A flexible patch based approach for combined denoising and contrast enhancement of digital X-ray images

Paolo Irrera, Isabelle Bloch, Maurice Delplanque

A flexible patch based approach for combined denoising and contrast enhancement of digital X-ray images

Paolo Irrera[a,b], Isabelle Bloch[a], Maurice Delplanque[b]

[a] *Institut Mines Telecom, Telecom ParisTech, CNRS LTCI, Paris, France*
[b] *EOS imaging, Paris, France*

Abstract

Denoising and contrast enhancement play key roles in optimizing the trade-off between image quality and X-ray dose. However, these tasks present multiple challenges raised by noise level, low visibility of fine anatomical structures, heterogeneous conditions due to different exposure parameters, and patient characteristics. This work proposes a new method to address these challenges. We first introduce a patch-based filter adapted to the properties of the noise corrupting X-ray images. The filtered images are then used as *oracles* to define non parametric noise containment maps that, when applied in a multiscale contrast enhancement framework, allow optimizing the trade-off between improvement of the visibility of anatomical structures and noise reduction. A significant amount of tests on both phantoms and clinical images has shown that the proposed method is better suited than others for visual inspection for diagnosis, even when compared to an algorithm used to process low dose images in clinical routine.

Keywords:
Low Dose X-ray imaging, Non Local Means, Noise containment, Contrast enhancement.

1. Introduction

Medical imaging based on X-rays is the main source of exposure to artificial radiation (Smith-Blindman et al., 2012), which, as highlighted in some recent studies, entails negative secondary effects for the patient health. Shuryak et al. (2010) have pointed out that all age groups run the risk of developing radio-induced cancers, and Ronckers et al. (2008) have shown that the patients affected by scoliosis have a higher probability of developing a cancer because they undergo more X-ray exams.

The way clinical image quality is perceived depends on how raw X-ray image quality is improved through the different steps of the image processing chain. In particular, the noise level on the outcome indirectly indicates if an image has been acquired in good conditions (Shepard et al., 2009). Therefore, it is important to define an algorithm robust to changes in the amount of signal at

the detector, i.e. stable to changes in the amount of skin entrance dose and to inter-patient variability. The dose could be for instance reduced and still achieve the same diagnostic goal for a given study. Alternatively, the same amount of input signal could be used despite an increase in patient's size.
In this paper we consider X-ray images acquired with a low dose, and as a typical example we process images acquired with the stereo-radiographic imaging system EOS (Wybier and Bossard, 2013), that allows simultaneously acquiring full body frontal and lateral images of a patient in weight-bearing position. The density of the tissues significantly changes according to different anatomical regions (see Figure 1a), which considerably affects signal values and noise levels. This intra-patient variability is another important factor that needs to be taken into account to optimally process the acquired data.
X-ray images present both components of noise and signal that cannot be clearly distinguished because the local contrast at the acquisition is low. Therefore, the image quality enhancement requires to both reduce the noise and increase the visibility of fine anatomical details. In some works (Sakata and Ogawa, 2009; Loza et al., 2014) the authors propose to restore the input image by using wavelet-based approaches and, then, to enhance it. Nevertheless, this type of approach can lead to a loss of spatial resolution that is not acceptable in clinical routine. The use of more advanced denoising filters that represent an image in a patch space could overcome the aforementioned issue. The patches are sub-images that capture local characteristics and, hence, the noise can be attenuated while preserving edges and texture. These filters have been also used in medical applications (Cerciello et al., 2012) showing promising results. Nevertheless, as pointed out by Lebrun et al. (2012) in a survey on this denoising technique, very fine texture, e.g. fine bone texture, may be flattened out. The use of highly performing noise reduction filters is then only a partial solution in radiography applications: the resulting images may risk to be perceived as unnatural by the clinicians due to a lack of information in the bone structures. The noise containment maps (Stahl et al., 1999) are an alternative approach that consists in defining, in a multiscale framework, whether a coefficient can be fully enhanced as associated with signal information, or not. Stahl et al. (1999) use density and activity measures to define the noise containment maps. However, the definition of these maps depends on global user-defined parameters, which is the main drawback of this method. Indeed, the capacity of containing the noise relies on these parameters. Finally, given the heterogeneity of digital X-ray data, parametric noise containment maps are sub-optimal.

Contributions. This work proposes a general framework for joint denoising and enhancement of X-ray full body images that addresses the aforementioned drawbacks of existing methods. The main contribution of this paper consists in showing how the output of a denoising filter applied to an X-ray image well approximates the ground truth image and can be then exploited to increase the visibility of anatomical details while containing the noise. We propose an extension of the Non Local Means (NLM) filter (Buades et al., 2005) that can be easily adapted to our noise model and is called X-ray Non Local Means (XNLM) filter. The result of this filter is exploited to define non-parametric

noise containment maps that are used in a multiscale framework to robustly limit the presence of the noise in the final solution. Note that the independence from manually set parameters is a crucial element, which makes the method robust to the heterogeneity of the data to be processed. While the tests presented in this paper have specifically been carried out on EOS images, the approach could likely be applied to any image exhibiting similar characteristics.

The main scope of the validation consists in quantifying how much our work can help clinicians in their diagnosis. The provided outcomes are meant to be suitable for diagnosis without any further manual user interaction. Manual windowing can optimize contrast and brightness, however these adjustments can cause noise to raise and further slow down diagnosis process. For these reasons, the results that optimize the trade-off between contrast and amount of noise should be automatically obtained.
The quality of clinical images is not easy to objectively assess and this aspect is studied in this work. Therefore, two new validation approaches are proposed. In particular, classical measures of contrast (average local variance (Chang and Wu, 1998) and contrast improvement index (Laine et al., 1995)) are revisited by associating them with anatomically meaningful regions. Moreover, the image quality evaluation is completed with clinical assessments according to the feedbacks of a radiologist.

Some aspects introduced in this paper are partially related with two of our former works. The first one (Irrera et al., 2013) has allowed showing that the denoising filter parameters need to be tuned for different anatomical regions in order to efficiently restore a full body X-ray image. However, while in this previous work the parameters were set by manually adjusting the shape of a curve, this is not the case for the XNLM filter here introduced, that now exploits automatic estimates of the noise levels. In the second paper (Irrera et al., 2014) the denoising process has been combined with a multiscale decomposition with the aim of reducing the spatial resolution loss on EOS images used in follow-up examinations. This is very different from the noise containment approach proposed here, which has the advantage of not being limited to a specific clinical case as it is free from critical parameter setting.

The paper is organized as follows. Section 2 explains how to estimate a curve that gives the noise standard deviation as a function of the signal and, then, how to exploit it to formulate the XNLM filter. Section 3 outlines how to estimate the noise containment maps and to increase the visibility of anatomical details. Section 4 presents some results on both phantom and clinical images. Section 5 concludes the paper, and summarizes the achieved objectives and perspectives.

2. X-ray Non Local Means filter

2.1. Overview of the Non Local Means filter

The Non Local Means (NLM) filter estimates the intensity value of a pixel x_i by means of a weighted average that depends on the similarity between patches (Buades et al., 2005). The result of the filter is good as long as the information

in the image is redundant, i.e. similar structures can be found at different spots. This hypothesis is valid for X-ray images. Given the input image I, the gray level of the filtered image $\hat{I}$ at a pixel x_i is formally defined as follows (Buades et al., 2005):

$$\hat{I}(x_i) = \frac{\sum_{j=1}^{|\Omega|} \varsigma(i,j) I(x_j)}{\sum_{j=1}^{|\Omega|} \varsigma(i,j)} \tag{1}$$

where $\varsigma(i,j)$ is the weight associated with $I(x_j)$ in the estimation of $\hat{I}(x_i)$. The domain Ω represents the search space for similar patches. In practice, this is a window of half-size w (i.e. $|\Omega| = (2w+1)^2$) centered at pixel x_i. The weight $\varsigma(i,j)$ quantifies the distance in the patch space between spatially near pixels. Formally, let P_i and P_j denote patches of half-size p (i.e. $|P| = (2p+1)^2$) centered, respectively, at x_i and x_j. The corresponding weight is defined as:

$$\varsigma(i,j) = \exp\left(-\frac{d(P_i,P_j)}{h^2}\right) \tag{2}$$

where $d(P_i,P_j)$ is the distance between the patches P_i and P_j and h is a smoothing parameter controlling the decay of the exponential function. In the original formulation (Buades et al., 2005), d was a Gaussian-weighted Euclidean distance. However, in more recent works (Darbon et al., 2008; Coupé et al., 2008), the convolution with the Gaussian is avoided and the normalized Euclidean distance is used without noticeable effect on the final quality:

$$d(P_i,P_j) = \frac{\|I(P_i) - I(P_j)\|_2}{2|P_i|} \tag{3}$$

where $I(P_i)$ are the intensities of I in the patch centered at pixel x_i and, thus, $\|.\|_2$ is the Euclidean norm in $\mathbb{R}^{|P|}$.
The NLM filter in its naive formulation is computationally expensive. The complexity is of order $\mathcal{O}(N|\Omega||P_i|)$ where N is the number of pixels of I. Darbon et al. (2008) have addressed this issue by making the computation independent of the patch size. Basically, the weights are computed from the discrete integration between the difference of the image I and its shifted versions by using integral images. Therefore, the computational load is reduced to $\mathcal{O}(N|\Omega||4|)$.

We propose an extension of the classical NLM filter that takes into account the noise model which is described in the following section.

