Faculdade de Ciências e Tecnologia da Universidade de Coimbra

Development of co-registration techniques for comparison of whole-body PET/CT studies

"Desenvolvimento de técnicas de co-registo para comparação de exames PET/CT de corpo-inteiro"

André Filipe Cidra Duarte Rocha Libório

Mestrado Integrado em Engenharia Biomédica

2010/2011

Faculdade de Ciências e Tecnologia da *Universidade* de Coimbra

Development of co-registration techniques for comparison of whole-body PET/CT studies

Author: André Filipe Cidra Duarte Rocha Libório

Scientific Orientation:

Ph.D. Nuno Chichorro Ferreira

Ph.D. Francisco Amado Caramelo

This dissertation is presented to University of Coimbra with the purpose to conclude the Master Degree in Biomedical Engineering. All work was done in IBILI *("Instituto biomédico de investigação da luz e imagem")* and in *ICNAS ("Instituto de Ciências Nucleares Aplicadas à saúde")* at University of Coimbra

À minha mãe e família

Aos meus amigos

Acknowledgments

Chegada é a altura de se fazer uma espécie de balanço de 5 anos memoráveis nesta cidade carregada de mística, que passaram num ápice e me deixam muitas boas memórias. Parafraseando a letra da tão conhecida "balada da despedida": "**Segredos desta cidade / levo comigo p'rá a vida**", digo que estas palavras têm para mim um significado verdadeiro e transmitem um sentimento especial que só o percebe em pleno quem estuda/ou nesta cidade. Espero e acredito que a minha ligação a ela não fique por aqui!

Deixo, desde já, o agradecimento mais importante à minha mãe por tudo o que fez por mim e por me ter dado as condições para chegar ao fim desta etapa. Agradeço também à minha irmã por me receber calorosamente em casa à sexta-feira, mesmo sabendo que possivelmente lhe iria infernizar à cabeça durante todo o fim-de-semana. Aproveitando o agradecimento à família, deixo desde já também o meu sincero agradecimento aos meus tios. Entre eles (que são muitos) não posso deixar de destacar a minha tia Maria de Fátima pela sua grande amizade e por sempre ter acreditado em mim mesmo quando nem eu acreditava. Agradeço também ao meu tio João e à minha tia Lurdes por no ano de adaptação a uma realidade tão diferente, que foi o ano de ingresso, tão bem me receberem em sua casa fazendo-me sentir em minha própria casa. Deixo ainda um agradecimento muito especial à minha tia Prazeres e ao meu tio Carlos que, nestes dois últimos anos, juntamente com os meus primos me inseriram no seu núcleo familiar fazendo-me sentir muito bem e, de certeza, num futuro irá deixar saudades.

Pela realização deste trabalho, agradeço aos meus orientadores Nuno Ferreira e Francisco Caramelo pela disponibilidade e apoio. No entanto, queria destacar a orientação prestada pelo professor Nuno Ferreira. Agradeço-lhe sinceramente pela paciência que teve comigo, pela disponibilidade constante de, se fosse preciso, se sentar comigo em frente ao computador e juntos desenvolver códigos e pela grande ajuda prestada na redacção desta tese.

A todos os meus amigos (eles sabem quem são) deixo também o meu agradecimento pela amizade e companheirismo em todos estes anos.

Por último queria também deixar marcado neste texto o meu agradecimento mais que especial ao meu avô João, que, infelizmente por circunstâncias da vida, não pôde assistir ao cumprimento deste meu objectivo, mas sei que certamente estaria a torcer para que o alcançasse e neste dia partilharia da minha felicidade.

A todos vós o meu **OBRIGADO**!

Abstract

The objective of this thesis is to develop an algorithm that co-registers, in a single volume, whole-body Human PET/CT Scans taken at different instants in time, to facilitate the comparison of studies in situations where the patient had to multiple PET/CT exams. The idea is to isolate, from the initials CT volumes, the rigid body parts that can be transformed with a rigid transformation – the bone structures. Bone structures were obtained with segmentation (the process of dividing a volume into subsections, here based on intensity values) and a further labeling, taking into account the connectivity between neighboring voxels.

Once the bone structures were labeled in both volumes, the next step was the establishment of its correspondences between the regions in the two volumes. This was based on the number of voxels associated with each region (size of region) and with the Euclidian distance between the centers of mass of regions from different volumes. Once this correspondence was achieved, the rigid transformation model is constructed and applied to each regions of the volume, which was defined as the Target volume.

The results of the developed method are still far from those required for clinical applications. We hope that, with future developments of the presented algorithm, will lead to the improvement of medical image analysis in this field.

Table of Contents

1) INTRODUCTION

1.1) CONTEXT

One of the most important issues traversing all societies has always been the development of healthcare solutions. A large amount of tools and devices capable of providing these healthcare solutions and are in continuous progress. The evolution of technology is behind all this progress, and is a result of the natural ambition of mankind to want more and better. In fact, Human's inherent lack of conformity has always lead the search for new and better solutions to the problems of humanity and, because of that, the healthcare solutions provided to general population have been increasing its quality over the years [1].

In clinical centers, the way clinicians diagnose and plan or execute therapy procedures is a result of a broad evolution of techniques. Most of diagnosis and therapy methods are based on images that cannot be seen in naked eye because of information that are being imaged is in the interior of body. These images are acquired with specific imaging devices such as CT (Computed Tomography), PET (Positron Emission Tomography), MRI (Magnetic Resonance Imaging), US (Ultrasound), and so on [2].

Images are provided to the clinical staff and the diagnosis is dependent on the performance of the technician that is on duty. Given the large amount of images and information that needs to be interpreted, there is an increasing interest in developing computerized systems to facilitate the interpretation of information. In some medical centers, CADx (Computer-aided Diagnosis) are already a reality supporting the medical diagnosis. These systems use a priori knowledge of physiological/anatomical conditions of the body and use elements of artificial intelligence and digital image processing. A typical application is the detection of tumors (breast, lung and colon mainly).

In order to exploit the maximum information contained in images, the computer-aided analysis provides a new set of information to clinicians and specialists involved in healthcare. From this set can be highlighted:

- extraction of quantitative parameters, computed in some region of interest (ROI),
- precocious detection of any variation in images from the same subject taken at different instants in time, important for disease stage monitoring and support the assessment/reassessment of prescribed therapies procedures,
- fusion of information from different complementary imaging modalities in a single image,
- comparison of inter-subject images in a way to analyze and extract similar pathological parameters – Anatomical/Functional atlas construction.

The use of computing capabilities in the medical imaging field may be extended to simulation and planning, where images can aid the construction of a *"virtual patient"* for therapeutically procedures testing and comparison. These *"virtual patients"* are analogous to aircraft simulators used to train airline pilots and may have a huge importance in future developments in healthcare [3].

1.2) MOTIVATION/OBJECTIVE

An oncologic patient usually has the need to be scanned with PET/CT technology several times over his diagnosis and treatment. These scans were not performed (unfortunately) with the patient exactly at the same position. The comparison of several scans of the same patient would be much more effective if it was possible to automatically move specific regions of the patient's body, using natural movements determined by its specific anatomy, until two or more scans were perfectly aligned with each other. It would then be possible to analyze in greater detail the changes that occurred in the patient from examination to examination, and it would be easier to detect earlier signs of the disease that were difficult to see in the initial exams. On the other hand, a tool of co-registration based on whole body atlas facilitates the automatic classification of different regions of the body, namely the bone regions, opening the door to other applications, such as the analysis and detailed comparison of studies from different patients and patient groups, as well as the creation of semantic databases for normal and pathological situations.

Another potential application is the possibility of reducing the radiation dose that the patient receives in each examination, if it would be possible to use CT images from previous scans in the calculation of the CT of the patient in the current position, for the purpose of attenuation correction of PET data and lesion localization.

The main objective of this work is therefore to develop a tool to automatically coregister two PET / CT scans of the same patient, acquired in different times. The registration is done with the possibility of aligning the body based on the articular character of body structures that are obtained by segmentation (bone structures).

1.3) ORGANIZATION OF DOCUMENT

The structure of this document follows the methodology employed during the development of this work. Chapters 2, 3 and 4 bring the theoretical foundations related with this work, respectively: the PET/CT device and its characteristics; a review of some Image Processing techniques and a detailed description of a specific one: Image registration. In section 5 a brief review of the state of art in medical image registration is provided. The methodology applied is presented in section 6 and its results were presented in section 7. In section 8, the results from the previous section are discussed, as well as further possible developments in the methodology used.

2) PET/CT

2.1) INTRODUCTION

The visualization and interpretation of the behavior of atomic particles and their interactions is a theme that have always intrigued the scientific community, and some of their exponents like Einstein, Maxwell, Bohr and Louis de Broglie for example made incredible advances in realizing particle origins, understanding its existence and interpreting their relations. Since then, the scientific community has continued studying the "unseen world" and tried to get the maximum benefit from it, developing new applications that somehow act in human profit.

One example is the use of scientific discoveries in the development of health care solutions. In 1895, Wilhelm Conrad Röntgen discovered, almost by accident while making his experiences, that it is possible to evaluate the interaction of particles (radiation) with the human body. Röntgen flabbergasted the medical community inventing the "X-rays" several used in diagnosis, treatment planning and procedures. Since then many other image modalities with different characteristics appeared, based on transmission, reflection or refraction of radiation. Examples are CT (Computed Tomography), PET (Positron Emission Tomography), SPECT (Single Positron Emission Computed Tomography), MRI (Magnetic Resonance Imaging) and US (Ultrasounds), which yield anatomic and/or functional information according to the application desired [4].

Anatomic imaging allows looking inside of the body non-invasively and provides a well-defined shape characterization of anatomic parts of the body. Interpreting the variation of intensity of a known X-Ray beam that crosses the body (CT), applying a magnetic field as a way to analyze the diamagnetic characteristics of the human tissue (MRI) or taking advantage of the reflections and transmissions of ultra-sound waves in certain tissues (US) are solutions to obtain anatomical images. These imaging techniques are suitable to detect structural alterations but frequently it is difficult to infer if these changes are malignant or benign. By way of contrast, functional imaging usually yields information on the metabolic behavior of tissues and permits to assess the malignancy of suspicious tissues, however with less

anatomical accuracy. The basic idea of functional imaging is to label with an "unstable" particle (isotope), a molecule of interest and map its spreading in the body by detecting photons which arise from the interaction between the labeled molecule and the body with a scanner. The most sensitive functional imaging modality, Positron Emission Tomography (PET), takes advantage of the characteristics of the unstable isotope, which leads to the emission of 2 photons in opposite directions, resulting from the annihilation of a positron and an electron. The most clinical used examples of anatomic imaging are CT, MRI and US, and in functional imaging are PET and SPECT (Figure 1) [5].

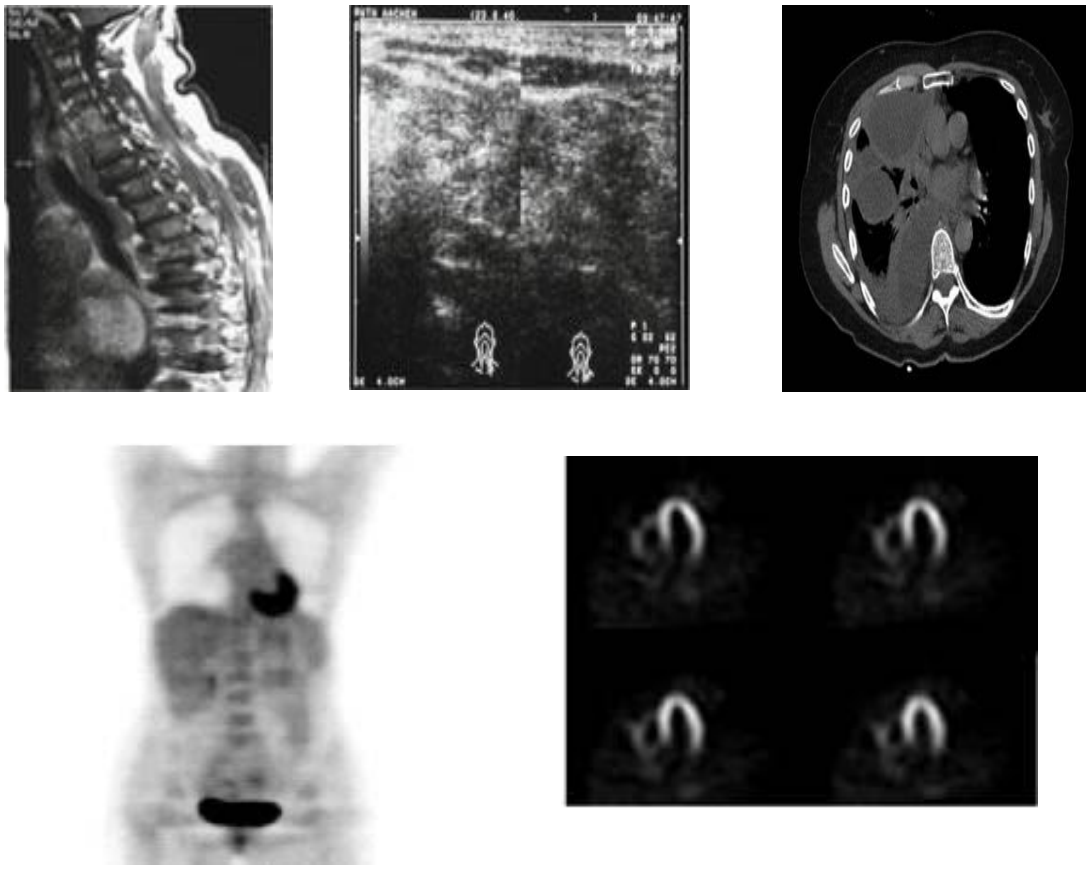


Figure 1 - Examples of imaging methodologies routinely used in clinical practice. From left to right: (top) Magnetic Resonance Imaging (MRI), Ultrasound (US) and Computer Tomography (CT) [taken from 1]; (bottom) Positron Emission Tomography (PET) (taken from [6])

Having said that, one obvious question arises: why not combine functional and anatomical imaging techniques in a way that retains both virtues?

This question seems easy to implement in clinical practice, but as will be seen, it is not. The combination of modalities could be done in two ways: by comparing PET and CT exams side-by-side (low accuracy) or by computationally fusing both exams (right way to do). The

idea of exams fusion is that each voxel of CT data has a certain corresponding voxel in PET data and the computer must find the transformations that allocate these voxels. This could be a very complex process because of variability of patient position and oscillation over time of the position of internal organs, requiring labor-intensive mapping techniques, sometimes without satisfactory results (mainly at small lesions adjacent to mobile organs) [7].

The fusion of PET and CT can be, as it was said, computationally very heavy. The fusion of both exams that were essentially based on software methods such as transformation models – **Software fusion approach**, has evolving to a new approach. In a way to avoid insensitive-labor algorithms to fuse PET and CT data and increase the fusion accuracy, the solution was the construction of a single (hybrid) device with embedded CT and PET scanners that sequentially take and almost instantly and co-register both exams – **Hardware fusion approach**. These exams are obtained with the patient lying almost in the same position at the imaging table, thus facilitating the alignment process because of minimization of temporal and spatial differences. Table 1 summarizes the comparison of Software and Hardware Fusion approaches and summarizes some of the Software fusion approach drawbacks that Hardware fusion approach intend to overcome [7, 8].

Table 1 – Software fusion versus Hardware fusion (adapted from [8])

Software Fusion Approach	Hardware Fusion Approach
Access to archives to obtain images	Images Immediately available
Very different patient position	Patient almost in same position
Different scanner imaging table profiles	Same imaging table
Internal organs movements	Little movement of internal organs
Disease progression over time	Scan taken close in time
Limited registration accuracy	Better registration accuracy
Sophisticated transforming algorithms	No further alignment required

These above-cited reasons led many research groups to point their efforts towards developing a device that meets medical requirements and in 1998 was clinically introduced the first PET/CT prototype composed by a spiral CT with PET capability (Siemens Medical Solutions) [9]. The device had PET and CT scanners separated by 60 cm mounted in a single support (see Figure 2). Many other prototype examples arose from several other medical imaging

companies, allowing the instrumentation progress and the improvement of the device's characteristics [10].

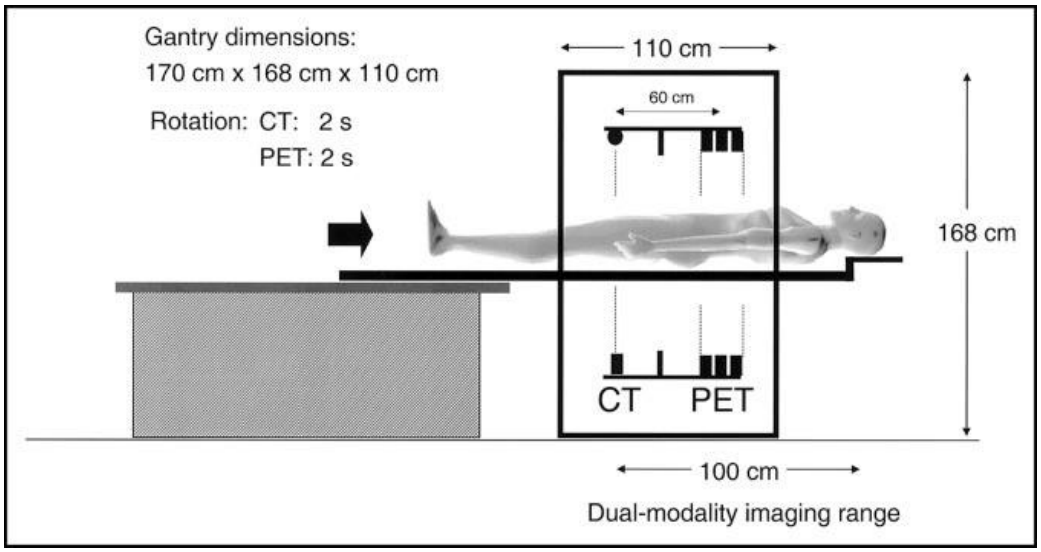


Figure 2 – Schematization of the first prototype of hybrid PET/CT (taken from [10])

One of the greatest advantages of hybrid devices (beyond the possibility of combining anatomical and functional information) is the use of CT based attenuation correction, which is very useful reducing the overall duration of examination in 40%, because it avoids the lengthy PET transmission scan. The PET attenuation correction is obtained integrating the linear attenuation coefficients along the path of the LOR (line-of-response) in CT transmission, taking into account the different energy of the PET photons (511 KeV) relative to CT (40-140 KeV). Another advantage is the possible of use CT information to correct the influence of scattered coincidences, the photons that suffer path deviations because of interaction with tissues. [5, 8]

Hybrid devices have 10%-15% of statistically significant improvement over the separate interpretation of exams [11]. Accuracy of medical diagnosis, in staging and re-staging of disease, as well as planning and monitoring of treatments are improved as a consequence of technologies and methodologies improvements.

2.2) COMPUTER TOMOGRAPHY (CT)

Tomography derives from the Greek word *tomos* which means "section" or "part" and *graphy* that is related to "writing" or "representation", and together means "section writing" or "section representation", related with the ability of achieving a higher vision capability

than human eyes. Computer tomography (CT) is a technique that arose after radiography. It is also based on measuring the attenuation suffered by an x-ray beam after it has crossed different tissues, but the data are acquired in a different way (e.g., sequence of projections instead of a single projection) and are processed by sophisticated computational algorithms to form an image. The contrast enhancement relative to normal radiographs and the digital acquisition and processing of images are among the main improvements introduced by this technology [12].

In fact, conventional radiography is incapable to distinguish differences in contrast inferior to 5% [12]. Radiographies are 2D projections and all tissues crossed by the x-ray beam influence the image and, in particular, structures aligned parallel to the x-ray beam are not correctly imaged. Another drawback of radiography is that the large-area x-ray beam suffers much scatter in the human body, interfering with imaging in structures with similar contrast.

CT made the transition from the analog (characteristic of conventional radiography) to the digital field and in consequence changed completely the way things are done in a medical imaging center. Image digitalization opens new opportunities to image manipulation, storage, transmission and display and quickly seduced medical community. Medical Images centers have transformed into multidisciplinary centers where different fields (physics, electronics, informatics, medicine) act together to improve image quality and as consequence try to improve the health care solutions [12, 13].

