.3 Camera Intrinsic Parameters

The intrinsic or internal parameters of an imaging system are described by the focal length and the principal point coordinates of the camera. The following Eq. (4.14) allows for unit conversion (physical/camera units) and describes the camera matrix.

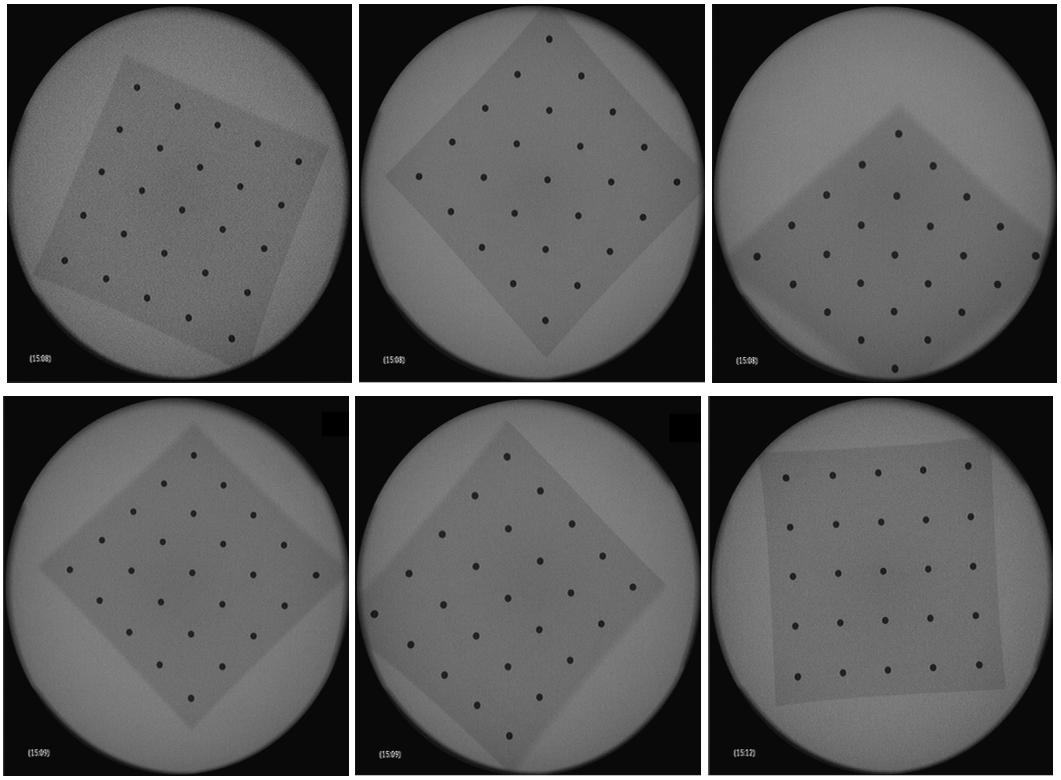


Fig. 4.5: Selected image processing ROI (region of interest) for the acquired images of the calibration body. The defined ROI corresponds to the FOV (field of view) of the X-ray images.

$$\alpha \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} f_x\, 0\, o_x\, 0 \\ 0\, f_y\, o_y\, 0 \\ 0\, 0\, 1\, 0 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}, \tag{4.14}$$

where f_x, f_y stand for the camera focal lengths, and o_x, o_y describe the optical center coordinates (in pixel units). The focal length is now considered to have different values in x and y due to the fact that the camera pixel may not be exactly squared. These parameters are computed via a calibration method described in Section 4.2.4.

4.2.4 Calibration Method

The process of undistorting the images and finding the intrinsic parameters of the camera starts with the selection of the pattern used for calibration. Three different patterns are supported: classical black and white chessboard, symmetrical circle

pattern, and asymmetrical circle pattern. The goals of the used method for the calibration of the system can be defined as follows:

- Distortion matrix computation;
- Camera matrix computation;
- XML file usage for input settings and output results;
- Re-projection error calculation.

A symmetrical circle pattern is used in this case (the calibration pattern is des-cribed in Section 4.2). The algorithm starts by reading the setup settings from an XML file (`calib_input.xml`):

```
Settings s;
const string inputSettingsFile
  = argc > 1 ? argv[1] : "calib_input.xml";
FileStorage fs(inputSettingsFile, FileStorage::READ);
```

The `calib_input.xml` file defines all the input settings as it is described below:

```
// number of inner corners per row and column
<BoardSize_Width> 5 </BoardSize_Width>
<BoardSize_Height> 5 </BoardSize_Height>

// square size
<Square_Size> 1000 </Square_Size>

// if a non-zero value is chosen, the input images
// are flipped around the horizontal axis
<Input_FlipAroundHorizontalAxis> 0 </
   Input_FlipAroundHorizontalAxis>

// used calibration pattern
// (symmetric circle grid)
<Calibrate_Pattern> "CIRCLES_GRID" </Calibrate_Pattern>
```

The used calibration pattern `CIRCLES_GRID` uses the OpenCV function `findCirclesGrid`. The images and the size of the board are passed to this function, with

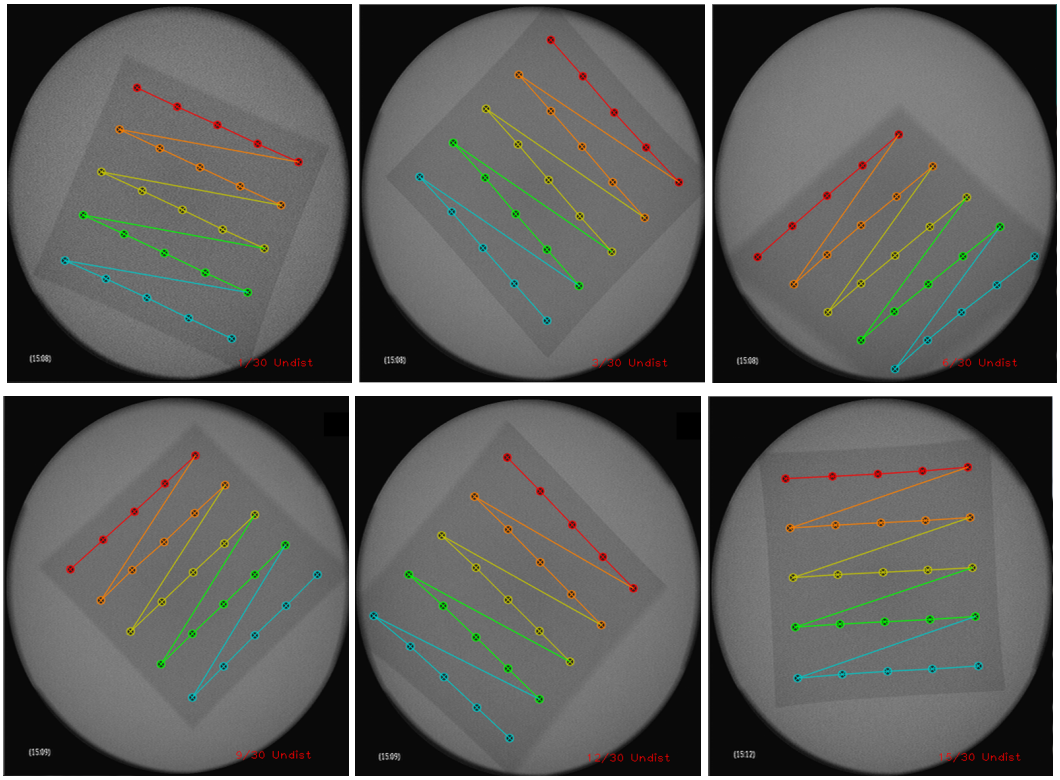


Fig. 4.6: Pattern identification on the calibration images. The algorithm runs the input images, finding the major patterns (the circles) and the path between them.

the position of the pattern being returned by it. The function `findCirclesGrid` finds the centers of the grid cirles and is defined by the following parameters:

- `image` – the input image (it must be an 8-bit grayscale or color image);
- `patternSize` – number of circles per row and column;
- `centers` – output array of detected centers;
- `flags` – operation flag with the value `CALIB_CB_SYMMETRIC_GRID`;
- `blobDetector` – detection of blobs like dark circles on light background.

If the image contains a grid of circles as defined previously in the configuration file, the function `findCirclesGrid` locates the centers of the circles. The identification of the pattern circles can be seen in Figure 4.6. The designation of the XML file with the direction to the folder where the input images are located is also done by calling the file `calib_input_dir.xml` via:

```
// input XML file with the images directory
<Input> "calib_input_dir.xml" </Input>

// images directory (in calib_input_dir.xml)
<opencv_storage>
<images>
calib_imgs/0001.pgm calib_imgs/0002.pgm calib_imgs/0003.pgm
calib_imgs/0004.pgm calib_imgs/0005.pgm calib_imgs/0006.pgm
calib_imgs/0007.pgm calib_imgs/0008.pgm calib_imgs/0009.pgm
calib_imgs/0010.pgm calib_imgs/0011.pgm calib_imgs/0012.pgm
calib_imgs/0013.pgm calib_imgs/0014.pgm calib_imgs/0015.pgm
calib_imgs/0016.pgm
</images>
</opencv_storage>
```

The following parameters define the number of input images to be used; the ratio fx/fy (the ratio between the two directions of the focal length is not changed); tangential distortion coefficients (considered as zero); and the principal point position (considered as not being changed during the global optimization step):

```
// number of images
// that was used
<Calibrate_NrOfFrameToUse> 16
</Calibrate_NrOfFrameToUse>

// ratio fx/fy
// (focal length in xx and yy)
<Calibrate_FixAspectRatio> 1
</Calibrate_FixAspectRatio>

// tangential distortion coefficients
<Calibrate_AssumeZeroTangentialDistortion> 1
</Calibrate_AssumeZeroTangentialDistortion>
// principal point position
<Calibrate_FixPrincipalPointAtTheCenter> 1
</Calibrate_FixPrincipalPointAtTheCenter>

// output log file
<Write_outputFileName> "calib_output.xml"
</Write_outputFileName>
```

Initial tests were done for a non-fixed ratio fx/fy, considering tangential distortions coefficients, and with a non-fixed principal point. The algorithm was initialized with 16 test images and the achieved re-projection error was **1.3 px**.

The reprojection error is computed[2] by the total sum of squared distances between the observed feature points and the projected (using the current estimates for the camera parameters) object points. The achieved error revealed itself to be too high. Tests with more input images were done but the obtained results did not improve. Finally, the ratio fx/fy was set to 0, the tangential distortions coefficients were considered null, and the principal point was established has being fixed. These changes resulted in a considerable improvement of the re-projection error: **0.7px**.

A camera matrix was obtained. The following equation defines the mapping from a 3D point to the 2D image plane for this specific camera:

$$\alpha \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} 2061.389 & 0 & 203.655 & 0 \\ 0 & 2254.646 & 205.117 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}. \tag{4.15}$$

From this matrix, one can conclude that the intrinsic parameters (focal length and principal image point) of the camera are the following:

- Focal length xx-axis (f_x): **2061.389 px**
- Focal length yy-axis (f_y): **2254.646 px**
- Principal image point xx-axis (o_x): **203.655 px**
- Principal image point yy-axis (o_y): **205.117 px**

Once the calibration process is done, the camera is ready to perform the measurements described in Chapter 5, using the focal length and principal image point values given by the method that was described above.

[2] As defined in the *Camera Calibration and 3D Reconstruction* section of the OpenCV online documentation, accessed at: `http://goo.gl/dcNbg`

5 3D Object Pose Estimation

The pose of an object is defined by its position and orientation relative to a camera. The computation of this parameters is a important subject in areas such as robotic guidance, manipulation, photogrammetry, object recognition, and camera calibration [72–76]. This Chapter describes the work done in order to find the pose of a 3D object relative to a C-arm imaging system. More on object pose estimation can be found in [77–80].