2.2. Noise model estimation

An appropriate model of noise affecting the input image is a key element in the denoising procedure. In medical X-ray images, the noise is signal-dependent. Specifically, the noise at the detector is a combination of quantum and electronic noise that can be described, respectively, by Poisson and Gaussian distributions. However, as shown by Damet et al. (2014), in EOS images the quantum noise contribution is predominant and, even at very low doses, the electronic noise is negligible. Therefore, the noise model can be approximated by a Poisson

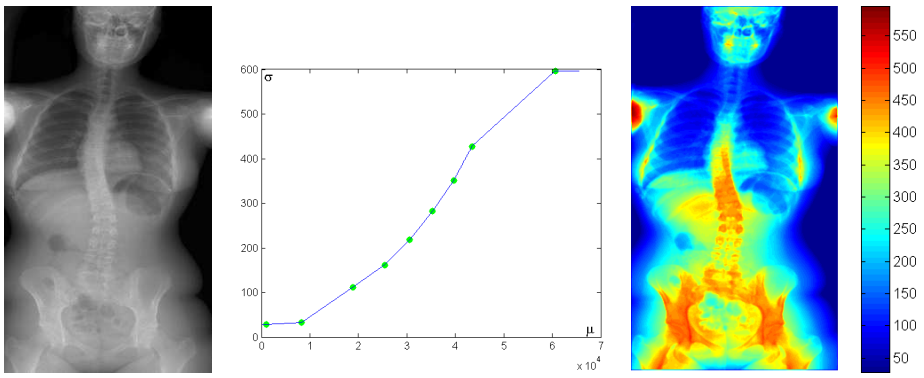
distribution. Nevertheless, such a model is rarely observed in X-ray images (Hensel et al., 2006), as the raw image is mapped into a logarithmic domain to compensate for the exponential attenuation of X-rays passing through the body. This operation gives sense to the gray levels that are then linearly dependent on the matter thickness and density. Consequently, the relation between signal and noise standard deviation does not follow anymore a root square function. Hensel et al. (2006) have verified that the noise in this case can be modeled with an additive zero mean normal-distribution η with signal-dependent standard deviation $\sigma_\eta(S)$. Then, if I and S are the observed data and the ground truth signal, respectively, then $I = S + \eta$ and the probability density function of the variable η is:

$$f(\eta; S) = \frac{1}{\sigma_\eta(S)\sqrt{2\pi}} \exp\left(-\frac{\eta^2}{2\sigma_\eta(S)^2}\right) \tag{4}$$

Then, in order to properly characterize the problem, $\sigma_\eta(S)$ needs to be estimated. Colom and Buades (2013) have recently proposed a noise estimation method that does not require the assumption of homoscedastic (i.e. signal-independent) noise and consists of a block-based formulation of the percentile method (Ponomarenko et al., 2007), which is very robust for rating the noise standard deviation from a single image (Lebrun et al., 2012; Colom and Buades, 2013). The first step of the percentile method consists in computing a high-pass filtered version of I, that we denote H_I. This allows for the elimination of the deterministic component due to the signal and, so, to find a predominant component of noise in many small windows of half-size b. Some examples of suitable high-pass filters are the discrete Laplacian, the Discrete Cosine Transform (DCT) and the Wavelet coefficients. Then, a local variance image V is computed from H_I using windows of half-size b. The key idea of the percentile is that signal components, e.g. edges, are present only in the rightmost part of the histogram of V. This allows for a biased value of the noise standard deviation to be obtained by computing a low percentile of the histogram. The bias is corrected by applying a linear correction factor that depends on the percentile value, the window half-size b and the choice of the high-pass filter. We use a 10% percentile, $b = 7$ and the first detail level of the multiscale decomposition (see Section 3.1.1) as high pass filter in all our tests.

In order to extend the method beyond the homoscedastic case, the input image is divided according to the intensity levels into n not overlapped and not necessarily connected regions $i = 1, 2, \ldots, n$ of equal size. It is assumed that in each region i, the noise follows a Gaussian distribution and is signal independent. The classical percentile method is then used to determine the noise standard deviation σ_i in a region i. A signal level μ_i is also computed as the median of the average gray levels of the blocks in the region i. Finally the n points $\{(\mu_1, \sigma_1), (\mu_2, \sigma_2), \ldots, (\mu_n, \sigma_n)\}$ (see Figure 1b) are linearly interpolated to obtain the noise curve.

The assumptions made by Colom and Buades (2013) are respected by the formula in Equation 4 and, hence, this method can be used to estimate a noise map for describing the noise strength in digital X-ray images. For example,



(a) (b) (c)

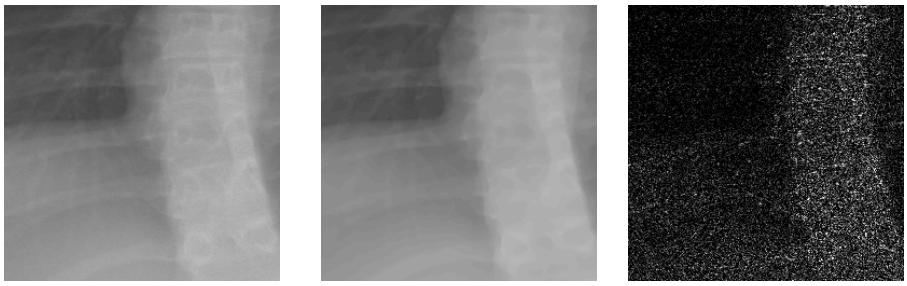
Figure 1: Noise estimation using the percentile method: (a) Input image I. (b) Interpolated noise curve from the n points (μ_i, σ_i). (c) Resulting map of noise standard deviation $(\sigma(S))$.

given an EOS image (Figure 1a)[1] the percentile method allows estimating a noise curve (Figure 1b) that is used to obtain the corresponding noise map (Figure 1c). It is worth noting that the noise level increases with the amount of absorption, which is expected after the logarithmic conversion (Hensel et al., 2006) and has been observed on all the processed images. Figure 1c shows that the estimated σ values change significantly and coherently according to the body tissue thickness. For example, the noise level in a low density level region such as the lungs is 4 times lower than that in the pelvis.

2.3. X-rays Non Local Means filter

The noise model so characterized is used to propose a patch-based filter adapted to X-ray image denoising. The noise is signal-dependent and the same parameters cannot be used to denoise the whole image. However, in a small region the σ_η value is approximately stable because the density of the tissues therein is practically constant, i.e. the map σ_η changes smoothly. It should be noted that the smoothing parameter h is proportional to the noise standard deviation σ (Buades et al., 2005), i.e. $h^2 = 2k\sigma^2$. The proportional factor k is meant to adjust the automatic estimation of σ. Coupé et al. (2008) have pointed out that the optimal value of k depends on the strength of the noise affecting the data, e.g. $k = 0.5$ for low noise levels and $k = 1$ for medium and high noise levels. However, this is mainly due to the tendency of over-estimating σ at low noise levels (Colom and Buades, 2013). Considering that the percentile method suffers less from this issue (Lebrun et al., 2012) and that, in the worst case, the only region affected by over-smoothing would be the background where the

[1]The image I is obtained after the following sequence of steps: system calibration, logarithmic mapping and LUT inversion. The latter operation associates high gray levels to high absorption regions.



(a) (b) (c)

Figure 2: Zoom of XNLM filter output: (a) input image I; (b) denoised image $\hat{I}$; (c) absolute difference $\|I - \hat{I}\|$.

X-rays are not absorbed, k has been fixed to 1. Formally, h is integrated in the patch distance definition in order to take into account the dependency on the X-ray absorption:

$$d_X(P_i, P_j) = \frac{\|I(P_i) - I(P_j)\|_2}{2{\sigma_i}^2|P|} \tag{5}$$

Note that in this equation the noise level is the one associated with the window Ω centered at the pixel x_i. This implies that $\sigma_i \simeq \sigma_j \ \forall j \in \Omega$, which is coherent with the assumption on the smooth transition of the tissue density values. The weights are then computed as follows:

$$\varsigma_X(i,j) = \exp\left(-d_X(P_i, P_j)\right) \tag{6}$$

In conclusion, considering the signal-dependent nature of the noise (Section 2.2), the NLM filter can be formulated to denoise X-ray images. The proposed filter is called XNLM, which stands for X-rays Non Local Means filter. It is worth noting that the automatic definition of σ avoids manual parameter tuning and, hence, it is more robust to changes in X-ray acquisition settings (see Section 4.1) and patient morphotypes. However, the patch and window size still need to be defined by the user. In this work patch and window half-sizes are, respectively, fixed to $p = 2$ and $w = 7$. Besides, note that the similarity between patches is quantified with an Euclidean distance which allows implementing the XNLM filter by using integral images (Darbon et al., 2008).

Figure 2c shows the absolute difference between a region of the noisy image I (Figure 2a) and the corresponding denoised image $\hat{I}$ (Figure 2b), which indicates that the noise is mainly removed while the structures are well preserved. Moreover, the filter adapts to different levels of absorption. For example the noise reduction is stronger in the spine region than in the pulmonary one. In the following sections we explain how the image $\hat{I}$ can be used to define noise containment maps that are then applied in a multiscale contrast enhancement approach.

3. Local noise containment maps

3.1. Overview of multiscale contrast enhancement methods

3.1.1. Multiscale analysis

The multiscale (MS) decomposition allows modifying separately fine and coarse details (Li et al., 2005) and is often used to enhance the contrast of digital X-ray images (Stahl et al., 1999; Dippel et al., 2002; Fan and Han, 2011). The Laplacian Pyramid (Burt and Adeldon, 1983) (LP) is a classical MS decomposition technique. An image is encoded in k band-pass images $\{D_0, D_1, \ldots, D_{k-1}\}$ and its low frequency residual L_k by recursive filtering:

$$\begin{cases} L_t = \downarrow (g(L_{t-1})) \\ \\ D_t = L_{t-1} - \uparrow (L_t) \end{cases} \quad t \geqslant 1 \tag{7}$$

where $L_0 = I$, $\downarrow$ and $\uparrow$ are, respectively, the down-sampling and up-sampling operators, and $g(.)$ is the smoothing filter, e.g. a binomial filter. The smooth transition between scales causes redundancy. This allows avoiding strong halo artifacts that appear when decomposing with an orthogonal basis (Dippel et al., 2002). The sub-sampling, used to double the spatial scale at each iteration, is the reason of aliasing artifacts due to non linear operations in the reconstruction (Li et al., 2005). Therefore, in this work an undecimated version of the LP, also known as Isotropic Undecimated Wavelet (IUWT) (Starck et al., 2007), is used.

3.1.2. Multiscale synthesis

An image decomposed using IUWT is rebuilt by adding all the detail images D_t to the low frequency residual L_k. Consequently, the enhanced image Y is obtained by adding the boosted detail images Λ_t:

$$Y = L_k + \sum_{t=0}^{k-1} \Lambda_t \tag{8}$$

where $\Lambda_t = f_t(D_t)$ and $f_t(.)$ is a non-linear boosting remapping function. Dippel et al. (2002) give an example of such a function and study how its parameters affect the X-ray image quality. Basically, these functions confine the enhancement in low activity regions, which are defined according to a contrast measure (Mantiuk et al., 2006). This allows limiting the halo artifacts, which are a known drawback of linear MS decomposition techniques (Li et al., 2005). Finally, the parameters of the boosting functions are set to well balance coarse and fine details (Dippel et al., 2002).