Computer tomography (CT) is an anatomical/morphological imaging technology based on transmission tomography. The source of radiation used in imaging is an X-ray tube, external to the body, which generates a beam whose transmission in the tissues is measured, yielding attenuation coefficients of tissues, which were related with density of tissues, yielding the characterization of structures and describing anatomical proprieties of body. High energy X-Rays penetrate the body and are detected by an arc of detectors in a fixed geometry. This way, cross-sectional transmission images are acquired and, by rotating the apparatus (source and detectors together) around the patient, body images can be reconstructed [4].

The basic principle behind the image reconstruction is, as has been said, the attenuation of an x-ray beam after crosses a body. The radiation beam produced at the source, with initial intensity I_0, gradually loses intensity until it reaches the detector with intensity I, because of the interaction with matter. Mathematically, this process could be seen as

$$I = I_0 e^{-\int_{i=1}^{n} \mu_i x_i} \qquad \textit{Equation 1}$$

where the term μ_i is the linear attenuation coefficient of a body element with thickness x_i (different elements have different coefficients across one line because, as is well known, the human body is not homogeneous). For example the attenuation coefficient of fat tissue differs significantly from the attenuation of bone or water and because of that it can be distinguished in images. In order to standardize the data and provide sufficient gray scale for display, the voxel data of the reconstructed volume are typically converted to CT numbers (Hounsfield units (***HU***)) as shown in the follow equation:

$$HU = \frac{\mu_{tissue} - \mu_{water}}{\mu_{water}} \times 1000 \qquad \textit{Equation 2}$$

where, μ_{water} represents the linear coefficient attenuation of water and μ_{tissue} represents the linear coefficient of the tissue. Table 2 shows the HU values of some physiologic tissues. Note that these values vary according to the bibliography used [6, 13, 14].

Table 2 – CT values for some human tissues (adapted from [14])

Tissue	CT number (HU)
Boné	1000
Liver	40 – 60
White matter	~20-30
Grey matter	~37-45
Blood	40
Muscle	10 - 40
Kidney	30
Cerebrospinal fluid	15
Water	0
Fat	-50 - 100
Air	-1000

Figure 3 shows some configurations of CT devices and focuses on the source and detector configurations. The two configurations at the bottom are the example of a “third and

fourth generation". Nowadays, these systems are the most used in clinical practice because they provide quality images in acceptable time. The configuration of the PET/CT device used to obtain the images processed in this work ("*Gemini GXL"*) is "fourth generation". It consists in the rotation of a single source with a stationary ring of detectors that surround patient. Then the measures taken by the detectors are transmitted to a computer that, with mathematical algorithms, will achieve the attenuation coefficients [12].

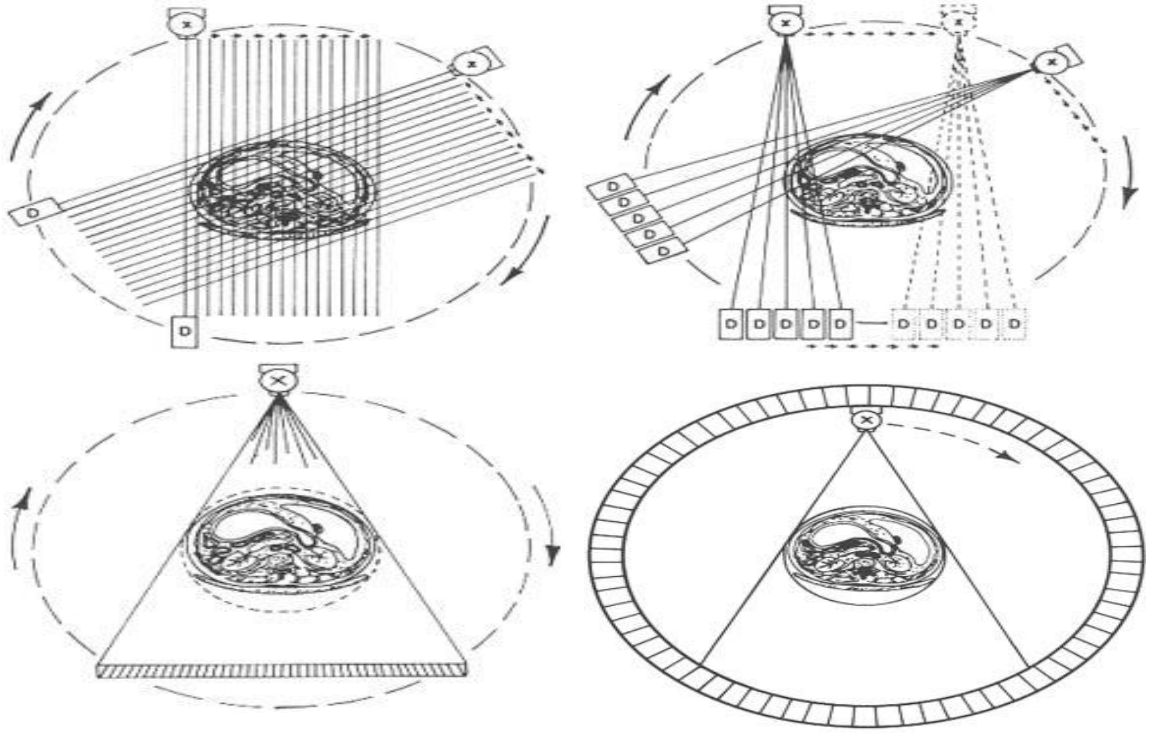


Figure 3 – The different possible configuration to do CT. From left to right: - (top) "first generation" and "second generation" ; - (bottom) "third generation" and "fourth generation" (adapted from [12])

To whole body examinations, spiral CT scanning using "third" and "fourth generation" configurations were used since 1989. X-rays tube rotates while imaging table moves without stopping (Figure 4). The helical movements of apparatus synchronized with imaging table longitudinal movement permits acquisition of 64 thin slices (0.625 mm) in as little as 0.35 s. Is evident the reduction of time consumption of scan and consequently the dose "charged" in patient [12, 13].

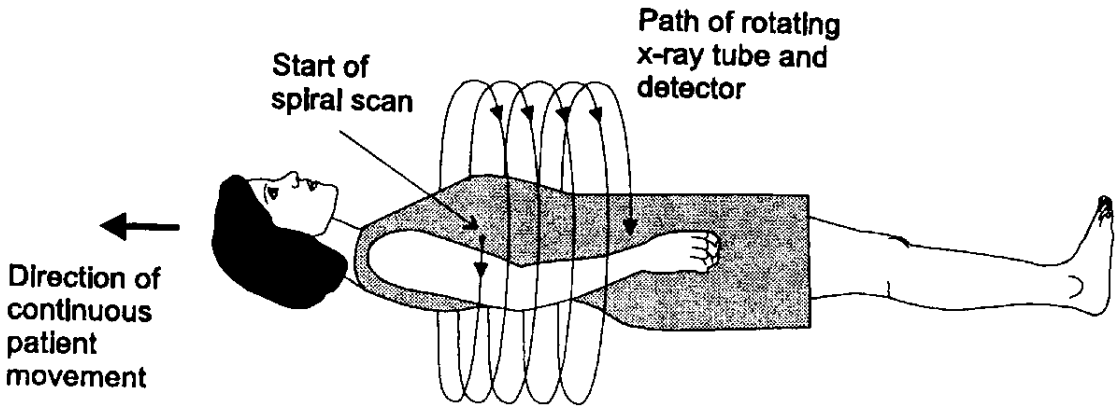


Figure 4 – Spiral CT scan (adapted from [12])

2.3) POSITRON EMISSION TOMOGRAPHY (PET)

PET is an imaging system where the distribution, in the body, of a specific substance is measured by taking advantage of an unstable state induced in one of the atoms of its molecules. The unstable state in particles is induced by the bombardment of stable elements with accelerated particles, and these elements transform into neutron-deficient isotopes that achieve stability through the nuclear transmutation of a proton in a neutron consequent emission of a positron (β^+) and a neutrino (ν). The interaction of a positron with its anti-particle (the electron) leads to the simultaneous emission of two 511 KeVs photons approximately in opposite directions.

In 1957, Hal Anger purposes a camera (gamma camera) capable of detecting high energy photons with external instrumentation [15]. The concept of gamma ray imaging idealized by Anger has survived over the years and is still the basis of more recent devices. One adaptation of this concept led to the development of a system that detect in coincidence two opposite direction photons – PET. Injection into the body of a substance labeled with an unstable isotope (radionuclide), known to be an important component of a specific biological/physiological mechanism, and by further detection of the annihilation photons, the estimation of where the photons are emitted and consequently its distribution can be monitored.

FDG (fluorine-18 fluoro-2-deoxy-D-glucose) is the most widely labeled substance (radiopharmaceutical) used in PET imaging, namely in oncology. FDG is an analogue of glucose,

integrating its metabolic pathway. Monitoring the glucose uptake of cells is discriminative of their activity and can be used for tumor detection. In fact, tumor cell are in continuous differentiation, so glucose uptake by these cells is superior when compared with "healthy" cells. By comparing glucose uptake of cells, detection of tumors can thus be done [9].

After being administrated in a patient, the labeled substance decays and the energy of the positron gradually decreases because of the interaction with matter, until it interacts with its anti-particle – the electron – when the energy is low, forming a transient state known as the positronium. The two gamma-rays (resulting from the are detected by a ring of detectors (composed of scintillation crystals and photo-detectors) in coincidence (thus detecting coincidence events) and, after specific analog/digital processing, stored in specific matrices called sinograms. With reconstruction algorithms it is possible to recover the underlying radioactivity distribution and, consequently, the tracer distribution that may be associated with a specific biological/physiological/… mechanism. Figure 5 summarizes the entire process [8]:

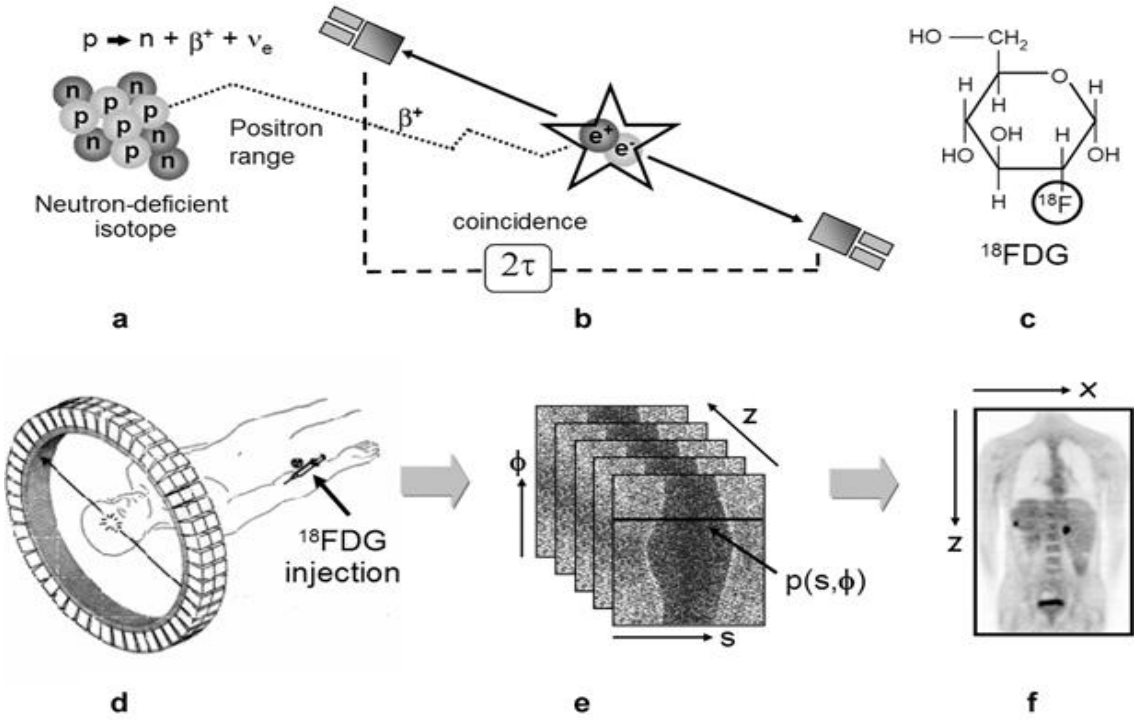


Figure 5 – Summary of the PET acquisition process. The radioisotope emits a positron (a) that interacts with his anti-particle (electron) annihilating and originating 2 photons that are detected in coincidence (b). The most used radiopharmaceutical used in clinical routine is FDG (c) and is injected directly in the patient circulatory system (d). Raw data is saved in sinograms that contain the information about the annihilations in a specific projection (e) and are used to generate the final reconstructed image (f) (adapted from [8])

The detection of gamma rays is done with scintillation crystals that convert the energy from photons into electrical signals that are subsequently processed by an electronic circuit.

The scintillation crystals absorb and convert gamma-rays into light photons that are further converted into an electric quantified signal by photomultipliers or other photodetectors (such as APD's - avalanche photodiodes - or SiPM – silicon photomultipliers). This electric signal is analyzed to evaluate time and energy (essential to prevent random and scattered events) and this process is done for each LOR (line-of-response) defined in the system. If a signal passes the energy threshold (511 KeV, characteristic of annihilation photons), time of occurrence is recorded and used for coincidence analysis (Figure 6) and the location of annihilation is estimated [5].

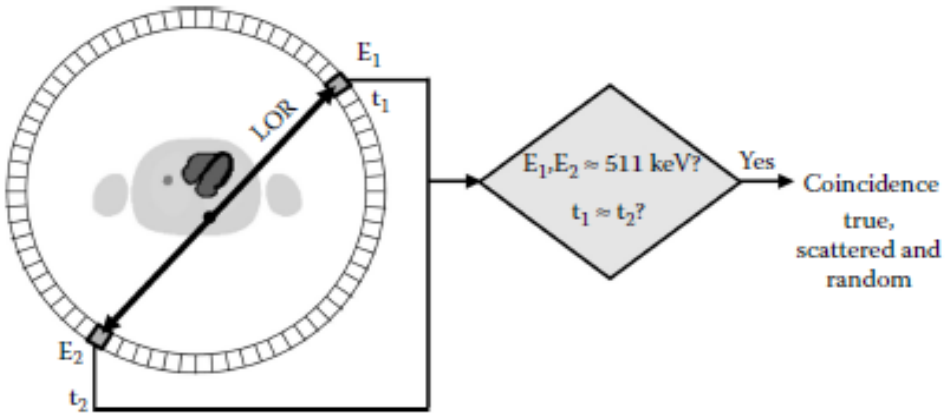


Figure 6 – Illustration of coincidence detection of 2 photons and the circuit with time and energy filtering. (Taken from [5])

The spatial resolution of PET can be seen as the minimum distance between two points on an object that could be seen as distinctive points and in common PET devices is about 4 - 5 mm. PET's spatial resolution is conditioned by the fact that the location of radioactive decay may differ (in the order of the fraction of millimeter for ^{18}F) from the location where the annihilation occurs, so what are being estimated are annihilation sites and not decays sites. Another factor that influences spatial resolution is the residual momentum that results from the non-completely dissipation of positron energy that interferes with the 180° angle difference of two photons and consequently in the photons' detection process. Advances in spatial resolution are being achieved by investigating alternative image reconstruction and data processing algorithms, improvements in instrumentation and correction of patients' motion with help of external optic detectors that assess the patient motion [5, 11].

2.4) PET/CT DEVICES AND MAIN APPLICATIONS

Nowadays there are five major companies selling clinical PET/CT devices: Philips Medical Systems, Siemens Medical Solutions, Hitachi Medical, GE Healthcare and Toshiba Medical Corporation (Figure 7). All equipment, excepting Hitachi Corporation's device (Figure 7 (E)), are capable of 16 slice CT imaging yielding images with reasonable anatomical detail. The time of PET and CT scans are also acceptable: less than 30 minutes. This is a vantage point for these devices, because, in most of cases, the urgency for correct diagnosis is too high and in an extreme case, could make the difference between life and death. From the standpoint of the medical imaging service, reducing the acquisition time of PET or CT both increases the number of studies and facilitates and optimizes their logistic process, either in patient's throughput and professional staff planning [5, 10].

The data used in this work were obtained with a PET/CT scanner owned by ICNAS – Gemini GXL (Philips Medical Systems, Illustrated in Figure 7 (C)).

PET/CT technology contributes to the continuous improvement of the medical service provided to the population, mainly in providing more accurate and timely diagnosis, and because of that it is widely needed in Medical centers, however, because of its cost, almost of them cannot afford it. PET/CT is not only valuable in oncology, where its utility is undeniable known namely in tumors detection. In fact is already an important tool in other clinical areas like in Cardiology – in coronary artery disease (CAD) detection, in radiation treatment planning – detecting metastasis spread and in neurology – detecting neurodegenerative or inflammatory disorders and epilepsy [10, 16].

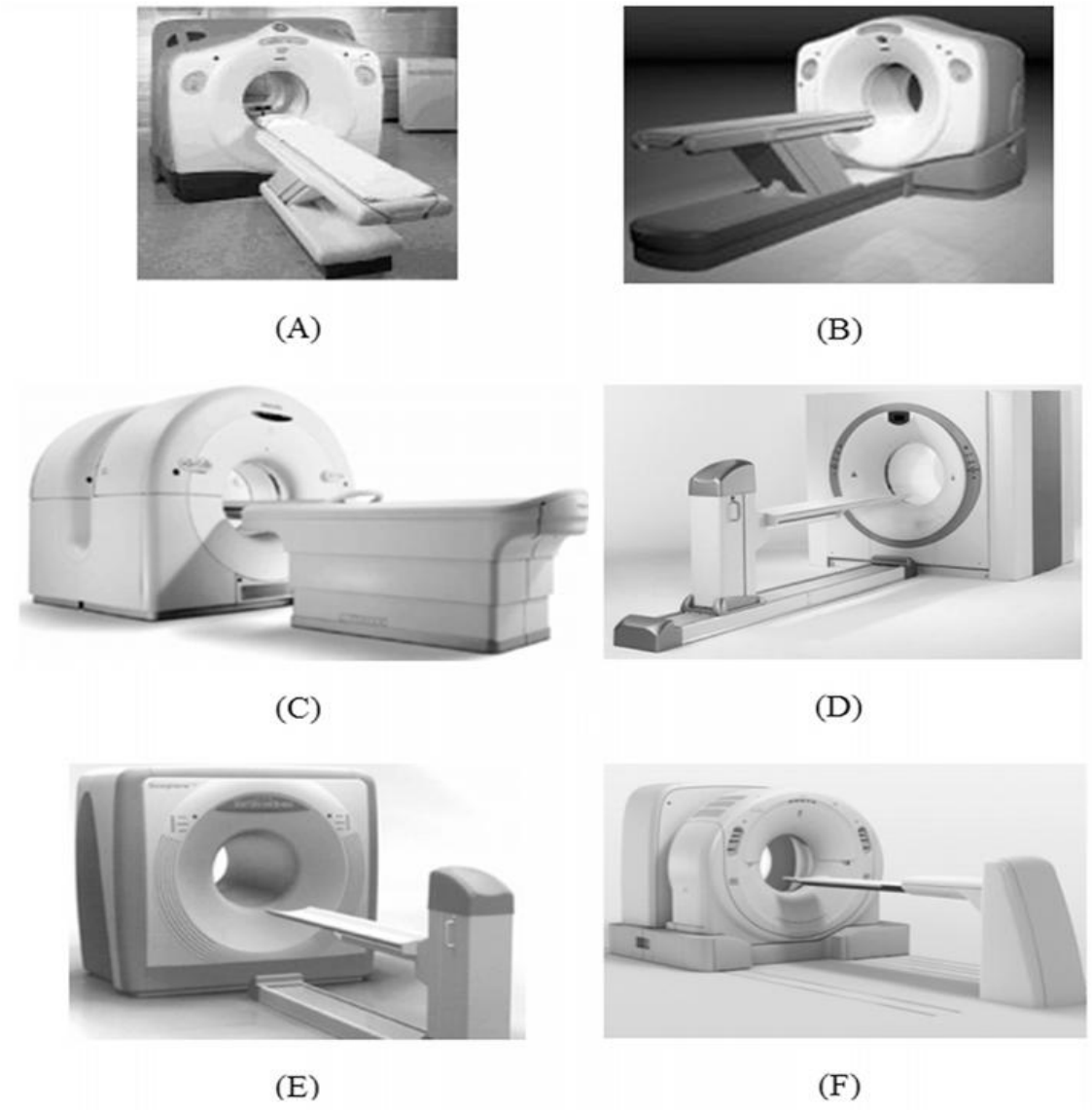


Figure 7 – Some of the most popular commercial PET/CT systems. (A) – Discovery LS (GE Healthcare), (B) - Discovery ST (GE Healthcare), (C) – Gemini (Philips Healthcare), (D) – Biograph (Siemens Healthcare), (E) – Septre P3 (Hitachi Medical Systems), (F) – Aquiduo (Toshiba Medical Corporation) (taken from [8])

3) IMAGE PROCESSING

3.1) INTRODUCTION

Medical images are in continuous progress and almost every day a new project using medical images arises from several scientific groups. As was already said, the first way achieved to produce medical images of the interior of the body was by X-Ray imaging. Since its discovery in 1895 until 2010, more than 5 billion medical imaging studies were done worldwide [17].