5.1 Object Geometry

A simple and regular shaped rigid body with 9 markers (Figure 5.1/5.4) was submitted to the pose computation process. The dimensions of the body are $5\times6\times4$ cm. In order to get to know its geometry, the object was scanned using a CT system. Image processing methods were used to compute the coordinates of the nine markers. These coordinates were initially expressed with respect to the CT coordinate system and later a coordinate system with origin in one of the markers was defined.

5.1.1 Image Acquisition

The object was scanned in a Siemens Healthcare SOMATOM Definition AS+ CT. The images were obtained using the following scanner settings: 120kVp (peak KVP output of the X-ray-generator used) and 85mA (X-ray tube current). The scan was composed of 118 slices (512×512, 16 bits) and exported in DICOM format. The pixel size is 0.3515625×0.3515625 mm^2. In Figure 5.2 some slices are shown.

5.1.2 Volume Rendering

The DICOM slices were used to compose a three-dimensional image (Figure 5.5) and a three-dimensional representation of the markers was extracted. For this pur-

Fig. 5.1: The object used for pose estimation. Dimensions of the object: 5×6×4 cm.

pose, ITK[1] was used [81]. The DICOM series are read by an ITK object called `GDCMSeriesFileNames` that explores the directory where all the CT slices are and generates a sequence of filenames for the study/series DICOM files. The volumetric image is then saved to other specified file, being the output a VTK reconstructed 3D image. Figure 5.3 shows a scheme of the markers isolation method.

```
typedef itk::GDCMImageIO ImageIOType;
ImageIOType::Pointer dicomIO = ImageIOType::New();
reader->SetImageIO( dicomIO );
```

5.1.3 Markers Isolation

Once the volume rendering was done, more ITK based algorithms were used in order to do the isolation of the markers. Given the rendered volume as an input, the first step was to transform the image into a binary one. This was accomplished by using a segmentation process which chooses the pixels according to an interval of intensity values defined by a lower and upper threshold:

$$I(X) \in [lower, upper]. \tag{5.1}$$

The position where the algorithm should start the process has to be defined as well as the replace value for all the values present in the range of (5.1). The lower value was defined as -1024; the upper as 1000; and the replace value as 255. This means that every value inside this range will become white and outside this range will turn

[1] ITK – Segmentation and Registration Toolkit (`http://itk.org/`)

Fig. 5.2: From left to right: slices #36, #52 and #65 originated from the CT acquisition. 118 slices where acquired with the CT scan, being then used to do a 3D volume render of the body. The 9 different markers can be seen as white spots.

to black. The chosen seed location was (0,0,0). The result is the growth of the region from the seed, being the background now white and all the markers (that had maximum intensity in the input image) black. The image is now binarized.

```
{InputImageType::IndexType index;}
index[0] = ( 0 ); index[1] = ( 0 ); index[2] = ( 0 );
connect -> SetLower( -1024 );
connect -> SetUpper( 1000 );
connect -> SetReplaceValue( 255 );
```

It is then necessary to use a negation filter. This filter will invert all the values of the image - everything that is in black turns to white and everything in white to black. The markers will become now white and the background black. After this, it is necessary to label all the objects (markers) using a connectivity component filter. A unique label is assigned to each distinct marker. The filter iterates each element of the image by columns and then by rows accordingly to the neighbor pixels – if the element is not in the correspondent area to the background of the image, it searches for neighbor elements of the present element; if no neighbors exist, it does the labeling of that unique element and goes on; if this last condition does not fit too, it searches for the neighbor with the lowest label and assigns it to the present element. At the end, it saves all the equivalences between neighbors.

```
typedef itk::ConnectedComponentImageFilter< InputImageType,
InputImageType > ConnectedComponentFilterType;
ConnectedComponentFilterType::Pointer labeller =
ConnectedComponentFilterType::New();
```

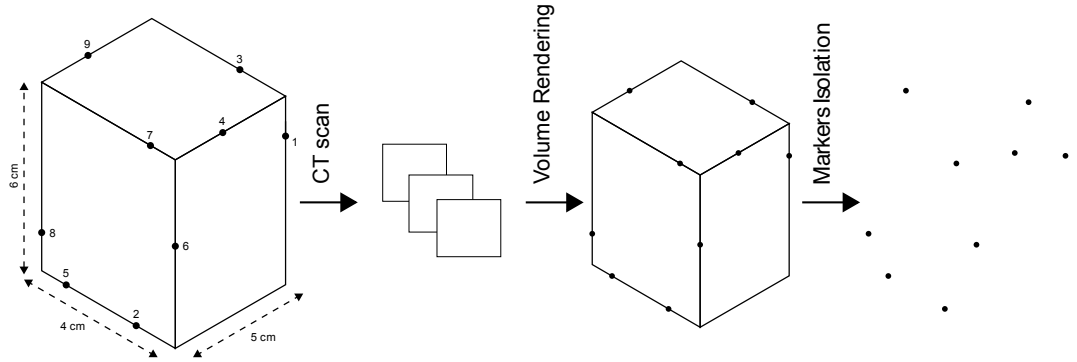


Fig. 5.3: Schematic representation of the method used to isolate the markers from the rest of the image. ITK based algorithms were used to perform the image processing operations that allowed the markers isolation.

A relabel component filter is then used so that consecutive labels are used. The labels associated with the objects are remapped such that the label numbers are consecutive, with no gaps between them. The labels are sorted according to the size of the object – the largest object will have label #1; the second #2; and the smallest will have the highest number (every marker has now a specific value).

```
typedef itk::ConnectedComponentImageFilter< InputImageType,
InputImageType > ConnectedComponentFilterType;
ConnectedComponentFilterType::Pointer labeller =
ConnectedComponentFilterType::New();
```

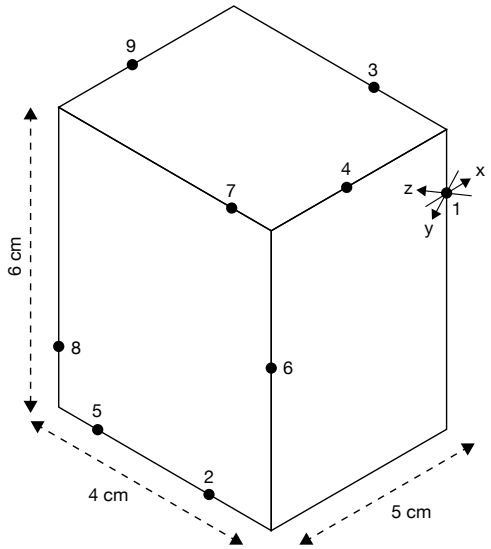
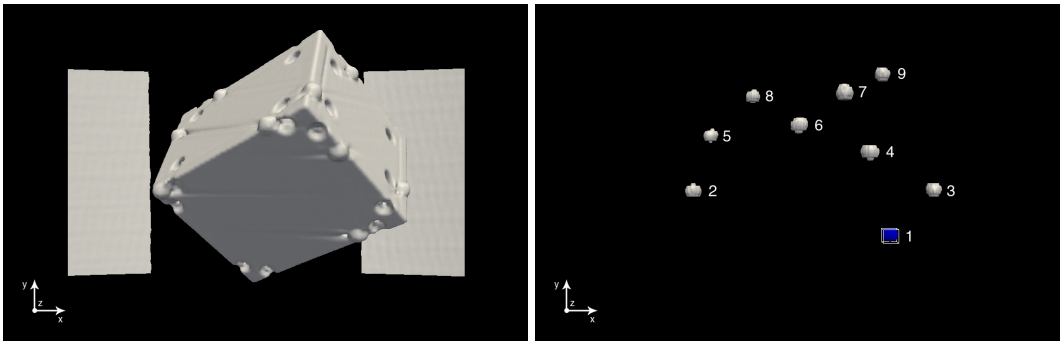


Fig. 5.4: Geometry of the body used for pose estimation. The body has dimensions $5\times6\times4$ cm and there are 9 spherical metal markers (1-9).



(a) (b)

Fig. 5.5: Output images for (a) the volume rendering; and (b) the object isolation. The nine markers are represented according to the assignment done to the rigid body. ITK image processing algorithms were used to do the volume rendering.

Finally, a binary threshold filter is used with lower and upper thresholds of, respectively, 1 and 9 (these values correspond now to the markers). Pixels that are inside this range of values will have value 255 (white), while the ones that are outside will have value 0 (black):

$$Out(x_i) = \begin{cases} 255 & \text{if } 1 \leq x_i \leq 9 \\ 0 & \text{otherwise} \end{cases}. \tag{5.2}$$

This is done in order to isolate the 9 biggest objects remaining in the image (corresponding to the markers).

```
threshold -> SetLowerThreshold( 1 );
threshold -> SetUpperThreshold( 9 );
threshold -> SetOutsideValue( 0 );
threshold -> SetInsideValue( 255 );
```

5.1.4 Coordinates Extraction

The coordinates (shown in Table 5.1) were computed using ParaView[2] data analysis methods. A connectivity filter was used, which assigns a region identifier to connect components of the input data set. This filter takes any data set type as input and produces an unstructured grid output. It is characterized by the three properties that are next described: ColorRegions, ExtractionMode, and Input.

[2]ParaView – Open Source Scientific Visualization (`http://paraview.org`).

- ColorRegions – controls the coloring of the connected regions; its default value is 1 and only the values 0 and 1 are accepted.
- ExtractionMode – controls the extraction of connected surfaces; its default value is 5 and the value must be one the following: ExtractPointSeededRegions1, ExtractCellSeededRegions2, ExtractSpecifiedRegions3, ExtractLargestRegion4, ExtractAllRegions5, ExtractClosestPointRegion6.
- Input – specifies the input to the filter; it restricts the selected dataset to be one of the type vtkDataSet (or a subclass of it).

A threshold filter was used to extract the portions of the input dataset whose scalars lie within a specified range. This filter operates on either point-centered or cell-centered data and has the following properties:

- AllScalars – if the value of this property is 1, then a cell is included in the output if the value of the selected array for all its points is within the threshold. This is only relevant when thresholding by a point-centered array.
- Input – the dataset must contain a point or cell array with 1 component and it must be of the type vtkDataSet (or a subclass of it).
- Scalars – contains the name of the scalar array from which to perform thresholding; it requires an array of scalars.
- ThresholdRange – specifies the upper and lower bounds of the thresholding operation; the default values for the bounds are 0,0 and the value must lie within the range of the selected data array.

Tab. 5.1: Coordinates of the markers: CT (xyz_{CT}) and object (xyz_o) coordinate system.

# marker	x_{CT} **(mm)**	y_{CT} **(mm)**	z_{CT} **(mm)**	x_o **(mm)**	y_o **(mm)**	z_o **(mm)**
1	27.60	-114.50	-616.00	0.00	0.00	0.00
2	-31.30	-113.50	-602.00	-58.85	1.00	14.00
3	40.95	-113.50	-602.00	13.35	1.00	14.00
4	19.35	-134.50	-592.00	-8.25	-20.50	24.00
5	-28.45	-95.05	-585.00	-56.05	19.00	31.00
6	0.53	-134.50	-585.00	-27.07	-20.00	31.00
7	12.50	-136.00	-576.00	-15.01	-21.50	40.00
8	-14.60	-95.7	-572.00	-42.20	18.75	44.00
9	26.20	-109.50	-567.00	-1.40	5.00	49.00

This was done once for each marker together with an outline filter. With this process it was possible to define a box for each marker and to know the bounds of this box in x, y and z. The markers coordinates values were obtained by the computation of the center point of the box (which is consider as a center point of the marker).

5.2 Pose Estimation

The position and orientation (pose) of an object relative to a camera can be computed when the geometric configuration of the object and the correspondent projections are known. The method used in the section is based on an algorithm developed by DeMenthon [82]. This algorithm is able to compute the rotation and translation parameters of an object relative to a camera if a correspondence between two sets (object points and its projections) is defined. At least four non-coplanar points have to be defined. A simulation model was elaborated in order to test the validity of the algorithm for the present method. Once the method was validated, the same tests were made with the pose object for different C-arm system orientations.