3.1.3. The noise

Since the detail images are obtained by recursively smoothing, the noise is progressively reduced at coarser scales. Nevertheless, the noise equally contaminates all frequencies. Therefore, by fully enhancing the band-pass images D_t, the noise is increased.

The noise containment maps address this problem by considering that the amount of noise in the output image Y can be limited by not enhancing the detail coefficients that are corrupted by noise. Formally, a noise containment map is a function $w_t(x_i) \in [0, 1]$. The 0 and 1 values are associated with a noise free and full noise coefficient, respectively. Therefore, the boosted output level at scale t is the result of the coefficient-wise weighted sum between the input band-pass image D_t and its fully enhanced version $f(D_t)$:

$$\Lambda_t = w_t D_t + (1 - w_t) f_t(D_t) \tag{9}$$

The advantage of noise containment maps compared to filters is that no information is lost, as nothing from D_t is erased. For this reason, this approach is often preferred to process X-ray images. However, it is not trivial to define correct w_t maps. While an overestimation of noise would cause a reduction of visibility of anatomical structures, an underestimation would overshoot the detail coefficients by hence giving an unnatural effect. Stahl et al. (1999) define w_t by using local density (gray level) and activity (e.g. local standard deviation) maps. Fan and Han (2011) implicitly estimate noise containment maps by comparing Peli's contrast measure (Peli, 1990) before and after boosting. The main drawback of these methods is that the results depend on some global user-defined parameters, e.g. the level of activity associated with noise and structures. This is particularly tricky in digital radiography for two reasons. First, there is a high intra-patient variability because the noise level changes significantly in a full body X-ray image (Section 2.2) and therefore globally defined parameters are sub-optimal. Secondly, there is a high inter-patient variability in age, size and acquisition conditions. Consequently, different parameter settings should be defined to address this large heterogeneity, which is very difficult to validate.

3.2. Definition of local noise containment maps

The band-pass images D_t^I obtained by decomposing an X-ray image I with the IUWT contain both signal and noise components. Figure 3b and 3f show the detail levels D_0^I in lung (Figure 3a) and lumbar spine (Figure 3e) regions, respectively. These examples confirm that the anatomical structures are merged with noise. On the other hand, the corresponding detail levels $D_t^{\hat{I}}$ (Figures 3c and 3g) are noise-free. This does not imply that all the relevant information is preserved, but it is easier to identify which coefficients should be fully enhanced. In other words, the restored image $\hat{I}$ is used as an *oracle* for the definition of the noise containment maps. The examples also exhibit that the difference between the original and restored band-pass images depends on the properties of the anatomical regions. This is due to the fact that the XNLM filter does not uniformly restore an image I but rather depending on the estimated σ values.

These observations are quantitatively studied by associating an energy measure with the band-pass images. Given a detail image D_t, only a few coefficients are significantly not null, i.e. the signal is sparse in the MS decomposition. This has led many researchers to use statistical models to describe detail coefficients. Further information on this aspect can be found in the work by Loza et al.

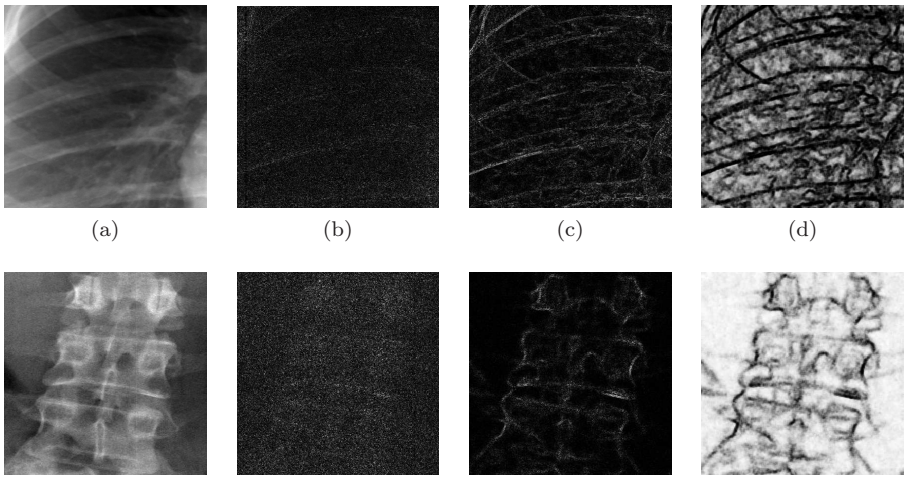


Figure 3: Estimation of the LNC-maps from analysis of noisy and denoised images. Top row: lung ROI. Bottom row: spine ROI. (a,e): I manually windowed for better visibility. (b,f): D_0^I. (c,g): $D_0^{\hat{I}}$. (d,h): LNC-maps, the interval of gray levels is $[0, 1]$.

Table 1: Percent energy loss δ_t in a lung region (R_1) and in a lumbar spine one (R_2).

	δ_0	δ_1	δ_2	δ_3	δ_4	δ_5
R_1	0.78	0.35	0.18	0.05	0.02	0.00
R_2	0.97	0.85	0.56	0.27	0.09	0.03

(2010) and references therein. This paper simply assumes that the coefficients can be modeled by a Laplacian distribution with zero mean and scale parameter β. Given the maximum likelihood estimation of β of a zero mean Laplacian distribution, the energy $\Xi(D_t)$ of the detail level D_t distribution is:

$$\Xi(D_t) = 2\left(\frac{1}{M}\sum_{i=1}^{N}|D_t(z_i)|\right)^2 \tag{10}$$

where M is the number of detail coefficients at scale t and z_i are the coordinates of the coefficient i. The percent energy loss at a band-pass image $D_t^{\hat{I}}$ is then computed as follows:

$$\delta_t = 1 - \frac{\Xi(D_t^{\hat{I}})}{\Xi(D_t^I)}. \tag{11}$$

Table 1 indicates that δ_t monotonically decreases to zero as a function of t. This happens because the XNLM filter removes only the high frequency component of the noise as it is patch-based (Lebrun et al., 2012). Besides, the detail band-pass images are obtained by applying an iterative smoothing filter (Equation 7), which progressively reduces the noise importance. From the δ_t values in Table 1,

where the lung and lumbar spine region are denoted by R_1 and R_2, respectively, we deduce that the percent energy loss decays faster in low density regions. For example, the noise in the lumbar spine region needs to be taken into account up to the third decomposition level, while in the lung region it is negligible at the second level. As a consequence, by comparing the input and noise-free detail coefficients it is possible to automatically decide up to which decomposition level the noise needs to be contained while avoiding relying on empirically defined parameters.

Figures 3c and 3g show that the energy is preserved in correspondence of edges and structures, such as the ribs or the internal parts of a vertebra, because the XNLM tends to preserve locally redundant information. As a consequence, we can estimate local noise containment maps (LNC-maps), where the term *local* highlights the fact that the measures come from an analysis in local patches.

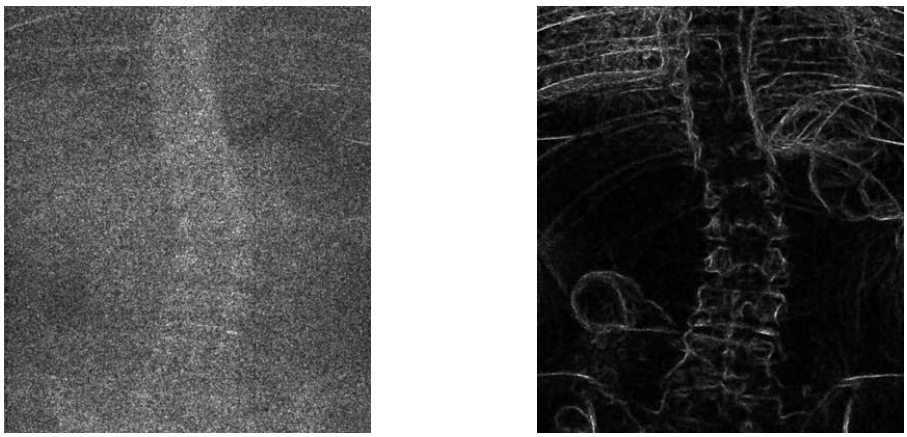
The LNC-maps w_t are computed by comparing a measure of local contrast of D_t^I and $D_t^{\hat{I}}$. This measure, denoted by C_t, is the average magnitude of coefficients in a patch of size $|P|$, i.e. the same patch size used by the XNLM filter. It is preferable to use the average within a small window rather than single coefficients in order to avoid taking into account aberrant values, especially at fine levels. The LNC-map at scale t is then computed as follows:

$$w_t = \delta_t \left(1 - \frac{2C_t^I C_t^{\hat{I}} + \epsilon}{{C_t^I}^2 + {C_t^{\hat{I}}}^2 + \epsilon} \right) \tag{12}$$

where C_t^I is the contrast measure from D_t^I, ϵ is a scalar factor used to avoid singularities, in this case fixed to 1, and δ_t is the percent energy loss at a level t (Equation 11). In Equation 12 w_t is computed as the product of a local term that measures the correlation between local contrast measures and a global one. The local term represents the probability of a coefficient being pure noise. The probability is low when the contrast values are similar because the XNLM filter detects a structure of interest preserving the most of its energy. The global term δ_t indirectly quantifies the amount of redundancy in the image and can help distinguishing a small morphotype from a big one, because the higher the signal at the detector the more visible are the structures. The relevance of this global term has been confirmed by gathering the feedback from experts who concluded that, by defining the LNC-map using only the local one, the outcome presented a slight smoothing effect. Finally, it is worth noting that the LNC-maps do not depend on user-defined parameters.
The resulting LNC-maps at scale 0 for the two ROIs taken as examples are shown in Figure 3. According to Equation 9, by using these noise containment maps the anatomical structures of interest, such as the ribs and the bronchi, the vertebral edges in the lungs (Figure 3d) and the lumbar spine (Figure 3h), are fully enhanced.

