



HAL
open science

A Landmark Detection Approach Applied to Robust Estimation of the Exposure Index in Digital Radiography

Paolo Irrera, Isabelle Bloch, Maurice Delplanque

► **To cite this version:**

Paolo Irrera, Isabelle Bloch, Maurice Delplanque. A Landmark Detection Approach Applied to Robust Estimation of the Exposure Index in Digital Radiography. *Innovation and Research in BioMedical engineering*, 2017, 38 (1), pp.42 - 55. 10.1016/j.irbm.2016.12.001 . hal-01467954

HAL Id: hal-01467954

<https://hal.science/hal-01467954>

Submitted on 21 Feb 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A landmark detection approach applied to robust estimation of the exposure index in digital radiography

Paolo Irrera^b, Isabelle Bloch^a, Maurice Delplanque^b

^a*LTCI, CNRS, Télécom ParisTech, Université Paris-Saclay, Paris, France*

^b*EOS Imaging, Paris, France*

Abstract

Purpose. The exposure index is an important measure used in digital radiography to control the dose at the detector. This value should be computed in regions of interest that are adapted to each patient's anatomy and pose.

Material and methods. We propose to define automatically these regions based on anatomical landmarks in the main structures of interest (head, thoracic spine, lungs, lumbar spine, pelvis, femurs, knees, tibiae). This task is achieved by combining the global information on the size and the positions of the anatomical structures on the one hand, with local analysis on the other hand.

Results. Experimental results, on a varied database of 82 full-body acquisitions, demonstrate the interest of the proposed approach, with less errors than existing approaches, in particular on frontal view acquisitions. The method is also robust to variations in patient's conditions and to the potential presence of metallic objects.

Conclusion. The approach proposed in this paper allows consistently estimating exposure index values associated with different X-ray acquisitions. This suggests that the application of the proposed method to clinical practice is promising.

Keywords: Automatic anatomical landmark detection, Digital Radiography, Exposure Index.

1. Introduction

Digital radiography has many advantages over screen-film detectors. For example, digital systems are able to generate well contrasted images at wider dose ranges than analogical ones (Shepard et al., 2009). Indeed, in screen-film imaging, the quality totally depends on acquisition conditions because the image is not post-processed. Overexposed images tend to look too dark and underexposed ones are too bright. On the other hand, digital systems allow obtaining images that are well balanced in terms of contrast by using post-processing methods. However, the relationship between image quality and X-ray dose is

lost. This was the reason of the *exposure creep* in digital X-ray radiography. Basically, since the noise level is the only image quality measurement that changes according to the amount of dose, the users may tend to prefer overexposed images that have better signal to noise ratio (SNR) than correctly exposed one. Nevertheless, this choice is clearly in conflict with the ALARA principle that strongly suggests clinicians to optimize the amount of X-ray exposure *As Low As Reasonably Achievable* according to the purpose of the exam.

The Exposure Index (EI) is a standardized image quality measure that has been proposed thanks to a joint initiative of the International Electrotechnical Commission (IEC 62494-1, 2008) and of the American Association of Physicists in Medicine (Shepard et al., 2009) in order to specifically address this issue. The EI quantifies the amount of dose at the detector, and, hence it must not be mistaken with patient radiation dose. Nevertheless, since it is proportional to the squared SNR (Seibert and Richard, 2011), it can be used to define the lower limit of radiation exposure depending on the intended use of the exam and the maximum acceptable amount of noise for clinicians.

Table 1: Exposure index measurements in anatomical regions of a full-body frontal view exam acquired with the EOS system.

	Exposure index value
Head	33.0
Thoracic spine	38.3
Lungs	104.5
Lumbar spine	25.4
Pelvis	28.8
Femurs	41.5
Knees	67.5
Tibiae	77.6

The standard IEC 62494-1 (2008) is extremely clear on the procedure to follow in order to estimate EI values from image gray levels and we refer to it for any information about, for example, X-ray beam characterization. Nevertheless, the manufacturers are free to choose a method to define the region of interest (ROI) where the EI is computed. It is worth noting that this aspect is not only important as the EI value depends on the selected ROI, but also not trivial to address.

It is then important to define the ROI used to compute the EI in such a way that the comparison between acquisition protocols on different patients is consistent. Furthermore, the variations of patients' poses or the presence of multiple anatomical structures in the field of view make it very challenging to get significant EI measurements. This is even more important for clinical exams requiring a full-body analysis of the musculoskeletal apparatus (Illés and Somoskeöy, 2012). As a typical example, we consider in our experiments images acquired with EOS system, which is dedicated to this type of analysis. Table 1 provides an example of EI measurements computed from an EOS frontal view

acquisition of the full body. The values given in this example show how heterogeneous is the information. For example, the EI value behind the lung region is four times higher compared with the measurement behind the lumbar spine, which is a region at higher density than the chest. As a consequence, a unique EI value computed at the center of the image gives a poor description of the image quality of an exam. It is therefore necessary to detect the anatomical regions of interest that appear in the image.

Irrera et al. (2015) have recently proposed a landmark-based approach that allows addressing the aforementioned issues. However, the evaluation of the method was conducted from manually annotated landmarks. In this work we propose an unsupervised approach that automatically detects these landmarks. Multiple aspects make anatomical structures detection challenging on planar 2D radiographic images: the image quality significantly changes from an exam to another, there are rotational issues due to the projection of the 3D volume on a 2D plane and the intensity values inside the same structure are not homogeneous given tissue superposition. The proposed method should then be able to address all these challenges while being efficient in terms of computational time because the EI has to be immediately displayed on the processed image. The validation is another significant contribution of this work as we consider eight anatomical regions, two acquisition views, patients of different ages and morphotypes, and acquisition protocols at several X-ray exposition levels.

The paper is organized as follows. Section 2 starts by introducing the EI algorithm and by presenting the method for computing EI values with a landmark-based approach (Irrera et al., 2015). We then present the proposed landmark detection approach and describe our clinical database. Section 3 evaluates the proposed method and discusses the obtained results. Section 4 concludes the paper, and summarizes the achieved objectives and perspectives.

2. Materials and Methods

2.1. Exposure index

The exposure index is a standardized measure that represents the amount of dose at the detector in a region that is of interest for the undergoing clinical exam (ROI). The amount of dose measured in Gy is estimated from intensity image values by means of a calibration function that depends on the system (IEC 62494-1, 2008).

The input to the exposure index algorithm is the acquired image corrected in offset, gain and dead pixel. It is worth noting that any further operation on the image that changes intensity values or noise distribution, for example contrast enhancement, must be avoided as the exposure index describes the image quality at the acquisition. The input image is denoted by \mathbf{u} . A ROI $\Psi \subset \Omega$, where Ω is the whole pixel space, indicates the region of the image that is considered meaningful for the undergoing exam. The ROI selection methods presented in the standard IEC 62494-1 (2008) and by Shepard et al. (2009) are based on gray level histogram thresholding, positioning of rectangular or circular shapes at the center of the envelop of the patient or a combination of these

two techniques. The ROI selection may be problematic because of the presence of metallic objects, unexpected positions of body parts and variations in the patient thickness especially in case of full body images. Given the mentioned issues, the techniques used to define the ROIs are just given as examples and are not imposed by the standard IEC 62494-1 (2008). Moreover, the manufacturer is invited to give users the possibility of validating and, eventually, of modifying the automatically selected ROI. Therefore, according to the original formulation, the EI should rely on user-interaction to validate the measures, which can be seen as a remedy to the aforementioned issues related to the ROI definition. Nevertheless, our experience on the use of the EI in clinical routine invalidates this assumption: the users collect the EI and the corresponding deviation index values (Mothiram et al., 2013), whereas the control on the automatically computed ROIs is, to the best of our knowledge, neglected. It is then important to compute exposure index values without requiring user interaction.

From the distribution of the gray levels of the pixels $x_i \in \Psi$, a value of interest (VOI) v is extracted. Since v has to represent the central tendency of the histogram of \mathbf{u} , that we denote with \mathbf{q} , it is equal to the median of $\mathbf{q}(x_i)$, $x_i \in \Psi$ (Shepard et al., 2009). The EI value is then computed as follows:

$$EI = c_0 g(v) \quad (1)$$

where c_0 is a constant fixed at $100\mu Gy^{-1}$ according to the standard IEC 62494-1 (2008) and $g(\cdot)$ is a calibration function. The value returned by $g(v)$ is the Kerma in the air at the receptor associated with the ROI Ψ and it is expressed in μGy . The function $g(\cdot)$ depends on the X-ray system and must be defined in the X-ray standard beam geometry and calibration conditions specified in the standard IEC 62494-1 (2008). The definition of this function for EOS system is not addressed in this paper as it is beyond our original scope.

The EI value on its own is not really significant as what really matters is to assess if the amount of noise in a given image is tolerable according to the anatomical region of interest and the medical purposes of the exam. The manufacturers are in charge of defining target exposure (EI_t) index values according to ideal ALARA dose conditions. However, the users must be given the possibility to update these target values. The deviation index (DI) from the correct exposure EI_t is then computed as follows:

$$DI = 10 \log_{10} \left(\frac{EI}{EI_t} \right) \quad (2)$$

where $DI < 0$ and $DI > 0$ respectively indicate underexposure and overexposure. In practice, the acquisition setting would need to be adjusted only if $|DI| > 1$, i.e. if the estimated EI value increases by +25% and decreases by 20% with respect to the target EI value.