5.2.1 Simulation Model

A pinhole camera model was simulated in order to test the validity of the algorithm. The principles behind pinhole cameras were already discussed in Section 4.1.2. Some model points were generated and their projections computed:

$$x_c, y_c, z_c = \begin{pmatrix} 0 & 0 & 500 \\ -100 & 0 & 600 \\ 100 & 0 & 500 \\ 0 & -100 & 500 \end{pmatrix}, \quad f = \overline{OP} = 1000px\,, \tag{5.3}$$

$$x' = \frac{x_c}{z_c} \cdot f, \quad y' = \frac{y_c}{z_c} \cdot f \quad \Rightarrow \quad x', y' = \begin{pmatrix} 0 & 0 \\ -167 & 0 \\ 200 & 0 \\ 0 & -200 \end{pmatrix}, \tag{5.4}$$

where x_c, y_c, z_c are the coordinates of the model points in the object coordinate system; f is the focal length and was defined as 1000 px; x' and y' are the computed projections for each point. Figure 5.6 shows a schematic representation of the adopted model, with the model points A, B, C, D and their projections A', B', C', D'.

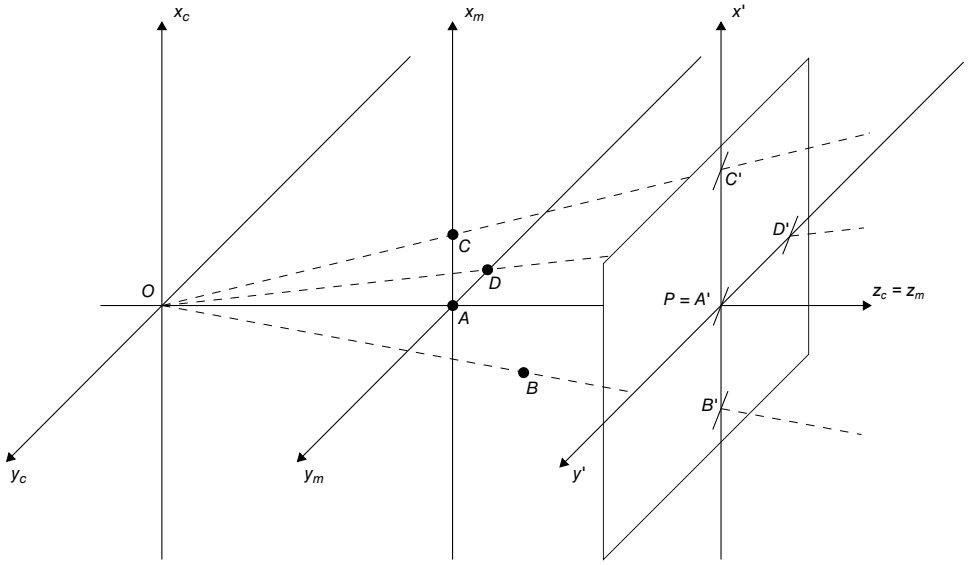


Fig. 5.6: Representation of the pinhole camera model used for the simulation. x_c, y_c, z_c: coordinate system of the camera; x_m, y_m, z_m: coordinate system of the model; x', y': coordinate system of the projection plane; O: optical center; P: principal point/image center; A, B, C, D: model points; A', B', C', D': projection points.

0° Rotation

A rotation of 0° was initially simulated[3]. A vector `modelPoints` with the object points (in object coordinate system) was created, while the correspondent projections were defined in `srcImagePoints`:

```
#define FOCAL_LENGTH 1000
#define FOCAL_LENGTH_X 1000
#define FOCAL_LENGTH_Y 1000

std::vector<CvPoint3D32f> modelPoints;
modelPoints.push_back(cvPoint3D32f( 0,0,0 ));          // #1
modelPoints.push_back(cvPoint3D32f( -100,0,100 ));     // #2
modelPoints.push_back(cvPoint3D32f( 100,0,0 ));        // #3
modelPoints.push_back(cvPoint3D32f( 0,-100,0 ));       // #4

std::vector< CvPoint2D32f > srcImagePoints;
srcImagePoints.push_back( cvPoint2D32f( 0,0 ));        // #1
srcImagePoints.push_back( cvPoint2D32f( -167,0 ));     // #2
srcImagePoints.push_back( cvPoint2D32f( 200,0 ));      // #3
srcImagePoints.push_back( cvPoint2D32f( 0,-200 ));     // #4
```

[3]As defined by Javier Barandiaran's POSIT tutorial, accessed at `http://goo.gl/LqcwP`

The achieved rotation and translation between the model and camera coordinate systems is described by the matrix M_{0°:

$$M_{0^\circ} = \left(\begin{array}{c|c} R_{0^\circ} & t_{0^\circ} \\ 0_{\times 3} & 1 \end{array}\right), \quad R_{0^\circ} = \begin{pmatrix} 0.999 & 0 & -0.015 \\ 0 & 1 & 0 \\ 0.001 & 0 & 0.999 \end{pmatrix}, \quad t_{0^\circ} = \begin{pmatrix} 0 \\ 0 \\ 499.926 \end{pmatrix}. \quad (5.5)$$

Theoretically, a 0° rotation is represented by the following rotation matrix:

$$R_{t,0^\circ} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (5.6)$$

Comparing the rotation matrices from the above two equations (R_{0° and $R_{t,0^\circ}$), it is possible to see that the achieved rotation matrix (Equation 5.5) reflects indeed a 0° rotation, and that the verified translation is as expected (half of the SID value, correspondent to the distance between the model points and the optical center of the camera). An image with the projected points was created and it can be seen in Figure 5.7a. Figure 5.7b shows the output of the algorithm, where it is possible to see that the algorithm's expected projections match with the actual ones.

90° Rotation

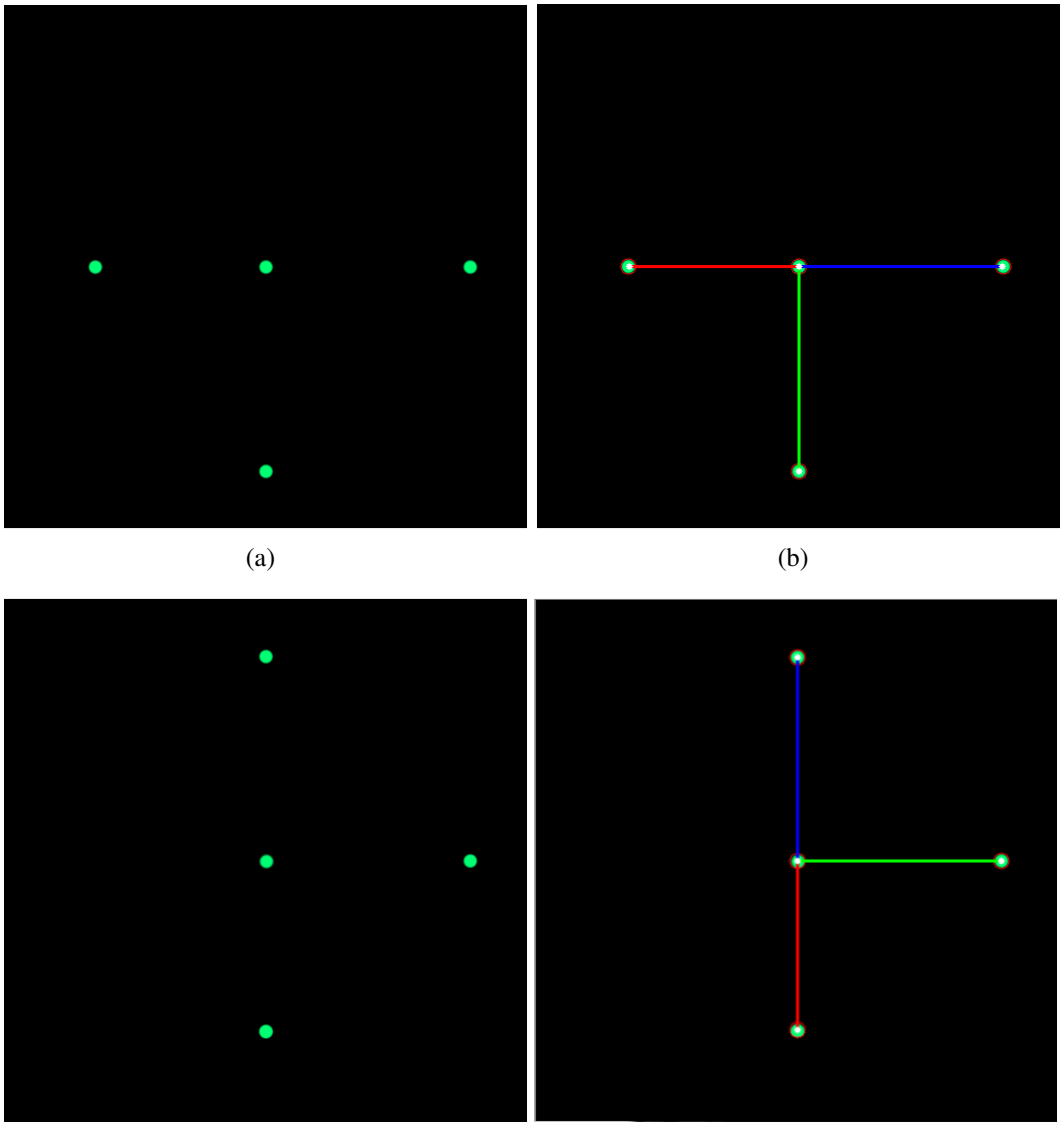
A 90° rotation around the *zz*–axis was simulated (from the original model represented in Figure 5.6). Considering the focal length as being equal to 1000 px, the obtained rotation/translation between the model–camera coordinate systems was:

$$R_{zz_{90^\circ}} = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & -1.4_{\times 10^{-7}} \\ 1.4_{\times 10^{-7}} & 0 & 1 \end{pmatrix}, \quad t_{zz_{90^\circ}} = \begin{pmatrix} 0 \\ 0 \\ 500 \end{pmatrix}. \quad (5.7)$$

In theory, a 90° rotation around the *zz*–axis is represented by the following matrix:

$$R_{zz_{90^\circ}} = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (5.8)$$

In this case, the obtained rotation matrix is also similar with the theoretical one. In Figure 5.7d it is possible to see that the expected projections match the actual ones.



(a) (b)

(c) (d)

Fig. 5.7: Input (a,c) and output (b,d) images of POSIT. In (b) and (d), the real projections are identified as red circles and the expected ones as white spots. The 0° case is described by (a) and (b); while (c) and (d) are relative to a 90° rotation.

5.2.2 FOV Definition

After the tests performed in the last Section (5.2.1) and before moving for the pose estimation with C-arm X-ray images, the standardization of the Field of View (FOV) was done. All the used X-ray images were cropped according to this definition. Figure 5.8 shows a scheme of the FOV that was established, defined as having the

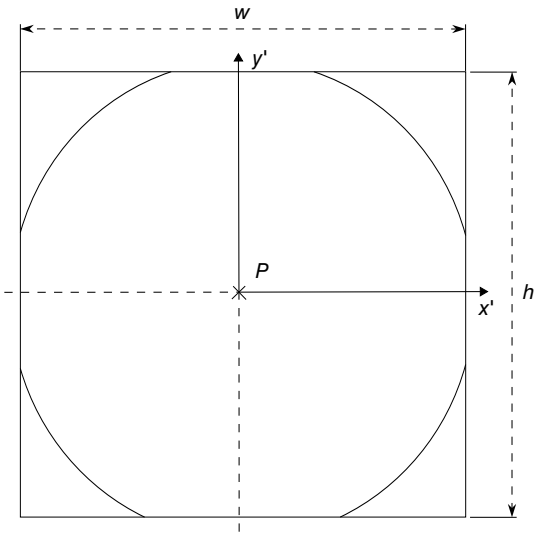


Fig. 5.8: Representation of the Field of View (FOV): $w = 408px$ and $h = 408px$. The principal point P has coordinates $(204, 204)$.

dimensions 408×408 px. The principal point P has coordinates 204×204 px. The need for the creation of a standard FOV arises from the fact that the original X-ray images are not coincident with the detected X-ray area of the C-arm projection plane. In the following Section (5.2.3), the pose estimation method applied to the acquired X-ray images is described.