One important progress in the medical imaging field was the transition of images provided by medical devices to digital form. In fact, medical imaging in general, and this improvement (digitalization) in particular, led to the joint efforts of two distinct areas such as medicine and engineering, and medical image processing became a critical field in health care solutions [17].

Medical image processing could be defined as the branch that studies the best transformations to apply to medical images in a way to optimize their quality or its interpretation simplicity. One important note is that images, nowadays, are acquired with several signal processing techniques and the term “medical image processing” are only related with techniques/methods applied after image formation, also called “digital image processing”. The improvement of quality is achieved by a broad of set of algorithms and tools that were applied to rectify, analyze and visualize images. Tasks such filtration of noisy or degraded images, detection or visual enhancement of certain details in images, characterization of shapes and structures and also the possibility of comparison of different subjects and modality images in a single image, are some of the major strengths of image processing field (Figure 8) [5].

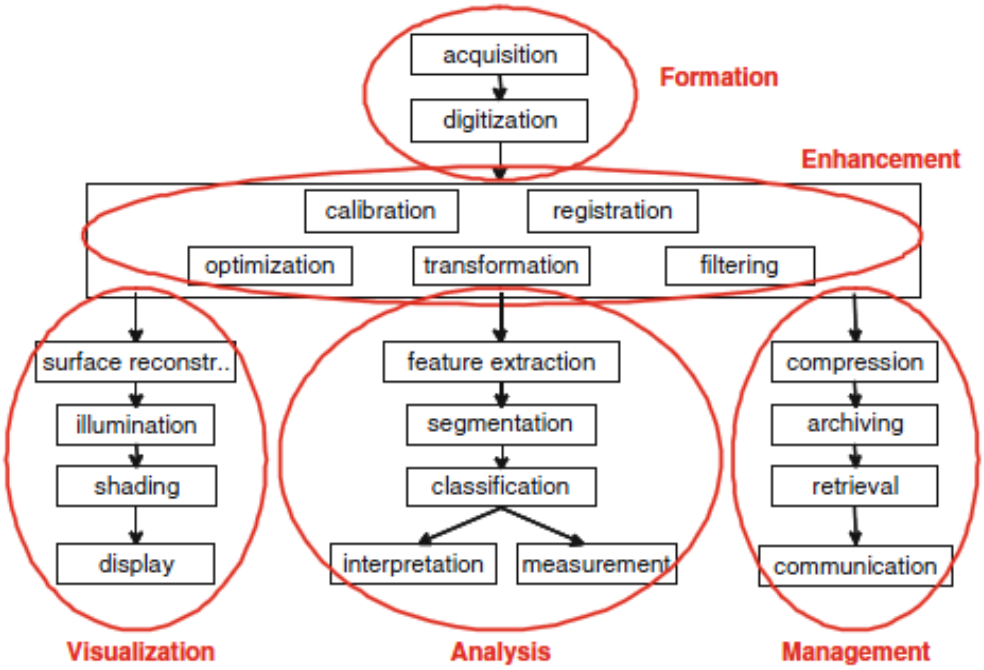


Figure 8 – Main areas covered by image processing techniques (taken from [4])

Digital images are 2D or 3D matrices (depending on the image modality characteristics) that characterize an object using elementary elements - pixels or voxels (depending if 2D or 3D) that allocate in a coordinate system a respective parameter quantization (normally a gray level intensity).

The processing techniques can be discriminated in three main levels: Low-level, medium-level and high-level.

In low-level processing, both input and output are images, so the tasks associated within this level are the most simple ones, such as convolving input image with a filters/mask for denoising, operations in contrast (histogram equalization for example) and improvement of some specific details.

Medium-level processing is a more sophisticated task. In this level, the input are images and the output are now interesting features extracted from the input images, such as regions of interest (ROIs), the borders of an object and the identification of some objects. Segmentation algorithms are commonly used in this level to isolate some object(s) from the background, describe object characteristics and classify (recognize) objects.

High-level processes, as the name suggests, are processes with high computational operation level that try to describe and model the more complex cognitive functions normally

associated with vision. It involves the recognition of sets of characteristics of objects with high accuracy and, in almost all cases, automatically [5].

Usually, image processing projects that are idealized to do registration of images are composed by several image processing techniques applied to images with the objective of de-noising (filter), segmenting and finally registering images. Filtration is the step of the project devoted to improving the quality of images, facilitating the post-processing techniques that will be applied next. Segmentation is the step of selecting from an image regions with similar proprieties. The registration is the step where 2 or more different images are aligned and fused in a single image [18].

3.2) DENOISING

Noise is an undesirable companion of all medical imaging and is responsible for at least some component of the complexity that these processes have. In CT and PET images, one of the typical sources of noise is the presence of scattered radiation reaching the detectors, rather than traveling in a straight line from the X-ray source (in the case of CT) and from the original annihilation point (in the case of PET). A photon may thus be scattered in tissues, interacting in the detector with an angle that does not represent the anatomic structures through which it passed. This results in an overall decrease in natural image characteristics (e.g., contrast) and as a consequence the discrimination of different tissues is more difficult. A review of major techniques is provided by [19, 20].

One method to remove noise is by convolving the original image with a mask that takes into account neighborhood pixels' intensities, and smooth abrupt changes that often are associated to noise. Examples of these filters are the mean filter and Gaussian filter. Mean filter uses a weighted contribution of the neighborhood pixels to the middle pixel/voxel, as well as the Gaussian mask that comprises elements determined by a Gaussian function with curve characteristics adapted to the contribution desired of neighborhood. This convolution brings the value of each pixel into closer harmony, taking into account the intensity values of its neighbors, avoiding abrupt transitions between neighborhood pixels. However smoothing filters tend to blur an image, because differences between the true pixel intensity value and the surrounding neighbors were attenuated [21].

Another different approach used to reduce noise is related with the possibility of analyzing an image in the frequency domain. High frequencies are associated with abrupt transitions of intensity between pixels, fact that can imply noise interference on the image. Establishing a cutoff frequency corresponding to the transition between "true image" and noise, and applying this cutoff frequency to the frequency spectrum of the image, noise interference can be drastically decreased. The Fourier Transform allows the transition from the time domain to the frequency domain.

3.3) MORPHOLOGICAL OPERATORS

Mathematical morphology is based on the analysis and description of the shape of objects and is applied in digital image processing in areas such as edge detection, image enhancement, noise removal and image segmentation. This technique is based on the comparison of an input image and a structuring element using simple operators such as intersection, union, inclusion and complement. These operations work by traversing the structure element through the input image and analyzing the intersection between the translated structure elements coordinates and the image coordinates. Dilate and erosion are the fundamental operators used in morphological image processing.

Working with binary images (characterized by 0 value for background and 1 to foreground), dilation can be roughly defined as an enlargement of the foreground object boundary guided by the structure element. Over each pixel/voxel of a boundary of foreground elements an expansion occurs that was characterized by the structuring element. In other words, the basic effect of the operator on a binary image is to gradually enlarge the boundaries of regions of foreground pixels. Thus areas of foreground pixels grow in size while holes within those regions become smaller. The size and shape of the structuring element determines what is expanded, extracted or deleted from an image.

The erosion operator is very similar to dilation, however instead of expanding the boundaries of a region, it has the opposite effect of shrinking the boundaries. Erosion generally decreases the size of objects and removes small anomalies by subtracting objects with a radius smaller than the structuring element.

3.4) SEGMENTATION

In digital image processing, segmentation is the process of detecting and classifying some interesting regions in the original image. The purpose is usually to extract some regions in images that exhibit high correlation with objects or areas of the real world. Depending on the level of human interactivity, segmentation can be automatic, manual and semi-automatic. Manual segmentation is time consuming, dependent on human interpretation variability (non-repeatable). In contrast, automatic segmentation is highly reproducible and is less time consuming. However, it is more sensible to noise, which may induce errors in the segmentation. Semi-automatic approaches combine automatic and manual components, so in most cases studied it is the most reliable way to do segmentation [18].

Segmentation procedures can be classified as pixel- (or voxel-) and region-based depending on input data and on the characteristics of features that are desired.

3.4.1) PIXEL/VOXEL-BASED SEGMENTATION

Pixel/voxel-based segmentation procedures consider each element of image (pixel/voxel) as an isolated element, regarding only their individual intensity and neglecting the relations between pixels/voxels and their neighborhood. The most used and simpler pixel-based segmentation procedure is known as thresholding. Thresholding can be lucidly described as the separation of an image applying a limit value on pixels/voxels intensity. For example, having an image of random distributed values between 0-100 and we only want those pixels/voxels which values are superior to 50, a threshold can be applied to image and those pixels/voxels that are superior to the limit value (50) are assigned with 1 (one) value and the others assigned with 0 (zero). This type of segmentation is specially valorized when pixels/voxels intensities of a certain type of tissue is constant and known. The determination of the threshold value requires prior knowledge of the characteristics of the region of interest, which can be achieved with more or less complex methods. As has been said, biological tissues have constant and approximately known Hounsfield values, so they can be discriminated in images by applying thresholds to pixels/voxels intensities (Figure 9) [4].

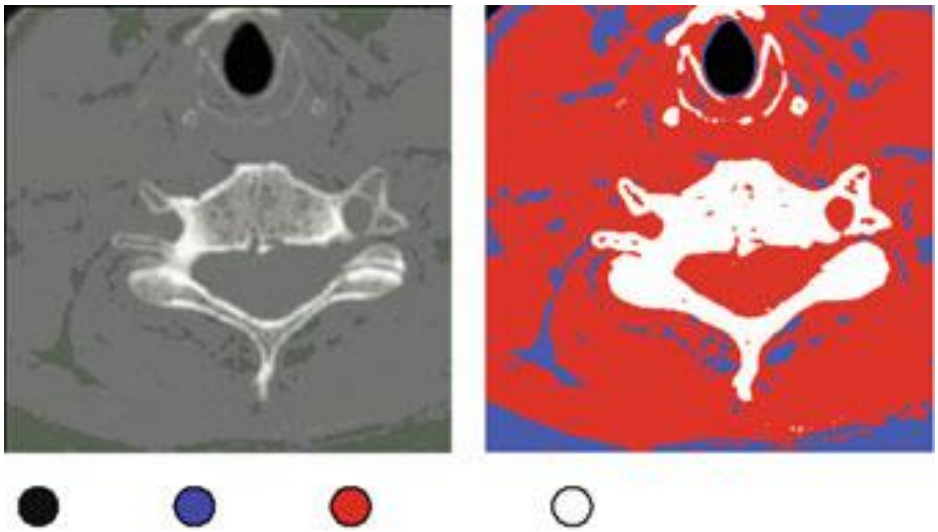


Figure 9 – Segmentation of different tissues in a CT image. In different colors are discriminated the different types of tissues (black for background, Blue for fat tissue, Red for water and white for bone structures). (Adapted from [4])

Another pixel-based segmentation method is clustering. This is a statistic method based on the distance between registered points in a feature space (feature in this context is a statistic measure done in the image and each point of the original image is assigned as a vector of features) and a cluster center arbitrarily placed in feature space. Then two steps are done iteratively until a stipulated convergence is guaranteed [4]:

- each point in feature space is aligned with the nearest cluster center
- recalculation of cluster centers is based on current assignment

An existing agglomeration of measured features in the feature space can be an indicator of similarity of regions in original images, so this can be used for segmentation purposes. Clustering is important in studies in which more than one value is associated with each element (e.g. color processing techniques).

3.4.2) REGION-BASED SEGMENTATION

Region based segmentation takes into account individual elements of images and its possible relations with neighboring elements. These relations guide the segmentation and can be related with distance or a similarity measure, depending on the case being studied. In this type of segmentation two main subtypes can be distinguished: Agglomerative and divisive.

Agglomerative region-based segmentation, as the name suggests, can be defined as the growing of a specific region indexing neighboring elements that satisfies a certain relation factor. This type of segmentation starts with a seed element, which can be either manually or

automatically set and iteratively grows until no more merges can be carried out. However, this procedure is susceptible to some influences such as:

- number and position of seed points,
- the order of annexing of neighborhood pixels/voxels,
- the relation factor and threshold that were applied for annexing decision.

Divisive region-based segmentation is almost the inverse of agglomerative. In this approach, regions are split until a certain degree of homogeneity is achieved. The main advantage is that seed points are not necessary because split is done in the whole region [4].

Related to region-based segmentation, one important task can be introduced: labeling. Classification can be viewed as the assignment of connected regions, which are obtained by segmentation because of exhibition of a relationship factor, to a specific class of object giving to different classes, different indexes as Figure 10 shows [22].

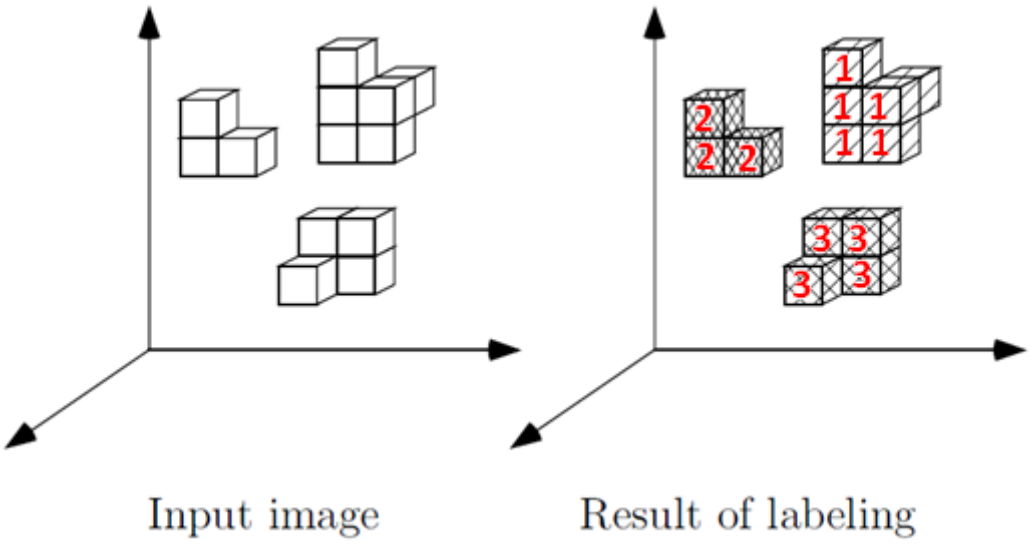


Figure 10 – Example of objects labeling [adapted from [22])

4) IMAGE REGISTRATION

4.1) INTRODUCTION

In a nuclear medicine center, namely in a medical imaging department where image analysis and interpretation is "the dish of the day", it is very common the need of multiple images for analysis and interpretation, as a basis for diagnosis and, in some cases, plan surgery and therapy. In most cases, analyzing a single image isn't enough, so the need to appeal to the field of image registration, where two or more images are aligned and overlapped, emerges. The registration field is very important since it allows the possibility of combining different features from different image modalities (inter-modality) with different characteristics (PET, CT, MR, US, etc.) in one image or align two same modality images (intra-modality) taken at different times [23]. Image registration is nowadays almost crucial at biomedical research (in anthropological measurements and comparison for example) and in healthcare (in medical centers in improvement of diagnosis, evaluation of disease stage, plan surgery, etc.). The point of medical image registration is to establish correspondences of shape, size and structures in different images and to accurately locate these features in the body with as much detail as possible, and in most of cases, to obtain new information that could not otherwise be identified in individual images [23, 24].

In fact, different medical images modalities have different intrinsic characteristics and give specific information about some topic of interest. For example:

- CT (Computerized Tomography) images yields information about the loss of energy (attenuation) that radiation suffers as it crosses matter (human body tissue in this case) allowing to obtain linear attenuation coefficients which are related with matter density and atomic composition (anatomical images);
- PET (Positron Emission Tomography) images return information about the distribution of a substance (radiopharmaceutical) in the body, allowing the identification of a tumor, for example (functional images);

- Magnetic Resonance (MR) which uses a magnetic field to get information about proton density, flow and relaxing times and with that information create an image (anatomical and functional images);
- Ultra-sounds (US) that exploits the relation between the loss of energy of a mechanical wave in a tissue with depth of penetration, obtaining with that discrimination of the crossed tissues (anatomical images) [5].

Having the chance to combine images from distinct modalities yields complementary information to the technician, acting in clinical benefit in more accurate and fast diagnosis, surgical plan and in therapy phase.

Medical centers have multiple exams from same subject, either from same modality or not, because patients rarely are scanned once. Accomplishing the evolution of a certain disease is the main cause of the existence of this amount of images from same patient. Another case is dynamic imaging where several images are taken sequentially in time [23]. This increasing amount of images in the medical center's database can be related to a common reference, potentially adding relevant clinical information and eventually giving rise to emerging new applications in clinical fields such as neurology, orthopedy, etc. One possibility is the construction of databases (atlases) that contains all important specific intrinsic characteristics of a certain condition and with that, establishes parameters that indicate the presence of the cited condition. Then the comparison of exams from different patients appears since the possibility to compare a given exam with a specific database may also support a diagnosis decision [4].

The term registration could be roughly defined as the achievement of an optimal transformation to apply (locally or globally) to a volume (target) in order to establish a spatial correspondence of common features with another volume (reference). Put in this way, the process may seem simple, but it is usually not the case. The registration process is usually very complex and need to take in account many steps such as: **set correspondences, select the transformation model (rigid, non-rigid), optimization, volume resampling and validation**. These are the main steps of the registration technique that will be described below [5, 25].

4.2) SET UP CORRESPONDENCE BETWEEN POINTS

As it has been said, the goal of the registration procedure is to achieve a suitable geometric transformation that describes the variation of the spatial position of features. Therefore is crucial to select correctly features in volumes, in other words, select which point in one volume (target) corresponds to the same point in the other volume (reference), allowing the computation of a spatial relationship between these points. In this field, two types of methodologies to achieve correspondences can be distinguished: "hard" and "soft" [25]. Using a "Hard" methodology means that one special external "marker" is placed in a certain local in patient before imaging (called a *fiducial* marker) so that the marker can be easily identified by the computer. This methodology is more uncomfortable for patients and in case of exams taken at different times it may not be located at same place, yielding wrong alignment. The other methodology ("soft") acquires features only by information extracted from volumes. This one requires more expert knowledge about volumes characteristics and sophisticated algorithms to identify those features. From the data volumes could be extracted, depending on the desired application, some features such as points with high curvature, regions derived from segmentation (figure 11), points with similar voxel intensity, closed contours, corners, etc. illustrated a correspondence between two right femurs in two different volumes. The femurs are extracted from two same patient CT scans [5, 24, 25].