2.2. Estimation of exposure index values from anatomical landmarks

Irrera et al. (2015) have proposed to replace the ROI definition with a landmark based approach by associating a set of anatomical landmarks with struc-

tures of interest. We summarize here how EI values are computed from these sets of points.

A local EI measure can be computed considering the distribution of gray levels at the pixels $x_i \in P_i$, where P_i is a patch centered at the landmark l_i . Circular patches of radius equal to 128 pixels are used, i.e. about 23 mm on an EOS image. Note that on an adult patient such a patch approximately covers the area of a vertebra. The circle has to be large enough to avoid the measure to be excessively affected by noise and, at the same time, small enough to guarantee the gray level distribution to be approximately mono-modal. In this way the median value of a distribution will describe in a more representative way the actual amount of signal and, hence, of the EI value in a region. Therefore, local measures are much less biased by presence of outliers than those extracted from the histogram of the whole image.

Formally, we denote a ROI by \mathcal{A}_r . In this ROI, the landmark detection method described next provides a cluster \mathcal{L}_r of landmarks l_j ($l_j \in \mathcal{L}_r$). Then the corresponding values of interest $v(l_j)$ and the local EI values $e(l_j)$ are computed using Equation 1. Finally, the EI value in the ROI \mathcal{A}_r is computed by using the following weighted sum of local EI values:

$$EI_r = \frac{\sum_{l_j \in \mathcal{L}_r} \omega(l_j) e(l_j)}{\sum_{l_j \in \mathcal{L}_r} \omega(l_j)} \quad (3)$$

where the weights $\omega(l_j)$ assess the accuracy of the measure provided by the landmark $l_j \in \mathcal{L}_r$ by giving higher importance to $e(l_j)$ values that are computed from homogeneous gray level distributions because the corresponding value $v(l_j)$ better represents the overall level of X-ray absorption in the region inside the patch P_j . The entropy is used to define the weights $\omega(l_j)$. The entropy of a random variable X is defined as follows:

$$H(X) = \sum_{x \in S_X} -P(x) \log(P(x))$$

where S_X is the set of elements of the discrete distribution that are taken into account to compute the entropy and, here, corresponds to the gray levels at the pixels $x_i \in P_j$ where P_j is the patch centered at the landmark l_j . Then, the weights are defined by using the following exponential function:

$$\omega(l_j) = \exp\left(-\frac{H(l_j)}{\alpha_H}\right) \quad (4)$$

where $H(l_j)$ is the entropy computed at the landmark $l_j \in \mathcal{L}_r$ and α_H is a constant smoothing parameter set to 2 in our experiments.

This approach presents two advantages with respect to the ROI detection methods presented in the standard IEC 62494-1 (2008). First, it takes into account the presence of different anatomical structures in the field of view, which is important for clinical studies that cover large fields of view, such as the

spine. Second, by exploiting redundancy of local estimates, the landmark-based approach is robust to detection errors.

In (Irrera et al., 2015) landmarks were positioned manually and errors were simulated in order to show the interest of the method. In this paper, we propose a technique to detect these landmarks, and, hence, to automatically compute exposure index values. The considered anatomical ROIs are the following ones: head (\mathcal{L}_1), thoracic spine (\mathcal{L}_2), lungs (\mathcal{L}_3), lumbar spine (\mathcal{L}_4), pelvis (\mathcal{L}_5), femurs (\mathcal{L}_6), knees (\mathcal{L}_7) and tibiae (\mathcal{L}_8).

2.3. Salient points

The landmarks to be detected occupy only a small portion of the whole pixel space. Therefore, an exhaustive search for landmark occurrences is unnecessary and time consuming. To reduce the search space, the detection and recognition task is initialized from a set of sparse points that we call salient points.

From a very abstract point of view, the information in X-ray images can be associated with changes in the intensity levels as related to tissues of different densities. Therefore, when we look at an EOS image, the attention is essentially captured by these variations of signal. For example, in the abdomen, we first look at the lumbar vertebrae and then at the surrounding tissues because the bones absorb more signal and have irregular shapes. As a consequence, the relevant information can be associated with differences of intensity levels. Moreover, the information is scale-dependent and, hence, the search for salient points should be led at a scale that contains the features of interest. For example, a coarse scale could be sufficient to capture the envelop of the spine, whereas an analysis at finer scales would be required if the goal is to capture the internal structures of the vertebrae.

In practice, given the observed image \mathbf{u} of size $R \times C$, non-overlapping sub-images \mathbf{k}_i of size $S \times C$ are considered, where $S \ll R$, extracted from each vertical positions $y_i \in \{1, 2, \dots, R\}$. A sub-image \mathbf{k}_i is then projected to a 1D signal \hat{k}_i where each element corresponds to the column-wise average. Afterwards, the signal is smoothed by using a linear average kernel of size S . This sequence of operations encodes the information related to X-ray absorption as a function of the sub-window height S that gives the information of scale. It is worth noting that if $S = 1$, a signal \hat{k}_i simply coincides with the horizontal profile of the image \mathbf{u} at line y_i .

The salient information is expressed by means of local maxima or minima of the function \hat{k}_i that are associated with peaks of signal at the detector and of absorption, respectively. The salient information of each sub-image \mathbf{k}_i will be then linked with a set of point \mathcal{P}_i , where the x-coordinates x_j are the positions of salient points on \hat{k}_i and $y_j = 0.5(y_i + y_{i+1}), \forall j$, i.e. the y-coordinates are simply the centers of the sub-images \mathbf{k}_i .

Figures 1a and 1b show two examples of functions \hat{k}_i that capture the same region but at different scales, i.e. $S = 128$ and $S = 64$, respectively. In this example we consider the region nearby the vertebra T1 (extracted from the image in Figure 2b), the relevant information is hence located at the center

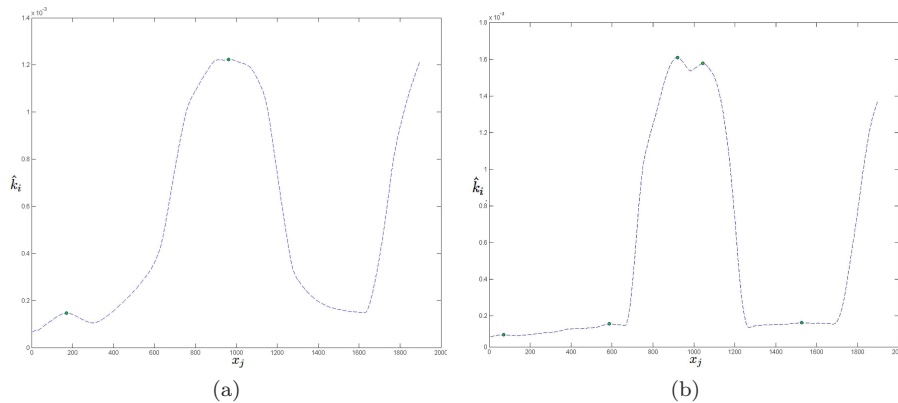


Figure 1: A signal \hat{k}_i that encodes region nearby the vertebra T1 on a frontal view image at different scales: (a) $S = 128$ (see salient points $\mathcal{P}_{(h,128)}$ in Figure 2b) ; (b) $S = 64$ (see salient points $\mathcal{P}_{(h,64)}$ in Figure 2c). The scale has an impact on the amount of captured information and outliers.

(i.e. $x_j \simeq 1000$) while some outliers associated with the shoulders appear in the leftmost and rightmost sides of the signals \hat{k}_i . We consider as outliers the salient points that are associated with regions that are not covered by any of the anatomical clusters \mathcal{L}_r taken into account. By setting $S = 128$, only one value is detected at the center of the cervical spine, whereas, by using $S = 64$, two points inside the vertebra are found. Therefore, coarser scales have the advantage of being sparser and, hence, present less outliers, but may fail in capturing some relevant details. On the other hand, the finer the scale the more precise the analysis is, but the number of outliers also increases. As a consequence, the choice of the value S should depend on the estimated minimal degree of precision that is required to well describe a structure of interest. In the given example, $S = 128$ is the preferable setting because there is no need to associate two points with the vertebra, for applications to EI estimation. The scale values associated with the anatomical structures to capture have been empirically defined in the method that is described in the following sections.

Figure 2 shows some examples of salient points on the frontal and lateral view acquisitions for a given exam at different scales S . The red stars indicate salient points $p_j \in \mathcal{P}_{(l,s)}$ associated with peaks of signal at the detector (index l) and computed at the scale S (index s), whereas the green circles are the salient points $p_j \in \mathcal{P}_{(h,s)}$ that correspond to strong absorption regions (index h) at the scale S (index s). Note that the salient points are computed over the raw image \mathbf{u} , but in Figure 2 we display the points on the output image, where the gray levels are inverted with respect to \mathbf{u} , i.e. high and low intensity values correspond to strong and low absorption, respectively.