3.3. Proposed boosting technique

The detail coefficients are modified according to an activity map in order to increase the contrast. In particular, while the detail coefficient values associated



(a) (b)

Figure 4: Activity map: (a) Computed using C_t^I. (b) Computed using $C_t^{\hat{I}}$.

with low activity are increased, those at high activity are preserved. We use an activity map that is a variant of Peli's formula (Peli, 1990) that measures the contrast on complex images:

$$A_t = \frac{C_t^{\hat{I}}}{L_t^{\hat{I}} + \tilde{\epsilon}} \tag{13}$$

where $C_t^{\hat{I}}$ is the contrast measure computed from $D_t^{\hat{I}}$ (Section 3.2), L_t is the low frequency residual at scale t and $\tilde{\epsilon}$ is a scalar value, fixed to 100 for all the tests, that is used to avoid singularities and dependency from very low coefficient magnitudes. With respect to the original formulation (Peli, 1990), we introduce two variants. First, the contrast is computed within a small local window and not by considering coefficient-wise magnitude because as shown by Li et al. (2005) a smooth gain enhances better the contrast with less halo artifacts. Secondly, $C_t^{\hat{I}}$ is considered rather than C_t^I. Figures 4a and 4b show two examples of activity maps at scale 0 computed using C_t^I and $C_t^{\hat{I}}$, respectively. While Figure 4a shows very similar activity values for noise and structures, in Figure 4b the underlying anatomical structures are well highlighted and this type of activity maps is preferred here.
Given the activity maps at each scale, the relative gain functions are computed as follows:

$$G_t(z_i) = \begin{cases} (\gamma_t - 1)\left(1 - \frac{A_t(z_i)}{\alpha_t}\right)^2 + 1 & \text{if} \quad A_t(z_i) < \alpha_t \\ 1 & \text{otherwise} \end{cases} \tag{14}$$

where γ_t is the maximal gain at scale t, α_t is the activity cut off value beyond which the coefficients $D(z_i)$ are not enhanced. The cut off values are defined as

follows:

$$\alpha_t = \min\left(c\bar{A}_t, \max(A_t)\right) \tag{15}$$

where $\bar{A}_t$ is the average activity at scale t, and c is a constant value, defined by the user, that avoids to take into account aberrant activity values and is set to 25 in the validation tests.
As for the maximal gain parameters, they are initially set by linearly ranging their values from 6 at the finest scale to 3 at the coarsest one. Empirically, these are ideal parameters when the amount of noise is very low. However, in our context of interest, this case rarely occurs. The actual gain values are then semi-automatically computed as follows:

$$\gamma_t = g_t - (g_t - g_{k-1})\bar{w}_t \tag{16}$$

where g_t are the initial ideal gain values, k is the number of scales (here $k = 6$) and $\bar{w}_t$ is the average value of the LNC-map at t. The initial gain values tend to be preserved if the noise is absent at a given scale, i.e. $\bar{w}_t \simeq 0$. On the contrary, when $\bar{w}_t \simeq 1$, then γ_t is set equal to the gain at the coarsest scale, i.e. the lowest one. The initial values g_t are defined by taking into account the property of full body X-rays; they should probably be changed to apply the algorithm to other images.
Finally, the problem of defining the maximal gain parameters is simplified. Indeed, only one set needs to be fixed for all the images and, then, the values are automatically adjusted according to the image content. The fully enhanced band-pass images are obtained by applying the coefficient-wise multiplication between input detail coefficients and the gain maps (Equation 14).

3.4. Overview and computational load of the proposed method

The main steps of the proposed method are summarized here. The noise standard deviation image σ is estimated from the input I by using the percentile method (Section 2.2) and used in Equation 5, which computes the patch distances. The denoised image $\hat{I}$ is obtained (Section 2.3) and then both I and $\hat{I}$ are encoded into band-pass images D_t and low frequency residuals L_k by using the IUWT (Section 3.1). Then, the gain functions (Equation 14) are computed by using the activity maps (Equation 13) and the semi-automatically defined gain parameters. The LNC-maps (Section 3.2) are used in Equation 9 to weight the contribution of the input noisy coefficients and the fully boosted coefficients. Finally, the enhanced image is obtained by adding the output detail levels (Equation 8). Table 2 reports the setting of the main parameters in the most relevant steps of the proposed method. All the parameters are either automatically estimated, or fixed experimentally once for all, i.e. the same values are used in all our tests.

As for the computational load, it takes on average 2 minutes to completely process a full body EOS image, which contains about 14 millions of pixels, in a not optimized MATLAB based environment. The main computational load is due to the XNLM filter, while the other operations are computationally comparable to those of a conventional MS contrast enhancement method multiplied by

Table 2: Main steps and related parameters of the proposed method.

Percentile method	- High pass filter = D_0 of the IUWT; - Percentile value = 10%; - Square block half size = 7.
XNLM filter	- Square patch P half size = 2; - Square window Ω half size = 7; - $h = 2\sigma^2$, σ estimated by the percentile method.
NC maps	- Automatically estimated from D_0^I and $D_0^{\hat{I}}$.
Boosting	- Max gain value = 6.0 - Min gain value = 3.0 - Number of scales = 6 - Cut off value = average level activity $\times 25$

two because we decompose two images and not one. Finally, since the XNLM filter is coded using integral images, an estimation is that the proposed framework could be efficiently implemented in a C++ oriented environment and using parallel programming.

4. Experimental results

4.1. Dataset

The proposed method is validated on X-ray EOS data that are encoded on 16 bits. EOS is a stereo-radiographic system based on a slot scanning principle and it uses the multiwire gaseous detector that has been conceived from Georges Charpak's[2] researches (Charpak et al., 1968). The detector amplifies the X-rays that pass through the body of a scanned patient by means of electronic avalanche in the gas, which allows keeping the dose low. Moreover, a wide range of different tissues is covered, from cartilage to bones in thick regions. The acquisition parameter setting changes depending on the type of exam and morphotype of the patient. These parameters are the peak kilo voltage output of the X-ray images generator (kV), X-ray tube current in mA (mA), and exposure time C $\in \{1, 2, \ldots, 8\}$, where bigger values stand for longer exposure time. The input signal strength associated with a parameter setting is indicated in Air Kerma

[2] 1992 Physics Nobel Laureate

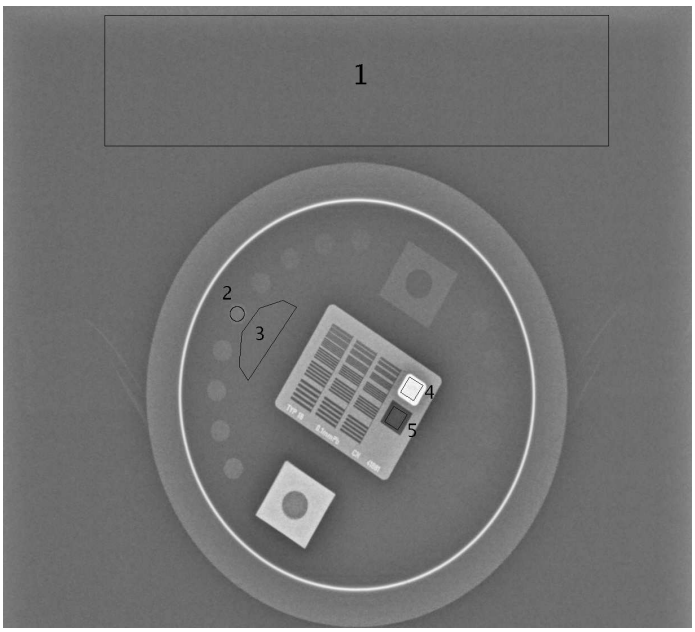


Figure 5: Standard PHD5000 phantom with annotated ROIs used to compute the following measures: SNR from region 1; CNR from regions 2 and 3; DYN from regions 4 and 5. See text for the detailed definitions.

dose value, i.e. the absorbed energy by unit mass of air, which is measured in μGy.

Both phantom and clinical EOS images are used for the validation of the proposed method, denoted by LNCE (local noise containment enhancement), that is compared with three other methods. Two of these approaches are simplifications of the proposed one as they consist of fully enhancing the input (D_t^I) or the denoised ($D_t^{\hat{I}}$) band-pass images, and are denoted by NE (noise enhancement) and DE (denoised enhancement), respectively. Note that the method DE is similar to the one that we have proposed in Irrera et al. (2013) with the difference that manual tuning of the filter parameters is avoided as presented in Section 2.3. The third method is a EOS proprietary algorithm, denoted by EOSE, that is used by default for exams in clinical routine. It contains the noise at the two finest levels of a IUWT decomposition according to predefined thresholds on the noise level. Therefore, EOSE exploits parametric noise containment maps. This method is more relevant for our study than others from the literature because the parameters of the gain functions used to modify the detail coefficients have been chosen by taking into account the opinion of EOS users. Moreover, note that the choice of the default EOS algorithm is justified by the interest in heterogeneous clinical data. The evaluation on specific types of exams, e.g. follow-up in pediatrics, is beyond the scope of this paper.

Table 3: Samples of the phantom dataset described by the thickness of the PMMA block in cm, the entrance dose in μGy and the SNR, CNR and DYN values computed over the input image.

cm	μGy	snr	cnr	dyn
10	10	47.25	1.35	49%
10	67	116.47	4.29	53%
10	71	126.02	4.23	53%
10	75	110.73	4.24	53%
10	215	190.71	7.47	52%
15	65	108.51	2.44	40%
15	87	115.30	3.01	41%
15	89	116.84	2.98	42%
15	100	123.53	2.80	42%
15	102	132.75	2.95	41%
15	136	165.43	3.78	41%
20	16	37.63	0.53	30%
20	109	117.90	1.96	33%
20	454	180.01	2.59	33%
25	109	97.13	1.16	26%
25	136	98.24	1.26	26%
30	33	27.39	0.20	19%
30	566	128.46	1.14	22%

Table 4: Average SNR, CNR and DYN improvements on phantom images.