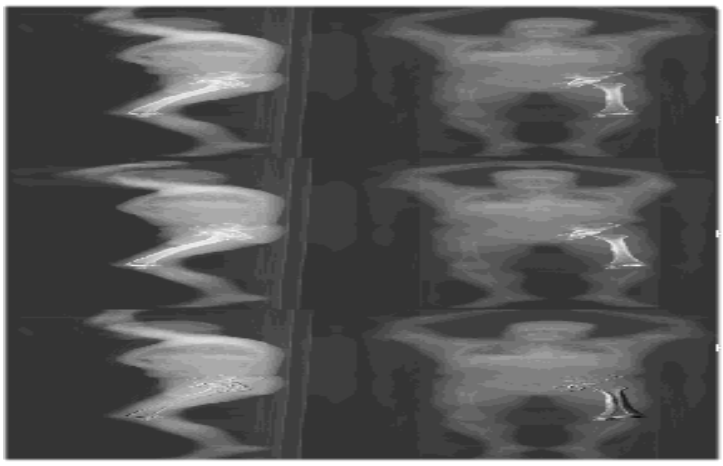


Figure 11 – Correspondence of two regions of interest, from two different intra-subject CT volumes, derived from segmentation (right femur). Images are total projections in coronal (left) and sagittal (right) directions. Upper and middle images are from volume 1 and 2 respectively and bottom image is the difference between volumes.

4.3) TRANSFORMATION MODEL

With the correspondence of features established, it is time to proceed to the elaboration of a **transformation model**. In this step it is fundamental to take into account the coordinate systems from the two volumes and if necessary to convert the coordinate system of the target volume to the coordinate system of the reference volume, to avoid misregistration. This step may introduce some blurring in the data, however.

There are two main classes of registration transformation techniques: **rigid** and **non-rigid**. In the following diagram (Figure 12) the most popular techniques applied in the medical imaging field are represented. Notice the existence of several more techniques that we will not describe here.

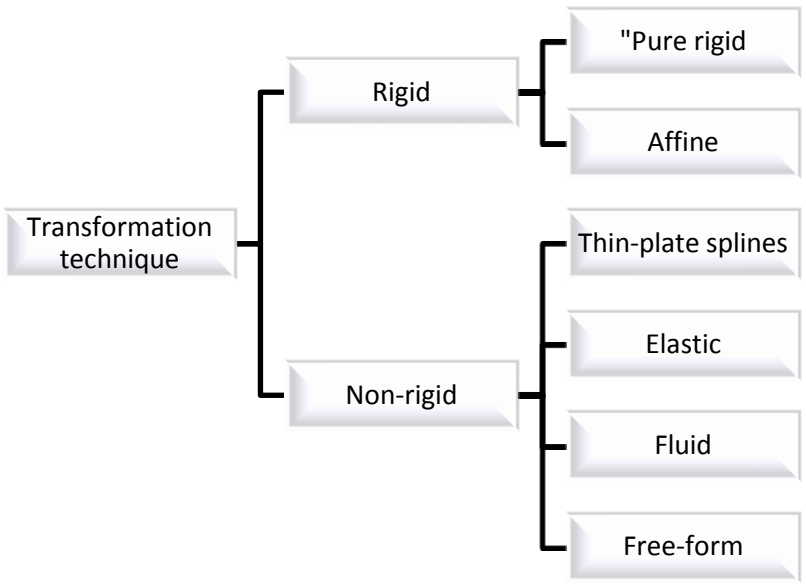


Figure 12 – Diagram of some of the transformation techniques

4.3.1) RIGID TRANSFORMATION

The **rigid** transformation, as the name suggests, is a transformation type that only applies rotations and translations to volumes, so no size and shape changes in volume data are allowed. This technique only changes the position and orientation of objects (features) defined by the *degrees of freedom* (DOFs). For 3D data, 6 DOFs are allowed: 3 rotations (one per axis x, y, z) and 3 translations (one per axis too), always maintaining constant distances between points. This corresponds to applying to the coordinates of the volume elements a matrix $\mathbf{T}_{\mathbf{rigid}}$ with 6 variables (DOF) that represents the transformation:

$$T_{rigid} = \begin{pmatrix} \cos\beta\cos\gamma & \cos\alpha\sin\gamma + \sin\alpha\sin\beta\cos\gamma & \sin\alpha\sin\gamma - \cos\alpha\sin\beta\cos\gamma & t_x \\ -\cos\beta\sin\gamma & \cos\alpha\cos\gamma - \sin\alpha\sin\beta\sin\gamma & \sin\alpha\cos\gamma + \cos\alpha\sin\beta\sin\gamma & t_y \\ \sin\beta & -\sin\alpha\cos\beta & \cos\alpha\cos\beta & t_z \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} \quad \textit{Equation 3}$$

This equation represents the "pure" rigid transformation model. The terms α, β, γ represents the rotations in axes x, y and z respectively and t_x, t_y and t_z are related with translations along the three axes [23].

The rotation and translation operators are affine operators. In other words, they preserve the collinearity and the degree of points distance, and so straight lines remain straight after the transformation and also spatial relations between lines are preserved [3].

Affine transformation is similar to the rigid transformation but more sophisticated, encompassing the possibility of scaling and skewing volumes and increasing the number of DOF to 12, yielding more liberty to align volumes. Let $\mathbf{T_{scale}}$ and $\mathbf{T_{skew}}$ represent the matrix of scale and skew transformation respectively. The affine transformation $\mathbf{T_{affine}}$ could be expressed as [23]:

$$T_{affine}(x, y, z) = T_{skew} \cdot T_{scale} \cdot T_{rigid} \cdot (x, y, z)^T \quad \textit{Equation 4}$$

4.3.2) NON-RIGID TRANSFORMATION

In many cases, applying simply rigid transformations is not enough to describe the spatial differences of features, so it is necessary to appeal to **non-rigid** transformation models. In most registration studies reported, the use of non-rigid transformation is substantially large in inter and intra-subject registration. The main vantage of non-rigid transformations compared with rigid transformations is the fact that non-rigid transformations techniques accommodate better soft tissues deformations. For example in intra-subject registration, the need of a model that prevents the influence of time in biological tissues is advisable, and in inter-subject registration, this type of registration is exceptionally important given the human tissue's size and shape variability across subjects. In contrast with rigid transformations models, non-rigid transformations techniques are still ongoing investigation and many research groups are continually developing new and more sophisticated methods [23, 5].

As has been said, the registration process has three components: the transformation that relates the two volumes, a measure that evaluates the accuracy of alignment and a step of optimization of the transformation parameters. Non-rigid techniques also encompass these

three steps. The main difference between rigid and non-rigid registration is the amount of degrees of freedom (DOFs). Adding DOFs, affine transformations that are purely linear are then nonlinear and have more "liberty" to fit one volume to other, increasing the success of registration process [23, 4].

The essence of non-rigid transformation is the possibility of combining linear functions (polynomials or basis functions) in a way that the volume could be "wildly" deformed, so the selection of the best function is crucial. Most commonly, basis functions are trigonometric or wavelets. The choice of trigonometric basis functions relies on the possibility of processing the volume in the frequency domain: each basis function could define a specific frequency and act more selectively [4, 23]. There's a wide range of candidates to represent basis functions, depending on the application. In algorithms specially developed to be applied in medical centers, the most commonly applied non-rigid registration techniques are splines (mainly thin-plate splines), elastic registration, fluid registration and free-form deformation.

4.3.2.1) THIN-PLATE SPLINES

Spline basis techniques are the most popular in the medical imaging field to model spatially variant geometric variations in volume registration [24]. This approach was first proposed by Goshtasby [26] and later applied to medical field. The term spline is frequently associated with construction of ships' hulls, those long strips of wood or metal that molds the bottom external surface of ships. The curvature is obtained placing weights in specific points (control points). This assumption could be translated to the mapping of volumes, assuming the spatial displacement as plates deformations and with the control points' information (inputted after achieved the correspondence), that were strictly aligned, achieve the best transformation algorithm that warps and aligns the volumes. Note that between control points, the technique provides a smoothly varying displacement. Given a set of control points $\{p_1, p_2, \ldots, p_n\}$, the overall condition that describes the transformation of point *x* in target volume to *f (x)* is:

$$f(x,y,z) = a_1 + a_2 x + a_3 y + a_4 z + \sum_{i=1}^{n} c_i \, \phi(\|(x,y,z) - p_i\|) \qquad \textit{Equation 5}$$

where *φ (r)* is the radial basis function (RBF), the *a* coefficients are related with the affine part of transformation and c_i is the mapping coefficients. This RBF can have as much unknown variable as desired, noticing that the more variables we set, the more adjustable the model is, and consequently more computationally demanding, so a compromise has to be taken. To

solve the equation, as many control points as unknown variables are needed, so the applicability of the model is limited to the existence of a set of points [5, 23, 24].

One of the most common radial basis functions used in medical image registration [23, 27] is thin-plate splines and was introduced by Duchon [28] and Meinguet [29] continues the Duchon's study. For 3D volumes, the radial basis function is defined as: $\varphi(s) = |s|$

A wide range of radial functions exists including Gaussian and multiquadrics, with varying weights of deformations on neighborhood points. This method is particularly attractive to researchers because of its intuitive physical interpretation and relative algebraic simplicity, besides the possibility to incorporate constraints such as rigid bodies or directional constrains into the model [23].

4.3.2.2) ELASTIC REGISTRATION

Elastic registration was first proposed by Bajcsy to compare a CT brain image with an atlas set. The main idea is to match the two volumes as if it were a physical process viewed as the stretch of an elastic material like rubber. This process is governed by two forces: one internal and another external and can be described with partial differential equations by Navier as

$$\mu \vec{\nabla}^2 \vec{\mu} + (\lambda + \mu) \vec{\nabla} (\vec{\nabla} \cdot \vec{\mu}) + \vec{F} = \vec{0} \qquad \textit{Equation 5}$$

where $\vec{\mu} = \vec{\mu}(\mathrm{x}, \mathrm{y}, \mathrm{z})$ represents the displacement field, $\vec{F} = \vec{F}(x, y, z)$ represents the external force actuation and the constants λ and μ are intrinsic material elasticity characteristics constants (Lamé constants).

The external force acts in elastic body and is responsible to guide the transformation. Normally the gradient of a similarity parameter is used as external force.

The partial differential equation is solved in order to the displacement field μ that is dependent on the external force applied, and a process of finite differences and successive relaxations is used, resulting in a discrete displacement field at each voxel. Another way is determining the displacement field only on known external force nodes and the unknown are obtained by interpolation [5, 23].

4.3.2.3) FLUID REGISTRATION

Fluid registration appeared from the necessity to overcome some drawbacks from elastic registration. In elastic registration, the ability to highly deform locally is limited, since the deformation energy caused by stress increases with the strength of deformation. In fact, fluid registration is very important in cases of inter-subject registration, since it possibilities larger deformations and larger variability of DOFs. The process of fluid registration is guided by the Navier-Stokes equation:

$$\mu \vec{\nabla}^2 \vec{v} + (\lambda + \mu) \vec{\nabla} (\vec{\nabla} \cdot \vec{v}) + \vec{F} = \vec{0} \qquad \textit{Equation 6}$$

This equation is similar to the elastic registration equation, but now instead of the displacement, the velocity is involved. The relationship between velocity and displacement is given by:

$$\vec{v} = \frac{\partial \vec{\mu}}{\partial t} + \vec{v} \cdot \vec{\nabla} \vec{u} \qquad \textit{Equation 7}$$

The way to solve the fluid registration equation is almost equal to elastic registration but the computational time slight increases because of velocity usage [23].

4.3.2.4) FREE-FORM DEFORMATION

Firstly introduced by Sederberg and Parry [30], the free-form deformation (FFD) consists in deforming the interior space of objects. A way for better understanding this technique is to imagine a transparent plastic parallelepiped with some objects embedded, as shown in figure 13 [2].

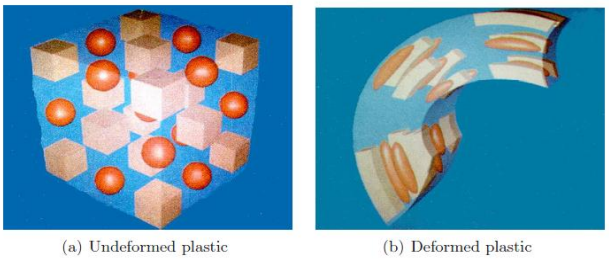


Figure 13 – An analogy for free form deformation. A plastic parallelepiped with objects embedded. Deforming the plastic implies the deformation of objects embedded too. (Taken from [2])

In the basis of this process is the establishment of a local-coordinate system, where some seed points are inserted. These points are used as coefficients of a trivariate tensor product of a polynomial, and then the new position (after deformation) is calculated in the

local coordinate system. Finally, the polynomial equation is evaluated. This technique is frequently used in computer graphics but can be extended to several registration methodologies used.

4.4) OPTIMIZATION PROCESS

The registration process approximates one volume (target volume, V_t) to other (reference, V_r) by applying the correct transformation function (T_o). The task of selecting the correct transformation is done by defining cost functions ρ that evaluate the registration accuracy by measuring similarity between volumes. The goal is to maximize the volumes' similarity:

$$T_o = \arg_{T_o} \max \rho(V_r, T(V_t)) \qquad \textit{Equation 8}$$

This is an iterative step - in other words, it starts from an initial guess and successively converges to an optimal solution established by the cost function [4, 5].

4.4.1) SIMILARITY/COST FUNCTIONS

The similarity functions are usually divided in two main classes: Feature-based and Voxel-based similarity measures and the choice of a particular class typically reside in the intended application. In these main classes we have:

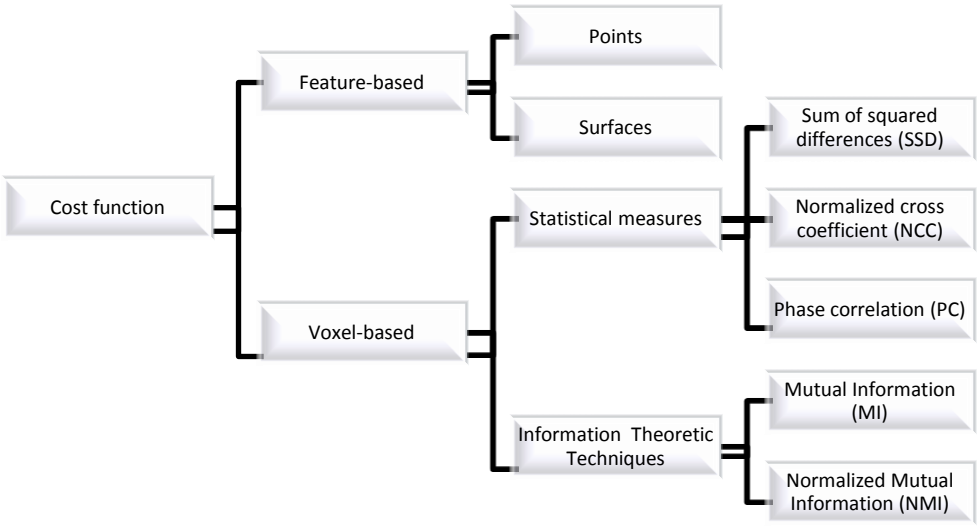


Figure 14 – Diagram illustrating examples of some similarity functions

4.4.1.1) FEATURE-BASED REGISTRATION

Feature-based registration establishes one-to-one correspondences, minimizing the distance between them. The success of this step is intuitively entirely dependent on the correction of the ***"set up correspondences between points"*** step, so if correspondences are not well established, obviously the quality of the registration will not be guaranteed.

Features are, as it was already said, components of interest extracted from volumes. The main components extracted from volumes are points or surfaces. Given a set j of corresponding points in the two volumes (target volume t and reference r), a simple measure of similarity between points could be achieved calculating the Euclidian distance D between them:

$$D = -\sum_{j} \left\| r_j - T(t_j) \right\|^2 \qquad \textit{Equation 9}$$

Optimizing this simple equation supposes the *a priori* knowledge of the points' correspondences and needs as much correspondence points as unknown variables. To solve this equation, simple matrix manipulation is enough, using the method of Singular value Decomposition (SVD). The main disadvantage of using points in the registration process is the correct identification of correspondences, as discussed above.

Another alternative to feature-based registration is to use surfaces instead of points. Surfaces can be extracted from images by applying processing techniques like segmentation, labeling, region growing, etc. These processing techniques are applied and anatomical structures extracted from the volumes. Segmentation, labeling and region growing are computed in both volumes and can be achieved automatically or with human interaction. The surfaces are treated as sets of points and are registered minimizing the distance between corresponding points of both sets. There's a nuance in this process: the correspondence between all points is not known and need to be estimated at the same time. There are some algorithms to do this type of registration, such as the "head and hat" applied first by Pelizzari et al [31] and, the most common, Iterative closest point (ICP) proposed by Besl and McKay [32].

The ICP algorithm assumes relations between each point of one set and the closest point in the other set. The similarity is measured by

$$S = -\sum_{j} \left\| x_j - T(t_j) \right\|^2 \qquad \textit{Equation 10}$$

where $x_j = \min\{\left\| r_j - T(t_j) \right\|^2\} \quad r_j \in r$, so x_j is the closest point in the reference's set of points [4].

To do the registration of points, first the x_j were computed and then the registration using the similarity function S was applied. This process is iterative and only stops when convergence is achieved.

4.4.1.2) VOXEL-BASED REGISTRATION

Since the proposed work uses Landmark-based techniques, only a brief introduction will be made to voxel-based registration techniques.

The greatest advantage of voxel-based registration, when compared with feature-based, is the possibility of its application in inter-subject registration processes. Instead of using corresponding points in different volumes as in feature-based registration, voxel-based registration uses directly the voxel's intensities, preventing the miscorrespondence of points that may appear in feature-based registration [23]. In this type of optimization, the algorithm measures the amount of shared information in both volumes.

To evaluate the degree of shared information between volumes based on the intensity of voxels, statistical functions and Information Theoretic Techniques are used. There are many statistical functions that measure the similarity between 2 volumes but the most applied are: **Sum of Square Differences (SSD), Normalized Cross Correlation (NCC) and Phase correlation (PC)** [23, 4].

The first one (SSD) is simply the calculation of the sum of squared differences between the intensity of voxels of the two volumes at the same location, so it is restricted to mono-modality comparisons. The second one (NCC) assumes a linear relationship between both volumes voxels' intensities, so is not restricted to one modality. Different modalities with linear relationship between voxel intensities can be compared. The phase correlation (PC) can simply be applied using the Fourier Transform and is based on the Fourier shift property. A constant spatial difference between two volumes results in a phase difference in the frequency

domain, and the PC of two volumes for all their mutual coordinates can be calculated applying the inverse Fourier Transform to the normalized cross-power spectrum (CPS)[i]. $PC = \Im^{-1}\{CPS\}$ [5, 24].

All these statistics methods are more suitable for intra-subject registration [2].

The **Mutual Information (MI)** technique is based on the idea of extracting the most information that is similar in both volumes as possible. It assumes the superposition of the same structures in aligned volumes using Entropy to measure the total information superposed. Entropy is a measure of disorder and can be associated with the degree of alignment of two volumes: two misregistration volumes have high entropy. In fact Mutual Information (MI) is related with the entropy of the two volumes:

$$MI = H(T) + H(R) - H(T,R) \qquad \textit{Equation 11}$$

where H(T) and H(R) represent the target and reference volumes' entropy and H(T,R) represents the joint entropy that is a similarity measure of both volumes. Mutual Information depends on the degree of alignment and only exploits the presence of similar information in both volumes, without exploring the possible relations between the intensities of the volume voxels. In order avoid these drawbacks, the concept of **Normalized Mutual Information** is introduced [4]:

$$NMI = \frac{H(T) + H(R)}{H(T,R)} \qquad \textit{Equation 12}$$

Returning to the optimization problem, knowing now its pillars, the objective is to ensure the best fit between the 2 volumes by minimizing an associated energy function, so an iterative step is needed until the best alignment is reached. The algorithms usually take a serial of guesses from an initial start position and the initial condition has to be sufficiently close to the target volume to converge to the correct answer. The registration algorithm proceeds by taking progressive guesses and recalculating the cost function until a minimum or maximum (depending on the function used) are reached [33].

One of the most important issues related to optimization algorithms is the possibility of convergence to a wrong optimal solution also called "local optimum". Because of

[i] Cross-power spectrum of two volumes is defined as: $CPS = \frac{F_2 F_1^*}{\left|F_2 F_1^*\right|}$, where F_1 and F_1^* represents the Fourier Transform of volume 1 and his conjugate respectively and the same for F_2 (volume 2).

interpolation artifacts and local good match between features or intensities, the algorithm may yield multiple optimum solutions. Related with the function that transforms the target volume, exists a space of parameters of possible solutions and the main task is to find the global optimum value that best fits the two volumes (maximizes the similarity) [24].

