31st Annual International Conference of the IEEE EMBS
Minneapolis, Minnesota, USA, September 2-6, 2009

Comparative Study of Two Image Space Noise Reduction Methods for Computed Tomography: Bilateral Filter and Nonlocal Means

Juan C. Ramirez Giraldo, *Student Member, IEEE,* Zachary S. Kelm, Luis S. Guimaraes, Lifeng Yu, Joel G. Fletcher, Bradley J. Erickson, Cynthia H. McCollough

Abstract—**Optimal noise control is important for improving image quality and reducing radiation dose in computed tomography. Here we investigated two image space based nonlinear filters for noise reduction: the bilateral filter (BF) and the nonlocal means (NLM) algorithm. Images from both methods were compared against those from a commercially available weighted filtered backprojection (WFBP) method. A standard phantom for quality assurance testing was used to quantitatively compare noise and spatial resolution, as well as low contrast detectability (LCD). Additionally, an image dataset from a patient's abdominal CT exam was used to assess the effectiveness of the filters on full dose and simulated half dose acquisitions. We found that both the BF and NLM methods improve the tradeoff between noise and high contrast spatial resolution with no significant difference in LCD. Results from the patient dataset demonstrated the potential of dose reduction with the denoising methods. Care must be taken when choosing the NLM parameters in order to minimize the generation of artifacts that could possibly compromise diagnostic value**.

I. Introduction

There is a growing concern regarding the amount of radiation dose associated with medical imaging modalities such as those based on X-rays (i.e. radiography and computed tomography) or γ-rays (i.e. SPECT and PET) [1,2]. Unfortunately, in X-ray based modalities such as CT, there is a fundamental tradeoff between dose and image quality: the larger the dose the better the image quality (i.e. lower noise, better low contrast detectability, fewer artifacts). As a result, it is imperative to carefully optimize and manage exam protocols to ensure that the dose delivered to the patient in CT examinations is as low as possible, while ensuring appropriate image quality.

There exist several different strategies one might use to lower the dose in CT. The first one is related to improvement in the hardware components of the scanners, such as more sensitive and efficient detectors [3], or by the use of different acquisition geometries [4]. Recently, CT scanners have included automatic tube current modulation that works both angularly and along the z-axis direction to provide significant reductions in dose [5]. As another route, the filtered backprojection process that CT scanners offer for image reconstruction allows the user to employ different kernels that target different image qualities in the reconstructed images. Typical kernels are roughly divided into smooth and sharp kernels. The smooth kernels reduce the noise but decrease spatial resolution, whereas sharp kernels maintain better spatial resolution but have higher noise [6]. A third possibility for improvement of image quality is the use of iterative reconstruction algorithms [7], which allow for accurate modeling of system geometry, physical effects like beam hardening, scatter, and incomplete data sampling. Although iterative reconstruction holds the promise of becoming the trend in future CT scanners, it is currently limited by the very large number of computations needed, usually requiring several hours to reconstruct a single dataset in specialized hardware.

Denoising algorithms applied to either the projection or image space data have been proposed for lowering the dose in CT. Projection space methods to reduce noise in the sinogram, prior to image reconstruction, include the use of the iterative penalized weighted least square method [8], multi-dimensional adaptive filtering [9], and bilateral filtering [10]. The advantage of projection space denoising to image space, is that, after an appropriate data transformation, it is possible to approximate a Poisson-like model for the photon counting process. In contrast, there exists no certain noise model in the image space for CT imaging. Still, projection space denoising methods requires access to raw projection data and image reconstruction capabilities, limiting its use in clinical practice unless included in the commercial scanner workflow.

Various edge preserving algorithms have been proposed for image denoising in CT, including non-linear adaptive filtering [11] and wavelet based [12] methods. Here we focus our attention on the use of two edge preserving filters: the bilateral filter (BF) [13] and the more recently introduced nonlocal means (NLM) algorithm [14]. These methods are fast (compared to iterative methods), easy to implement, and do not require raw projection data. Furthermore, we compare the spatial resolution and Low Contrast Detectability (LCD) of the BF and NLM with conventional weighted filtered backprojection algorithms (WFBP) available with commercial scanners.

The paper is organized as follows: section II explains the theory of the BF and NLM algorithms; section III describes the experiments performed and includes details of the datasets used. Section IV presents the results with several

Manuscript received April 23, 2009.

J.C. Ramirez Giraldo, L.S. Guimaraes, L. Yu, J.G. Fletcher and C.H. McCollough are with the CT Clinical Innovation Center, Dept. of Radiology, Mayo Clinic, Rochester, MN, USA.