On frontal acquisitions, the points $p_j \in \mathcal{P}_{(l,s)}$ are mainly located in the lungs with a limited amount of outliers in the region that surrounds the groin. Note that a quite good approximation of the anatomical cluster \mathcal{L}_3 is obtained at

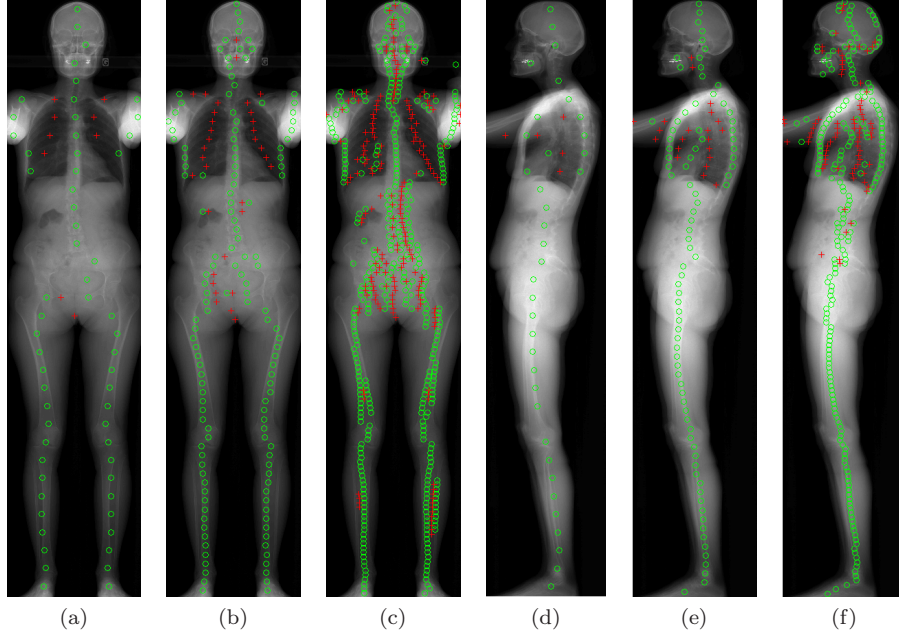


Figure 2: Sets of salient points $\mathcal{P}_{(l,s)}$ (red) and $\mathcal{P}_{(h,s)}$ (green) on a frontal view acquisition at (a) $S = 256$; (b) $S = 128$; (c) $S = 64$ and on the corresponding lateral view at (d) $S = 256$; (e) $S = 128$; (f) $S = 64$. See colors on the on-line version.

scale $S = 128$. On the other hand, $\mathcal{P}_{(l,256)}$ does not give a precise enough description and many points $p_j \in \mathcal{P}_{(l,64)}$ are located in the spine and the pelvis, but peaks of signal in these regions are not robust as they mainly depend on the morphotype.

The points $p_j \in \mathcal{P}_{(h,s)}$ cover the whole field of view as peaks of absorption are mostly due to the presence of bone tissues. The majority of the points $p_j \in \mathcal{P}_{(h,s)}$ well fits the spine and the bones in the legs both with $S = 256$ and $S = 128$, which implies that the clusters \mathcal{L}_r with $r \neq 3, 5$ (i.e. all the ROIs except the lungs and the pelvis) could be well estimated by regularizing these salient points. On the other hand, it is more complicated to get a proper initialization in the pelvic region because in strong absorption regions the signal tends to be homogeneous. In this case, it is preferable to rely on an over-complete representation such as the one obtained by setting S to 64 .

On lateral acquisitions, the salient points $p_j \in \mathcal{P}_{(l,s)}$ do not provide useful information. They are indeed randomly concentrated in the lung and cardiac regions that we do not include in our analysis of lateral view images. On the contrary, all the landmarks $l_j \in \mathcal{L}_r$, with $r \neq 2$ (i.e. except for the thoracic spine) are well initialized by using the salient points $p_j \in \mathcal{P}_{(h,128)}$, that offer the best compromise between completeness of the representation and reduction of outliers. Furthermore, note that the salient points $p_j \in \mathcal{P}_{(h,128)}$ in the chest well

respond to the rib cage, which helps defining the cluster of landmarks associated with this region.

These remarks are valid for all the patients in our database (see Section 2.7) and, furthermore, no noticeable change of trend has been observed on patients with high body mass index or on acquisitions at very low amount of dose.

2.4. Control points

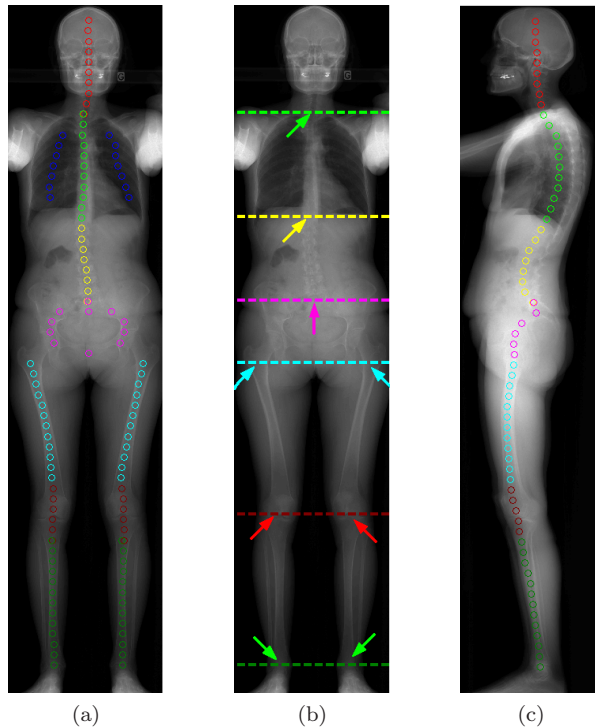


Figure 3: According to the manually annotated landmarks on the frontal view acquisition (a), 6 horizontal lines (b) can be identified to initialize the method (from top to bottom): T1 (green); T12 (yellow); L5 (magenta); proximal femur (cyan); knees (red); ankles (green). This initialization is also valid on the corresponding lateral view acquisition (c). See colors on the on-line version.

Once the salient points have been detected, they need to be classified according to the anatomical regions of interest in order to compute the exposure index values. This can be achieved by roughly separating the field of view into disjoint sub-windows according to the considered anatomical ROIs \mathcal{A}_r . The coarse identification of these areas helps to both narrow the search space for the landmarks $l_j \in \mathcal{L}_r$ and to verify if an estimated solution is coherent with the spatial relations between clusters. The manually annotated landmarks on frontal view acquisitions (Figure 3a) indirectly provide the ideal division into sub-windows (Figure 3b). For example, the thorax is identified as the region between the upper and lower landmarks of the cluster \mathcal{L}_2 . In order to initialize the

detection on full body images, it is then necessary to identify 6 horizontal lines that pass through the following control points c_i : the vertebra T1, the vertebra T12¹, the vertebra L5, the proximal femur, the distal femur and the ankle. The problem is initialized by simply estimating the y-coordinates y_j of the control points $c_j \in \mathcal{C}$, i.e. the projection \mathcal{C}_y on y-axis where \mathcal{C} is a set composed by the 6 control points. Since the image is subdivided along the vertical direction, which is the common axis between the frontal and lateral acquisitions, the same initialization can be used on the two views.

Figure 3b shows that the control points can be described by simple spatial relations such as relative position (e.g. the vertebra T1 is above the vertebra T12) and distance. The relative positions of the control points is fixed, but, in order to encode the distances, a subset of 9 full-body EOS exams is used for learning. These samples well represent the variability of a bigger database composed by 82 patients, and the 9 patients have different ages, genders and morphotypes. In this section we describe an approach that allows defining the distances with respect to two already detected control points. In the following section we are going to describe how to label a subset of salient points to the control points.

Table 2: Rough estimation of $y_j \in \mathcal{C}_y$ for the proximal femur (y_4), L5 (y_3), T12 (y_2) and the distal femur (y_5) by considering proportions ϱ_j from neighbors (see Equation 5). The values for T1 (y_1) and the ankle (y_6) are assumed to be known.

	$y_4 \mid \{y_1, y_6\}$	$y_3 \mid \{y_1, y_4\}$	$y_2 \mid \{y_1, y_3\}$	$y_5 \mid \{y_4, y_6\}$
$\bar{\varrho}_j$	0.43	0.72	0.55	0.51
$\sigma(\varrho_j)$	0.02	0.02	0.02	0.01
$\min(\varrho_j)$	0.41	0.70	0.53	0.49
$\max(\varrho_j)$	0.45	0.75	0.58	0.53
α_j	0.39	0.67	0.51	0.47
β_j	0.47	0.77	0.60	0.55
$\ \beta_j - \alpha_j\ $	11.39 cm	4.96 cm	3.49 cm	6.43 cm

The position of a target control point will be intuitively more stable to changes in the data by exploiting the fact that the human body parts are approximately proportionate to each others. Formally, a horizontal line at y_j can be described according to the proportion between the distances from two other horizontal lines of known positions. For example, let us assume that y_1 and y_6 have already been identified, and that y_4 has to be found. The following proportionality factor can then be associated with y_4 :

$$\varrho_4 = \frac{y_4 - y_1}{y_6 - y_1}. \quad (5)$$

¹We do not strictly need to associate a control point with an exact vertebra but rather to distinguish the chest from the abdomen. Therefore, the control point T12 may correspond to, for example, T10, without any consequence on the results.