In 1997, C. Studholme and his team developed a hierarchical approach to overcome the multiple optimum solutions problem. In this new approach, volumes are first registered at low-resolution and the solution of that is used to initiate the algorithm at a higher resolution, and so on. This approach increases the robustness of algorithm, however a problem could exist if the global optimum is absent at low resolutions local optimum search. To avoid this possibility, more sophisticated techniques were developed to make the global optimum more probable, such as genetic algorithms (GA), simulated annealing (SA), particle swarm optimization (PSO) and evolutionary strategies (ES) for example [24, 34].

4.5) IMAGE RESAMPLING AND VALIDATION

The optimization process is an iterative step of image registration and needs to, at each guess, transform a target volume into the space of another (reference). This process requires resampling one volume to the corresponding voxels in the reference volume's grid and determining the new values of voxel's intensity. This process is called resampling.

The calculation of new positions and intensities is done using interpolations, and there's a several range of choice in that, and the most used are nearest neighbor, bi-cubic, bi-linear, and cubic B-splines. The different type of interpolations have different complexity and accuracy, also related with computational load, and to made the solution applicable in practice a compromise has to be done. In most of cases, the bilinear interpolation is the best choice because the accuracy is acceptable and so the computational time [5].

In medical registration is subtended the vital importance of a good alignment and the consequences of a bad one. After processed, volume of images will be interpreted by a specialist and, errors in the alignment process can cause errors in the diagnosis.

Medical software industries have developed standard norms, protocols and quality control processes in order to validate the registration process. **Validation** is the way that medical community has to ensure that the software respects all the requirements, like

accuracy with an associated tolerance, time consumption and robustness. Validation of algorithms uses measures from phantoms and proved volunteers' or patients' exams, and this measures are considered "gold standard" or "ground truth" measures [23].

The following image summarizes the image registration process, with all of the steps:

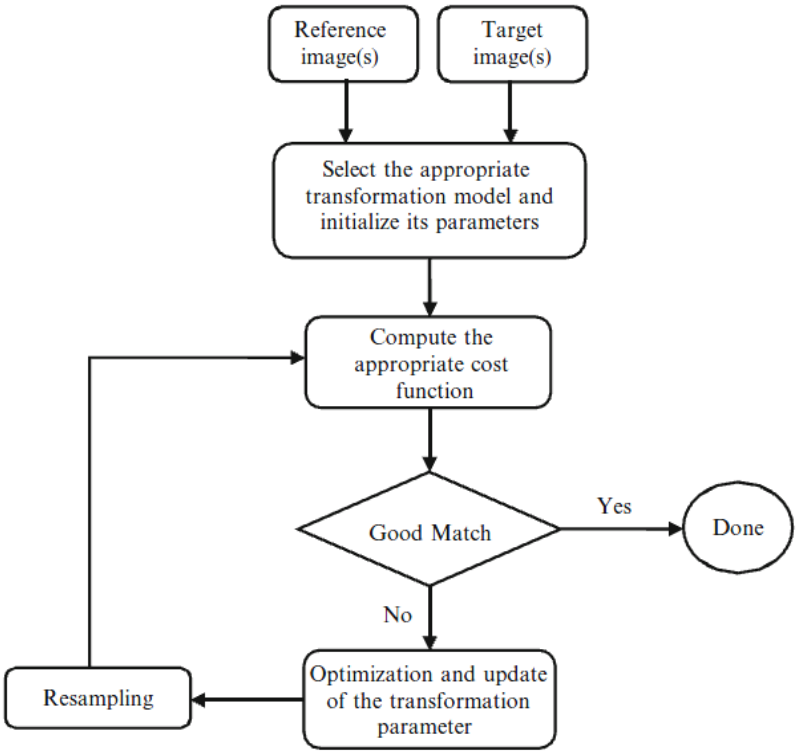


Figure 15 – summary of Image Registration process. (Taken from [24])

5) RELATED WORK

In recent past years, many attempts in medical images have been done in inter- and intra-subject registration methodologies. The importance of comparing different images has been recognized as an area of great interest and many scientific groups have developed algorithms to try to accurately align these images. To do the alignment, sophisticated algorithms are arising using rigid or non-rigid transformations or even a combination of both types of transformations.

Initial approaches using rigid transformations on CT images are applied only at some parts of human body, essentially in regions that not suffer size and shape variations between scans. The skeleton is an example of these regions, where simple rigid transformations usually solve the registration problem. Other initial registration attempts in small animals are already documented in literature, however human whole body registration is a field where there is still much to be done. In whole-body registration, the deformable field is very large and requires sophisticated algorithms, and, because of that, the results are still below those required for acceptance by the scientific and medical community.

Little et al. [43] present in 1997 a technique that uses bone segmentation and manual labeling of surrounding tissues. Having the bones accurately segmented, rigid deformations are applied to them and non-rigid deformation techniques are applied to the surrounding tissues that were manually segmented. They have applied this technique to MR images of the head and neck. Arsigny et al. [44] also realized the need to combine non-rigid and rigid transformations to model whole body registration. They were the first group realizing that these types of transformations yield non-invertible results and to focus their efforts in finding a way in which a combination of local transformations would yield an overall invertible transformation. They have tested their algorithm in histologic images and state that their method could by suitable for articulation registration but do not present examples.

For lung, breast and abdomen images, registration algorithms with non-rigid transformations seem more suitable because of the deformable characteristics of the tissues. Camara et al. [36] made the pioneer work in abdomen and thoracic images registration with reasonable accuracy. Camara use the Free Form Deformation (FFD) algorithm guided by a gradient vector flow combined with Mutual Information (MI) to do this task. The FFD technique is already used by Rueckert et al. in his work [37]. They developed an algorithm to

align MR breast images and apply the registration process both globally and locally. Global motion was modeled with affine transformations and the model then refined with local breast transformations using a more sophisticated algorithm, non-rigid for instance (Free Form Deformation (FFD) based in B-splines).

Xia Li et al. purposed what they call “the first automatic inter-subject whole body registration”. Their algorithm is applicable to both humans and animals, however in humans it was only tested in upper torso parts. The registration process that Xia Li and their team developed has two main steps. They first apply a rigid point-based method algorithm to the entire skeleton, which was previously segmented from the original CT volume. In a first attempt, they apply the thin-plate spline (TPS) technique to align the volumes. With the results from the first attempt achieved, they use it as an initiator to a second and more accurate attempt, using a voxel intensity-based method (Normalized mutual information) [35].

The hierarchical approach for image registration procedures was firstly introduced by N. Kovacevic et al. in 2003 [41]. In this publication they realize that the registration process is more accurate if done sequentially. So, firstly the major organs are affine registered and the algorithm progressively refines the registration process to single bones and organs. As an initial attempt, the expectation of the registration accuracy was not high. In fact, the algorithm fails in major different postural scans and was only tested in simulated data (without natural noise) [41].

The main differences in distinct images are related with restriction of movement established by joints’ characteristics. So Papademetris et al. [45], in 2005, “attach” this articulation information to the registration process. They characterize each joint with information about the axis of rotation and degrees of freedom (DOFs). They blend each piecewise rotation derived from each joint in an overall transformation and applied their method to register lower limbs in mouse images. They hypothesized the fact that their algorithm could be a good initiator for an intensity-based algorithm but do not present any example [45].

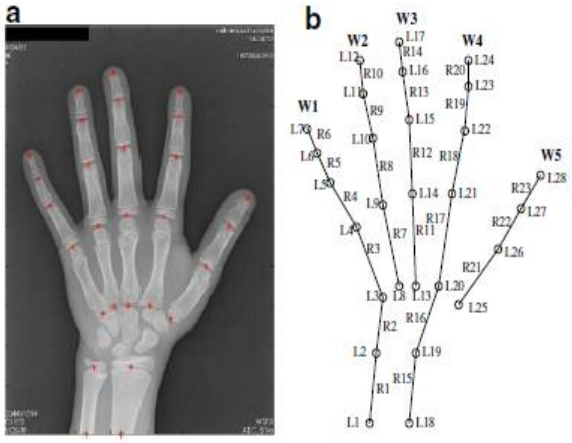


Figure 16 - a) Landmarks marked in the radiography of hand; b) Scheme of landmark selection that construct the wire model in the realistic way. (Taken from 42])

In 2005, Fernandez et al. [39] were the first group assuming the human body, not as an entire "block" but, as many "sub-blocks" linked by specific structures – the articulations. Fernandez "split" body in sub-blocks, and these sub-blocks are relatively independent from each other. The final transformation is a linear combination of the transformation of individual sub-blocks. He applies his algorithm to hand radiographs and achieved a good global alignment of the different hand images, however the global continuity of bones is compromised when images with huge shape and size variations appear [39].

Elastic deformation algorithms are also a way register medical images. Miguel Martín-Fernández and his team published in 2009 [42] an improved approach of his previous work from 2005, this time testing elastic algorithms for hands radiographs registration. Their algorithm is based on automatic landmark detection in images using a cascade of image processing techniques (Figure 17). Landmarks are located in fingers, radio and ulna (28 in total, represented in Figure 16). A wire model is built by joining with straight segments consecutive features in a realistic way and permits an exact correspondence between these points in two different images applying an affine transformation. The remaining pixels of the image are elastically registered. The results reported in Miguel Martín-Fernández paper suggest that using wire deformable models in individual bones of hands and elastic models to deform surrounding tissues yields better results than previous work reported using rigid transformations and TPS (Thin Plate Splines) algorithms. They based this statement on the comparison of values of MI (Mutual

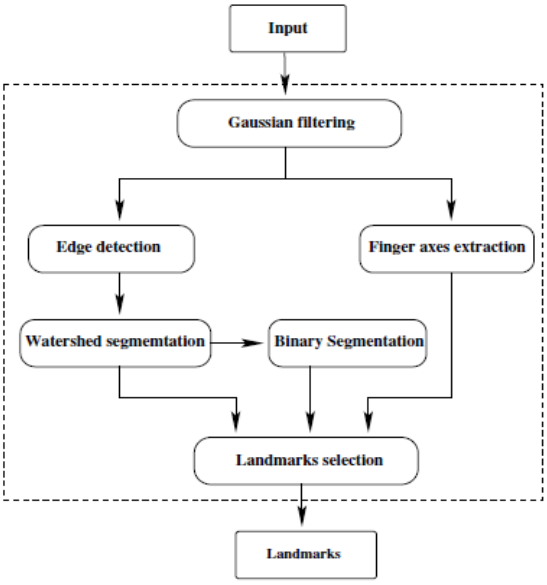
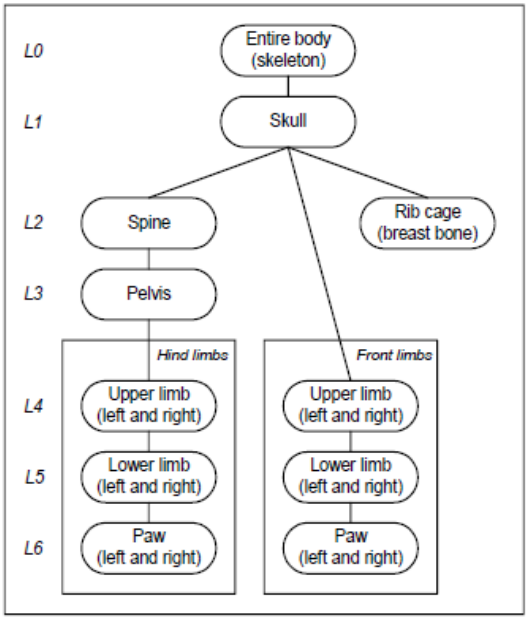


Figure 17 – Landmark extraction for registration (taken from [42])

Information) and JC (Jaccard Coefficient) in both registrations, concluding that these coefficients are best in elastic-based deformation [42].

Baiker and his team published in 2007 [38] an article where they present a fully automated whole-body registration approach for mice micro-CT images. In this study, they use atlas information from an anatomical mouse atlas developed by Segars et al. [40]. Their study is based on a hierarchical anatomic approach where each joint is spatially defined and characterized in terms of liberty of movements (DOFs) (Figure 18 at right). The Fitting model is performed by first applying an Iterative Closest Point (ICP) algorithm to coarsely align the two volume surfaces by minimizing the Euclidean distance. Then the articulated registration is done traversing a hierarchical tree (Figure 19 at left) in a top-down manner, minimizing the global error with the Levenberg-Marquardt minimization algorithm [38].



Joint types	Modeled joint	DoFs of the articulated bone
Ball joint	Shoulder Wrist Hip Ankle	3 translations 3 rotations 3 scalings
Hinge joint	Elbow Knee	3 translations 1 rotation 3 scalings

Figure 18 - (Left) Hierarchical anatomical tree for the skeleton; (Right) Joint types of the animal skeleton and the respective DOFs. (Taken from [38])]

One of the most complete and realistic approaches to the whole body registration problem was also developed by Baiker and his team for micro-CT mice whole body phantom images, continuing their previous work. They used the MOBY phantom, developed by Segars [40], and segment high contrast structures such skeleton, skin and lungs in the gerated images. The information about joints characteristics in a hiearchical way was inpputed as in the previous work (in a top-down manner). The main differences between the previous work are related with the transformation model used to fit the skin and this time realistic joints restritments are defined for each joint. In this work, a TPS algorithm is applied from atlas to

the reference image to interpolate the soft tissues. A summary of this method is presented in Figure 19. They validated the method by fitting the model to 41 micro-CT datasets scanned in different postures, constituing the most extensive validation reported in whole-body mice images registration [46].

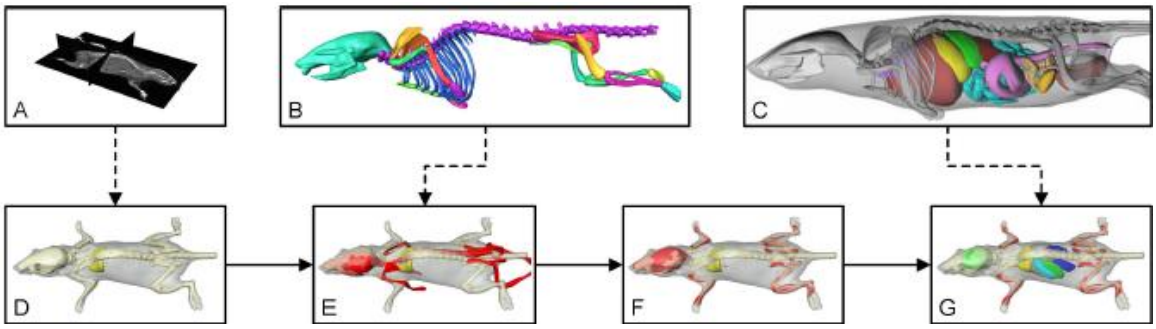


Figure 19 – Summary of the registration process. The skeleton, skin and lungs are extracted from reference (A), then labeled skeleton (B) and segmented lungs (D) are registered (E and F) using an anatomically realistic deformable model. Using the atlas (C), the major organs are mapped (G). The dashed arrows indicate the input volumes. (Adapted from [46])

These are the main studies reported in literature for image volumes registration that were based in medical image processing. Systems capable to model and simulate human body and the majority of its movements are a reality for other applications such as game animation, ergonomics, sports, etc. There are some interesting software products developed, that simulate human body in a realistic way, such SIMM (Software for Musculoskeletal Modeling) developed by Musculographics Inc. and Character Studio developed by Autodesk Inc. and Discreet Logic Inc. However, the possibility of incorporating the information from these software products in the medical field, such as in registration of images, is still little explored [2].

6) METHODS

Having defined the main concepts of a registration procedure, in this section we will describe in detail the registration algorithm developed.

All algorithms were developed in the IDL (Interactive Data Language) environment.

PET/CT data were acquired with the Gemini GXL (Philips Healthcare) scanner owned by ICNAS (University of Coimbra) and the subject identity protection was assured in all visualized and tested exams: apart from the voxel data, only a few specific fields related with the voxel size of the PET and CT volumes were supplied to do this work. The data were extracted from files in the "*Imagio*" format, a proprietary format provided by Philips. Clinical whole body PET/CT scans from 7 different patients, with at least 2 studies from each patient, were available to test the algorithm. These included 5 scans using the conventional Field-Of-View (top of the head to mid thighs) and 2 true whole-body scans (top of the head to bottom of feet) from the same patient. The two true whole-body scans were used more intensively in the test, since they also allowed to study coregistration in the legs: in this case, the CT volumes were three dimensional arrays of integer values in HU (Hounsfield Units) with 512×512×378 and 512x512x379 voxels, each voxel with dimensions 1,17188×1,17188×5 mm^3. The PET volumes were three dimensional arrays of floating point values with 144×144×465 cubic voxels, with dimensions 4×4×4 mm^3.

However, working with multiple CT volumes with such large dimensions requires high computer capabilities (memory for example), and to reduce computational burden, the volumes were halved in each dimension. Halving was done with the *congrid* function from the IDL library. This function allows resizing 3D volumes for a specific dimension and uses linear interpolation to do the resizing. So, the CT volumes used for processing had 256×256×189 voxels, more suitable to test and develop the algorithms, decreasing drastically the time consumption and memory requirements.

The summary of the registration process can be seen in the following diagram. There are three main steps (Segmentation and Labeling, Matching and Registration) each one comprising a set of processing techniques.

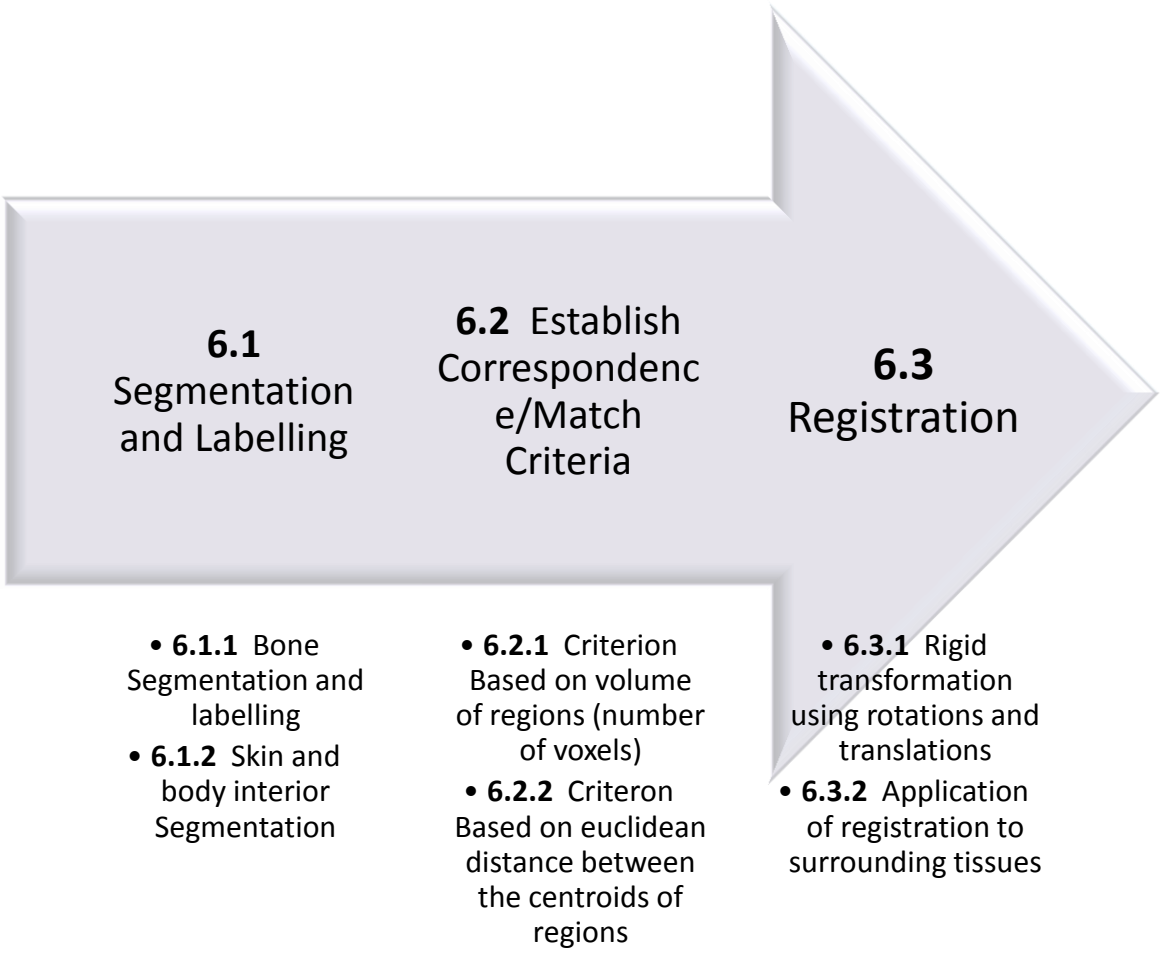


Figure 20 – Summary of the methodology developed

6.1) SEGMENTATION AND LABELING

In this step, the original CT volume is processed in a way to extract features of interest, namely bone structures, body interior and skin.