5.2.3 Object Model

The method used in Section 5.2.1 was applied to the object model points scenario. Several tests were done for different C-arm orientations in order to test the method's validity to measure how much the C-arm was rotated between two distinct orientations. This section describes the results of three selected tests. The values for the focal length and principal point obtained with the calibration (Chapter 4) were used as constant parameters:

```
#define FOCAL_LENGTH 2158.0175
#define FOCAL_LENGTH_X 2061.389
#define FOCAL_LENGTH_Y 2254.646

#define cX  205.155
#define cY  206.617
```

For each test, four object points and its correspondent projections were selected. The coordinate system of the object was translated from the first to the third marker.

Tab. 5.2: Object model and respective projections coordinate for each C-arm orientation. Four non-coplanar object points (3,4,7,8) where chosen to perform the pose estimation operation.

#	object			0°		10°		30°		50°		90°	
	x	y	z	x'	y'	x'	y'	x'	y'	x'	y'	x'	y'
3	0	0	0	133	180	155	179	188	177	199	177	153	182
4	-21.6	-21.5	10	134	270	160	271	201	271	221	275	182	286
7	-28.4	-22.5	26	129	313	151	313	184	314	196	319	153	335
8	-55.4	17.8	30	240	317	251	318	250	321	217	326	63	347

The matrix that gives the object–camera coordinate systems transformation is given by the rotation and translation parameters that relate the object's pose to the camera. The achieved rotation matrices can be seen in Tables 5.3, 5.5 and 5.4. Figures 5.10, 5.11 and 5.12 show the input and output images for each orientation. It can be seen that, in all cases, the expected projections match the actual ones. In Section (5.2.4), the validation method that was implemented in order to find the validity of these results is described.

Tab. 5.3: Pose Estimation for 0° and 10° C-arm orientation. The achieved rotation matrices are shown for each orientation.

	0°			10°		
	-0.743312	0.555513	-0.372682	-0.750081	0.450934	-0.483773
R	0.646059	0.497863	-0.578568	0.676801	0.505180	-0.535475
	-0.135857	-0.670831	-0.728962	0.002929	-0.729068	-0.684118

Tab. 5.4: Pose Estimation for 0° and 90° C-arm orientation. The achieved rotation matrices are shown for each orientation.

	0°			90°		
	-0.743312	0.555513	-0.372682	-0.105376	-0.691182	-0.714957
R	0.646059	0.497863	-0.578568	0.666663	0.477397	-0.572410
	-0.135857	-0.670831	-0.728962	0.736957	-0.536954	0.410479

Tab. 5.5: Pose Estimation for 30° and 50° C-arm orientation. The achieved rotation matrices are shown for each orientation.

	30°			50°		
	-0.710796	0.174023	-0.681531	-0.585016	-0.139689	-0.798901
R	0.673766	0.502170	-0.542093	0.671720	0.512442	-0.534973
	0.247908	-0.844510	-0.474191	0.484120	-0.849605	-0.205955

5.2.4 Validations

A method to test the validity of the results demonstrated in Section 5.2.3 was developed. An unit vector was defined along the *zz*–axis of the camera coordinate system associated with each rotation:

$$\vec{v}_c = (0,0,1). \tag{5.9}$$

This vector (for each rotation) was then transformed to the object coordinate system by being multiplied with the inverse of the rotation matrix relative to each specific rotation:

$$\vec{v}_c R_{0^\circ}^{-1} = \vec{v}_{o,0^\circ}, \tag{5.10}$$

where $\vec{v}_{o,0^\circ}$ stands for the vector v_c in the object coordinate system, for a 0° rotation. The same was done for the 90° case:

$$\vec{v}_c R_{90^\circ}^{-1} = \vec{v}_{o,90^\circ}, \tag{5.11}$$

where $\vec{v}_{o,90^\circ}$ stands for the vector v_c in the object coordinate system, for a 90° rotation. The angle between these two vectors was then computed. The angle θ between two 3D vectors $\vec{a}$ and $\vec{b}$ is given by:

$$\theta = \arccos\left(\frac{\overline{a}\cdot\overline{b}}{|a|\cdot|b|}\right), \tag{5.12}$$

where $\overline{a}\cdot\overline{b}$ stands for the dot product of the vectors $\vec{a}$ and $\vec{b}$:

$$\overline{a}\cdot\overline{b} = \sum_{i=1}^{n} a_i b_i = a_1 b_1 + a_2 b_2 + a_3 b_3, \tag{5.13}$$

and $|a|\cdot|b|$ stands for the product between the two vectors magnitudes:

$$|a| = \sqrt{a_1^2 + a_2^2 + a_3^2}\,, \tag{5.14}$$

$$|b| = \sqrt{b_1^2 + b_2^2 + b_3^2}\,. \tag{5.15}$$

A schematic representation of the described validation method for the 0°–90° case is shown in Figure 5.9. Table 5.6 shows the expected/achieved results for each test.

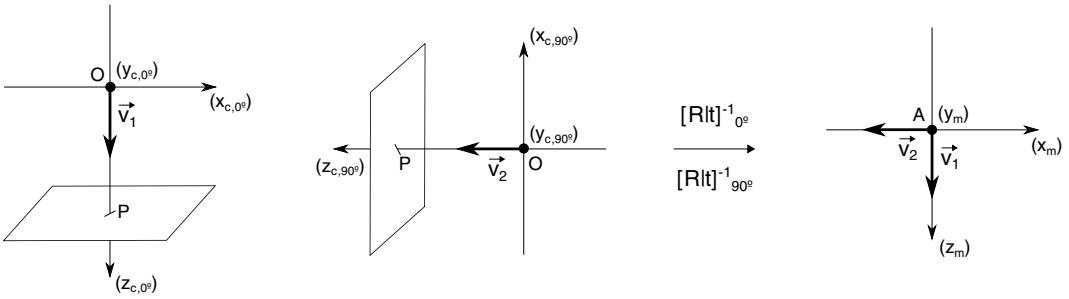


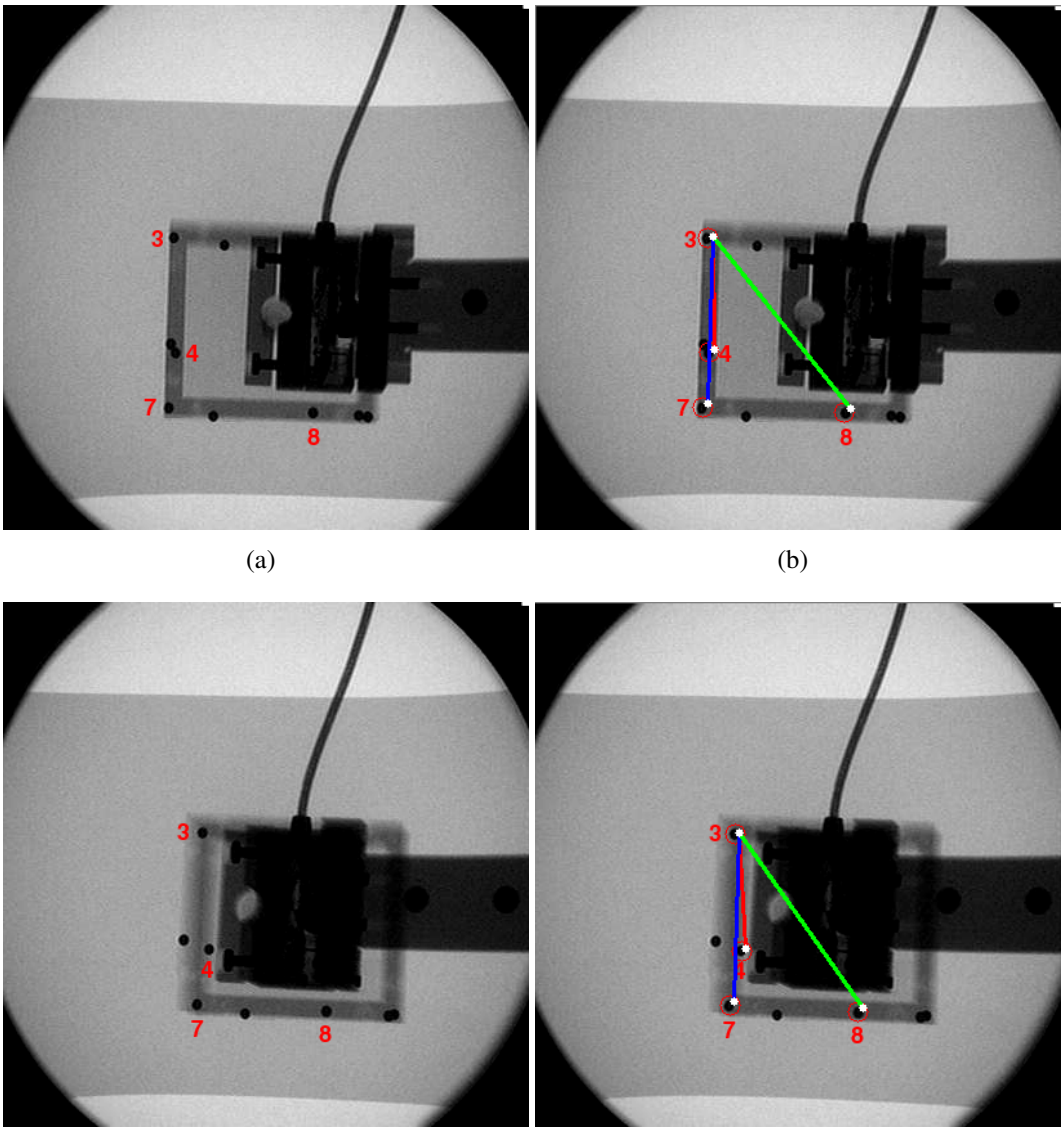
Fig. 5.9: Representation of the vector $\vec{v}$ in the C-arm coordinate system for a 0° orientation ($\vec{v_1}$) and 90° orientation ($\vec{v_2}$). The object model coordinate system is also represented with the two vectors, resulting from the computation of each vector with the inverse of the transformation matrix.