Table 2 reports the average ($\bar{\kappa}_j$), standard deviation ($\sigma(\kappa_j)$), minimal ($\min(\kappa_j)$) and maximal ($\max(\kappa_j)$) relative distances between two given horizontal lines $y_j \in \mathcal{C}_y$, that are assumed to be already detected (e.g. T1 and ankle), and a third one, computed over the training database. Then, each control point c_j is associated with an interval $[\alpha_j, \beta_j]$ that constrains the position of c_j on a new test image according to the manually annotated data of the training set. In particular, the width of the ranges is computed as follows:

$$\|\beta_4 - \alpha_4\| = \varrho_4 \max_i (y_6^i - y_1^i) \quad (6)$$

where y_j^i is the y-coordinate of the control point c_j on the patient i of the training set. The extension of the interval refers then to the worst case scenario, i.e. the maximal distance between the two referential horizontal lines over the training set. This implies that $\alpha_j = \min(\varrho_j) - \sigma(\varrho_j)$ and $\beta_j = \max(\varrho_j) + \sigma(\varrho_j)$. These values indicate that by exploiting the proportionality principle the initialization of the problem is more precise, i.e. a local analysis to define the exact position of the control points will be limited to a small region compared with the whole pixel space. This helps increasing then both efficiency and robustness. Quantitatively, by relying on the proportions ϱ_j , the same control points will be searched in areas that cover 11.33 cm, 4.96 cm, 6.43 cm and 3.49 cm. The problem can then be initialized by using the learned proportionality factors given the upper and lower control points, i.e. the vertebra T1 and the ankle. Section 2.5 describes how to detect these two points.

Some spatial relations in the horizontal direction can also be used to facilitate the localization of some landmarks $l_j \in \mathcal{L}_r$. For example, the clusters in the left and right lungs constrain the position of the landmarks $l_j \in \mathcal{L}_2$ (thoracic spine). Similarly, the clusters \mathcal{L}_4 (lumbar spine) and \mathcal{L}_6 (femurs) help defining the landmarks in the pelvic region.

The initialization from the frontal view is straightly used on the corresponding lateral view. However, it would be also useful to get the rough position of the most prominent points of the kyphotic and lordotic curves in the spine (see the clusters \mathcal{L}_2 and \mathcal{L}_4 in the image in Figure 3c). By computing the proportionality factor in this case, the interval associated with the most prominent point in the lumbar spine is equal to $[-0.01, 0.71]$ which indicates that the initialization is highly uncertain. This is due to the posture of the patient that influences the position of the most prominent point. Similarly, the changes in the posture do not allow for a significant initialization on the horizontal direction and, thus, the lateral acquisition can only rely on information extracted from the frontal view to initialize the positions of the control points.

2.5. Control point detection

The control points are sequentially detected among the salient points by taking into account the spatial relations (Section 2.4). The procedure starts by identifying two main control points that are the vertebra T1 and the ankle. These points are chosen among the others not only because they are located at the upper and bottom extrema limits of the image, but mostly because they are

the easiest ones to detect in a completely unsupervised way. It is worth noting that this logic requires the input images to contain the whole body of a patient.

We present here only the principles for the detection of control points. Please refer to (Irrera, 2015) for a complete description of the mathematical steps that allow extracting the control points.

The frontal view image is used in order to localize the control point for T1. This relies on two principal information: the position of patient’s central vertical axis and the salient points associated with peaks of signal. The central vertical axis does not necessarily pass through the center of the image because the image may have been collimated². However, by relying on the values of the collimation limits, the position of the patient’s central vertical axis in the image space is straightly deduced. This allows retaining the salient points that roughly correspond to the spine. In order to select the salient point associated with the control point T1, the second information is exploited, i.e. the salient points at peaks of signal roughly correspond with the cluster of landmarks $l_j \in \mathcal{L}_3$. Finally, T1 is detected by exploiting the following spatial relation: T1 is immediately over the lungs, and between the left and right lung.

The ankle landmark is the next to be detected, and, as for T1, the frontal view is used too. The junction between the lower bound of the tibia and the ankle cavity causes a strong horizontal discontinuity which can be well captured by a gradient-based descriptor. In practice, we compute the following gradient based measure, associated with each salient point $p_j \in \mathcal{P}_{(h,128)}$:

$$g_j = \sum_{(x_i, y_i) \in P_j} \frac{\|\Delta_h(x_i, y_i)\|}{|P_j|} \quad (7)$$

where $\Delta_h(x_i, y_i)$ is the horizontal component of the Sobel operator at the pixel of coordinates (x_i, y_i) that belongs to a circular patch P_j of radius equal to $S/2$ centered at the point p_j . The size of the patch is chosen such that the measure associated with a given point is independent from the ones computed at nearby points. Considering the width of the human leg, by setting $S = 128$ only one landmark on each S horizontal line is obtained and, then, the measure g_j computed at the landmark l_j can be expressed as a function of its y-coordinate y_j . In other words, the measures g_j computed at the anatomical landmarks $l_j \in (\mathcal{L}_6 \cup \mathcal{L}_7 \cup \mathcal{L}_8)$ that belong to one leg can be represented as a mono-dimensional function of the y-coordinate y_j . The control point for the ankle is located at the position of the last peak of the gradient-based measures.

The same measure is used to define the control point for distal femur because this structure also presents a strong horizontal discontinuity. The search area is in this case constrained by the position of the proximal femur and the ankle according to the relative distances presented in Section 2.4. This requires detecting the control point at proximal femur.

²The collimation in this case consists in narrowing the width of the field of view. This allows limiting the irradiated area to regions of interest for the undergoing clinical study. The collimation limits can be extracted from DICOM metadata fields.

The proximal femur is the most difficult control point in the leg to detect because of the proximity to another anatomical region, the pelvis, and of the X-ray absorption due to tissue superposition that is much higher compared with that in the knees and ankles. We consider a set of points $\mathcal{T}_0 \subset \mathcal{P}_{(h,128)}$ composed, for each line y_f , $f = 1, \dots, |\mathcal{P}_y|$, by the leftmost and rightmost points of x-coordinates x_l and x_r , respectively. The extrema are retained because they most probably correspond to the legs. Formally, x_l and x_r for a given y_f are defined as follows:

$$x_l = \min\{x_j \mid p_j \in \mathcal{P}_{y_f}\} \quad x_r = \max\{x_j \mid p_j \in \mathcal{P}_{y_f}\} \quad (8)$$

where \mathcal{P}_{y_f} is the set of points $p_j \in \mathcal{P}$ that lie on the line defined by $y = y_f$. The y-coordinate of the control point c_4 is then defined as follows:

$$y_4 = \arg \max_{y_j \in \mathcal{P}_y} (d(t(x_l, y_j), t(x_r, y_j))), \quad (9)$$

which means that the proximal femur is on the line y_j where the points $t(x_l, y_j)$ and $t(x_r, y_j)$ are the most distant from each other, which is coherent with the assumption made for the manual annotations. Nevertheless, the distribution of the points may change depending on the morphotype or the presence of metallic objects. Therefore, the position of the proximal femur is estimated multiple times, i.e. one time for each considered scale S . According to multiple tests conducted on part of the database, we remarked that either the estimates are similar regardless the chosen scale or two estimates are close and one is wrong. As a consequence, the final position is computed as the median of the estimates at the three considered scales, which is a basic way of exploiting information at multiple scales, but good enough according to the tests and application of interest.

The control points in the spine, except for the vertebra T1, are the most difficult ones to precisely define on frontal view acquisitions and, hence, are reserved for the last step. The difficulty comes from the complexity in distinguishing between a vertebra and another one, and consistent features are not trivial to define considering the rotational issues, the low contrast to noise ratio and the tissue superposition. According to our tests, the definition of consistent features that allow capturing the vertebra L5 (c_3), despite changes in morphotype and entrance dose, remains an open question, but, for the EI application, this does not cause a significant problem for two reasons. First, according to the values in Table 2, the interval where to look for c_3 is relatively narrow, i.e. 4.96 cm in the worst case scenario. Secondly, the EI values associated with the lumbar spine and pelvis are similar and, hence, potential misplacement of c_3 would only slightly affect the exposure indicators (Irrera et al., 2015). Finally, the position of L5 is just roughly inferred from the vertebra T1 and the proximal femur by using a proportionality factor equal to 0.72, i.e. at the center of the interval given by the model.

As for the control point T12 (c_2), it can be much more easily detected on the lateral view acquisition. The position of the control point for the vertebra

T12 (c_2) can be refined by analyzing the landmarks $l_j \in (\mathcal{L}_2 \cup \mathcal{L}_4)$. The passage from \mathcal{L}_2 to \mathcal{L}_4 is marked by a strong reduction of the local EI estimate due to a higher density in the lumbar region than in the chest. Therefore, the local EI estimates $e(l_j)$ are computed at the landmarks $l_j \in (\mathcal{L}_2 \cup \mathcal{L}_4)$, and the following function is defined from these values:

$$\delta(y_j) = e(l_{j+1}) - e(l_j), \quad j = 1, \dots, (|\mathcal{L}_2 \cup \mathcal{L}_4| - 1) \quad (10)$$

where $\delta(\cdot)$ depends on y_j because there is one point for each line in the set $\mathcal{L}_2 \cup \mathcal{L}_4$. Since each element of the function quantifies how much consecutive local EI values change, the y-coordinate of the control point for T12 is estimated as follows:

$$y = \arg \min_{y_j \in Y_m} (\delta(y_j)) \quad (11)$$

where Y_m is the set of values that constrain the position of $y = y_2$ according to the model of global relations (see Table 2).