6.1.1) BONE SEGMENTATION AND LABELING

The original CT data have information about tissue characteristics. Tissues are described according to the attenuation imposed to the radiation beam and since the intensities of voxels are in HU, the segmentation problem was simplified because there is considerable work reported concerning the segmentation of bone structures. Segmentation of bone tissues was obtained by the simple application of threshold limits to the CT data. From the literature and after some tests and visual inspection, a value of 400 HU was achieved for

the inferior limit and 1500 HU for the superior limit to do the bone segmentation. All voxels with values within these limits were assigned with one value, the other values were zeroed. Each voxel in CT volume was thus classified as a bone (value 1) or non-bone (value 0) structure. In other words, for a voxel *x* with intensity *I(x)*,

$$x \in \begin{cases} Bone\,(value\ 1) & if\ \ 400 \le I(x) \le 1500 \\ Non-bone\ (value\ 0) & otherwise. \end{cases} \qquad \textit{Equation 13}$$

This segmentation process yields bone structures from the entire CT volume, and the idea is to subdivide the "new" whole volume (skeleton) in regions with logical anatomic definition such as bones linked by articulations. In fact, the main reason for the differences in the spatial position of bone structures in CT exams taken at different instants of time is a consequence of the movements allowed by articulations. If individual bones could be equally labeled in the two volumes and the correspondence of these bones determined for these volumes, by applying suitable coregistration models, the alignment of bone structures can be done.

The labeling of the entire skeleton into individual parts is done applying the function *"label_region"*. The *label_region* function splits a volume into connected regions and gives an index number for each region. This way, the skeleton can be divided in individual bone segments.

Having divided the skeleton in connected regions, the smallest regions were then filtered, ignoring regions with less than a specific number of voxels (we used 500 as a limit). This filtering step allowed us to apply the subsequent corregistration steps only to the largest bone regions, ignoring the smallest groups of labeled voxels. Most of these labeled regions were a natural consequence of the noise in CT data and of the interpolations done when the volume was resized to a smaller size, which interfered with relations between neighboring voxels and consequently in their connections.

From each labeled and filtered region, information of interest is then extracted: the x,y, and z coordinates of the center of mass, the bounding box (a box that spatially confines the region), a pointer to an array of indices that indicates what individual voxels belong to the region and the mass/volume ratio (that characterizes the shape of the region). This information is saved in a structure variable containing these fields such as Figure 21.

```
A = {region, $
    count:0L, $                 ; total of pixels of region
    minx:0,     maxx:0,  miny:0,  maxy:0,  minz:0,  maxz:0, $
    cx:0.,  cy:0.,      cz:0.,      ; centroid in x, y and z respectively
    r:ptr_new(), $  ; array of pixel indexes that belong to the region (pixels of the original volume)
    rlocal:ptr_new(), $         ; array of pixel indexes that belong to the region (pixels of the bounding box
                                volume)
    area_volume_ratio:0. }
```

Figure 21 – Example of a structure variable with its respective fields. An array of these structures is used to save the characteristics of all the filtered labeled regions.

6.1.2) SKIN AND BODY INTERIOR SEGMENTATION

Skin is characterized as the frontier from background and foreground in CT data, so its segmentation is of great interest in order to provide a mask that confines in a meaningful way our regions of interest. Skin was segmented using both PET and CT data volumes and morphological operators such as Dilations, Erosions, Opening and Close.

The segmentation of skin from CT data has an added difficulty associated with the scanner's imaging table. In fact, the imaging table is an undesired component in the CT volumes because its density (in HU) is similar to the density of some soft tissues, so even after the segmentation process it could remain in the volume. The way found to easily remove the imaging table from the CT volume was using PET information. Since the imaging table does not contain radioactivity (it should exist only the patient's body), a threshold level was used in the PET volume data to determine the voxels where activity exists. Since in practice some activity may seem to appear in some voxels outside the patient's body due to the presence of scatter, the threshold of PET data was followed by labeling and the larger labeled region selected, corresponding to the whole body. As was said before, since the PET volume does not have the same dimensions of the CT volume and the voxels sizes are different, the PET volume needed to be accurately interpolated to the CT dimensions in order to obtain the corresponding voxels in the CT volume.

Skin was segmentation applying morphological operators described in section 3.3. The structuring element used in the segmentation is illustrated in figure 22, and is characterized as an elementary element, more suitable to extract fine details from three-dimensional volumes.

It is possible to use a combination of both operators such as first eroding and then dilating (corresponding to the "open" operator) and vice-versa (corresponding to the "close" operator). The result of repeatedly applying dilations and erosions is the elimination of specific image details smaller than the structuring element without distortion/deformation of the shape of unsuppressed regions.

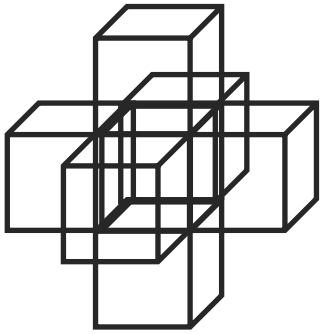


Figure 22 - Structuring element for skin segmentation

Applying morphological operators to volume, first filling internal gaps with close operator and then using a combination of erosions and dilations is possible to extract the skin. The procedure used was:

- First the close operator was used to fill internal gaps in the original volume with a cubic (3×3×3) structuring element.
- The original volume was then eroded and dilated with the structuring element illustrated in figure 22, creating 2 volumes (one eroded and one dilated)
- Skin was obtained subtracting the eroded volume from the dilated volume. Using this procedure it would be also possible to define an "internal skin" (= original volume - eroded volume) and an "external skin" (= dilated volume – original volume), if necessary.

6.2) ESTABLISH CORRESPONDENCE / MATCH CRITERIA

This step is characterized by establishing corresponding bone regions in two CT volumes from different studies of the same patient. Correspondence of two regions in the two volumes was achieved based on the number of voxels that belong to each region (an indicator of the similarity of the size of the two regions) and on the distance between the centroids of the regions (indicating proximity between the two regions).

6.2.1) CRITERION BASED ON THE VOLUME OF REGIONS (NUMBER OF VOXELS)

The list of regions of bone structures obtained in the previous step (see section 6.1.1.) was ordered by descending number of voxels using IDL's *sort* function. With that, larger regions appear first in the list than the other smaller regions. This step was important because it facilitated the comparison in some cases. In fact, similar bone regions in different volumes would have more or less the same number of voxels if the segmentation and labeling of bone structures were performed in similar conditions. However, given the vertical axial symmetry of the body, using only this matching criterion could lead to errors, such as switching the side of the corresponding region of the body (e.g., matching the left femur on one volume with the right femur in the other). To avoid this problem, a second criterion based on the region's center of mass was used in the analysis of correspondence between regions, as described in the next section.

6.2.2) CRITERON BASED ON THE EUCLIDEAN DISTANCE BETWEEN THE CENTROIDS OF REGIONS

The Euclidian distance between 2 points in 3D dimensions can be defined as the length of the line that joins the two points. In this criterion, each center of mass of a given region of a reference volume was compared with the centroids of the regions ii the other volume. Defining p as the center of mass of a region from volume 1 (target) with p_x, p_y and p_z being respectively the coordinates of the centroid in the x, y and z dimensions, and q as the center of mass of a given region from volume 2 (reference) with q_x, q_y and q_z being respectively the centroids in the x, y and z dimension, the Euclidian distance can be defined as:

$$D(p,q) = \sqrt{(p_x - q_x)^2 + (p_y - q_y)^2 + (p_z - q_z)^2} \qquad \textit{Equation 14}$$

Calculating the distances between the centers of mass of all regions from the two volumes, the following table could be constructed:

Table 3 – Table of Euclidian distances between regions

Region Number		**Volume1**					
		1	*2*	*3*	*(…)*	*n*	Minimum (i=1,2,3,…,n)
Volume 2	*1*	D(1,1)	D(2,1)	D(3,1)	…	D(n,1)	Min[D(i,1)]
	2	D(1,2)	D(2,2)	D(3,2)	…	D(n,2)	Min[D(i,2)]
	3	D(1,3)	D(2,3)	D(3,3)	…	D(n,3)	Min[D(i,3)]
	(…)	…	…	…	…	…	…
	m	D(1,m)	D(2,m)	D(3,m)	…	D(n,m)	Min[D(i,m)]
	Minimum (j=1,2,3,…,n)	Min[D(1,j)]	Min[D(2,j)]	Min[D(3,j)]	…	Min[D(n,j)]	

The correspondence was achieved by calculating the minimum distance between the center of mass of a given region from volume 1 and all centers of mass from the regions of volume 2. With that procedure it would be possible to appear regions that possessed multiple correspondences. For example, both region 1 and 3 of volume 2 corresponds to region 1 of volume 1.

To overcome this problem, we adopted the strategy of merging those regions that corresponded to an individual region of the other volume (remembering the aforementioned example, region 1 and 3 of volume 2 were merged). We developed a function that merged the voxels of two regions and recalculated the fields of the structure that described the region (discussed in section 6.1.1).

The table of distances between the regions of the two volumes was then calculated again, but now without regions with multiple correspondences and finally the indexes of the corresponding regions was determined to proceed with the registration algorithm described in the next section.

6.3) REGISTRATION

In this step, was developed a registration algorithm to align corresponding regions in the two CT volumes. It was considered, as a first approximation, only rigid transformation with 6 degrees of freedom (DOF), making clear that more sophisticated algorithms can be developed to better accommodate the deformations. These algorithms require more computational demands, however.

6.3.1) RIGID-REGISTRATION (USING ROTATIONS AND TRANSLATIONS)

The registration algorithm was developed based on the rigid transformation described in section 4.3.1. The volumes were transformed into masks (1 value for voxels that belong to the region region and 0 value for background voxels) before the application of the optimization method. Six DOFs were considered, 3 for translations and 3 for rotations along each axis. These DOFs were obtained using the Powell's optimization method available in IDL.

Powell's method minimizes a user-supplied objective function in an N-dimensional parameter space. The method achieves the optimal solution by a succession of one-dimensional minimum determination, which is relative simple to calculate. The process is iterative, starting with an initial guess X_0 (x_1, x_2, x_3, …, x_n), and with a linear combination of the minimum search at each dimension, reaches an optimal solution. For a more detailed description, see [33].

In this specific case, the dimension of the parameter space is related with the DOF allowed by the registration algorithm (6). After each iteration, the volume is transformed by applying the returned parameters (DOFs) to the reference volume, and the sum of square difference (SSD) of voxels intensities of the two volumes (reference and target) is calculated. Defining R as the Reference volume and T as Transformed volume for n points, the objective function is described as:

$$SSD = \sum_{i=0}^{n} (R_i - T_i)^2 \qquad \textit{Equation 15}$$

The algorithm stops when a set number of iterations is done or when the result of the objective function is acceptable (below a given tolerance threshold).

6.3.2) APPLICATION OF REGISTRATION TO SURROUNDING TISSUES

This step, duo the lack of time, was not implemented but will be described for a posterior work. Since we want to register the whole body CT scan information and not only the bone structures, was considered that the surrounding of bone structures "moves" coupled to bone. Using the information about skin and body interior segmentation, a mask can be constructed that spatially limits what is the body from background for each region. By applying the transformation obtained for bone regions to these masks containing the voxels that are

closest to the region and in the body interior, the entire volume can be transformed (considering the transformation of all regions).

7) RESULTS

7.1) SEGMENTATION OF THE SKELETON

As described in section 6.1.1, the results of the bone segmentation and labeling of the CT volumes (previously halved to reduce computational demands, as described in section 6), are shown in Figure 23. After a few tests with the several data volumes available, the range of values 400-1500 HU was considered a good threshold value for bone segmentation. The following figure shows results of bone segmentation for each volume (target and reference) of patient 7, the only patient where two studies with the complete body were available.

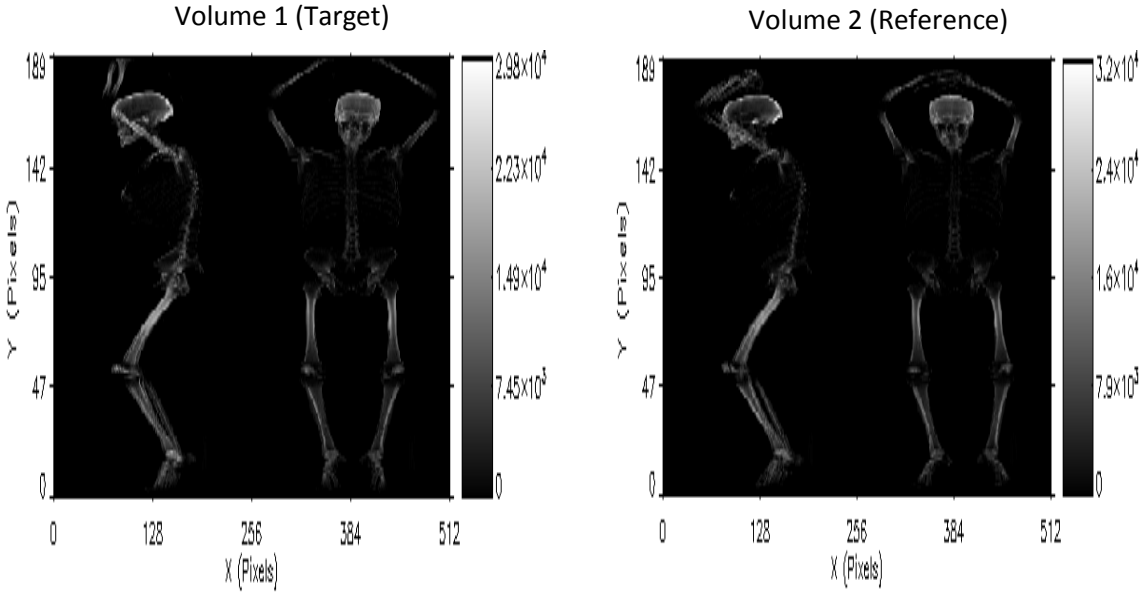


Figure 23 – Example of the results obtained for the bone segmentation of Target (right) and Reference Volume (left). The segmentation was performed in three dimensions. For visualization purposes, the sum of the sagittal and coronal segmented slices is shown for each volume.

7.2) BONE LABELLING

The labeling process, in which a skeleton is divided into individual regions, is applied after the segmentation of bone structures. The objective of this step is to isolate, from entire volumes, the individual bones of skeleton that were associated to the principal joints and are characterized by being those that vary more its spatial position over scans. First, regions are sorted by the number of voxels (dimension) and using the Euclidian distance between the centroids of regions, the first correspondence between regions is achieved. Note that, for example, if a region in volume 1 has multiple corresponding regions in volume 2, this fact leads to a process that merges those regions with multiple correspondences in volume 2. The results from labeling process of bone structures after these merge processes can be seen in figure 24 (volume 1 – target) and in figure 25 (volume 2 – reference).

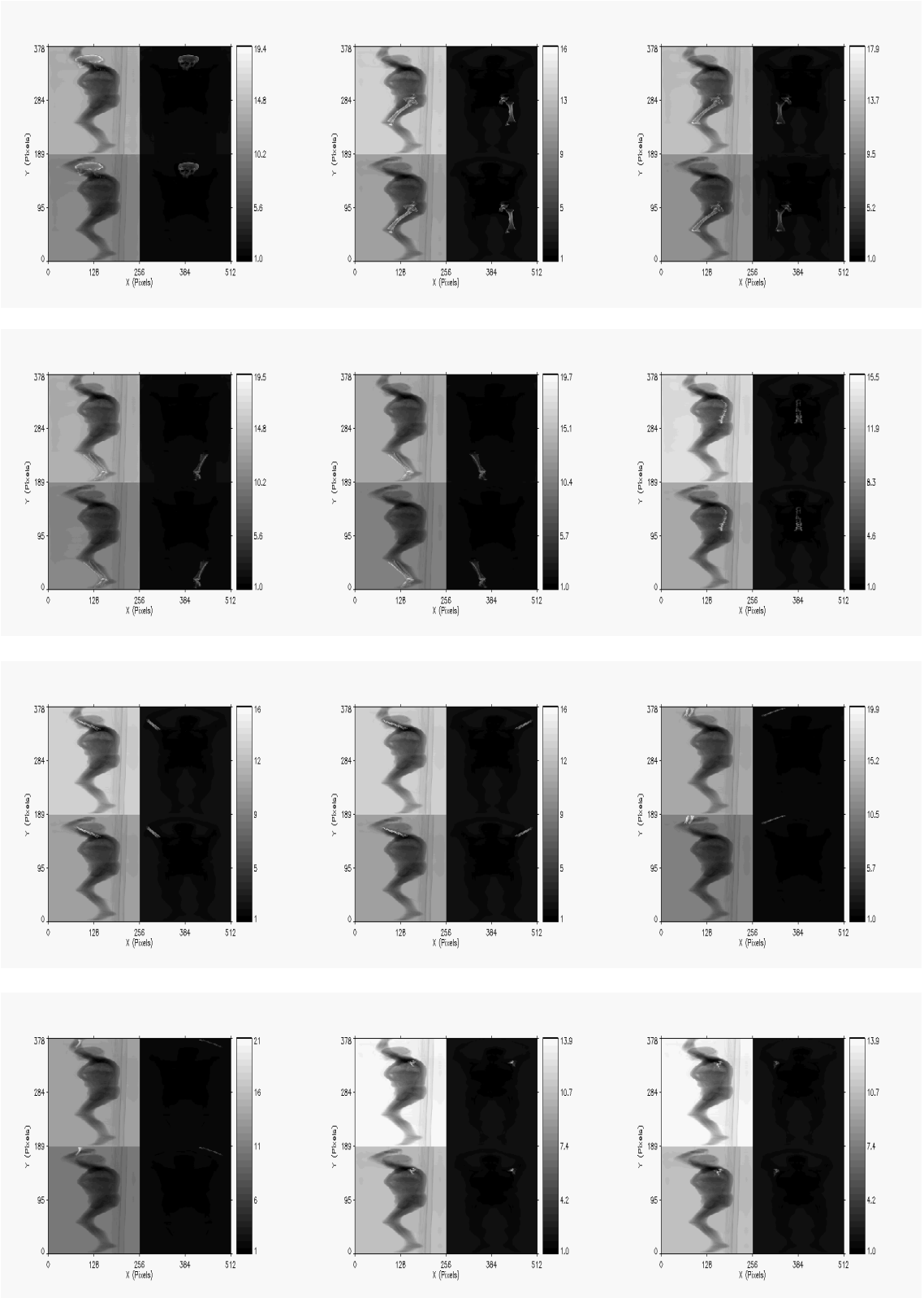


Figure 24 – Results of the Target Volume labeling for 12 regions. For each region, the top images are the sum of the sagittal (left) and coronal (right) projections of the labeled region in volume 1, with the region superimposed, and the bottom images are the sum of the sagittal (left) and coronal (right) projections of the labeled region in volume 2, also with the region superimposed.