Tab. 5.6: Validation tests results for 0°–10°, 0°–90° and 30°–50°.

test	**0°–10°**	**0°–90°**	**30°–50°**
achieved output	9.5546	92.2434	20.6009
expected output	10	90	20

```
// rotation matrices
R1 = [ a b c d e f g h i ];
R2 = [ a b c d e f g h i ];
// unit vector defined along the zz-axis
v = [0;0;1];
// inverse transformation
v_R1 = inv(R1) * v;
v_R2 = inv(R2) * v;
// angle computation
v_R1_sum = sum(v_R1);
v_R2_sum = sum(v_R2);
v_R1_v_R2_prod = v_R1 .* v_R2;
v_R1_v_R2_sum = sum(v_R1_v_R2_prod);
v_R1_magn = sqrt( v_R1(1,1)^2 + v_R1(2,1)^2 + v_R1(3,1)^2 );
v_R2_magn = sqrt( v_R2(1,1)^2 + v_R2(2,1)^2 + v_R2(3,1)^2 );
v_R1_v_R2_magn = v_R1_magn * v_R2_magn;
teta = acos(v_R1_v_R2_sum/v_R1_v_R2_magn);
teta = radtodeg(teta);
```

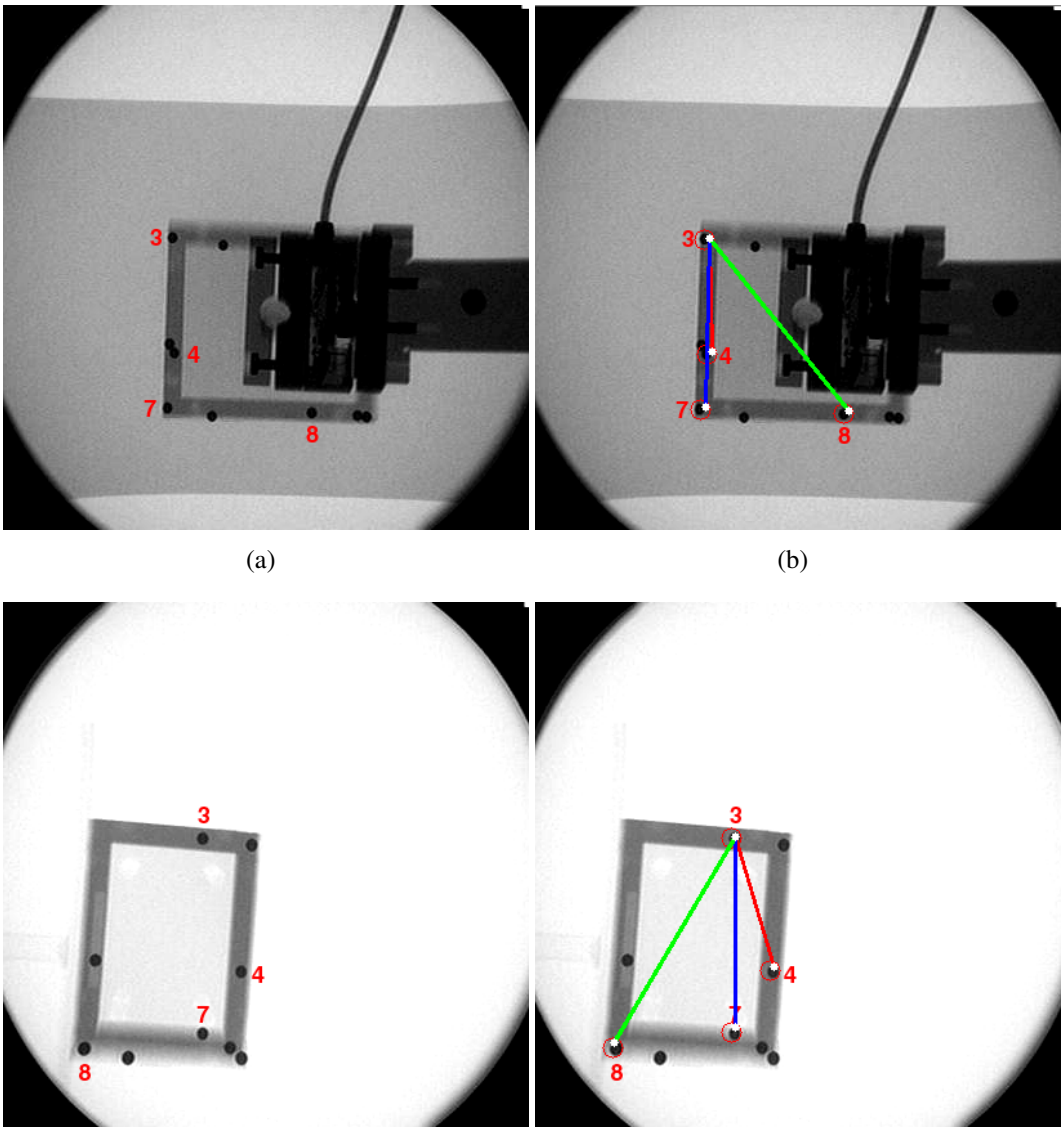
It is possible to see from Table 5.6 that the achieved results were very close to the expected ones, what proves the validity of the pose estimation method.



(a) (b)

(c) (d)

Fig. 5.10: Input/output for the pose estimation: 0° (a,b) and 10° (c,d). 4 non-coplanar points were chosen in order to compute the rotation matrix that relates the coordinate system defined in one of the markers with the coordinate system of the C-arm system. The output images show the expected projections as red circles and the real projections as white circles.



(a) (b)

(c) (d)

Fig. 5.11: Input/output for the pose estimation: 0° (a,b) and 90° (c,d). 4 non-coplanar points were chosen in order to compute the rotation matrix that relates the coordinate system defined in one of the markers with the coordinate system of the C-arm system. The output images show the expected projections as red circles and the real projections as white circles.

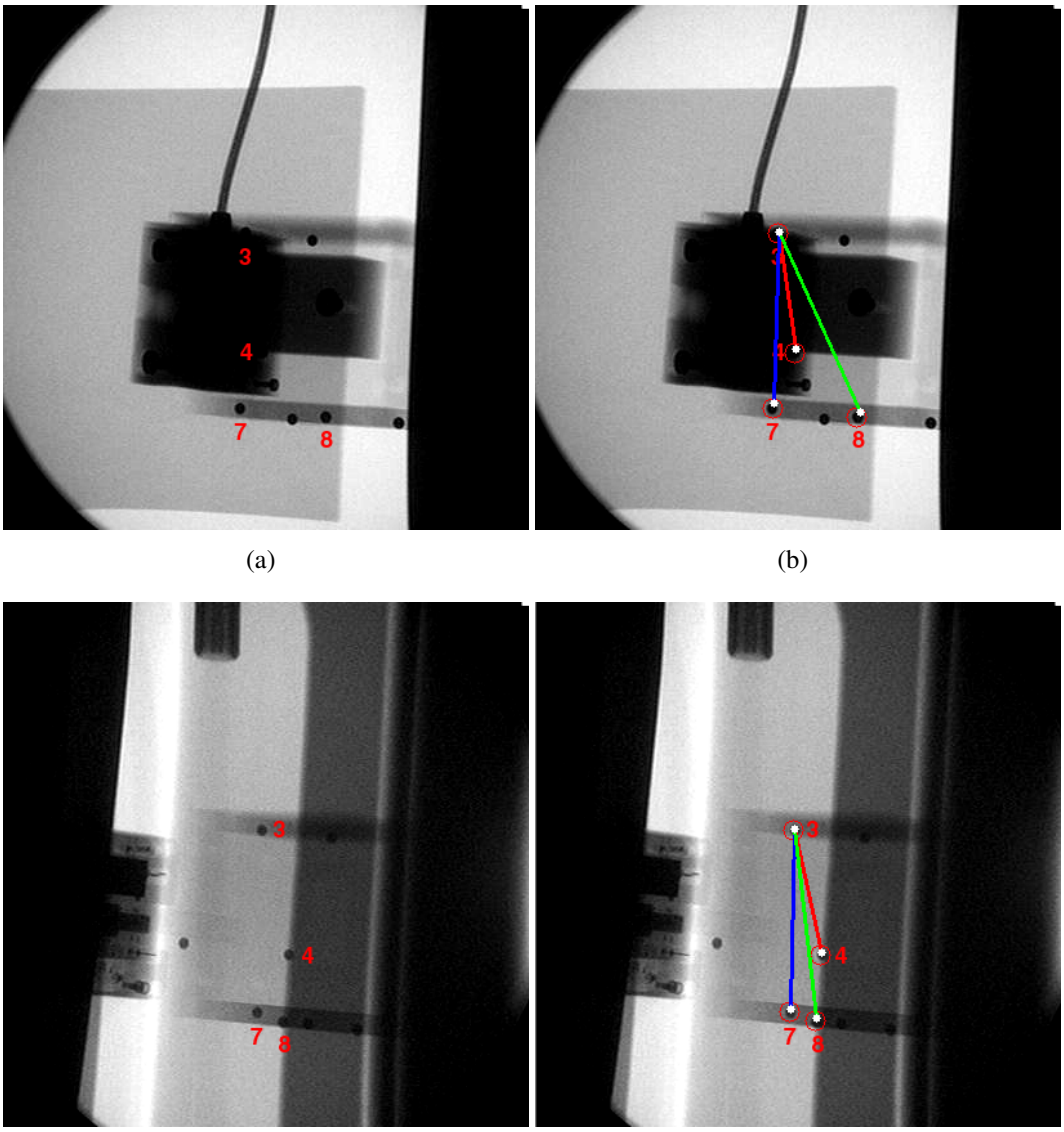


Fig. 5.12: Input/output for the pose estimation: 30° (a,b) and 50° (c,d). 4 non-coplanar points were chosen in order to compute the rotation matrix that relates the coordinate system defined in one of the markers with the coordinate system of the C-arm system. The output images show the expected projections as red circles and the real projections as white circles.

6 Conclusions

Image-guided procedures have today an important role in clinical interventions such as bronchoscopy and cardiology. They are characterized by low invasive actions, with C-arm X-ray imaging systems playing an important role in surgical navigation and structures tracking. C-arm systems have insufficiencies in the way they present information to the surgeon. A method that is able to give three-dimensional feedback on the location of a region of interest would be of supreme importance to the field. Pose estimation is a possible way to understand the relationship between different C-arm positions and orientations and, therefore, to obtain three-dimensional information from two-dimensional sets.

A pose estimation method was tested in order to find an object's position and orientation regarding a C-arm X-ray imaging system. In order to run the pose estimations tests, the system had to be geometrically defined and its calibration performed. The main steps of the presented work can be described as follows:

- Characterization of image-guided procedures and their applications;
- Definition of the main principles behind C-arm imaging systems;
- Geometrical modeling of a C-arm system and calibration of the system;
- Geometrical modeling of a simple and regular shaped object with 9 markers;
- Pose estimation of the object from its geometry and image projections.

The presented work proved itself to be a good first approach for the main to be solved problem. The relationship of different C-arm positions/orientations regarding a fixed referential is now defined. The calibration step was successfully performed. The intrinsic parameters of the C-arm system were obtained with a low re-projection error. The distortion coefficients were computed and the test images undistorted. The intrinsic parameters obtained via the system calibration were used

in the pose estimation operation. The used pose estimation method allowed the computation of the rotation matrices for different C-arm orientations. Finally, a validation test was performed to verify the accuracy of the results.

The obtained angle values are very close to the expected values. The automation of the described process should be performed. This could be done by developing a method that could automatically obtain the projections coordinates values and relate them with the geometry of the 3D object. The next step would be to define a new coordinate system in order to simulate a second position for the C-arm; and compute the geometric transformation between this new coordinate system and the one that has its origin in the first marker. With this, a relationship between the coordinate system for the two different positions would be known. If this goal is accomplished, the presented method will certainly serve as a basis for the creation of a technique that would allow the acquisition of real-time three-dimensional information regarding a clinical region of interest with a C-arm X-ray imaging system.

Bibliography

[1] C. Papalazarou, P. Rongen, and P. de With, "Feature-based depth estimation for monoplane X-ray imaging with limited C-arm motion," in *Medical Image Computing and Computer-Assisted Intervention (MICCAI) Workshop on Geometric Accuracy in Image Guided Interventions*, 2009.

[2] H. Livyatan, Z. Yaniv, and L. Joskowicz, "Robust automatic C-arm calibration for fluoroscopy-based navigation: a practical approach," in *Proceedings of the 5th International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pp. 60–68, Springer-Verlag, 2002.

[3] M. Clifford, F. Banovac, E. Levy, and K. Cleary, "Assessment of hepatic motion secondary to respiration for computer assisted interventions," *Computer Aided Surgery*, vol. 7, pp. 291–299, 2009.

[4] S. Goldberg, C. Grassi, J. Cardella, and J. Charboneau, "Image-guided tumor ablation: standardization of terminology and reporting criteria," *Journal of Vascular and Interventional Radiology*, vol. 16, pp. 765–778, 2005.

[5] P. Jannin, J. Fitzpatrick, D. Hawkes, and X. Pennec, "Validation of medical image processing in image-guided therapy," *IEEE Transactions on Medical Imaging*, vol. 21(12), pp. 1445–9, 2002.

[6] T. Acharya and A. Ray, *Image Processing - Principles and Applications*. Wiley-Interscience, 2005.

[7] W. Grimson, G. Ettinger, S. White, and T. Lozano-Pérez, "An automatic registration method for frameless stereotaxy, image-guided surgery, and enhanced reality visualization," *IEEE Transactions on Medical Imaging*, vol. 15, no. 2, 1996.