	NE	DE	LNCE	EOSE
$(\mathrm{SNR}_o - \mathrm{SNR}_i)/\mathrm{SNR}_i$	-81.0%	-28.9%	-69.5%	-72.9%
$(\mathrm{CNR}_o - \mathrm{CNR}_i)/\mathrm{CNR}_i$	-20.9%	281.9%	43.7%	-21.5%
$(\mathrm{DYN}_o - \mathrm{DYN}_i)/\mathrm{DYN}_i$	53.8%	53.0%	53.5%	32.8%

4.2. Validation on phantom images

We evaluate the image quality on the standard PHD5000 phantom (Figure 5) according to different signal strengths, and interposing between the X-ray tube and the phantom a polymethyl methacrylate (PMMA) block of various thicknesses. The samples of the dataset are reported in Table 3, where the thickness of the PMMA block is indicated in cm and the signal strength represented by the entrance dose in μGy. The measures of image quality computed over the input images are also indicated. These experiments simulate the acquisition conditions of typical clinical cases. For example, the samples $\{10cm; 10\mu\mathrm{Gy}\}$, $\{10cm; 71\mu\mathrm{Gy}\}$ and $\{10cm; 215\mu\mathrm{Gy}\}$ are acquired with $\{60kV; 83mA; 4C\}$, $\{83kV; 200mA; 4C\}$ and $\{100kV; 280mA; 4C\}$, respectively. These parameters are used in the following exams: follow-up of the full spine, diagnostic full spine and pelvis. The samples at $10cm$ correspond to the pediatric morphotype and those at $20cm$ to the normal adult one.

The signal to noise ratio (SNR) is equal to the average signal in a theoretically constant region (ROI 1 in Figure 5) divided by the standard deviation of its gray levels. The SNR is indirectly related to the amount of signal that reaches the detector: Table 3 shows that lower PMMA thickness and/or stronger input signal values (e.g. $\{10cm, 215\mu Gy\}$) produce higher SNR measures. The first line in Table 4.2 shows the average percent increase of SNR of the post-processed images with respect to the input data. The values are negative for each tested method, which is expected given the SNR definition. Indeed, while the standard deviation increases to optimize the contrast balance for the whole image, the average signal remains approximately the same. These results indicate that DE is the best option in terms of SNR, which is logical because it fully enhances

an image that has been filtered, while the other solutions only contain the noise (LNCE and EOSE) or simply do not take it into account (NE). However, since the SNR is measured in a void region, it gives no information on how regions of interest are processed.

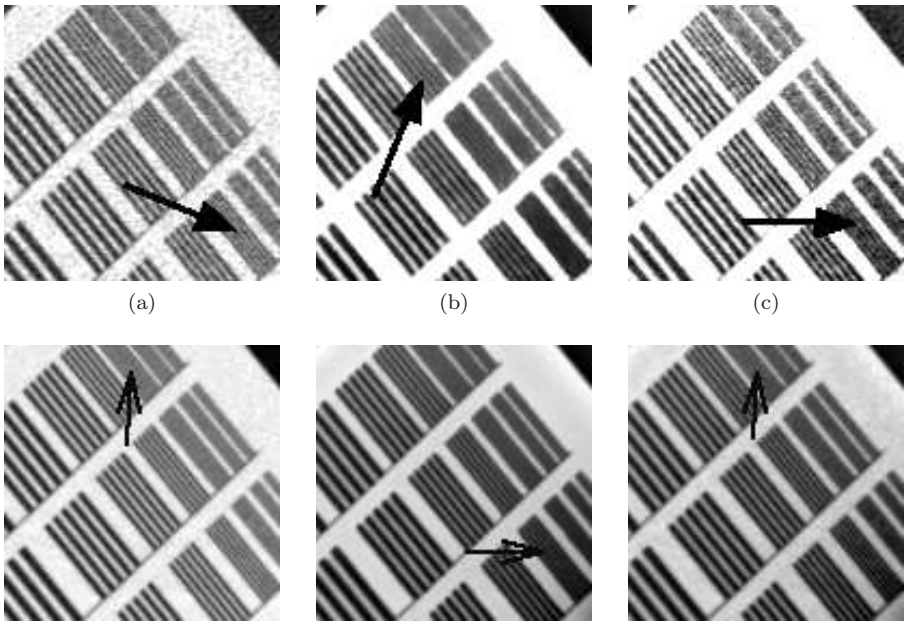
The contrast to noise ratio (CNR) gives a deeper insight on image quality than the SNR because it is computed as:

$$CNR = \frac{\bar{I}(\mathcal{R}_2) - \bar{I}(\mathcal{R}_3)}{\sigma(I(\mathcal{R}_3))} \tag{17}$$

where $\bar{I}(\mathcal{R}_j)$ and $\sigma(I(\mathcal{R}_j))$ are, respectively, the mean and standard deviation of the gray levels in a region $\mathcal{R}_j$ and $\mathcal{R}_2$ and $\mathcal{R}_3$ are, respectively, an object of interest (a disk, see region 2 in Figure 5) and nearby background (see region 3 in Figure 5). The CNR quantifies the compromise between increasing the visibility of an object of interest and boosting the noise. The second line in Table 4.2 shows the average percent increase of CNR of the post-processed images with respect to the input data. Positive values prove a good balance between contrast and noise, while negative ones mean over-shooting, i.e. the gain of contrast is not sufficient to compensate the increase of noise. Table 4.2 indicates that both NE and EOSE present over-shooting. While this is predictable with NE, because the noise is not contained at all, it is less expected for EOSE. According to our tests, there is not over-shooting when EOSE is used to process low signal or thick PMMA block instances (e.g. PMMA $\geq 20cm$), but it is present in the other cases. This problem is common with methods that use parametric noise containment maps because they do not provide sufficient flexibility. On the other hand, LNCE overcomes this drawback, because, exception made for one case ($\{10cm, 215\mu Gy\}$), there is no over-shooting. Nevertheless, the results obtained when the images are processed with DE are significantly better. While the CNR measures on the disks describe the increase of visibility of low and medium detail levels, no conclusions concerning fine structures can be deduced.

The dynamic range (DYN) is computed as $\bar{I}(\mathcal{R}_4) - \bar{I}(\mathcal{R}_5)$ where $\mathcal{R}_4$ is the most absorbing region (number 4 in Figure 5) and $\mathcal{R}_5$ is the least absorbing one (number 5 in Figure 5). It represents how well the gray level dynamic is exploited and is expressed as percentage on the total number of gray levels, i.e. 65535 in EOS images. Table 3 shows that the dynamic range depends on the thickness, while it is practically unaffected by the signal strength. The third line in Table 4.2 shows the average percent increase of DYN of the post-processed images with respect to the input data. These results show that the methods NE, DE and LNCE have very similar scores, because they use the same boosting technique (Section 3.3), and they perform better than EOSE.

The high resolution grid at the center of the phantom (Figure 5) is used to evaluate the spatial resolution of a system. The resolution is measured in line pairs per millimeter (lp/mm) and associated with the finest set of visible lines on the grid. It depends then on the signal strength and PMMA block thickness, in addition to detector properties. This section verifies whether the compared methods preserve the spatial resolution computed over the input image. The



(a) (b) (c)

(d) (e) (f)

Figure 6: Resolution grid of the phantom: top row = $10\mu Gy$ and bottom row = $215\mu Gy$; columns from left to right = IN, DE and LNCE. For each example, the arrows indicate the finest, clearly visible, set of lines.

input image manually windowed in the region of the resolution grid is compared to DE and LNCE. Note that the study of NE is unnecessary because, since the image is fully enhanced, it is not possible to reduce the resolution, and a part of our tests, not included in this paper, have indicated that EOSE does not cause a resolution loss. Figure 6 presents two of our experiments with $10cm$ PMMA: the columns, from left to right, show the resolution grid from the input image with manual windowing, DE and LNCE; the lowest and highest signal strengths of the sample at $10cm$ are considered, i.e. $10\mu Gy$ and $215\mu Gy$ (respectively the first and second rows in Figure 6). The arrows point at the finest set of visible lines. The images show that DE causes a spatial resolution loss, while LNCE does not. They also show that the grade of the decrease depends on the signal strength. Indeed, with $10\mu Gy$ the DE resolution is 1.4 lp/mm versus 1.8 lp/mm and with $215\mu Gy$ 1.8 lp/mm versus 2.0 lp/mm. This is due to the XNLM filter effectiveness in restoring the lines: the stronger the signal the easier to capture the similarity between patches is. Similar results are obtained by conducting the same tests with thicker PMMA blocks.

Finally, these results on phantom images indicate that by applying a filter and, then, enhancing the contrast of the resulting image (DE) the quality in terms of SNR, CNR and DYN is very good, but this entails also a loss of spatial

resolution that may be a problem for some clinical applications. On the other hand, the proposed algorithm (LNCE) overcomes this drawback and, at the same time, offers a good compromise between detail enhancement and noise containment.

4.3. Validation on clinical images

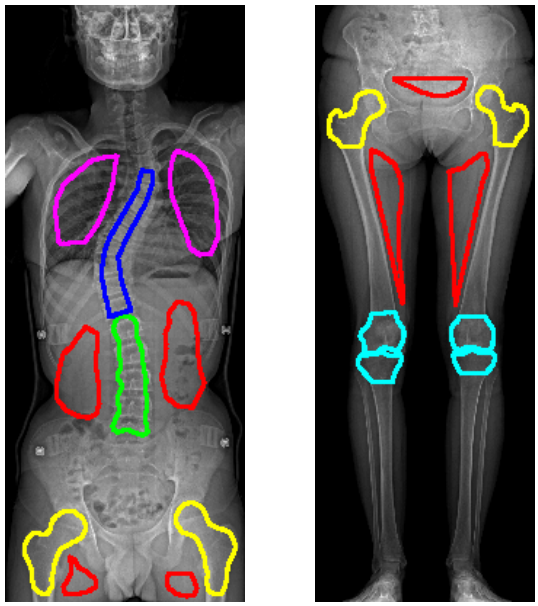
The clinical database consists of 130 images that have been randomly selected among anonymous data. These images present patients of different ages and sizes as well as a wide range of type of exams (full spine, lower limbs, diagnostic, follow-up). This is the logical transposition of the tests conducted on the phantom to the clinical context.