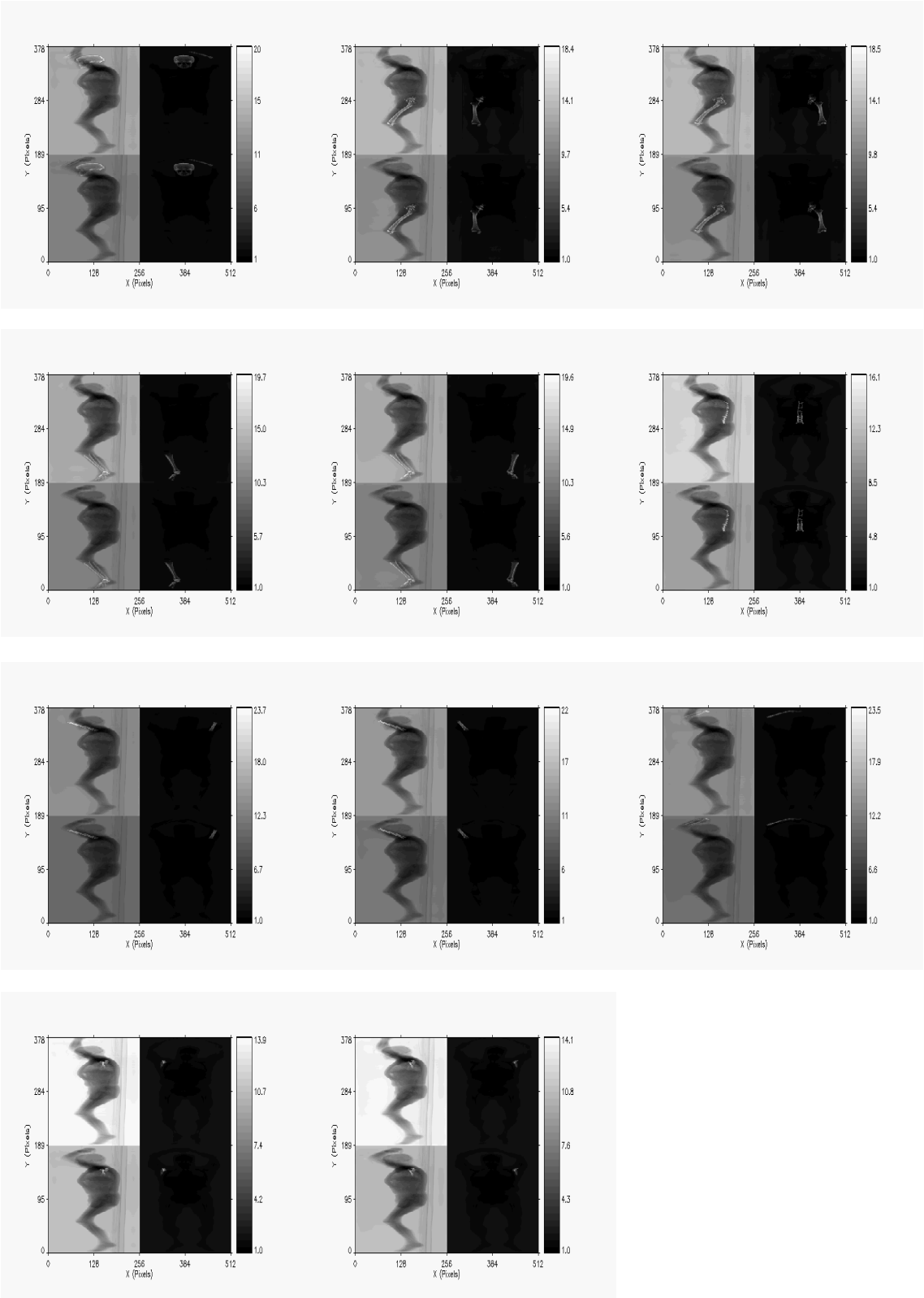


Figure 25 – Results of the Reference Volume labeling for 11 regions. For each region, the top images are the sum of the sagittal (left) and coronal (right) projections of the labeled region in volume 1, with the region superimposed, and the bottom images are the sum of the sagittal (left) and coronal (right) projections of the labeled region in volume 2, also with the region superimposed.

7.3) CORRESPONDENCES CRITERIA

Once the labeling of volumes in individual regions is achieved, the following procedure is the calculation of distances between regions of different volumes, in a way to establish the correspondence between similar regions. All regions are sorted in a list according to their dimension and the distances are calculated using the Euclidian distance. The criterion of correspondence used was the minimum distance calculated for regions from different volumes. The correspondence between the indices of the sorted regions in each volume can be seen in table 4. The same regions in different volumes were compared visually. This qualitative comparison can be done by analyzing figure 26.

Table 4 – Result of the correspondence of similar regions in different volumes for the dataset shown in figures 24 and 25 . The indices of regions were related to their size in voxels.

	Volume 1 (Target)	Volume 2 (Reference)
Region number	0	0
	1	4
	2	2
	3	1
	4	3
	5	6
	6	5
	7	7
	8	8
	9	9
	10	10

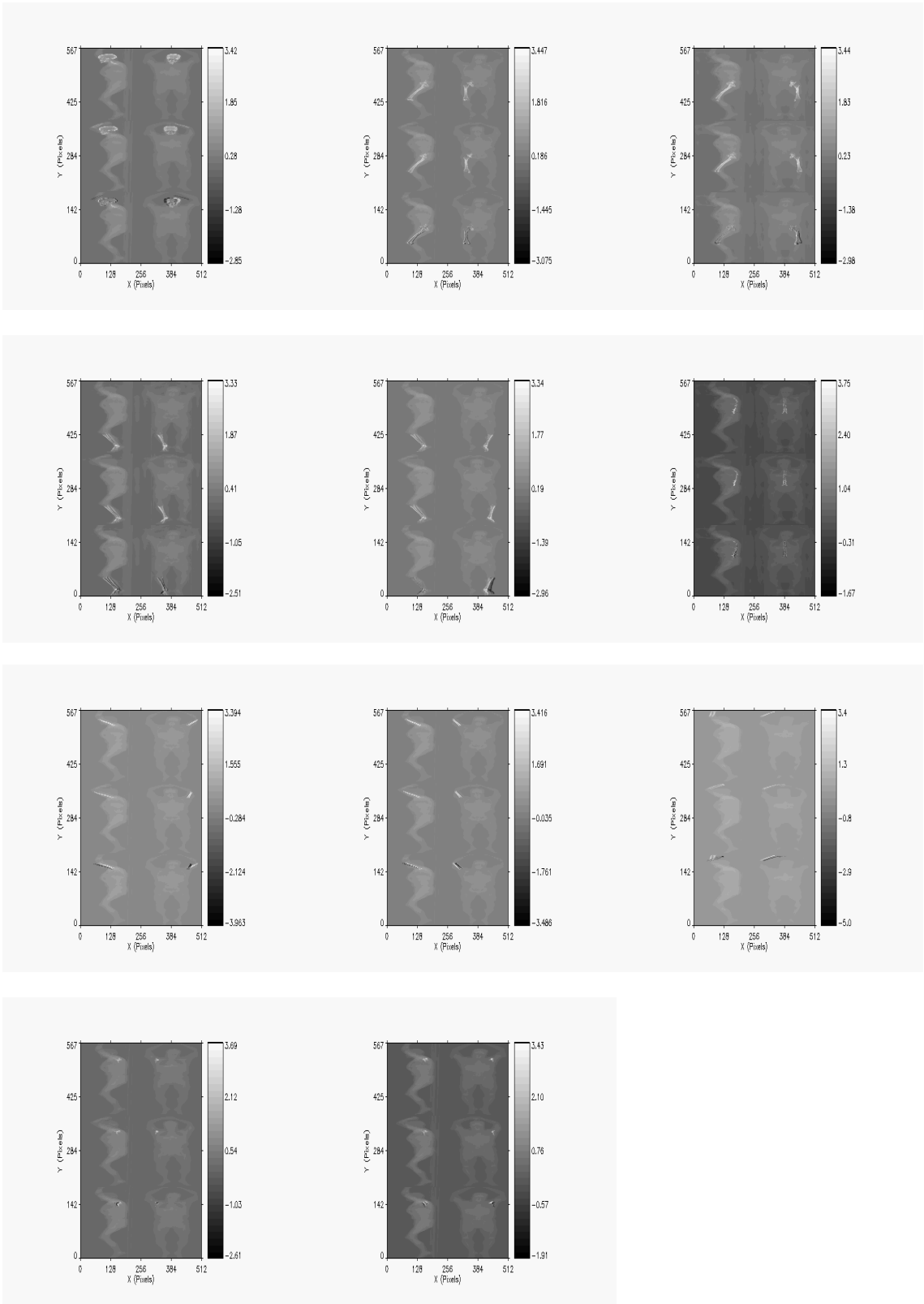


Figure 26 - Correspondence between similar regions in different volumes (11 regions). In each group of images, the region is shown superimposed on the sum of sagittal (left) and coronal (right) slices. In each group, the upper row shows volume 1 (target), the middle row represents volume 2 (reference) and the bottom image is the difference of the two upper rows.

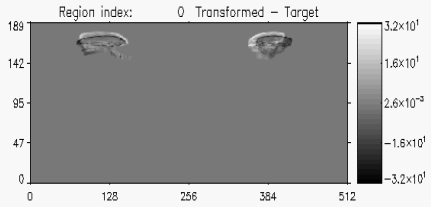
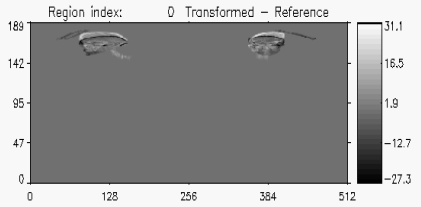
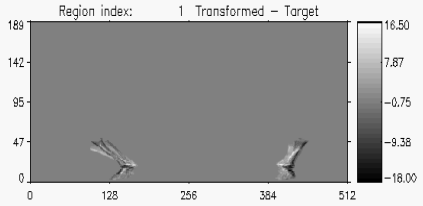
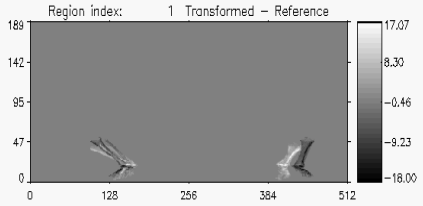
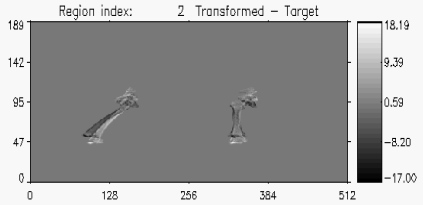
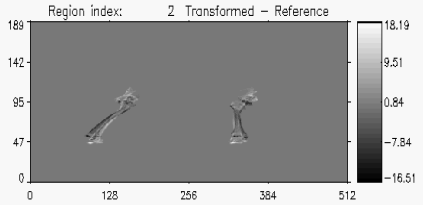
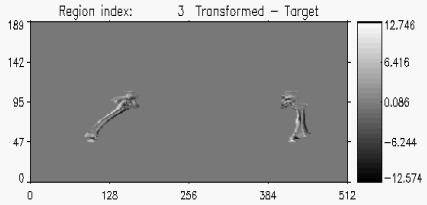
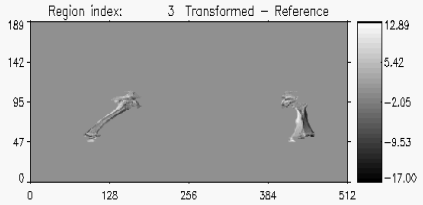
7.4) REGISTRATION PROCESS

The registration of the bone is done using a rigid transformation with 6 DOFs. As mentioned in section 4.3.1, the transformation matrix is defined as shown in Equation 3. Applying Powell's method, using as cost function (that we want to minimize) the sum of the squared differences between the intensities of the reference volume and the transformed target volume, 6 parameters were obtained for, as a first approximation, 7 major regions. The transformation parameters applied in these first 7 regions can be seen in the following table:

Table 5 - Parameters obtained from the Powell's method for seven selected regions

Volume 1- Parameters						
	Rotation in xx axis (°)	Rotation in yy axis (°)	Rotation in zz axis (°)	Translation in xx axis	Translation in yy axis	Translation in zz axis
Region 0	-4,21698	-7,64898	2,727396	2	2,367106	1,998797
Region 1	7,006034	7,004249	4,994975	1,991945	3,008403	1,979241
Region 2	9,636717	-5,74247	5,998388	0,82091	3,736166	3,625711
Region 3	3,692708	2,750552	4,962711	1,747392	1,995157	1,38078
Region 4	3,392483	5,008063	7,72982	0,374488	2,165856	1,863425
Region 5	-7,91886	2,903535	-0,37612	1,576333	2,084114	2,668864
Region 6	5,61624	4,999306	4,373835	2,592009	1,997922	1,987539

The visualization of the transformation applied to regions of volume 1 (target) is shown in figure 27. The transformed volume was compared with the target and reference volumes. For each comparison, the sum of the sagittal and coronal slices is displayed.

Region number	Transformed - Target	Transformed - Reference
0		
1		
2		
3		

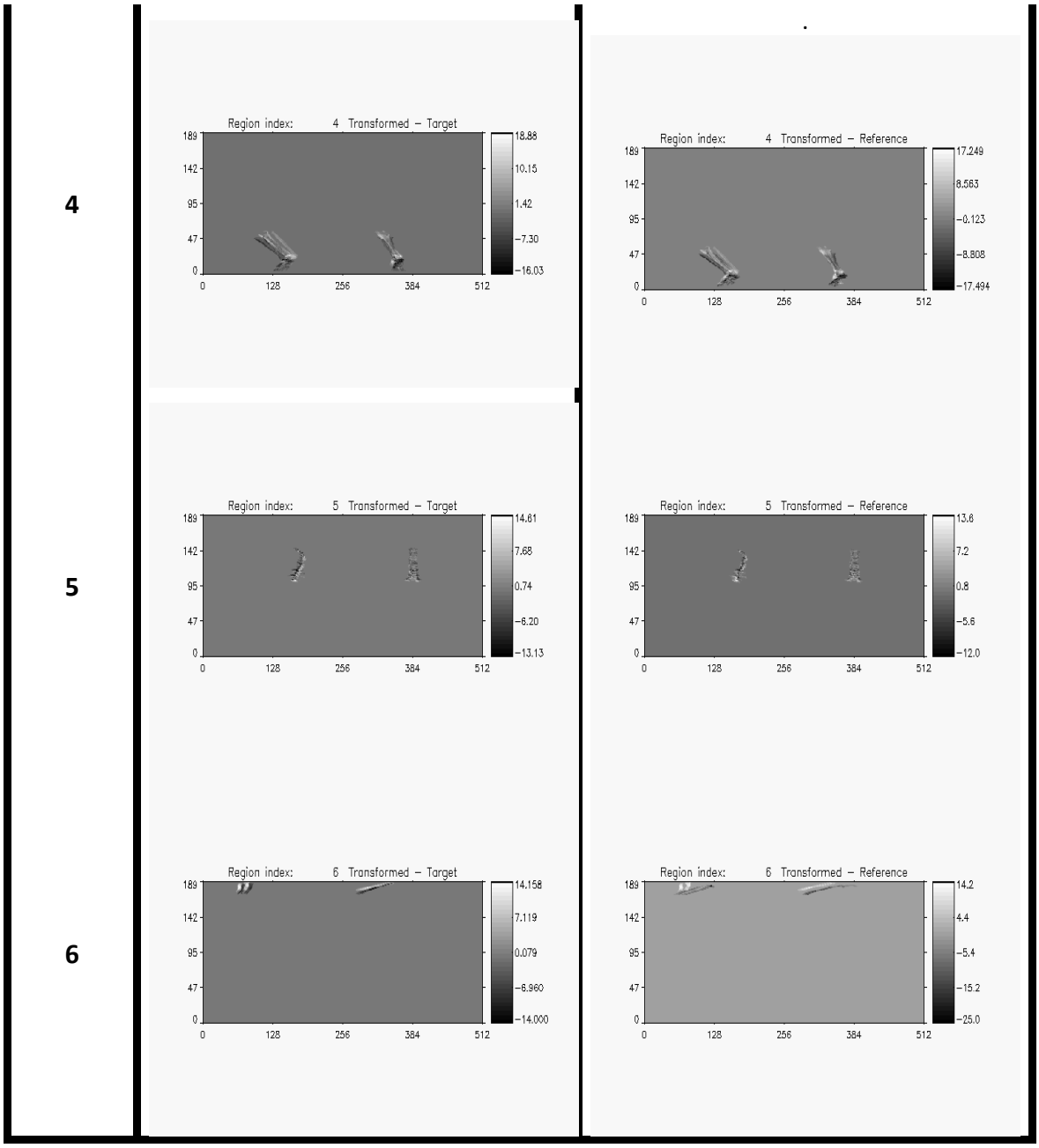


Figure 27 – Set of images illustrating the results obtained for each bone alignment. The images in the middle column are the sum of sagittal and coronal slices of the difference between the transformed image and the target image voxel's intensities (target image is darker) for a given region. The right column shows the corresponding results, but for the differences between the Transformed and Reference volume (Reference image is darker).

As can be seen in the column "Transformed – target", the target volume (darker in image) was transformed with the calculated parameters trying to fit the reference volume. The result alignment is not reliable as can be seen in the column "Transformed – Reference".

7.5) SKIN AND BODY INTERIOR SEGMENTATION

As described in section 6.1.2, another procedure was done for skin and body interior segmentation. The skin as well as interior of body is segmented from the entire volumes by applying morphological operators and the result of its segmentation can be seen in the following figure (figure 28). The skin can also be segmented for an individual region of the volume.

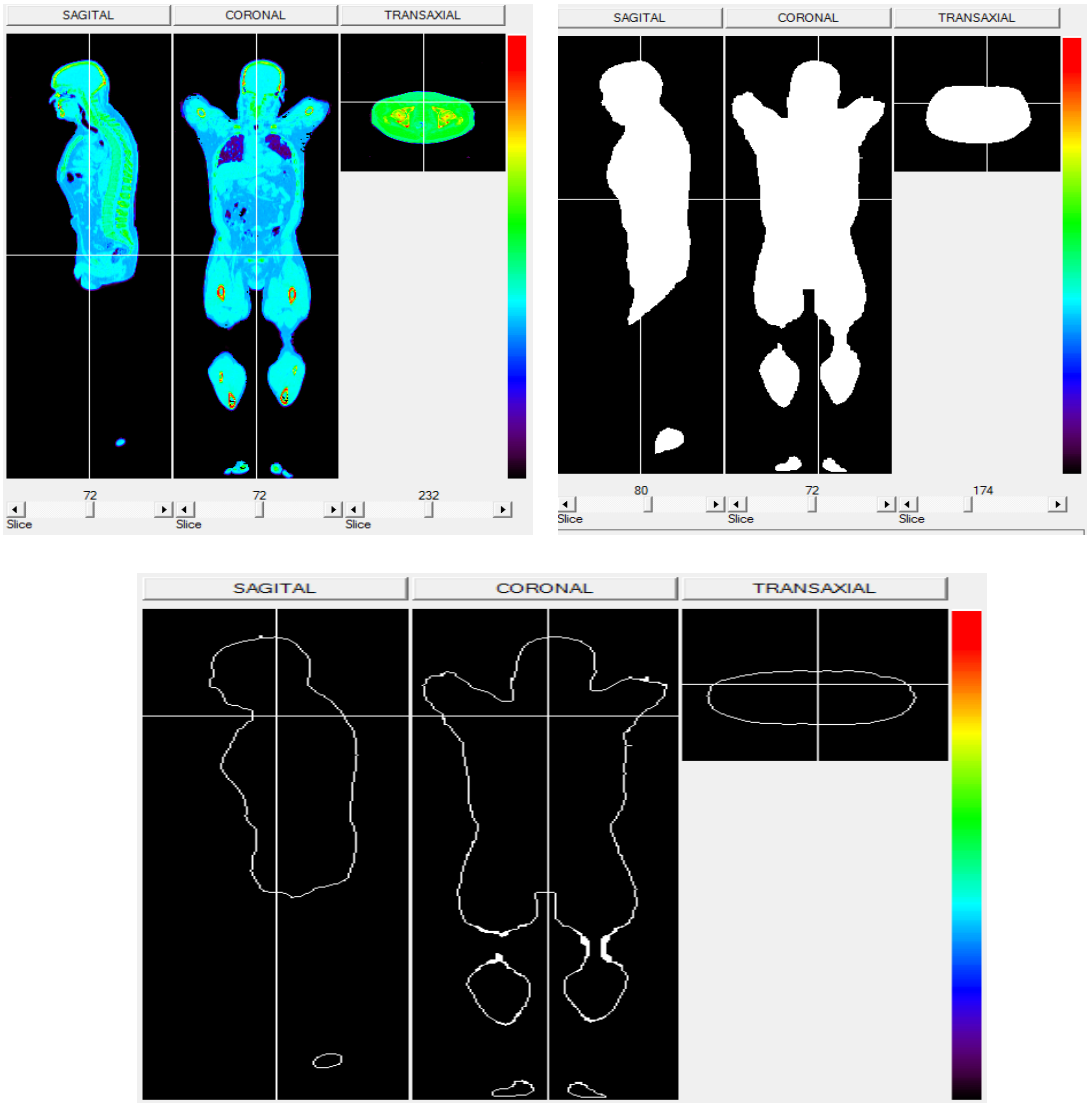


Figure 28 - Examples of segmentation of the skin and body interior of the entire CT volume (1 slice shown for each of the sagittal, coronal and transaxial directions). Top left figure represent the original volume, the top right figure represents the mask of body interior and bottom figure represents the skin segmented.