[8] E. Grimson, M. Leventon, L. Lorigo, and T. Kapur, "Image guided surgery," in *Fourth International Workshop on Algorithmic Foundations of Robotics (WAFR)*, pp. 443–450, 2000.

[9] A. Wolbarst and W. Hendee, "Evolving and experimental technologies in medical imaging," *Radiology*, vol. 238, pp. 16–39, 2006.

[10] M. Cardin, J. Wang, and D. Plewes, "A method to evaluate human spatial coordination interfaces for computer-assisted surgery," in *Medical Image Computing and Computer-Assisted Intervention*, vol. 3750, pp. 9–16, Springer Berlin Heidelberg, 2005.

[11] F. Papatheofanis, "The impact of minimally invasive surgery," *Decisions in Imaging Economics*, vol. 10, pp. 66–68, 1997.

[12] L. Joskowicz, C. Milgrom, A. Simkin, and L. Tockus, "FRACAS: a system for computer-aided image-guided long bone fracture surgery," *Computer Aided Surgery*, vol. 3, pp. 271–288, 1998.

[13] Z. Yaniv and K. Cleary, "Image-guided procedures: a review," tech. rep., Georgetown University, 2006.

[14] M. Mack, "Minimally invasive and robotic surgery," *The Journal of the American Medical Association*, vol. 285, no. 5, pp. 568–572, 2001.

[15] K. Cleary, M. Clifford, D. Stoianovici, and M. Freedman, "Technology improvements for image-guided and minimally invasive spine procedures," *IEEE Transactions on Information Technology in Biomedicine*, vol. 6, no. 4, 2002.

[16] R. Galloway, "The process and development of image-guided procedures," *Annual Review of Biomedical Engineering*.

[17] D. Gering, A. Nabavi, R. Kikinis, and N. Hata, "An integrated visualization system for surgical planning and guidance using image fusion and an open MR.," *Journal of Magnetic Resonance Imaging*, vol. 13, no. 6, pp. 967–975, 2001.

[18] D. Pham, C. Xu, and J. Prince, "A survey of current methods in medical image segmentation," tech. rep., John Hopkins University, 2000.

[19] C. Russ, *Image Processing Handbook*. Boca Raton, FL, USA: CRC Press, 4th ed., 2002.

[20] T. Peters, "Image-guided surgery: from X-rays to virtual reality.," *Computer Methods in Biomechanics and Biomedical Engineering*, vol. 4, no. 1, pp. 27–57, 2000.

[21] C. Vinhais, *Medical X-ray images of the human thorax: segmentation, decomposition and reconstruction*. PhD thesis, Faculdade de Engenharia, Universidade do Porto, 2007.

[22] H. Livyatan, Z. Yaniv, and L. Joskowicz, "Gradient based 2D/3D rigid registration of fluoroscopic X-ray to CT," *IEEE Transactions on Medical Imaging*, vol. 22, no. 11, pp. 1395–1406, 2003.

[23] J. Maintz and M. Viergever, "An overview of medical image registration methods," Tech. Rep. 22, Imaging Science Department – Imaging Center Utrecht, 1998.

[24] A. Bardera, M. Feixas, I. Boada, and J. Rigau, "Medical image registration based on BSP and quad-tree partitioning," in *Proceedings of the third international conference on Biomedical Image Registration*, WBIR'06, (Berlin, Heidelberg), pp. 1–8, Springer-Verlag, 2006.

[25] H. Chen, T. Cao, D. Liu, and C. Tang, "An efficient method to realize space registration in image-guided robot-assisted surgery," *ICIRA*, pp. 744–752, 2008.

[26] J. Ashburner and K. Friston, "Rigid body registration," tech. rep., The Wellcome Dept. of Imaging Neuroscience, 2003.

[27] D. Hill, P. Batchelor, M. Holden, and D. Hawkes, "Medical image registration," *Physics in Medicine and Biology*, vol. 46, no. 3, 2001.

[28] B. Zitová and J. Flusser, "Image registration methods: a survey," *Image and Vision Computing*, vol. 21, no. 11, pp. 977–1000, 2003.

[29] R. Khadem, C. Yeh, M. Sadeghi-Tehrani, and M. Bax, "Comparative tracking error analysis of five different optical tracking systems," *Computer Aided Surgery*, vol. 5, pp. 98–107, 2000.

[30] J. Bushberg, A. Seibert, E. Leidholdt, and J. Boone, *The Essential Physics of Medical Imaging*. Philadelphia, PA: Lippincott Williams & Wilkins, 3rd ed., 2012.

[31] J. West and C. Maurer, "Designing optically tracked instruments for image-guided surgery," *IEEE Transactions on Medical Imaging*, vol. 23, no. 5, 2004.

[32] J. Enderle, S. Blanchard, and J. Bronzino, *Introduction to Biomedical Engineering*. Elsevier, 2nd ed., 2005.

[33] C. Hartley, J. Moore, and G. Zouridakis, *Biomedical Technology and Devices Handbook*, ch. 2 – Ultrasonic Blood Flow and Velocity Measurement. The Mechanical Engineering Handbook Series, Taylor & Francis, 2004.

[34] R. Cobbold, "Foundations of biomedical ultrasound," *Oxford University Press*, 2007.

[35] F. Schmidt, *Development of a Time-Resolved Optical Tomography System for Neonatal Brain Imaging*. PhD thesis, University of London, 1999.

[36] W. Schmidt and M. Backhaus, "What the practising rheumatologist needs to know about the technical fundamentals of ultrasonography," *Best Practice & Research Clinical Rheumatology*, vol. 22, no. 6, pp. 981–999, 2008.

[37] C. Vinhais, "Ultrasound physics," in *IMAME Physics Lectures*, Department of Physics, Instituto Superior de Engenharia do Porto (Porto, Portugal), 2009–10.

[38] C. Vinhais, "Ultrasound transducers," in *IMAME Physics Lectures*, Department of Physics, Instituto Superior de Engenharia do Porto (Porto, Portugal), 2009–10.

[39] E. Amstad, S. Zurcher, A. Mashaghi, and J. Wong, "Surface functionalization of single superparamagnetic iron oxide nanoparticles for targeted magnetic resonance imaging," *Small*, vol. 5, no. 11, pp. 1334–1342, 2009.

[40] F. Khan, *The Physics of Radiation Therapy*. Lippincott Williams & Wilkins, 2009.

[41] T. Roberts, C. Macgowan, J. Moore, and G. Zouridakis, *Biomedical Technology and Devices Handbook*, ch. 8 – Magnetic Resonance Imaging. The Mechanical Engineering Handbook Series, Taylor & Francis, 2004.

[42] T. Szabo, J. Enderle, and J. Bronzino, *Introduction to Biomedical Engineering*, ch. 16 – Medical Imaging. Academic Press Series in Biomedical Engineering, Academic Press, 2011.

[43] G. Prasad, A. Joshi, A. Feng, and M. Barysheva, "Deformable organisms and error learning for brain segmentation," in *Proceedings of the Third International Workshop on Mathematical Foundations of Computational Anatomy - Geometrical and Statistical Methods for Modelling Biological Shape Variability*, pp. 135–147, 2011.

[44] J. Bronzino, *The Biomedical Engineering Handbook*, ch. 61 – X-Ray. The Electrical Engineering Handbook Series, CRC Press, 2000.

[45] C. Vinhais, "Computed tomography," in *PIMED Physics Lectures*, Department of Physics, Instituto Superior de Engenharia do Porto (Porto, Portugal), 2008.

[46] J. Bronzino, *The Biomedical Engineering Handbook*, ch. 62 – Computed Tomography. The Electrical Engineering Handbook Series, CRC Press, 2000.

[47] N. Binder, C. Bodensteiner, L. Matthäus, and A. Schweikard, "Image guided positioning for an interactive C-arm fluoroscope," 2006.

[48] L. Wang, P. Fallavollita, R. Zou, and X. Chen, "Closed-form inverse kinematics for interventional C-arm X-ray imaging with six degrees of freedom:

modeling and application," *IEEE Transactions on Medical Imaging*, vol. 31, no. 5, pp. 1086–1099, 2012.

[49] L. Wang, A. McKenzie-Brown, and A. Hord, *The Handbook of C-Arm Fluoroscopy-Guided Spinal Injections*. Taylor & Francis, 2006.

[50] N. Strobel, J. Boese, O. Meissner, and T. Brunner, *Imaging with Flat-Detector C-Arm Systems*. Siemensstrasse 1, 91301 Forchheim, Germany: Siemens AG, Healthcare Sector, MED AX, 2008.

[51] A. Jain, R. Kon, Y. Zhou, and G. Fichtinger, "C-arm calibration – is it really necessary?," in *Medical Image Computing and Computer-Assisted Intervention*, vol. 3749, pp. 639–646, Springer Berlin–Heidelberg, 2005.

[52] C. McCollough, "The AAPM/RSNA physics tutorial for residents – X-ray production," *Radiographics*, vol. 17, no. 4, pp. 967–984, 1997.

[53] D. Thomas, "ICRU report 85: fundamental quantities and units for ionizing radiation," *Radiation Protection Dosimetry*, 2012.

[54] R. Powsner and E. Powsner, *Essential Nuclear Medicine Physics*. Essentials Series, John Wiley & Sons, 2006.

[55] J. T. Bushberg, "The AAPM/RSNA physics tutorial for residents – X-ray interactions," *Radiographics*, vol. 18, no. 4, pp. 457–468, 1998.

[56] M. McKetty, "The AAPM/RSNA physics tutorial for residents – X-ray attenuation," *Radiographics*, vol. 18, no. 4, pp. 151–163, 1998.

[57] Philips, *Philips BV Endura Instructions for Use v2.1*, 2005.

[58] A. Brost, L. Yatziv, W. Gilson, and B. Meyer, "Accuracy of X-ray image-based 3D localization from two C-arm views: a comparison between an ideal system and a real device," in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, vol. 7261, 2009.

[59] C. Morales, E. Nofrerías, J. Mauri, and P. Radeva, "Modelling the acquisition geometry of a C-arm angiography system for 3D reconstruction," in *Proceedings of the 5th Catalonian Conference on Artificial Intelligence*, CCIA '02, pp. 322–335, Springer-Verlag, 2002.

[60] N. Fernandes, "Reconstrução de superfícies usando um scanner 3D de luz estruturada," Master's thesis, Physics Department, Instituto Superior de Engenharia do Porto, Porto, Portugal, 2012.

[61] Q. Meyer, "Camera calibration of C-arm devices," *Radiographics*, vol. 18, no. 4, pp. 151–163, 2006.

[62] E. Trucco and A. Verri, *Introductory Techniques for 3D Computer Vision*. Upper Saddle River, NJ, USA: Prentice Hall PTR, 1998.

[63] O. Faugeras, *Three-Dimensional Computer Vision*. Artificial Intelligence, MIT Press, 1993.

[64] S. Gorges, E. Kerrien, M. Berger, and Y. Trousset, "Model of a vascular C-arm for 3D augmented fluoroscopy in interventional radiology," in *Proceedings of the 8th international conference on Medical Image Computing and Computer-Assisted Intervention*, (Berlin, Heidelberg), pp. 214–222, Springer-Verlag, 2005.

[65] S. Chaganti, D. Kumar, S. Patil, and P. Alderman, "A language for effective communication between surgeons and radiographers in trauma theatre," *Annals of The Royal College of Surgeons of England*, vol. 91, no. 6, pp. 509–12, 2009.

[66] A. Rougée, C. Picard, C. Ponchut, and Y. Trousset, "Geometrical calibration of X-ray imaging chains for three-dimensional reconstructions," *Computerized Medical Imaging and Graphics*, vol. 17, no. 4/5, pp. 295–300, 1993.

[67] S. Birchfield, "An introduction to projective geometry (for computer vision)," 1998.

[68] Y. Morvan, *Acquisition, Compression and Rendering of Depth and Texture for Multi-View Video*. PhD thesis, Technische Universiteit Eindhoven, 2009.

[69] G. Bradski and A. Kaehler, *Learning OpenCV: Computer Vision with the OpenCV Library*. Cambridge, MA: O'Reilly, 2008.