Z.S. Kelm and B.J. Erickson are with the Medical Imaging Informatics Innovation Center, Dept. of Radiology, Mayo Clinic, Rochester, MN, USA. (Corresponding author: ramirezgiraldo.juancarlos@mayo.edu)

This project was partially supported by Grant No. R01EB007986 from the National Institute of Biomedical Imaging and Bioengineering.

examples comparing the BF and NLM with WFBP; finally, some general recommendations and discussion are included in Section V.

II. Theory

A. The Bilateral Filter

The formulation of the BF for noise reduction states that each pixel is replaced by a weighted average of its neighbors, as for example is done with the convolution of an image with a Gaussian filter. However, for edge preservation, the BF also takes into account the variation of pixel intensities. As a result, spatially close pixels are averaged only if their intensity values are similar, while dissimilar intensity values tend to be preserved [10,13].

First, consider that each pixel $x[k]$ is replaced by a weighted sum $W[k,n]$ of its neighbors and stored in $y[k]$.

$$y[k] = \frac{\sum_{n=-N}^{N} W[k,n] \cdot x[k-n]}{\sum_{n=-N}^{N} W[k,n]} \quad . \qquad (1)$$

The weight $W[k,n]$ is a product of the spatial and intensity weights, (2) and (3) respectively.

$$W_s[k,n] = \exp\left(-\frac{n^2}{2\sigma_s^2}\right) \quad , \qquad (2)$$

$$W_i[k,n] = \exp\left(-\frac{(x[k]-x[k-n])^2}{2\sigma_i^2}\right) \quad . \qquad (3)$$

From (2) and (3), notice that the two key parameters needed for the BF are σ_s and σ_i, which control the spatial and intensity weighting respectively.

B. The Nonlocal Means Filter

The NLM algorithm is a nonlinear spatial filter that adjusts the intensity value of each pixel using a weighted average of other pixels' intensities based on the similarity of their neighborhoods or "patches" to the patch around the pixel being adjusted [14]. Although these neighborhood comparisons can be performed between pixels located any distance apart in an image, for computation speed, the comparisons are usually limited to a specified search window around the pixel whose value is being adjusted. Mathematically, this process can be described as follows:

$$y[k] = \frac{1}{C_N} \sum_{n \in SW} W[k-n] \cdot x[k-n] \,, \qquad (4)$$

where $y[k]$ is the new pixel value at coordinate k, C_N is a normalization constant, SW is the search window centered at k, $x[k\text{-}n]$ is the original pixel value at the $(k\text{-}n)$ coordinate within the search window, and $W[k\text{-}n]$ is the weight given the intensity value of $x[k\text{-}n]$. The weights are calculated as:

$$W[k-n] = \exp\left(-\sum_{m \in P} K_G \left[\frac{x[k-m]-x[(k-n)-m]}{h}\right]^2\right), \qquad (5)$$

where P is the patch size around each pixel that is compared, K_G is a Gaussian kernel used to give patch pixels further from the center less weight in the similarity calculation, and h adjusts the intensity similarity-to-weight relationship. h can be considered a smoothing parameter, such that as it is increased, more dissimilar patches are given a higher weight, allowing for more averaging/smoothing. For the case when $n = 0$, $W[k]$ is set equal to the maximum of the other calculated weights within the search window in order to prevent giving too much weight to the original pixel value at k. C_N is then calculated to be the sum of the weights W within the search window. This algorithm, then, has the adjustable parameters of search window size SW, patch size P, and smoothing amount h.

III. Experiments

An image quality phantom (CATPHAN, Phantom Labs) was scanned using a dual-source CT scanner (Siemens Definition DS) with a routine abdominal protocol: 120 kV, 64x0.6 mm collimation, pitch = 0.8, and 240 (full dose) and 120 (half dose) effective mAs. Images were reconstructed using the scanner standard WFBP algorithm with B10, B20, B30, B40, and B45 kernels at 2 mm thickness and 2 mm position increment. The B45 image, which is expected to have higher spatial resolution, but also higher noise relative to the other kernels, was used as the starting image to apply the BF and NLM.

To compare the filters, we used the high-contrast spatial resolution HCSR (in lp/cm) module, as well as the low contrast module to assess low contrast detectability (LCD). It is already known that edge preserving methods such as the BF and NLM are effective in preserving strong edges. However, for CT it is crucial that they do not blur or eliminate low contrast objects, which are not necessarily comprised of strong edges.