8) DISCUSSION

This work presents an initial approach for the registration of whole-body PET/CT images even in cases where there are large movements of several parts of the body. Although its effective application in the clinical routine is far in the horizon, a successful effort towards achieving this goal would bring several significant advantages, allowing the detailed analysis and comparison of multiple studies of a patient, which would be crucial to better monitor and understand the progression of diseases and the effects of therapies, among other applications. Such methodologies, if successfully extended to allow the coregistration of whole body PET/CT studies from different patients to a common reference, would allow the detailed study and comparison of different patients, of groups of patients and of specific populations, potentially providing a new level of understanding of the progression and cure of pathologies. .

As was already said, CT and PET volumes have associated noise that results from the process of image acquisition. Noise is a factor that influences all the processing techniques applied in volumes, so it is important to minimize its impact on the results. We made some early attempts to reduce this noise using Gaussian, Mean and Median filters but did not reach a conclusion on what was the best effective approach to address this problem, in part because many subsequent steps in the processing still had to be developed.

Another point that influences the final result to be considered is the need of volume resizing (halving the number of voxels in each dimension) to test the algorithms in the available computer resources. In fact, this resizing procedure uses volume interpolations that can influence the distance between voxels and the sampling of the volume, with a negative impact on the segmentation and labeling of structures, as well as in all subsequent processing techniques that were carried out.

The bone structures were segmented in the entire skeleton by simply application of a threshold and were further labeled. However, the labeling process is influenced by several factors already cited in this section. One possible way to improve the labeling results is using information of the anthropomorphic human phantom (NCAT) in a process also called atlas-based segmentation. This process requires a-priori knowledge of bone structures and must be able to simulate human body in a realistic way, incorporating information about the

characteristics of joints (relative position in the body, movements allowed, relation with another joints, etc.), since joints are responsible for the main image differences in scans taken at different times.

Match criteria is one of the most important steps in image registration. In the developed algorithm, only the dimension of bone structures (number of voxels belonging to a given structure) and the Euclidian distance between structures from different volumes were considered to infer the correspondence. In this work was already tested a match criteria that is known as compaction (relation between superficial area and total volume of a region), however this match criteria did not produce conclusive results and was discarded. There are some other match criteria that can be used to establish correspondence between regions of different volumes, such as moments of third and fourth order. These were not tested but can be useful in a posterior, more sophisticated attempt.

We are dealing with intra-subject registration techniques, so the correspondence criteria are simpler than those required for inter-subject registration. If the objective of a posterior work is the construction of an atlas, other correspondence criteria have to be considered such as Mutual Information.

As it was seen in figure27 (section results) the transformation model algorithm developed was not robust in the bone alignment. In fact, the optimization method used returns results below expectations. It was tested for a volume containing a sum of three-dimensional gaussian functions (in order to create an irregular bone-like structure) and the results were encouraging so we expected that it would yield better results with real CT data. Powell's method is dependent on the order that the dimension minimization is calculated, so if it starts minimizing a rotation parameter the results could be different than those obtained if it started by optimizing translations. Other optimization methods can be more suitable to achieving the parameters, but involve the calculation of gradients, which is not feasible to run in the computer available for this work. Being one of the simpler exiting optimization methods, Powell's method took about a day to get results only for 7 of the regions, with the dimension of regions halved from the original volume. One fact is that the algorithm was not developed with the goal to be fast, the time consumption will be taken into account once the results are good enough to justify it. Another possible explanation is that the implementation used did not correctly take into account the anisotropic voxels of the CT data (the voxels were not cubic), leading to a transformation that tended to deform the structures in a not uniform way.

The registration processes for bone alignment were done, however one step remains to be done. After the transformation of bone structures and using the skin and internal body segmentation information, the entire CT volume (bone and surrounding tissues) can be transformed, and with interpolations the PET volume can also be transformed. However, because we are only dealing with rigid transformations, some discontinuities and overlaps in the transformed volume can appear. These drawbacks cause the loss of anatomical definition of the registered PET/CT volume and in consequence limit the applicability of the algorithm in clinical practice. As first approximation, the results are still far those required by medical center applications and need to appeal to more sophisticated registration algorithms, mainly non-rigid techniques to model soft tissues deformations (described in section 4.3.2). Using non-rigid registration techniques, the deformation can be done in a more freely way, allowing more DOF and consequently better results fitting the volumes.

Another problem found was that some regions that do not have direct correspondence are lost in the correspondence process. One way to solve this problem is, for these regions, to merge with the closest region.

The evaluation of algorithm performance is only done in a qualitative form, analyzing images with eye inspection. The objective is to achieve a way to quantitatively describe the algorithm performance.

This work did not reach the phase of the validation of a complete algorithm for whole body coregistration, but one possible strategy for validation would be as follows: transform a target volume ***T*** to a Reference volume ***R***. Once achieved the transformation, the volume ***Tr*** (transformed) is obtained and the same algorithm is applied to this new volume having now as Reference of registration volume ***T*** (original target), obtaining the Volume ***Tr'***. The comparison of volumes ***T*** and ***Tr'*** may constitute a measure for evaluating the performance of the algorithm developed. This procedure can be seen, just as an illustrative example, in Figure 29.

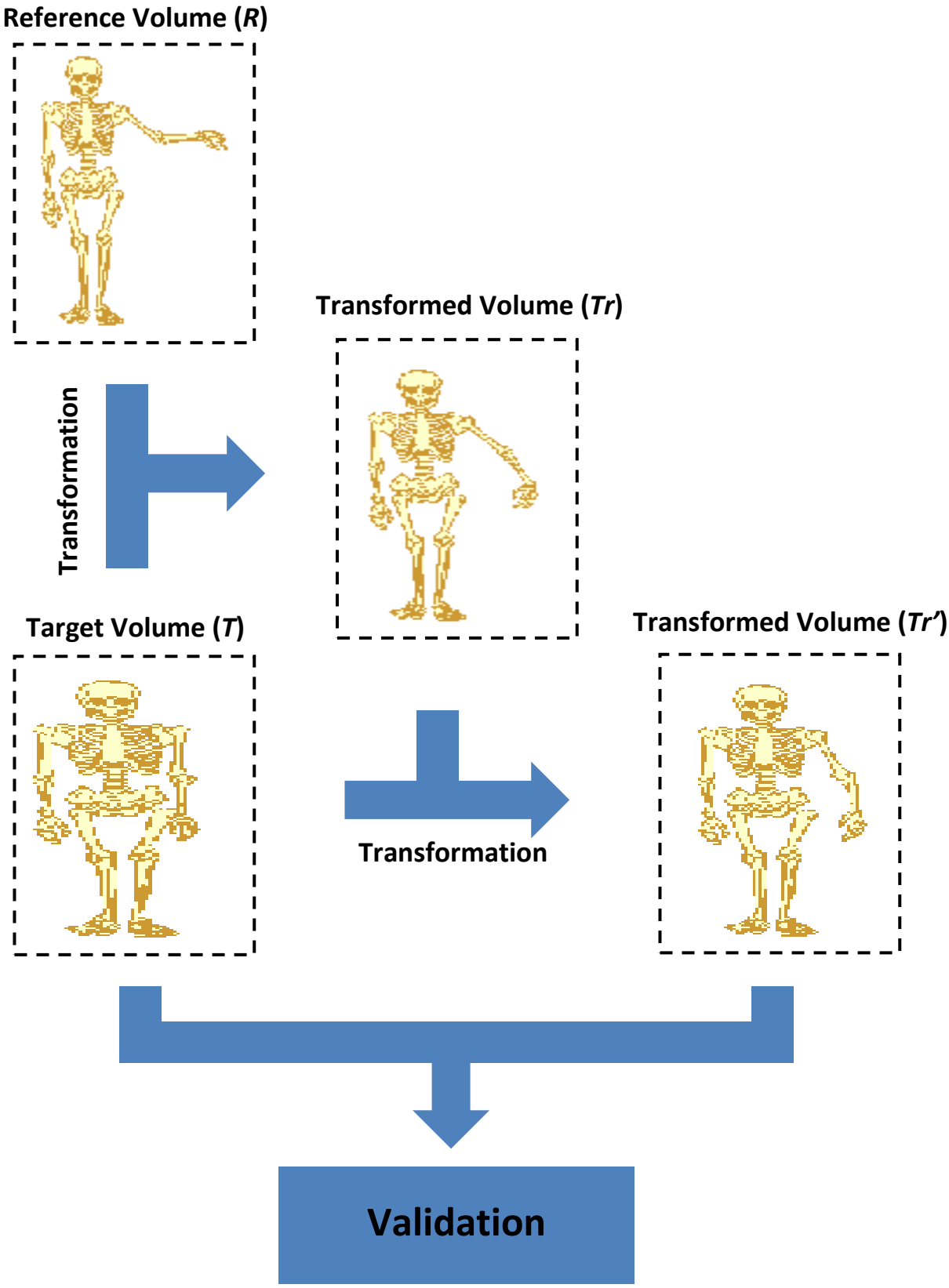


Figure 29 – Illustrative example of a validation strategy

9) CONCLUSION

In this thesis, a study of a registration method for bone structures alignment from PET/CT scans was performed. The proposed methodology starts with the extraction of bone regions from the 3D CT volume. Using match criteria, defined in section 6.2, the correspondence was achieved and regions of a target volume were rigidly transformed in order to align them with corresponding regions of a reference volume. The final results as well as the processing time of the registration methodology are still far for those required in medical environment, but the work developed can be a basis for further future studies.

As final remark, the registration of whole body 3D volumes is a field relatively unexplored. This fact is due, in part, to the complexity of the processing techniques and to the large computer demands required. However, it is a useful tool in a clinical setting, improving healthcare solutions provided to the community in general, and it is very likely that whole body 3D coregistration tools will be suitable for clinical application in the near future.

10) REFERENCES

[1] I. G. Maglogiannis, K. Karpouzis, and M. Wallace, "Image and Signal Processing for Networked E-Health Applications," Synthesis Lectures on Biomedical Engineering, vol. 1, no. 1, pp. 1-108, Jan. 2006.

[2] A. Maciel, "A biomechanics-based articulation model for medical applications," *Science*, Ph.D. thesis, 2005.

[3] M. L. GIGER, "Computer-aided diagnosis in medical imaging: A new era in image interpretation," WMA Business Briefing - Global Healthcare 2000, p. 75–79, 2000.

[4] T. M. Deserno, *Biomedical Image Processing.*, vol. 5, no. 3. Aachen, Germany: Biological and Medical Physics, Biomedical Engineering ISSN 1618-7210, 2011.

[5] J. J. Pedroso de Lima, *Nuclear Medicine Physics*, Boca Raton, U.S.A., Series in Medical Physics and Biomedical Engineering, 2011

[6] G. L. Zeng, *Medical Image Reconstruction: A Conceptual Tutorial*, Springer Heidelberg Dordrecht London New York, 2010

[7] E. Even-Sapir, Z. Keidar, and R. Bar-Shalom, "Hybrid imaging (SPECT/CT and PET/CT)--improving the diagnostic accuracy of functional/metabolic and anatomic imaging.," *Seminars in nuclear medicine*, vol. 39, no. 4, pp. 264-75, 2009.

[8] D. W. Townsend, "Basic science of PET and PET/CT," *Positron Emission Tomography*, pp. 1–16, 2006.

[9] T. Beyer et al., "A combined PET/CT scanner for clinical oncology.," *Journal of nuclear medicine : official publication, Society of Nuclear Medicine*, vol. 41, no. 8, pp. 1369-79, 2000.

[10] D. W. Townsend et al., "PET/CT today and tomorrow," The journal of nuclear medicine, vol. 45, no. 1 , 45:4S–14S, 2004

[11] T. Beyer, D. W. Townsend, J. Czernin, and L. S. Freudenberg, "The future of hybrid imaging—part 2: PET/CT," *Insights into Imaging*, 2011.

[12] William R. Hendee, E. Russell Ritenour, *Medical imaging Physics*, Fourth Edition, Wiley-Liss Inc. New York, ISBN 0-471-38226-4, 2002

[13] D. Delbeke, *Hybrid PET/CT and SPECT/CT Imaging*, Springer New York Dordrecht Heidelberg London, ISBN 978-0-387-92819-7, 2010

[14] http://www.medcyclopaedia.com/library/topics/volume_iii_1/h/hounsfield_unit.aspx (accessed on 04/08/2011)

[15] Hal O. Anger, Scintillation Camera, The Review of specific instrumentation, volume 29, number 1, 1958

[16] D. Delbeke, H. Schöder, W. H. Martin, and R. L. Wahl, "Hybrid imaging (SPECT/CT and PET/CT): improving therapeutic decisions.," *Seminars in nuclear medicine*, vol. 39, no. 5, pp. 308-40, 2009.

[17] Roobottom CA, Mitchell G, Morgan-Hughes G, "Radiation-reduction strategies in cardiac computed tomographic angiography". *Clin Radiol.* 65: 859–67, 2010

[18] C. Computing and M. Imaging, "Title of the Thesis 'Segmentation of Human Lung CT scans using Hybrid method' Submitted by Ramakrishna Paruchuri STATEMENT OF SOURCES," *Engineer*, 2006.

[19] Yunfeng Wu*, R. M. Rangayyan, Yachao Zhou, S. C. Ng, "Filtering electrocardiographic signals using an unbiased and normalized adaptive noise reduction system", *Medical Engineering & Physics*, 31(1): 17-26, 2009

[20] Wong WCK, Chung ACS. In: Lemke HU, Inamura K, Doi K, et al., editors. A Nonlinear and Non-Iterative Noise Reduction Technique for Medical Images: Concept and Methods Comparison. International Congress Series. Elsevier, p. 171–176, 2004.

[21] http://en.wikipedia.org/wiki/Noise_reduction#In_images (accessed on 27/08/2011)

[22] Junichiro Toriwaki, Hiroyuki Yoshida, *Fundamentals of Three-Dimensional Digital Image Processing*, Springer Dordrecht Heidelberg London New York, ISBN 978-1-84800-172-5, 2009

[23] Joseph V. Hajnal, Derek L.G. Hill, David J. Hawkes, *Medical Image Registration, Biomedical engineering series, IV Series,* Boca Raton, U.S.A., 2001

[24] Ayman S. El-Baz, Majid Mirmehdi, Rajendra Acharya U, Jasjit S. Suri, *Multi-Modality State-of-the-Art Medical Image Segmentation and Registration Methodologies,* Volume 1, Louisville, U.S.A., 2011

[25] J. Modersitzki, *Numerical Methods for Image Registration*, Oxford Science Publication, 2004

[26] Goshtasby A, *"Registration of image with geometric distortion,"* IEEE Trans Geosci Remote Sens 26(1):60–64, 1988

[27] F.L. Bookstein, "*Principal warps: thin-plate splines and the decomposition of deformations," IEEE Trans. on Pattern Anal. and Machine Intell*., 11(6):567–585, 1989.

[28] J. Duchon, *"Interpolation des functions de deux variables suivant les principes de la flexion des plaques minces"*. *RAIRO Analyse Numérique*, 10:5–12, 1976.

[29] J. Meinguet, "Multivariate interpolation at arbitrary points made simple," Zeitschrift für angewandte Mathematik und Physik, 30:292–304, 1979.

[30] Sederberg TW, Parry SR, *"Free-form deformation of solid geometric models,"* Proc SIGGRAPH.;20(4):151–160, 1986

[31] C.A. Pelizzari, G.T.Y. Chen, D.R. Spelbring, R.R. Weichselbaum, and C.-T. Chen, *"Accurate three-dimensional registration of CT, PET, and_or MR images of brain," J. Comput. Assist. Tomogr*., vol. 13, pp. 20–26, 1989.

[32] P.J. Besl and N.D. McKay, *"A method for registration of 3-D shapes," IEEE Trans Pattern Anal. Mach. Intell*., vol. 14, pp. 239–256, 1992.

[33] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery, *Numerical Recipes in C: The Art of Scientific Computing*, Cambridge University Press, Cambridge, England, 2nd ed., 1992.

[34] C. Studholme, D.L.G. Hill, and D.J. Hawkes, *"Automated 3D registration of MR and PET brain images by multi-resolution optimisation of voxel similarity measures," Med. Phys*., Vol. 24, pp. 25–35, 1997.

[35] X. Li, T. E. Peterson, J. C. Gore, and B. M. Dawant, "Automatic Inter-subject Registration of Whole Body Images", *Imaging*, pp. 18 - 25, 2006.

[36] O. Camara, G. Delso, I Bloch, "Free Form Deformations Guided by Gradient Vector Flow: A Surface Registration Method in Thoracic and Abdominal PET-CT Applications", WBIR 2003, pp. 224-233, 2003.

[37] D. Rueckert, L. Sonoda, C. Hayes, D. Hill, M. Leach and D. Hawkes, "Non-rigid registration using free-form deformations: application to breast MR images", *IEEE Trans. Med. Imaging* 18, pp. 712–721, 1999.

[38] M. Baiker et al., "Fully Automated Whole-Body Registration in Mice Using an Articulated Skeleton Atlas", *2007 4th IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pp. 728-731, 2007.

[39] M.A. Martín-Fernández, E. Muñoz Moreno, M. Martín-Fernández, C. Alberola- López, "Articulated registration: elastic registration based on a wire-model", in: J.M. Fitzpatrick, J.M. Reinhardt (Eds.), Medical Imaging 2005: Image Processing, SPIE Press, San Diego, CA, USA, pp. 182–191, 2005.

[40] W.P. Segars, B.M.W. Tsui, E.C. Frey, G.A. Johnson, and S.S. Berr, "Development of a 4D digital mouse phantom for molecular imaging research", *Molecular Imaging and Biology,* , vol 6(3), pp. 149-159, *2004*

[41] N. Kovacevic, G. Hamarneh, and M. Henkelman, "Anatomically Guided Registration of Whole Body," *Radiology*, pp. 870-877, 2003.

[42] M. Á. Martín-Fernández, R. Cárdenes, E. Muñoz-Moreno, R. de Luis-García, M. Martín-Fernández, and C. Alberola-López, "Automatic articulated registration of hand radiographs," *Image and Vision Computing*, vol. 27, no. 8, pp. 1207-1222, 2009.

[43] J. A. Little, D. L. G. Hill, and D. J. Hawkes, "Deformations Incorporating Rigid Structures," *Computer Vision and Image Understanding*, Vol. 66, No. 2, pp. 223-232, 1997.

[44] V. Arsigny, X. Pennec, and N. Ayache, "Polyrigid and Polyaffine Transformations: A New Class of Diffeomorphisms for Locally Rigid or Affine Registration," *Proc. of MICCAI'03*, pp. 829-837, 2003.

[45] X. Papademetris, D.P. Dione, L.W. Dobrucki, L.H. Staib, A.J. Sinusas, "Articulated Rigid Registration for Serial Lower-Limb Mouse Imaging," *MICCAI'05*, pp. 919-926, 2005.

[46] M. Baiker et al., "Atlas-based whole-body segmentation of mice from low-contrast Micro-CT data," *Medical Image Analysis*, 2010