[70] J. Heikkila and O. Silven, "A four-step camera calibration procedure with implicit image correction," in *Proceedings of the 1997 Conference on Computer Vision and Pattern Recognition (CVPR '97)*, pp. 1106–1112, IEEE Computer Society, 1997.

[71] B. Jahne, P. Geissler, and H. Haussecker, *Handbook of Computer Vision and Applications*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1st ed., 1999.

[72] J. Goddard, "Pose and motion estimation from vision using dual quaternion-based extended Kalman filtering," 1997.

[73] N. Binder, L. Matthäus, R. Burgkart, and A. Schweikard, "A robotic C-arm fluoroscope," *The International Journal of Medical Robotics and Computer Assisted Surgery*, vol. 1, no. 3, pp. 108–116, 2005.

[74] D. Ritter, J. Orman, C. Schmidgunst, and R. Graumann, "3D soft tissue imaging with a mobile C-arm," *Computerized Medical Imaging and Graphics*, vol. 31, no. 2, pp. 91–102, 2007.

[75] D. Ritter, M. Mitschke, and R. Graumann, "C-arm CT soft tissue 3D imaging with a mobile C-arm," *International Congress Series*, vol. 1268, pp. 3–6, 2004.

[76] D. Ritter, M. Mitschke, and R. Graumann, "Intraoperative soft tissue 3D reconstruction with a mobile C-arm," *International Congress Series*, vol. 1256, pp. 200–206, 2003.

[77] B. Rosenhahn, *Pose Estimation Revisited*. PhD thesis, University of Kiel, 2003.

[78] T. Phong, R. Horaud, A. Yassine, and P. Tao, "Object pose from 2D to 3D point and line correspondences," 1995.

[79] B. Kainz, M. Grabner, and M. Rüther, "Fast marker based C-arm pose estimation," in *Proceedings of the 11th International Conference on Medical Image Computing and Computer-Assisted Intervention, Part II*, pp. 652–659, Springer-Verlag, 2008.

[80] T. Alter, "3D pose from 3 corresponding points under weak-perspective projection," *AI Memo*, vol. 1378, 1992.

[81] L. Ibáñez, W. Schroeder, L. Ng, and J. Cates, *The ITK Software Guide*. Kitware, 2003.

[82] D. DeMenthon and L. Davis, "Model-based object pose in 25 lines of code," *International Journal of Computer Vision*, vol. 15, pp. 123–141, 1995.

A Philips BV Endura Specifications

The technical specifications that are shown in this Appendix are reproduced from the Philips BV Endura's Directions for Use manual (version 2.1) [57].

Tab. A.1: Philips BV Endura X-ray tube.

definition	**specifications**
type	fixed anode
model name	FO17
model number	9890 000 64801
input nominal power (big focus)	0.81 kW
input nominal power (small focus)	2.1 kW
max. anode thermic content	35.5 kJ (50.1 kHU)
max. cont. heat dissipation	200 W
max. heat dissipation	360 W (30.6 kHU)
target-material	tungsten
target-angle	12°
focal point nominal values	0.6, 1.4 CEI
equivalent filtration quality	0.6 mm eq. AI
nominal tension of the X-ray tube	110 kV

Tab. A.2: Philips BV Endura visualization accuracy.

definition	**value**
voltage offset	± (8% + 0.8 kV)
fluoroscopy current offset	± (8.5% + 0.01 mA)
radiography mAs offset	± (10% + 0.2 mAs)
dose rate accuracy	± 15%
dose rate accuracy (area)	± 25%

Tab. A.3: Philips BV Endura X-ray generator.

definition	**data**
type	high-frequency converter, DC generator
charging data (LDF)	focus 0.6 IEC kV margin 40–110 kV mA margin 0.10–3.00 mA maximum continuous load 120 s average charge 200 W (till 30 s connected) average charge 150 W (till 60 s connected)
charging data (HDF)	focus 0.6 kV margin 40–110 kV mA margin 0.24–7.20 mA maximum continuous load 20 s average charge 100 W (till 30 s connected) average charge 150 W (till 60 s connected)
charging data 1/2- and 1/4 dose (LDF/HDF)	focus 0.6 IEC kV margin 40–110 kV mA margin 0.2–7.2 mA frequency 12.5 Hz / 6.25 Hz pulse width 40 ms max. permanent load 10 min (LDF), 60 s (HDF)

Tab. A.4: Philips BV Endura environmental conditions.

condition	**value**
service temperature	10°C to 40°C (maximum allowed margin) 10°C to 35°C (ideal performance)
storage temperature	-25°C to 70°C
relative performance humidity	20–80%
relative storage humidity	5–95% (without condensation)
atmospheric pressure	700 / 1100 Hp
vibration	10–150 Hz, 0.35 mm peak
collision	25 g, 6 ms

Tab. A.5: Philips BV Endura maximum load factors.

definition	**fluoroscopy**	**radiography**
nominal tension of the X-ray tube	110 kV	105 kV
maximum electric current of the X-ray tube	7.2 mA	20 mA
maximum electric power	792 W	2100 W
nominal electric power	•	2000 W (0.1 s, 20mA, 100 kV)
min. current–time product	•	2.0 mAs (0.1 s, 20mA, 40–150 kV)

Tab. A.6: Philips BV Endura iris collimator.

definition	**9" version**	**12" version**
bundle min. diameter at the image intensifier entrance	< 5 cm	< 5 cm (14 cm)
symmetric field of max. radiation	240 mm 144 mm (fluoroscopy)	240 mm 144 mm (fluoroscopy)
adjustment	continuous	continuous

Tab. A.7: Philips BV Endura lead shutter.

definition	**value**
width adjustment	till < 4 cm slot in the image intensifier
rotation	360°
operation	remote control
shutters	2, independent
LIH indication	yes

B Software, Libraries and Toolkits

The used software, libraries and toolkits to achieve most of the tasks of the presented work are briefly described below.

ITK – The Insight Segmentation and Registration Toolkit
`http://itk.org`

ITK is an open-source software toolkit for performing registration and segmentation. Segmentation is the process of identifying and classifying data found in a digitally sampled representation. Typically the sampled representation is an image acquired from such medical instrumentation as CT, MRI or ultrasound scanners. Registration is the task of aligning or developing correspondences between data. For example, in the medical environment, a CT scan may be aligned with a MRI scan in order to combine the information contained in both.

ITK is implemented in C++. ITK is cross-platform, using the CMake build environment to manage the configuration process. In addition, an automated wrapping process generates interfaces between C++ and interpreted programming languages such as Java and Python. This enables developers to create software using a variety of programming languages. ITK's C++ implementation style is referred to as generic programming (i.e., using templated code). Such C++ templating means that the code is highly efficient, and that many software problems are discovered at compile-time, rather than at run-time during program execution. It also enables ITK to work on two, three, four or more dimensions. A simplified interface to ITK that does not expose templated code, SimpleITK, is available in multiple languages.

ParaView – Open Source Scientific Visualization
`http://paraview.org`

ParaView is an open-source, multi-platform data analysis and visualization application. ParaView users can quickly build visualizations to analyze their data using qualitative and quantitative techniques. The data exploration can be done interactively in 3D or programmatically using ParaView's batch processing capabilities.

ParaView was developed to analyze extremely large datasets using distributed memory computing resources. It can be run on supercomputers to analyze datasets of terascale as well as on laptops for smaller data. ParaView is an application

framework as well as a turn-key application. The ParaView code base is designed in such a way that all of its components can be reused to quickly develop vertical applications. This flexibility allows ParaView developers to quickly develop applications that have specific functionality for a specific problem domain.

OpenCV – Open Source Computer Vision
`http://opencv.willowgarage.com`

Reproduced from [69]:*OpenCV is an open source computer vision library. The library is written in C and C++ and runs under Linux, Windows and Mac OS X. There is active development on interfaces for Python, Ruby, Matlab, and other languages. OpenCV was designed for computational efficiency and with a strong focus on realtime applications. OpenCV is written in optimized C and can take advantage of multicore processors. If you desire further automatic optimization on Intel architectures [Intel], you can buy Intel's Integrated Performance Primitives (IPP) libraries, which consist of low-level optimized routines in many diff erent algorithmic areas. OpenCV automatically uses the appropriate IPP library at runtime if that library is installed.*

One of OpenCV's goals is to provide a simple-to-use computer vision infrastructure that helps people build fairly sophisticated vision applications quickly. The OpenCV library contains over 500 functions that span many areas in vision, including factory product inspection, medical imaging, security, user interface, camera calibration, stereo vision, and robotics. Because computer vision and machine learning oft en go hand-inhand, OpenCV also contains a full, general-purpose Machine Learning Library (MLL). This sublibrary is focused on statistical pattern recognition and clustering. The MLL is highly useful for the vision tasks that are at the core of OpenCV's mission, but it is general enough to be used for any machine learning problem.

C Pose Object CT Images

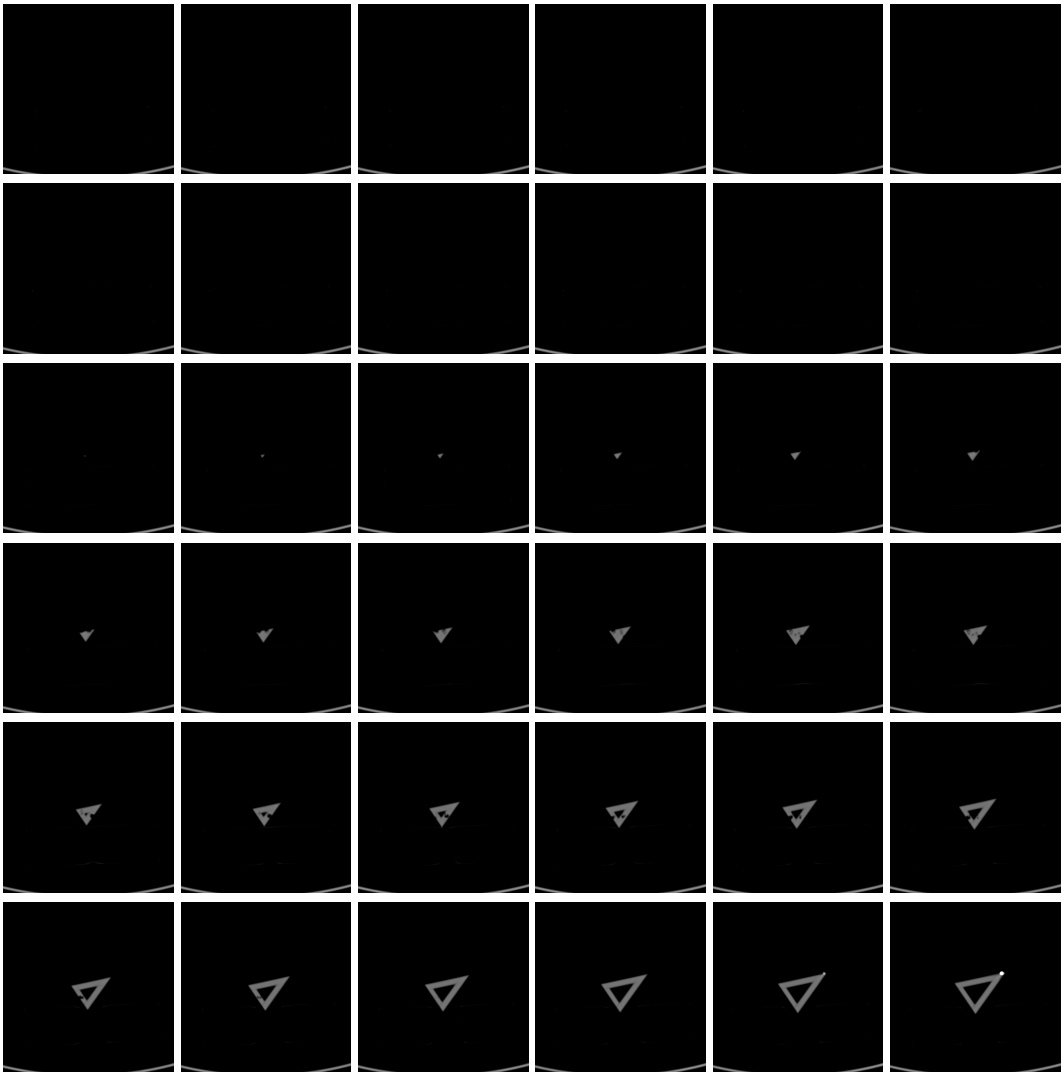


Fig. C.1: Acquired computed tomography images. Slices 0–36.