The position of the control points allow classifying the salient points into landmark clusters \mathcal{L}_r that will be used to compute exposure indices as explained in Section 2.2. The position of the points $l_j \in \mathcal{L}_r$ can be refined by relying on models of the corresponding anatomical regions. It is out of the scope of this paper to detail how landmark clusters are refined for each of the eight anatomical regions on both views, and the concepts presented above represent the core of the hierarchical detection and recognition algorithm. Details can be found in (Irrera, 2015).

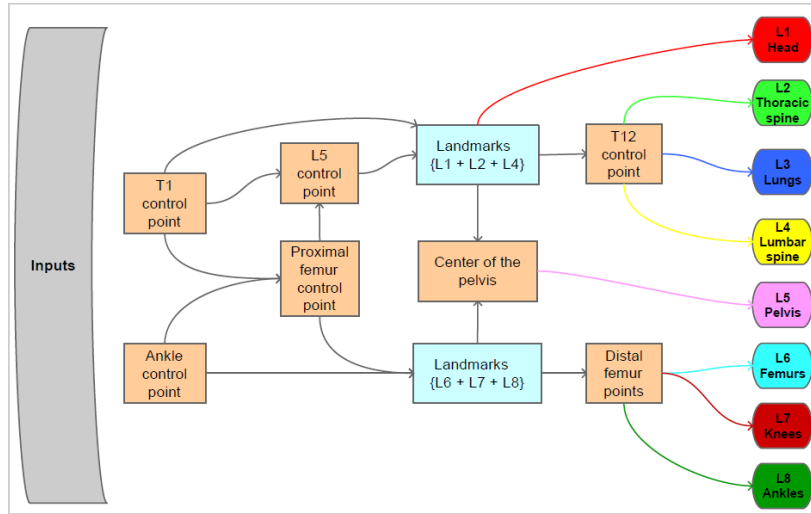


Figure 4: Hierarchical detection on frontal view acquisition images.

Figure 4 summarizes the described steps to hierarchically detect the anatomical landmarks on frontal view acquisition images. The inputs are the anatomical

constraints between the structures and the salient points positions. The landmarks in the upper part of the body and in the legs are first detected and then the definition of the control points for T12 and distal femur allows separating them into \mathcal{L}_1 - \mathcal{L}_2 - \mathcal{L}_3 - \mathcal{L}_4 and \mathcal{L}_6 - \mathcal{L}_7 - \mathcal{L}_8 , respectively. The cluster \mathcal{L}_5 at the last step because it is the most difficult region for the definition of consistent salient points. It is worth noting that some operations can be executed in parallel because they concern different anatomical regions. The detection on lateral view acquisition images consists in sequentially narrowing the solution space as on frontal view acquisitions. Moreover, control points are not re-detected (see Section 2.4). Only T1 is refined as previously explained.

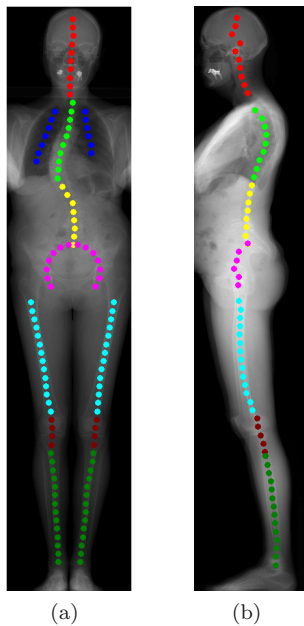


Figure 5: Example of automatic landmark detection: (a) frontal view; (b) lateral view.

Figure 5 provides an example of landmark detection that shows how the method is able to project the whole pixel space to a set of labeled landmarks while being robust to bone distortion such as thoracic scoliosis.

2.6. Landmark cluster from the patient's envelop

The proposed method that combines salient points with control points (Sections 2.3 and 2.5) is compared to an easier detection technique. The latter consists in a landmark-based formulation of the classical method that considers ROIs at the center of the image. A transposition of this idea into the landmark-based formulation leads to place the landmarks $b_j \in \mathcal{B}_r$ at the center of the mask. Given the patient's envelop (Figure 6a), the Euclidean distance transform (Figure 6b) is computed. The values of this map increase with the distance from the borders of the mask and are equal to zero outside of it. On each line,

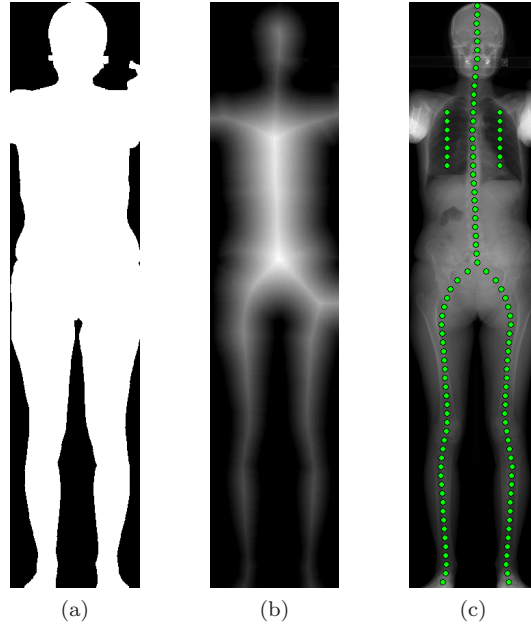


Figure 6: Landmarks defined from the patient’s envelop: (a) patient’s envelop; (b) Euclidean distance map; (c) landmarks $b_j \in \mathcal{B}_r$.

the position of the maximum of the Euclidean distance transform profile on a line y_j corresponds to the x-coordinate of a landmark b_j . Two points per line, i.e. two local maxima, may also be considered if symmetrical to the central vertical axis of the mask (e.g. the legs on frontal view acquisitions). Figure 6c shows an example of the resulting landmarks. Note that the landmarks in the lungs ($b_j \in \mathcal{B}_3$) are located proportionally to the positions of those in the thoracic spine ($b_j \in \mathcal{B}_2$) and of the borders of the patient. Formally, if p_f is the nearest point on the border at the left (right) of the landmarks $b_f \in \mathcal{B}_2$ on a line $y = y_f$, the landmark $b_j \in \mathcal{B}_3$ in the left (right) lung will be placed at the coordinates $0.5(p_f + b_f)$. Finally, the landmarks are assigned to the specific anatomical structures depending on the positions of the control points. This simpler approach is compared to the other one in Section 3 in order to quantify the meaningfulness of the initialization by means of salient points.

2.7. Database

We evaluate the method explained in the previous sections according to the DI values computed as follows:

$$DI = 10 \log_{10} \left(\frac{EI}{EI_{gt}} \right) \quad (12)$$

This is very similar to Equation 2, which compares the exposure index value to a previously recorded target. The difference is that Equation 12 assesses the

accuracy of the automatically estimated EI values with respect to the ground truth values, i.e. those computed from manually annotated landmarks. The precision in terms of positions of the landmarks is not discussed here because, as justified in Section 2.2, the proposed landmark-based approach allows for a certain degree of localization error.

The database is composed by 82 full-body anonymous exams that are categorized into the following subsets:

- \mathcal{D}_a : 24 diagnostic exams of adult patients aged between 54 and 82.
- \mathcal{D}_b : 29 diagnostic exams of young patients aged between 8 and 16.
- \mathcal{D}_c : 29 preview exams of the patients in the subset \mathcal{D}_b .

This database is representative of the variety of types of patients that undergo an imaging system dedicated to osteoarticular disorders, such as EOS. For example, 13 of the 29 young patients in our database are affected by idiopathic scoliosis. Moreover, many patients in our database contain metallic objects. Indeed, 30 of the 53 patients have at least one of the following metallic objects: gonad shields, spine rods (post spine surgery exams) and hip prosthesis (total hip arthroplasty exams). It is important to take into account metallic objects as, if not neglected during the computation, they cause an underestimation of exposure index values. Finally, the computed exposure index values should well represent the morphotype of patient. We then consider both normal and overweight patients. Since the DICOM field *weight* was not filled in the exams that compose our database, we cannot provide precise Body Mass Index (BMI) values. Nevertheless, we use a quantitative approach to establish the amount of over-weight patients in our database ($\text{BMI} > 30$). Given that all the patients in the same category of age (adolescents or adults) are scanned with default acquisition parameters, we can use the ground truth exposure index values in order to identify overweight patients. We take the median of the exposure index values associated with the pelvis of the frontal view as reference values, one for each group of age. We label as overweight the patients for which the deviation index, computed as in Equation 2, is lower than -1. Following this approach, we count 4 on 29 and 8 on 24 overweight patients among the adolescents and adults, respectively.