In digital radiography two measures are often used to quantify the quality of clinical images: the average local variance (ALV) (Chang and Wu, 1998) and the contrast improvement index (CII) (Laine et al., 1995). The ALV is the average of variance values associated with small windows (5×5 pixels in this paper). The ALV measure is computed in three disjoint regions of the pixel space: smooth (ALVS), detail (ALVD) and edge (ALVE). Generally, low values of ALVS and ALVE are preferred because they quantify the noise enhancement and the presence of halo artifacts, respectively. Similarly, high ALVD means good details processing. In the original approach (Chang and Wu, 1998) the regions are defined by applying user-defined thresholds to the local variance image of the input. However, threshold values are not easy to define given the signal-dependent noise. As a consequence, the ALV relates to contrast but it is not a pure measure of it, and therefore should be used along with visual comparison.

Rather than using a generic detail region, this paper proposes to use anatomically significant regions. For each image of the clinical dataset, regions associated with the following diagnostically significant structures are manually defined: lumbar spine ($\mathcal{A}_1$), thoracic spine ($\mathcal{A}_2$), proximal femur ($\mathcal{A}_3$), lungs ($\mathcal{A}_4$) and knees ($\mathcal{A}_5$). This approach permits to quantify the contrast in anatomical ROI and, hence, to understand how the performances of an algorithm depend on the features of anatomical structures of interest. However, the contrast measures over these regions are influenced by the presence of noise too. Therefore, the noise contribution is quantified by measuring the ALV in a void region ($\mathcal{A}_0$), i.e. with no signal of interest. Two examples are shown in Figure 7. Using these measures, we define the anatomy contrast (AC) as follows:

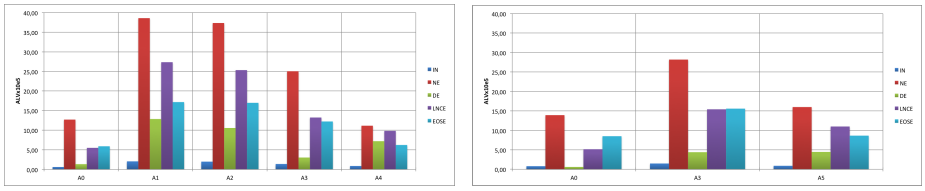
$$\mathrm{AC} = \frac{1}{K}\left(\sum_{i>0} \mathcal{A}_i\right) - \mathcal{A}_0 \tag{18}$$

where K is the number of previously defined anatomic classes that appear in the image, and $\mathcal{A}_i = 0$ if the relative object is outside the field of view. The first term of the equation relates to the amount of contrast in anatomical ROI. However, since it is biased by the presence of noise, another term that relates to noise only, i.e. $\mathcal{A}_0$, is subtracted. It would be interesting to compute this measure for each considered anatomic ROI i, but this would require defining



(a) (b)

Figure 7: Examples of manually segmented regions used for computing anatomic ALV: red = $\mathcal{A}_0$, green = $\mathcal{A}_1$, blue = $\mathcal{A}_2$, yellow = $\mathcal{A}_3$, magenta = $\mathcal{A}_4$, cyan = $\mathcal{A}_5$



(a) (b)

Figure 8: ALV measures in anatomical ROI computed from the images in (a) Figure 7a; (b) Figure 7b. For each of the anatomical ROI on the x-axis, the ALV measures are computed over IN, NE, DE, LNCE and EOSE.

$\mathcal{A}_0$ in a void region of mean signal comparable to that of ROI i. Since such a region may not always be present, e.g. for the lungs in the most of cases, this test has not been performed.

The CII quantifies the improvement in terms of contrast on an enhanced image Y with respect to the initial one. Since it only requires using the same techniques for quantifying the contrast in input and enhanced images, it can be used straightly with the proposed anatomical ALV. First, the CII is computed for each anatomical ROI in the image to evaluate. The average of these values is then computed to obtain the Signal CII (SCII). Similarly, the Noise CII (NCII)

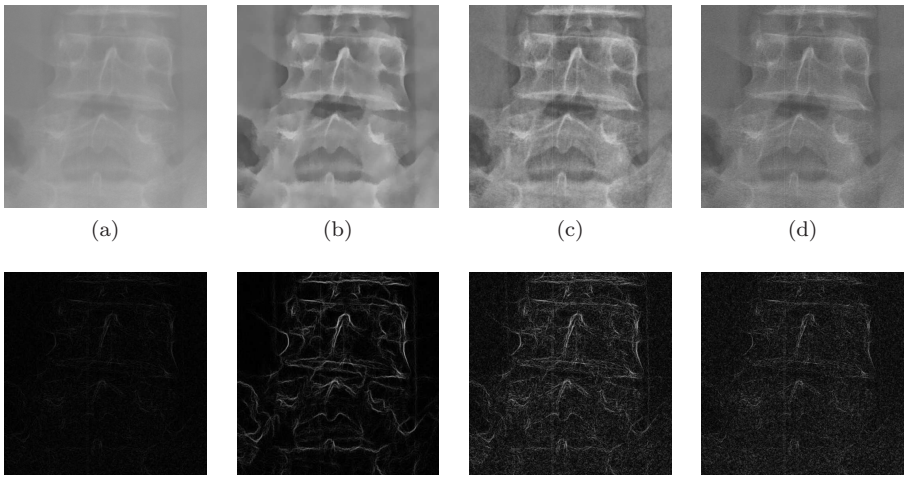


Figure 9: Visual comparison of vertebrae L4 and L5 in $\mathcal{A}_1$: top row = images and bottom row = related Sobel gradient magnitudes; columns from left to right = IN, DE, LNCE amd EOSE.

is computed as NCII = ${\mathcal{A}_0}^Y/{\mathcal{A}_0}^I$. Finally, we define an unbiased CII (UCII) as: UCII = SCII − NCII. This measure allows quantifying the over-shooting in clinical images. A negative value of UCII means that the noise is more enhanced than the signal, i.e. the higher visibility of structures comes at the price of an excessive boosting of noisy coefficients, while positive values of UCII indicate a relative higher enhancement of relevant information. Finally, a visual comparison permits to verify the coherence between computed measures and to validate them.

Figure 8 reports the ALV measures computed from the regions in Figure 7 and are used to compare the four image enhancement methods considered.

In the ROI $\mathcal{A}_0$, a significant increase of the noise with NE can be noted on both full spine (Figures 8a) and lower limbs (Figures 8b) images. Indeed, considering the full spine sample (Figure 8a) the ALV in region $\mathcal{A}_0$ increases with respect to the input by: $\times$20.84 (NE), $\times$2.20 (DE), $\times$9.02 (LNCE) and $\times$9.67 (EOSE). These results are coherent with the analysis conducted on phantom SNR (Section 4.2): DE limits the noise impact better than the noise containment method, whereas NE should be discarded because it excessively boosts the noise.

Into the lumbar spine, the ALV increases with respect to the input by: $\times$18.83 (NE), $\times$6.27 (DE), $\times$13.35 (LNCE) and $\times$8.38 (EOSE). The first observation is that the relative improvement of ALV given by NE is similar to the one revealed in region $\mathcal{A}_0$. Therefore, even if NE gets the highest score in $\mathcal{A}_1$, it is not reliable because it depends only on noise. On the contrary, the other algorithms do not show this trend. LNCE is better than DE and EOSE. However, the analysis needs to be completed with visual comparison to counter-

balance the relative contributions in terms of signal and noise. Figure 9 shows, in the first row, the visual rendering of the vertebrae L4 and L5 over IN, DE, LNCE and EOSE, and, in the second row, the magnitude of the Sobel gradient computed over these images. Refining the input image demands extensive user interaction to actually make structures of interest appear therefore slowing down the process of diagnosis. This is also an issue from an image quality point of view because it is very difficult to find the optimal balance between noise boosting and structure visibility with manual contrast setting. Moreover, given the low visibility of the edges, tasks such as automatic segmentation of spine structures become very challenging, whereas the post-processed images limit the need for manual interaction. On the other hand both DE and LNCE (Figures 9b and 9c) have better brightness than EOSE (Figure 9d), which may demand user interaction to be properly studied. The medium and low frequency structures of the spine are preserved with DE, but high frequencies are lost in the vertebral body, as shown by magnitude of gradient of DE and LNCE in Figures 9f and 9g, respectively. While DE returns a quite regular image that could be more adapted for instance to the automatic detection of the vertebral body, LNCE suits better for diagnosis because it reflects the textured nature of the bones. The higher ALV value obtained with LNCE is a further proof in this discussion. Finally, note that the gradient magnitude is extremely low in the tissues surrounding the vertebral body, while it is not inside of it. This means that the LNCE enhances actual information of interest and the contribution of noise to the obtained ALV value is negligible. In the thoracic spine, the ALV increases with respect to the input by: $\times$18.87 (NE), $\times$5.34 (DE), $\times$12.80 (LNCE) and $\times$8.58 (EOSE). These values are very similar to those obtained in the lumbar spine region because the type of structure belongs to the same category, i.e. vertebrae.

In the lungs, the ALV increases with respect to the input by: $\times$12.82 (NE), $\times$8.28 (DE), $\times$11.30 (LNCE) and $\times$7.16 (EOSE). The lungs are a low density area and rich in medium and low frequency structures, which explains why the performances of DE and EOSE are comparable. Moreover, it is logical that NE and LNCE have similar performances as the noise in this region is almost absent.

In the proximal femur of the full spine exam, the ALV increases with respect to the input by: $\times$18.15 (NE), $\times$2.20 (DE), $\times$9.59 (LNCE) and $\times$8.86 (EOSE). The same trend can be noted as for region $\mathcal{A}_1$, but the ALV score obtained with DE is significantly lower than LNCE. Two aspects explain this result. First the proximal femur is poorer in medium and low frequency structures than the vertebrae. Indeed, the only structures of this type are the femur edges and the cotyloid cavity, i.e. where the femur meets the pelvis. Secondly, as observed by analyzing the high resolution grid of the phantom (Section 4.2), the low amount of signal makes it more difficult to detect redundancy and, so, to preserve high frequency details. Therefore, $\mathcal{A}_3$ is a typical region where LNCE is preferred if a diagnostic image quality is desired.