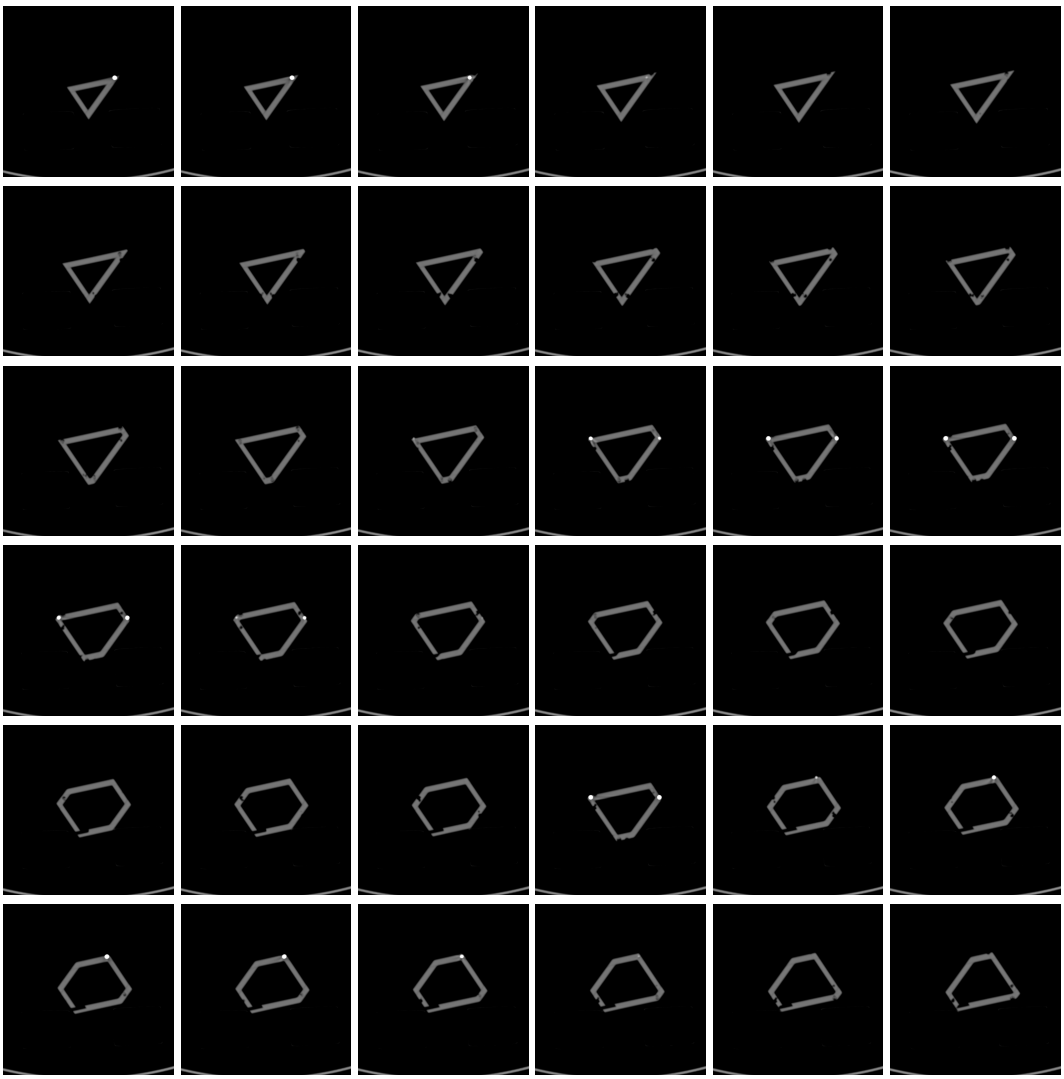


Fig. C.2: Acquired computed tomography images. Slices 37–72.

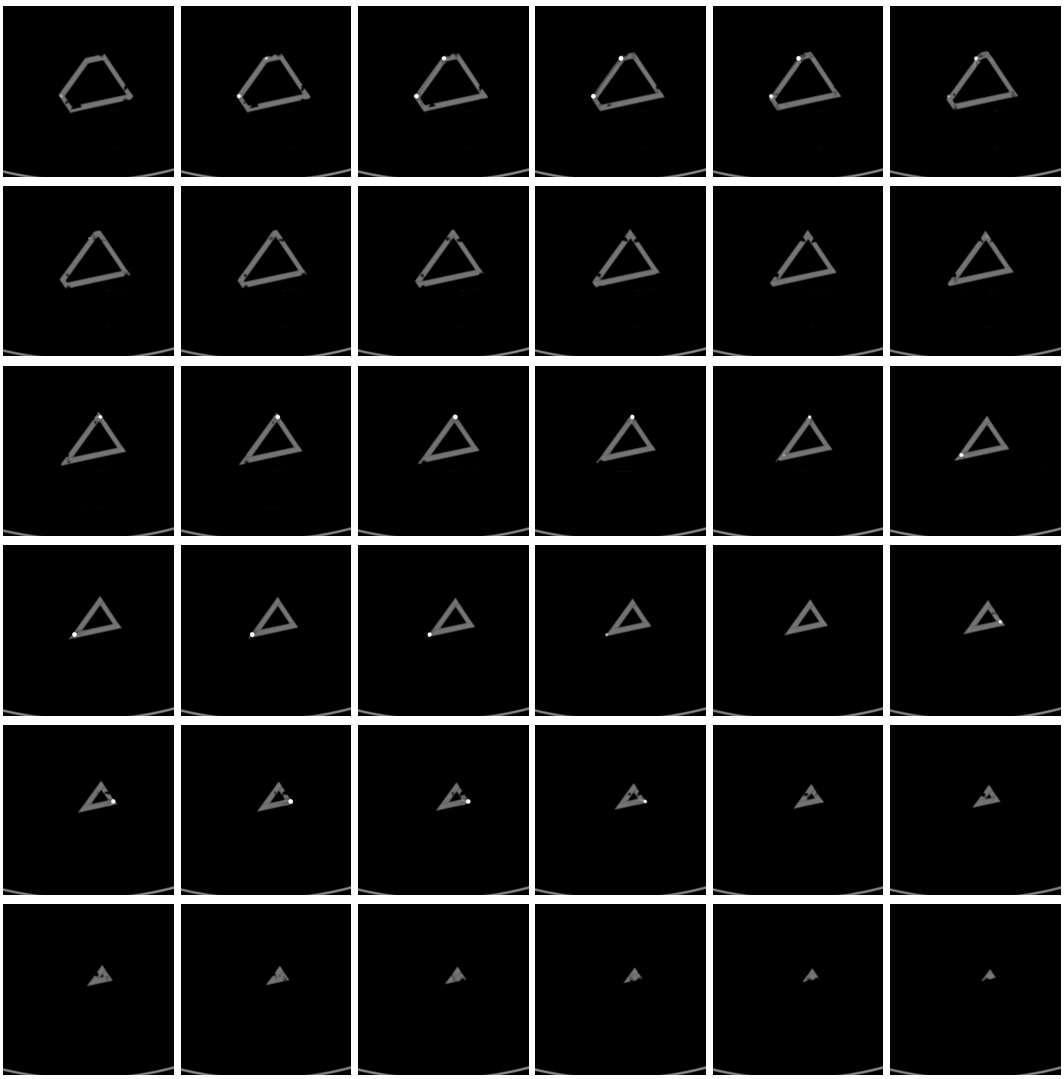


Fig. C.3: Acquired computed tomography images. Slices 73–108.

D Calibration Images Set

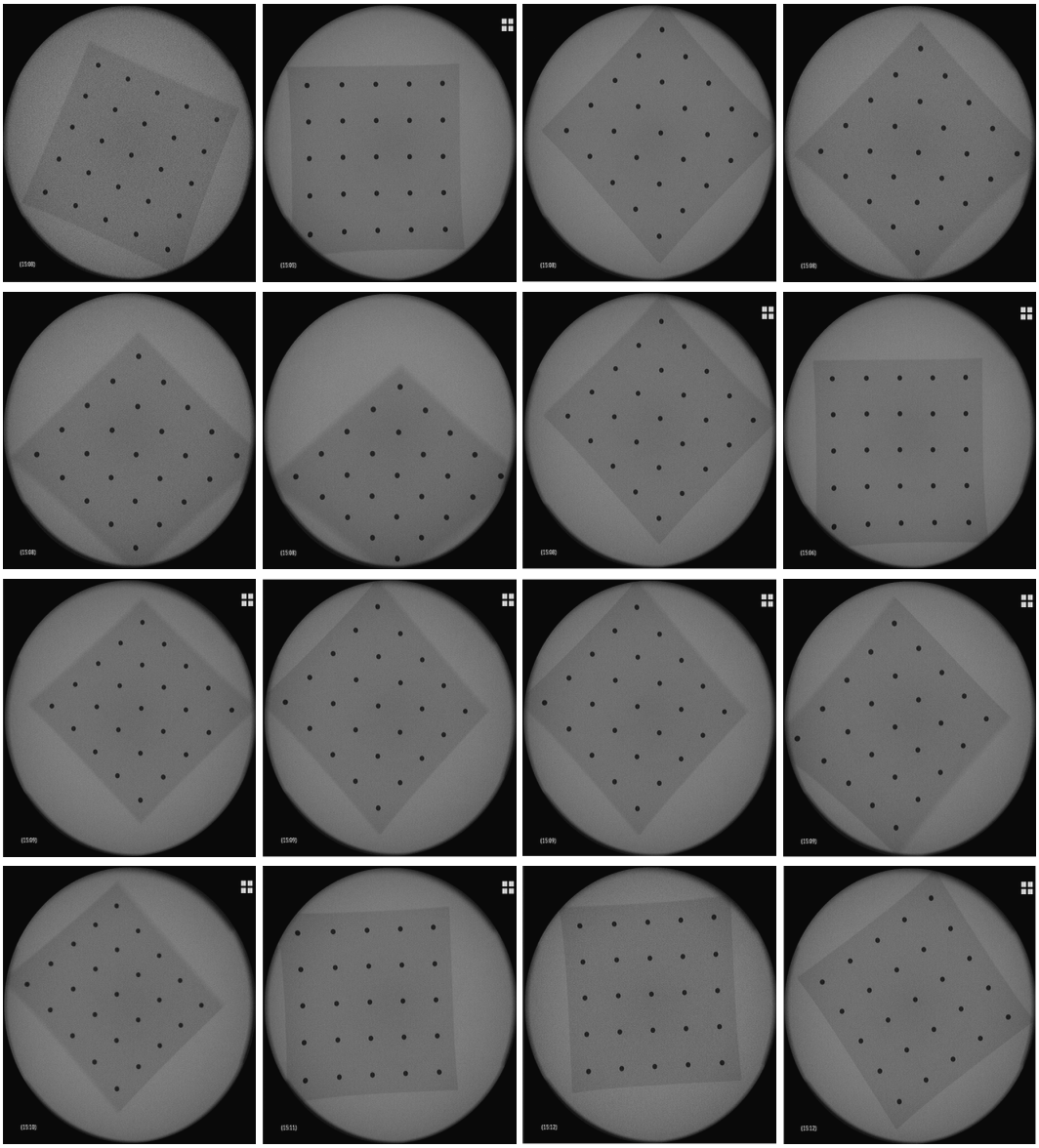


Fig. D.1: Calibration images set. 16 images were used to perform the C-arm calibration.