Sections 3.1 and 3.2 analyze the results corresponding to frontal and lateral view acquisitions, respectively. The descriptions are conducted by referring to the anatomical ROIs \mathcal{A}_r . The accuracy of the estimated EI values is evaluated according to the root mean square error (RMSE) defined as follows:

$$RMSE_{\mathcal{A}_r} = \sqrt{\frac{\sum_{z=1}^{N_p} (DI_{r,z})^2}{N_p}} \quad (13)$$

where N_p is the number of patients in a given database and $DI_{r,z}$ is the DI value obtained for the patient z and the region \mathcal{A}_r according to Equation 12.

The ideal result is a DI equal to zero, but a margin of error is accepted and different degrees of errors are, thus, considered: $|DI| \in [0, 0.25)$, $|DI| \in [0.25, 0.5)$, $|DI| \in [0.5, 0.75)$, $|DI| \in [0.75, 1)$ and $|DI| \in [1, +\infty)$ mean, respectively, negligible, low, medium, high and extreme errors. We then measure the percentages ε_0 , ε_1 , ε_2 , ε_3 and ε_4 that indicate the amount of samples in the database whose estimation errors are negligible, low, medium, high and severe, respectively, when compared with the ground truth values. We will use $\varepsilon_0 + \varepsilon_1$ as a measure of the amount of acceptable errors. Furthermore, specific observations are made on the three subsets \mathcal{D}_a , \mathcal{D}_b and \mathcal{D}_c in order to understand if the quality of the estimation may change depending on the age of the patient or the amount of radiation exposure.

Sections 3.3 and 3.4 deal with cases that, according to the state of the art, make it difficult to obtain consistent EI measures, i.e. particular conditions of the patient and presence of metallic objects, respectively.

The proposed method is very efficient as the analysis is limited to some salient points that are sparsely distributed on the pixel space: by running our MATLAB code on a conventional machine (Intel Core 2.20 GHz, 4 GB RAM), it takes about 6.50 s and 3.50 s at most to detect and recognize the landmarks on both frontal and lateral full-body acquisitions views, respectively.

3. Results and Discussion

3.1. Quantitative evaluation of frontal view acquisition

Tables 3 and 4 report the results corresponding to all the frontal view acquisitions in the database and to the three subsets \mathcal{D}_a , \mathcal{D}_b and \mathcal{D}_c , respectively.

Table 3 shows that the estimation of the sets of EI values on frontal view acquisitions with the proposed approach is very robust and overcomes some drawbacks of the method where the landmarks are detected depending on the position with respect to the borders of the patient envelop.

The improvement is particularly significant in the thoracic spine (\mathcal{A}_2), lung (\mathcal{A}_3) and femur (\mathcal{A}_6) regions of interest. In particular, as we will better explain in Section 3.3, the proposed detection technique allows respecting the curve of the spine when the patient is side-bending or in presence of a scoliosis, which is not the case if the landmarks are simply placed on the central vertical axis. This allows reducing of 50% the RMSE when the set of points \mathcal{L}_2 is used instead of \mathcal{B}_2 . Moreover, while by using the cluster \mathcal{B}_2 the amount of acceptable errors ($\varepsilon_0 + \varepsilon_1$) gets to 52.4%, the proposed approach allows reaching 93.9%. The same conclusion results from the analysis of error measurements associated with the lung area. This is due to the fact that a distorted thoracic spine directly affects the shape of the lungs. In the femur region the reduction of estimation error is particularly significant due to the robustness of the proposed method to changes in patient morphotype. This assessment is justified in Section 3.3.

The two methods perform similarly in the other tested regions. In the lumbar spine (\mathcal{A}_4) and pelvis (\mathcal{A}_5) the detection errors poorly affect the EI estimation because of the high density (Irrera et al., 2015). In the head (\mathcal{A}_1), knee (\mathcal{A}_7)

Table 3: Evaluation on frontal view acquisitions in the whole database: comparison of the automatic EI estimates associated with the clusters \mathcal{B}_r and \mathcal{L}_r , $r = 1, \dots, 8$, in terms of RMSE and percentages of negligible (ε_0), low (ε_1), medium (ε_2), high (ε_3) and severe (ε_4) estimation errors.

Frontal		RMSE	ε_0	ε_1	ε_2	ε_3	ε_4
\mathcal{A}_1	\mathcal{B}_1	0.32	65.9%	25.6%	4.9%	1.2%	2.4%
	\mathcal{L}_1	0.29	67.1%	26.6%	4.9%	1.2%	1.2%
\mathcal{A}_2	\mathcal{B}_2	0.61	20.7%	31.7%	28.0%	13.4%	6.1%
	\mathcal{L}_2	0.30	61.0%	32.9%	4.9%	1.2%	0.0%
\mathcal{A}_3	\mathcal{B}_3	0.56	45.1%	30.5%	15.9%	3.7%	4.9%
	\mathcal{L}_3	0.27	85.4%	9.8%	2.4%	1.2%	1.2%
\mathcal{A}_4	\mathcal{B}_4	0.11	96.3%	3.7%	0.0%	0.0%	0.0%
	\mathcal{L}_4	0.09	97.6%	2.4%	0.0%	0.0%	0.0%
\mathcal{A}_5	\mathcal{B}_5	0.35	73.2%	18.3%	4.9%	1.2%	2.4%
	\mathcal{L}_5	0.25	86.6%	8.5%	2.4%	1.2%	1.2%
\mathcal{A}_6	\mathcal{B}_6	0.68	39.0%	36.6%	8.5%	3.7%	12.2%
	\mathcal{L}_6	0.19	86.6%	13.4%	0.0%	0.0%	0.0%
\mathcal{A}_7	\mathcal{B}_7	0.16	65.9%	25.6%	4.9%	1.2%	2.4%
	\mathcal{L}_7	0.12	85.4%	14.6%	0.0%	0.0%	0.0%
\mathcal{A}_8	\mathcal{B}_8	0.09	100.0%	0.0%	0.0%	0.0%	0.0%
	\mathcal{L}_8	0.06	100.0%	0.0%	0.0%	0.0%	0.0%

and tibiae (\mathcal{A}_8) the anatomical structures are mostly at the center of the envelop as shown by the results.

Table 4 reveals that the estimation accuracy does not depend on the patient age (subset \mathcal{D}_a versus \mathcal{D}_b) nor on the amount of dose (subset \mathcal{D}_b versus \mathcal{D}_c). Nevertheless, in the head region, the accuracy is higher on \mathcal{D}_a (RMSE equal to 0.11) than on \mathcal{D}_b (RMSE equal to 0.33). This can be related to the position of the head for the patients in the databases \mathcal{D}_b and \mathcal{D}_c , which is bent forward by hence leading to a partial superposition of the head with the neck. As it will become more evident from the analysis on the lateral view acquisitions (Section 3.2), the partial occlusion of an anatomical structure by another one can cause some problems for the estimation of the EI values. In any case, a RMSE error equal to 0.33 is still a low error and, thus, this practically is a marginal issue.

3.2. Quantitative evaluation of lateral view acquisition

Tables 5 and 6 report the results associated with all the lateral view acquisitions in the database and the three subsets \mathcal{D}_a , \mathcal{D}_b and \mathcal{D}_c , respectively.

The results on lateral view acquisitions indicate that a level of accuracy similar to that registered on frontal view acquisitions can only be obtained in the regions \mathcal{A}_4 and \mathcal{A}_5 .

Table 6 shows that in the region \mathcal{A}_1 the results are biased by bad performance for the subsets \mathcal{D}_b and \mathcal{D}_c . Indeed, on the subset \mathcal{D}_a the method performs well,

Table 4: Evaluation over the frontal view acquisitions in the subsets \mathcal{D}_a , \mathcal{D}_b and \mathcal{D}_c : comparison of the automatic EI estimates associated with the clusters \mathcal{B}_r and \mathcal{L}_r , $r = 1, \dots, 8$ in terms of RMSE.

Frontal		RMSE		
		\mathcal{D}_a	\mathcal{D}_b	\mathcal{D}_c
\mathcal{A}_1	\mathcal{B}_1	0.11	0.36	0.37
	\mathcal{L}_1	0.11	0.33	0.35
\mathcal{A}_2	\mathcal{B}_2	0.69	0.62	0.51
	\mathcal{L}_2	0.28	0.32	0.28
\mathcal{A}_3	\mathcal{B}_3	0.33	0.60	0.67
	\mathcal{L}_3	0.28	0.19	0.33
\mathcal{A}_4	\mathcal{B}_4	0.11	0.10	0.13
	\mathcal{L}_4	0.06	0.09	0.10
\mathcal{A}_5	\mathcal{B}_5	0.17	0.46	0.32
	\mathcal{L}_5	0.15	0.16	0.36
\mathcal{A}_6	\mathcal{B}_6	0.59	0.78	0.65
	\mathcal{L}_6	0.19	0.18	0.19
\mathcal{A}_7	\mathcal{B}_7	0.12	0.10	0.22
	\mathcal{L}_7	0.09	0.06	0.17
\mathcal{A}_8	\mathcal{B}_8	0.09	0.07	0.10
	\mathcal{L}_8	0.07	0.04	0.08

with RMSE equal to 0.12, and high errors are only found in the subsets \mathcal{D}_b and \mathcal{D}_c . This is not due to the change of morphotype but rather to the fact that, in these subsets, the scans start from half the head and, hence, the low amount of landmarks make the estimation unstable. Moreover, in some exams, the hands of the patients are placed on the neck and interfere then with the structures of interest. On the contrary, the images in the subset \mathcal{D}_a are acquired from over the top of the head and there is no occlusion with the hands, which are the conditions that allow for a robust estimation of the EI values. Therefore, the ROI \mathcal{A}_1 should be taken into account only if it is important for the exam and, hence, acquired with the patient positioned in a suitable way.