Finally in the knee area, the ALV increases with respect to the input by: $\times$17.78 (NE), $\times$4.97 (DE), $\times$12.22 (LNCE) and $\times$9.61 (EOSE). Figure 10 shows

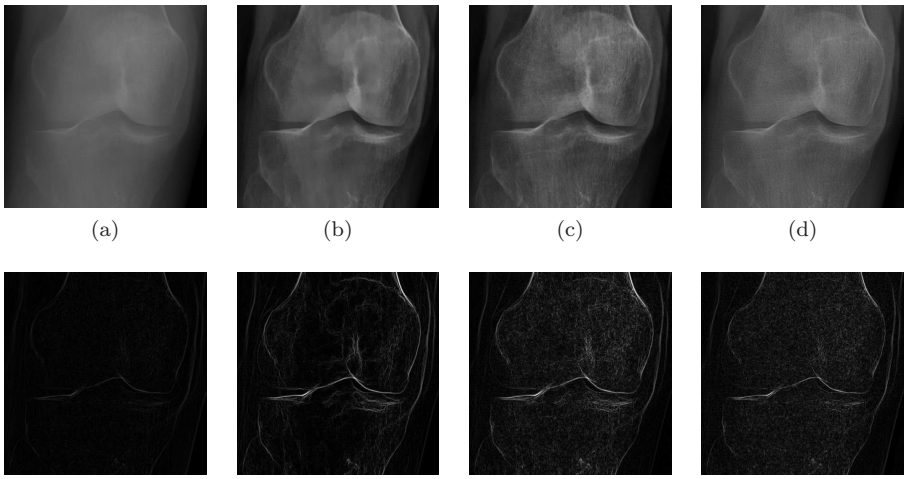


Figure 10: Visual comparison of knee in $\mathcal{A}_5$: top row = images and bottom row = related Sobel gradient magnitudes; columns from left to right = IN, DE, LNCE and EOSE.

Table 5: Average AC and UCII values computed over clinical images processed with the compared algorithms.

	IN	NE	DE	LNCE	EOSE
AC	2.01	25.14	3.79	11.45	5.69
UCII	0	−3.09	1.80	1.37	−1.16

that the most important information in the knee is the bone texture. Note that DE is not adapted to this region as shown by the significantly lower ALV value obtained with DE compared to the one obtained with LNCE.

The observations derived from the samples in Figure 7 hold for the whole set of data on which we have computed ALV measures. In proof of this, the average AC and UCII measures on the totality of the dataset are reported in Table 5. The highest value of AC is obtained with NE. However, this is mainly due to a strong noise contribution in anatomical ROI. This is confirmed by the value of UCII that is largely smaller than zero meaning that the over-shooting is too important to actually take NE into consideration. The algorithms DE and LNCE allow avoiding over-shooting and can be, in our opinion, alternatives to answer different clinical needs. Indeed, DE could be more adapted for tasks such as automatic bone segmentation, because it is more regular than LNCE while preserving edges. However, a noise containment oriented method is more suitable to be presented to clinicians for medical imaging aided diagnosis, which is confirmed by the higher AC value when using LNCE. Finally LNCE outperforms parametric noise containment methods. Indeed, with respect to EOSE, the anatomical contrast doubles while avoiding over-shooting. Moreover, the analysis on the regions $\mathcal{A}_i$ show that LNCE adapts well to the heterogeneous

features of tissues present in a full body image, which supports the robustness of the proposed method.

4.4. Quantitative evaluation by a radiologist

The last part of the validation is dedicated to quantify the contribution of the proposed method to clinical routine. According to the results from the previous sections, the study focuses on noise containment based approaches, i.e. compare LNCE to EOSE. The application of DE to object segmentation is not studied as it would require comparing different segmentation techniques, which is beyond the scope of this paper. For this purpose we have worked in collaboration with a radiologist at the AP-HP, Cochin hospital in Paris (France). We have considered a set composed by 10 patients, between 54 and 81 years old, having Body Mass Index (BMI) between 20.28 and 28.93 corresponding to normal and overweight adults. None of the patients presents particular pathology. Therefore, only the visibility of a set of diagnostically relevant structures of the human skeletal apparatus has been evaluated.
In practice, the radiologist has assigned a vote to each structure that is representative of its degree of visibility. The votes go from 0, meaning that the structure is not visible at all, to 5, meaning that the structure is perfectly visible. Given the interest in a full body diagnosis, different anatomical regions have been examined and the following list of structures retained:

- **Frontal cervical spine** ($\mathcal{C}^f$): spinous process (1), vertebral body (2), intervertebral disc (3), zygapophyseal joint (4), transverse process (5), pedicle (6).
- **Lateral cervical spine**($\mathcal{C}^l$): median atlanto axial joint (1), intervertebral disc (2), vertebral body (3), transverse process (4), uncus (5), pedicle (6), lamina (7), posterior interapophyseal joints (8), spinous process (9).
- **Frontal thoracic spine** ($\mathcal{T}^f$): spinous process (1), vertebral body (2), intervertebral disc (3), transverse process (4), costovertebral joints (5), pedicle (6).
- **Lateral thoracic spine** ($\mathcal{T}^l$): spinous process (1), vertebral body (2), intervertebral disc (3), transverse process (4), posterior interapophyseal joints (5), pedicle (6).
- **Frontal lumbar spine** ($\mathcal{L}^f$): spinous process (1), vertebral body (2), intervertebral disc (3), transverse process (4), pedicle (5), posterior interapophyseal joints (6), sacrum (7).
- **Lateral lumbar spine** ($\mathcal{L}^l$): spinous process (1), vertebral body (2), intervertebral disc (3), transverse process (4), pedicle (5), posterior interapophyseal joints (6), intervertebral foramina (7), lamina (8).
- **Pelvis** ($\mathcal{P}$): sacrum (1), sacroiliac joint (2), cotyloid cavity (3), pubic symphysis (4), hip bones (5), hip joint (6), ischium (7), femur (8).

- **Knee** ($\mathcal{K}$): patella (1), femoro-tibial joint (2), medial and lateral intercondylar tubercles (3), condyles (4), fibula (5), tibial plateau (6).

Table 6: Average votes assigned by a radiologist to anatomical structures.

	$\mathcal{C}_1^f$	$\mathcal{C}_2^f$	$\mathcal{C}_3^f$	$\mathcal{C}_4^f$	$\mathcal{C}_5^f$	$\mathcal{C}_6^f$			
EOSE	3.8	3.9	3.0	2.5	1.0	2.7			
LNCE	4.5	4.4	3.5	3.1	1.5	3.2			
	$\mathcal{C}_1^l$	$\mathcal{C}_2^l$	$\mathcal{C}_3^l$	$\mathcal{C}_4^l$	$\mathcal{C}_5^l$	$\mathcal{C}_6^l$	$\mathcal{C}_7^l$	$\mathcal{C}_8^l$	$\mathcal{C}_9^l$
EOSE	2.9	4.9	4.9	3.6	2.1	4.0	4.0	4.4	4.2
LNCE	3.6	4.9	5.0	4.3	2.7	4.3	4.1	4.6	4.6
	$\mathcal{T}_1^f$	$\mathcal{T}_2^f$	$\mathcal{T}_3^f$	$\mathcal{T}_4^f$	$\mathcal{T}_5^f$	$\mathcal{T}_6^f$			
EOSE	2.6	3.8	3.5	1.4	2.1	2.7			
LNCE	3.3	4.6	4.2	1.9	2.9	3.7			
	$\mathcal{T}_1^l$	$\mathcal{T}_2^l$	$\mathcal{T}_3^l$	$\mathcal{T}_4^l$	$\mathcal{T}_5^l$	$\mathcal{T}_6^l$			
EOSE	1.6	4.5	4.2	1.8	2.1	3.0			
LNCE	2.2	4.8	4.4	2.1	2.8	3.7			
	$\mathcal{L}_1^f$	$\mathcal{L}_2^f$	$\mathcal{L}_3^f$	$\mathcal{L}_4^f$	$\mathcal{L}_5^f$	$\mathcal{L}_6^f$	$\mathcal{L}_7^f$		
EOSE	4.3	4.1	4.0	1.4	3.4	1.8	2.6		
LNCE	4.8	4.6	4.0	2.1	4.3	2.3	2.8		
	$\mathcal{L}_1^l$	$\mathcal{L}_2^l$	$\mathcal{L}_3^l$	$\mathcal{L}_4^l$	$\mathcal{L}_5^l$	$\mathcal{L}_6^l$	$\mathcal{L}_7^l$	$\mathcal{L}_8^l$	
EOSE	2.6	4.2	4.7	2.2	3.7	2.6	3.6	3.1	
LNCE	3.5	4.9	4.7	2.9	4.4	3.4	3.9	3.8	
	$\mathcal{P}_1$	$\mathcal{P}_2$	$\mathcal{P}_3$	$\mathcal{P}_4$	$\mathcal{P}_5$	$\mathcal{P}_6$	$\mathcal{P}_7$	$\mathcal{P}_8$	
EOSE	2.2	3.0	4.0	4.0	5.0	4.4	4.3	4.9	
LNCE	2.4	3.9	5.0	4.6	5.0	4.9	5.0	5.0	
	$\mathcal{K}_1$	$\mathcal{K}_2$	$\mathcal{K}_3$	$\mathcal{K}_4$	$\mathcal{K}_5$	$\mathcal{K}_6$			
EOSE	3.5	4.90	5.0	5.0	5.0	5.0			
LNCE	4.1	4.90	5.0	5.0	5.0	5.0			

In total, 56 structures have been taken into account. The average votes for each structure evaluated over LNCE and EOSE are reported in Table 6. Regional qualitative measures can be defined from these votes to derive summarized evaluations. An anatomical region is considered perfectly visible as long as all the structures therein get a vote of 5/5. Then, a percentage is assigned to the anatomical ROI by comparing the total of the votes to the ideal optimum. Table 7 reports the average scores and points gained by the proposed method. These results show that the proposed method enhances the visibility of osteoarticular structures. Note that the two algorithms have similar performances for structures that are surrounded by thin soft tissues, e.g. the knees.

Table 7: Regional qualitative measures for each evaluated anatomical ROI.

	$\mathcal{C}^f$	$\mathcal{C}^l$	$\mathcal{T}^f$,	$\mathcal{T}^l$	$\mathcal{L}^f$	$\mathcal{L}^l$	$\mathcal{P}$	$\mathcal{K}$
EOSE	56%	78%	54%	57%	62%	67%	80%	95%
LNCE	67%	85%	69%	67%	71%	79%	90%	97%
Gain	**+11**	**+7**	**+15**	**+10**	**+9**	**+12**	**+10**	**+2**

The spine is a complex area because of the strong superposition of tissues and the projection of a rotated 3D volume on a 2D plane. Since in all areas of

the spine the votes of LNCE are greater than EOSE ones, the proposed method is more appropriate in facilitating the visibility in challenging cases. However, some objects remain difficult to see. For example, in most of the cases the transverse process is only slightly visible ($\mathcal{C}_5^f$, $\mathcal{T}_4^f$, $\mathcal{T}_4^l$ and $\mathcal{L}_4^f$). Globally, about 10 points are gained with respect to EOSE in the spine. The image quality in the pelvis improves significantly with the proposed method. Indeed, by using EOSE the image is of good quality (80%), but LNCE renders almost a perfect image (90%). This is due to the fact that the structures in the pelvis are quite regular and, therefore, easy to extract with the proposed method, even if the signal is low. The only problematic object is the sacrum ($\mathcal{P}_1$) because it is covered by a lot of soft tissues. Finally, for what concerns the knee, the improvement is negligible and only the patella ($\mathcal{K}_1$) is easier to delimit with LNCE. Indeed, the score is almost perfect because the absorption is very low. Then, in this region, it would be more interesting to compare the two algorithms in pathological cases that, for example, concern the health of the bone tissue.

In conclusion, this diagnostic feedback highlights the value of the method, which can hence be considered as promising for a potential use in clinical routine.

5. Conclusion

In this paper we proposed a method for improving the quality of X-ray digital images. This method has proved to be robust to changes in the input signal strength, patient morphotype and features of the multiple tissues that are in the field of view. We have first adapted the classical NLM filter to digital X-ray images and, by relying on an automatic noise level estimation, our approach does not require manual parameter tuning. The filtered images have been used as oracles to define non parametric noise containment maps that, when applied in a multiscale contrast enhancement framework, allow optimizing the trade-off between improvement of the visibility of anatomical structures and noise reduction. A significant number of tests on both phantoms and clinical images have shown that the proposed method is better suited than others for visual inspection for diagnosis, even when compared to an algorithm used to process EOS images in clinical routine.

Despite the good results, the choice of adapting the NLM for the patch-based filter could be discussed. Indeed, recent works show that better denoising quality can be achieved with the Non Local Bayes filter (Lebrun et al., 2013), which currently represents the state of the art denoising algorithm. However, this method is computationally expensive because it demands to learn twice a local Gaussian model for each patch (Lebrun et al., 2013). Therefore, the eventual improvement on the quality of the estimated noise containment maps should be counterbalanced with the increase in computational load, which is an important aspect in clinical routine. Moreover, it has been pointed out in Section 4.4 that some structures remain very difficult to see. This is related to the conditions of the acquisition and no major improvement could be achieved by using more complex methods. For these reasons, future works will focus on aspects related to the clinical evaluation rather than improving the method. For

example, it would be interesting to quantify how much the entrance dose can be reduced and still get an image quality convenient for diagnosis. Furthermore, our results suggest that the approach DE could be used for applications such as automatic segmentation or registration, but this remains to be more accurately evaluated.

Acknowledgment

This work was done during the PhD of P. Irrera which was supported by an ANRT grant (2012/0202). The authors would like to thank the Medical Doctor N. Regnard for the qualitative evaluation of clinical images and the Professor A. Feydy for providing and agreeing to use the data from AP-HP, Cochin hospital in Paris (France) for the research purposes presented in this paper.

References

Buades, A., Coll, B., Morel, J. M., 2005. A Review of Image Denoising Algorithms, with a New One. Multiscale Modeling and Simulation 4, 490–530.

Burt, P., Adeldon, E., 1983. The Laplacian Pyramid as a Compact Image Code. IEEE Transactions on Communications 31 (4), 532–540.

Cerciello, T., Bifulco, P., Cesarelli, M., A.Fratini, 2012. A comparison of denoising methods for x-ray fluoroscopic images. Biomedical Signal Processing and Control 7 (6), 550 – 559.

Chang, D. C., Wu, W. R., August 1998. Image Contrast Enhancement Based on a Histogram Transformation of Local Standard Deviation. IEEE Transactions on Medical Imaging 17 (4), 518 –531.

Charpak, G., Bouclier, R., Bressani, T., Favier, J., Č. Zupančič, 1968. The use of multiwire proportional counters to select and localize charged particles. Nuclear Instruments and Methods 62 (3), 262 – 268.

Colom, M., Buades, A., 2013. Analysis and Extension of the Percentile Method, Estimating a Noise Curve from a Single Image. Image Processing On Line 3, 332–359.

Coupé, P., Yger, P., Prima, S., Hellier, P., Kervrann, C., Barillot, C., 2008. An optimized blockwise nonlocal means denoising filter for 3-D magnetic resonance images. IEEE Transactions on Medical Imaging 27 (4), 425–441.

Damet, J., Fournier, P., Monnin, P., Sans-Merce, M., Ceroni, D., Zand, T., Verdun, F. R., Baechler, S., 2014. Occupational and patient exposure as well as image quality for full spine examinations with the EOS imaging system. Medical Physics 41 (6), 1–12.

Darbon, J., Cunha, A., Chan, T. F., Osher, S., Jensen, G. J., 2008. Fast nonlocal filtering applied to electron cryomicroscopy. In: International Symposium on Biomedical Imaging: From Nano to Macro. IEEE, pp. 1331–1334.

Dippel, S., Stahl, M., Wiemker, R., Blaffert, T., April 2002. Multiscale Contrast Enhancement for Radiographies: Laplacian Pyramid Versus Fast Wavelet Transform. IEEE Transactions on Medical Imaging 21 (4), 343 –353.

Fan, B., Han, H., October 2011. Medical image enhancement based on modified lut-mapping derivative and multi-scale layer contrast modification. In: 4th International Congress on Image and Signal Processing (CISP). pp. 696 – 703.

Hensel, M., Lundt, B., Pralow, T., Grigat, R.-R., 2006. Robust and Fast Estimation of Signal-Dependent Noise in Medical X-Ray Image Sequences. Bildverarbeitung für die Medizin 2006, 46–50.

Irrera, P., Bloch, I., Delplanque, M., April 2013. A Denoising Method for Whole-Body Low-Dose X-Ray Images with Adaptable Parameter Control. In: International Symposium on Biomedical Imaging. pp. 1228–1231.

Irrera, P., Bloch, I., Delplanque, M., April 2014. Contrast enhancement of Micro-dose X-ray images. In: International Symposium on Biomedical Imaging. pp. 489–492.

Laine, A., Fan, J., Yang, W., September/October 1995. Wavelets for Contrast Enhancement of Digital Mammography. IEEE Engineering in Medicine and Biology Magazine 14, 536–550.

Lebrun, M., Buades, A., Morel, J., 2013. A Nonlocal Bayesian Image Denoising Algorithm. SIAM Journal on Imaging Sciences 6 (3), 1665–1688.

Lebrun, M., Colom, M., Buades, A., Morel, J. M., 5 2012. Secrets of image denoising cuisine. Acta Numerica 21, 475–576.

Li, Y., Sharan, L., Adelson, E., 2005. Compressing and Companding High Dynamic Range Images with Subband Architecture. ACM Transactions on Graphics 24 (3), 836–844.

Loza, A., Bull, D., Canagarajah, N., Achim, A., 2010. Non-Gaussian Model-based Fusion of Noisy Images in the Wavelet Domain. Computer Vision and Image Understanding 114 (1), 54–65.

Loza, A., Hill, P., Verkade, P., Bull, D., Al-Mualla, M., Achim, A., April 2014. Joint denoising and contrast enhancement for light microscopy image sequences. In: International Symposium on Biomedical Imaging. pp. 1083–1086.

Mantiuk, R., Myszkowski, K., Seidel, H.-P., July 2006. A perceptual framework for contrast processing of high dynamic range images. ACM Transactions on Applied Perception 3 (3), 286–308.

Peli, E., Oct 1990. Contrast in complex images. Journal of the Optical Society of America A 7 (10), 2032–2040.

Ponomarenko, N., Lukin, V., Zriakhov, M., Kaarna, A., Astola, J., 2007. An automatic approach to lossy compression of aviris images. In: IEEE International Geoscience and Remote Sensing Symposium, IGARSS. pp. 472–475.

Ronckers, C., Doody, M., Lonstein, J., Stovall, M., C.E., L., March 2008. Multiple Diagnostic X-rays for Spine Deformities and Risk of Breast Cancer. Cancer Epidemioligy, Biomarkers and Prevention 17 (3), 605–613.

Sakata, M., Ogawa, K., October 2009. Noise Reduction and Contrast Enhancement for Small-Dose X-ray Images in Wavelet Domain. In: IEEE Nuclear Science Symposium Conference Record. pp. 2924 – 2929.

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.

Shuryak, I., Sachs, R. K., Brenner, D. J., 2010. Cancer Risks After Radiation Exposure in Middle Age. Journal of the National Cancer Institute 102, 1628–1636.

Smith-Blindman, R., Miglioretti, D., Johnson, E., June 2012. Use of Diagnostic Imaging Studies and Associated Radiation Exposure for Patients Enrolled in Large Integrated Health Care Systems, 1996-2010. The Journal of American Medical Association 307 (22), 2400–2409.

Stahl, M., Aach, T., Dippel, S., Buzug, T., Wiemker, R., Neitzel, U., February 1999. Noise-Resistant Weak-Structure Enhancement for Digital Radiography. In: Hanson, K. M. (Ed.), SPIE Vol. 3661: Medical Imaging 99: Image Processing. San Diego, USA, pp. 1406–1417.

Starck, J.-L., Fadili, J., Murtagh, F., Feb 2007. The Undecimated Wavelet Decomposition and its Reconstruction. IEEE Transactions on Image Processing 16 (2), 297–309.

Wybier, M., Bossard, P., 2013. Musculoskeletal imaging in progress: the EOS imaging system. Joint Bone Spine 80 (3), 238–243.