The evaluation on the whole database indicates that the EI value in the ROI \mathcal{A}_2 is not easy to estimate automatically. Indeed, the landmarks $b_j \in \mathcal{B}_2$ produce a RMSE equal to 0.89, and 23.2% and 41.5% of samples whose estimation errors are negligible and acceptable, respectively. The landmarks $l_j \in \mathcal{L}_2$ improve the performances given the RMSE equal to 0.64, and 41.5% and 58.6% of samples whose estimation errors are negligible and acceptable, respectively, but is not completely satisfying yet. Nevertheless, the results substantially change according to the subset here too. In details, the landmarks $b_j \in \mathcal{B}_2$ give a low quality estimation regardless the morphotype and the amount of dose, whereas the RMSE associated with the landmarks $l_j \in \mathcal{L}_2$ for the exams in \mathcal{D}_a is very low (0.22). Figures 7a and 7b show that the difference between the results asso-

Table 5: Evaluation on lateral view acquisitions in the whole database, of the accuracy of the automatic EI estimates associated with the clusters \mathcal{B}_r and \mathcal{L}_r , $r = 1, \dots, 8$, in terms of RMSE and percentages of cases whose estimation errors are negligible (ε_0), low (ε_1), medium (ε_2), high (ε_3) and severe (ε_4).

Lateral		RMSE	ε_0	ε_1	ε_2	ε_3	ε_4
\mathcal{A}_1	\mathcal{B}_1	0.68	62.7%	17.6%	5.9%	3.9%	9.8%
	\mathcal{L}_1	0.57	66.7%	17.6%	2.0%	5.9%	7.8%
\mathcal{A}_2	\mathcal{B}_2	0.89	23.2%	18.3%	19.5%	13.4%	25.6%
	\mathcal{L}_2	0.64	41.5%	17.1%	13.4%	14.6%	13.4%
\mathcal{A}_4	\mathcal{B}_4	0.16	90.2%	7.3%	2.4%	0.0%	0.0%
	\mathcal{L}_4	0.16	86.6%	13.4%	0.0%	0.0%	0.0%
\mathcal{A}_5	\mathcal{B}_5	0.28	72.0%	22.0%	2.4%	2.4%	1.2%
	\mathcal{L}_5	0.15	95.1%	3.7%	0.0%	1.2%	0.0%
\mathcal{A}_6	\mathcal{B}_6	0.44	50.0%	28.0%	14.6%	4.9%	2.4%
	\mathcal{L}_6	0.42	62.2%	22.0%	7.3%	4.9%	3.7%
\mathcal{A}_7	\mathcal{B}_7	0.75	25.6%	18.3%	17.1%	19.5%	19.5%
	\mathcal{L}_7	0.58	36.6%	22.0%	20.7%	11.0%	9.8%
\mathcal{A}_8	\mathcal{B}_8	0.47	56.1%	23.2%	13.4%	1.2%	6.1%
	\mathcal{L}_8	0.45	58.5%	20.7%	13.4%	1.2%	6.1%

ciated with \mathcal{D}_a and \mathcal{D}_b (\mathcal{D}_c) is not due to the morphotype but it rather depends on the position of the arms: while for the patient in Figure 7a the arms cover the heart but do not occlude the vertebrae, in Figure 7b the superior vertebrae (T1-T4) are partially occluded by the shoulders. The thoracic spine is always a region of interest and, thus, the estimation of the associated EI value should not depend on the position of the arms. A possible solution may consist in attributing higher importance to the local estimates $e(l_j)$ associated with landmarks l_j placed in the inferior part of the thoracic spine.

Finally, the results are not good for any anatomical region in the leg. This is mainly due to the fact that the leg nearer the detector is partially occluded by the one nearer the X-ray source. Given these conditions, it is not sure that an EI value estimation in this region from lateral view is actually meaningful given the superposition of the legs in the image. The performance may be improved by taking into account patient pose in exposure index estimation. This would however require automatically detecting patient pose, which is out of the scope of this work.

3.3. Robustness to patient conditions

We address in this section the challenges brought by the presence of deformed anatomical structures (e.g. scoliosis) and by patient obesity. The measures computed from the landmarks $l_j \in \mathcal{L}_r$ and $b_j \in \mathcal{B}_r$ are compared.

Figure 8a displays the cluster \mathcal{L}_2 (green points) and \mathcal{B}_2 (red points) on the anatomical region \mathcal{A}_2 of a patient affected by scoliosis. The landmarks $b_j \in \mathcal{B}_2$

Table 6: Evaluation on lateral view acquisitions in the subsets \mathcal{D}_a , \mathcal{D}_b and \mathcal{D}_c , of the accuracy of the automatic EI estimates associated with the clusters \mathcal{B}_r and \mathcal{L}_r , $r = 1, \dots, 8$ in terms of RMSE.

Lateral		RMSE		
		\mathcal{D}_a	\mathcal{D}_b	\mathcal{D}_c
\mathcal{A}_1	\mathcal{B}_1	0.11	1.00	0.84
	\mathcal{L}_1	0.12	0.83	0.70
\mathcal{A}_2	\mathcal{B}_2	1.16	0.72	0.79
	\mathcal{L}_2	0.22	0.70	0.80
\mathcal{A}_4	\mathcal{B}_4	0.17	0.17	0.15
	\mathcal{L}_4	0.11	0.19	0.16
\mathcal{A}_5	\mathcal{B}_5	0.21	0.34	0.27
	\mathcal{L}_5	0.14	0.14	0.17
\mathcal{A}_6	\mathcal{B}_6	0.42	0.42	0.47
	\mathcal{L}_6	0.35	0.40	0.48
\mathcal{A}_7	\mathcal{B}_7	0.50	0.91	0.75
	\mathcal{L}_7	0.47	0.68	0.55
\mathcal{A}_8	\mathcal{B}_8	0.58	0.53	0.26
	\mathcal{L}_8	0.59	0.46	0.25

generate a DI index value equal to 1.36 because the central axis does not correspond to the thoracic spine but rather to the surrounding structures. On the contrary, the landmarks $l_j \in \mathcal{L}_2$ well respect the curvature of the spine and, hence, the EI value is robustly estimated as shown by the DI value that is equal to -0.16 . In general, when a patient is affected by a scoliosis in the thoracic region, the measure computed at $b_j \in \mathcal{B}_2$ generates severe errors, whereas the proposed method overcomes this issue. This analysis justifies the results presented in Table 3 for the anatomical ROI \mathcal{A}_2 . On the other hand, if the scoliosis affects the lumbar spine the EI value is not considerably biased.

Figure 8b shows the femoral region of a patient with high body mass index. Since the legs are joined, the envelop of the patient does not well describe the underlying anatomy and the landmarks $b_j \in \mathcal{B}_6$ start separating into two branches only at half of the femur. As a consequence, the DI value rises up to 1.69. The proposed method does not depend on the morphology of the patient and the EI value is perfectly estimated (DI equal to 0.01). These considerations are coherent with the results presented in Table 3 corresponding to the anatomical ROI \mathcal{A}_6 .

3.4. Robustness to metallic objects

The metallic objects should be neglected in the estimation of the EI values because they are not associated with anatomical information and significantly bias the measure given the strong X-ray absorption. Therefore, they should be removed from the whole pixel space Ω , just like the regions outside the body of

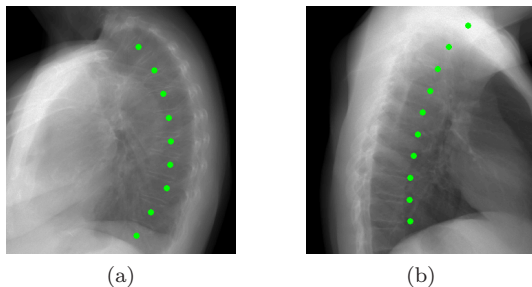


Figure 7: The detection of the landmarks $l_j \in \mathcal{L}_2$ on lateral view images depends on patient's pose: (a) example from the subset \mathcal{D}_a ; (b) example from the subset \mathcal{D}_b .

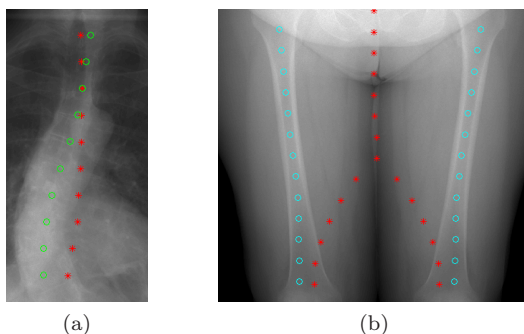


Figure 8: Particular cases due to patient conditions. In Figure (a) the landmarks $b_j \in \mathcal{B}_2$ (red points) do not follow the curvature of the spine due to scoliosis, whereas the landmarks $l_j \in \mathcal{L}_2$ (green points) allow overcoming this drawback. In Figure (b) as the legs are adjacent, the landmarks $b_j \in \mathcal{B}_2$ (red points) do not capture the femurs, whereas the landmarks $l_j \in \mathcal{L}_2$ (blue points) well match the anatomical structures.

the patient where the X-rays are not attenuated. However, considering that the EI value $e(l_j)$ is computed at a landmark l_j , the non-anatomical objects do not necessarily need to be segmented, but the weight ω_j should be set to 0 if the local estimate $e(l_j)$ corresponds to a metallic object. A local estimate $e(l_j)$ is an outlier only if more than 50% of the circular patch P_j centered at l_j is covered by a metallic object. In other words, the presence of metallic objects in the field of view does not necessarily imply that some of the local estimates $e(l_j)$ need to be discarded and, even if it is the case, the number of outliers is most likely low. According to these observations, a simple threshold-based method is used to reject the contributions from potential metallic objects.

Formally, the minima of the local estimates $\min(e(l_j))$, for the landmarks $l_j \in \mathcal{L}_r$, $r \in \{2, 4, 5, 6\}$, are computed. The median of these values is retained as the reference value that corresponds to anatomical exposure in strong absorption regions, which is why only the anatomical ROIs \mathcal{A}_r , $r \in \{2, 4, 5, 6\}$, are taken into account. The median value is computed because we assume that the outliers are at most present in two of the anatomical ROIs \mathcal{A}_r , $r \in \{2, 4, 5, 6\}$, as biased local estimates due to metallic objects are rare. The threshold τ_m is fixed at 80%

of the reference value, i.e. a $DI_{\tilde{c}}-1$ with respect to the reference is considered as acceptable. The landmarks $l_j \in \mathcal{L}_r$, $r = 1, \dots, 8$, such as $e(l_j) < \tau_m$ are considered as outliers and, hence, their weights $\omega(l_j)$ are fixed to zero.

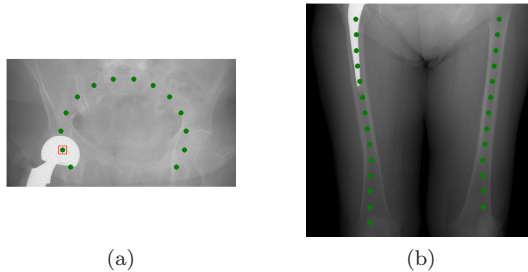


Figure 9: Robustness to the presence of the metallic objects. In Figure (a), among the landmarks $l_j \in \mathcal{L}_5$ (green points), the point that corresponds to the femoral prosthetic head is labeled as outlier (red square) and then rejected. In Figure (b), none of the landmarks $l_j \in \mathcal{L}_5$ is detected as outlier, but this does not affect the EI measure (see text for details).

Figure 9 shows two sub-images extracted from the same image with a femoral prosthesis. In details, Figure 9a displays the landmarks $l_j \in \mathcal{L}_5$ on the pelvic region and the one placed in the middle of the prosthetic femoral head (red square) is rejected. The DI value with respect to the ground truth is then equal to -0.07 . Nevertheless, even if the outliers are taken into account, the measure is not significantly affected as the DI value becomes -0.12 , which shows how the landmark-based method is practically not biased by the presence of outliers as long as their number is limited. Figure 9b shows that the presence of metallic objects does not necessarily imply presence of outliers. Indeed, the EI values associated with the left and right legs are practically identical, i.e. 41.24 and 40.27, even if all the landmarks are taken into account. It is worth noting, however, that in this example only 2 of the 5 landmarks placed in the prosthesis give wrong EI estimates because, as for the others, the metallic object is not represented by the median values of the local gray level distributions. As a consequence, the number of outliers is too low to influence the EI estimated at the femoral region. On the other hand, the proposed technique to reject outliers could be useful when their number in a cluster \mathcal{L}_r is relatively high with respect to the number of landmarks $|\mathcal{L}_r|$. This would be the case, for example, of the estimation of the EI value associated with the anatomical region \mathcal{A}_7 in presence of a knee prosthesis.

Finally, these results well summarize that our method is able to overcome the issues due to the presence of metallic objects, which is one of the advantages of the landmark-based formulation.

4. Conclusion

The exposure index can be computed by relying on clusters of landmarks associated with anatomical regions (Irrera et al., 2015). According to this formu-

lation, the image quality estimation becomes a matter of detection and recognition of anatomical landmarks. We have proposed an unsupervised method that addresses this task by combining the global information on the size and the positions of the anatomical structures on the one hand, with local analysis on the other hand. The search of the position of the landmarks is sparsely conducted by visiting salient points that correspond to peaks of absorption or signal at the detector. The results indicate that the EI values computed at the automatically detected landmarks correspond very well to those associated with manual annotations on frontal view acquisitions. On lateral view acquisitions, the estimates in the thoracic spine may be biased by the superposition with the shoulder in the superior part of the region. Therefore, in a next version of the method, higher importance should be given to the contributions from the lower thoracic vertebrae. An unstable estimation may also occur in the legs due to their superposition in a lateral image. We should then check with the users whether the estimation of the image quality in these regions is meaningful despite the superposition or not. Note that the validation has been conducted on a heterogeneous database composed by full-body images of patients of different morphotypes, acquired at different radiation exposure amounts, and including post-surgery and pathological cases. The proposed method has shown to be very robust to all these different cases, except for the aforementioned structures acquired from the lateral view for which the estimation needs to be improved.

The proposed detection method relies on the hypothesis that the anatomical structures present in the image are known. This information can be normally retrieved from the DICOM fields. For example, all the tested images cover the full body. This allows initializing the method by detecting a set of control points from which the search for the solution can be sequentially narrowed and then locally refined. Nevertheless, in unusual cases, such as when the DICOM field anatomical study is missing or wrong, the proposed method would not work. A possible solution to make the method robust to these non-nominal cases may consist in detecting the control points with a classification- or regression-based supervised approach. The control points could be then located despite the missing prior knowledge and the landmarks detected as proposed in our method. For example, Gauriau et al. (2015) have proposed an approach that combines cascade of supervised regressors to statistical shapes, which may be adapted to the problem of control points detection. While this work is applied to CT images, some recent works (Aubert et al., 2016; Ebrahimi et al., 2016) have also investigated supervised structure detection on X-ray images acquired with EOS system.

The detection is sequentially conducted in a predefined order. However, as proposed by Fouquier et al. (2012), it may be preferable to conduct the search according to information extracted from the image to process. Moreover, while an already detected cluster of landmarks can constraint the sets corresponding to other anatomical structures, our method does not offer a backward check on the already detected landmarks. This option could however prevent potential errors to propagate. The algorithm could also be improved by adaptive setting of the parameters, such as the scale at which the salient points are computed.

Ethical statement

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Disclosure statement

The authors have no conflicts of interest to disclose.

References

- Aubert, B., Vazquez, C., Cresson, T., Parent, S., Guise, J. D., April 2016. Automatic spine and pelvis detection in frontal x-rays using deep neural networks for patch displacement learning. In: 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI). pp. 1426–1429.
- Ebrahimi, S., Angelini, E., Gajny, L., Skalli, W., April 2016. Lumbar spine posterior corner detection in x-rays using haar-based features. In: 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI). pp. 180–183.
- Fouquier, G., Atif, J., Bloch, I., 2012. Sequential model-based segmentation and recognition of image structures driven by visual features and spatial relations. *Computer Vision and Image Understanding* 116 (1), 146 – 165.
- Gauriau, R., Cuingnet, R., Lesage, D., Bloch, I., 2015. Multi-organ localization with cascaded global-to-local regression and shape prior. *Medical Image Analysis* 23, 70–83.
- IEC 62494-1, 2008. Medical electrical equipment - Exposure index of digital X-ray imaging systems - Part 1: Definitions and requirements for general radiography.
- Illés, T., Somoskeöy, S., 2012. The EOSTM imaging system and its uses in daily orthopaedic practice. *International orthopaedics* 36 (7), 1325–1331.
- Irrera, P., June 2015. Low dose X-ray image processing. Ph.D. thesis, Telecom ParisTech, 2015-ENST-0031.
- Irrera, P., Bloch, I., Delplanque, M., 2015. A Landmark-Based Approach for Robust Estimation of Exposure Index Values in Digital Radiography. In: 18th International Conference on Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, Munich, Germany. Vol. 2. pp. 627–634.
- Mothiram, U., Brennan, P. C., Robinson, J., Lewis, S., Moran, B., 2013. Retrospective evaluation of exposure index (EI) values from plain radiographs reveals important considerations for quality improvement. *Journal of Medical Radiation Sciences* 60 (4), 115–122.

- Seibert, A., Richard, M., 2011. The standardized exposure index for digital radiography: an opportunity for optimization of radiation dose to the pediatric population. *Pediatric Radiology* 41 (5), 573–581.
- Shepard, S. J., Wang, J., Flynn, M., Gingold, E., Goldman, L., Krugh, K., Leong, D. L., Mah, E., Ogden, K., Peck, D., Samei, E., Wang, J., Willis, C. E., 2009. An exposure indicator for digital radiography: Aapm task group 116 (executive summary). *Medical Physics* 36 (7), 2898–2914.