To approach a more realistic scenario than a phantom model, we used a de-identified study from a patient scanned with a dual-source CT scanner (Siemens Definition DS) with the routine abdominal protocol: 120 kV, 64x0.6 mm collimation, pitch = 0.8, and 240 (full dose) effective mAs. A half dose dataset was simulated by inserting Poisson noise in the projection space before reconstruction. The noise insertion method used does take into account tube current modulation as well as the bowtie filter. Both datasets were reconstructed using the B45 kernel and 2 mm slice thickness.

For the purpose of more closely examining image sharpness, we used line profiles from low and high contrast regions.

The full and half dose abdominal images were filtered with varying parameter values for each algorithm. A radiologist was asked to select the best parameter values for each case. The chosen values were then used to process both the phantom and abdomen datasets at full and half dose. Because optimal parameter values are extremely difficult to define in

CT, we used radiologist input to evaluate appropriate diagnostic quality. An excessively smoothed medical image may look better to the untrained eye, but could have potentially diagnostic information blurred away.

IV. Results

A. Selected parameter values

For the bilateral filter, the parameter values selected were window size 5x5 pixels, $\sigma_s = 0.6$ for both half and full dose but with different intensity weighting parameter σ_i=0.026 and σ_i=0.021 respectively. Note that the BF implementation normalized image intensities from 0 to 1. For both the half and full dose cases, the parameter values selected for the NLM processed images at full and half dose were search window SW = 5x5 pixels, patch size P = 3x3 pixels, and smoothing value h=100.

B. Phantom

The phantom dataset confirmed the known tradeoff between spatial resolution and noise when using different kernels in WFBP; the lower the noise (e.g. B10), the lower the spatial resolution (see Fig.1). However, both the BF and NLM algorithms were able to keep similar spatial resolution as the B45-WFBP, while lowering the noise. As expected, the full dose images had lower noise than their half dose counterparts for the WFBP, BF, and NLM. When the low contrast module was studied, we found no significant difference in LCD among the WFBP, BF and NLM. Results indicated average LCD up to 3 mm and 7 mm for the 10 HU and 5 HU target diameters respectively for the full dose case. The average LCD were 6 mm and 9 mm for the 10 HU and 5 HU target diameters respectively for the half dose case.

C. Patient dataset

According to a radiologist, when using optimized parameters for both BF and NLM, the image quality of the processed images compared favorably with their corresponding original image (either full dose or half dose), mainly due to noise reduction without a significant increase in artifacts or blurriness. Moreover, the image quality of the half dose processed images approached that of the full dose images, as seen in Fig 2.

To assess the sharpness of the images, we compared the line profiles in a high contrast region (HCR) and a low contrast region (LCR), shown in Fig 3. The profiles for the HCR were very similar, indicating that strong edges were well preserved even though the BF and NLM filters reduced the image noise. For the LCR, the full dose profiles clustered together, as did the half dose profiles. Notice that the profiles for the BF and NLM follow the variations of the WFBP data, although a slight amount of smoothing is present as expected.

V. Discussion

We found that both the BF and NLM methods were able to improve the tradeoff between high-contrast spatial resolution and noise when compared against analytical WFBP methods. There was no significant difference for LCD, which is crucial to guarantee that low contrast objects (weak edges) are not smoothed in the filtering process.

The half and full dose patient datasets processed with the BF and NLM showed a significant improvement in image quality compared to their unfiltered counterparts, based on radiologist observation and supported by line profile examination. Half dose images approached full dose quality, as judged by a radiologist for the patient dataset studied. The same observation was found quantitatively with the phantom data since the half dose dataset filtered with BF and NLM decreased the noise to a level comparable to full dose, while keeping good spatial resolution and no significant change in LCD.

In spite of the fact that NLM was able to reduce the noise while preserving edges and contrast information, we found that, depending on the parameter values used, it often produced an organized noise pattern artifact which could significantly compromise image quality and diagnosis.

In contrast to nonmedical filtering applications, the noise reduction in CT should carefully balance the amount of smoothing to still maintain diagnostic value. We found it very challenging to recommend a single set of parameters which would guarantee optimal results for the BF and NLM filters. We foresee that in an ideal scenario for image space denoising, a system would be required that is able to provide radiologists with real-time feedback, analogous to changing window levels, but rather changing 'sharpness-smoothing' levels in the images. Practical implementation of such a scenario would require, however, an optimized implementation of the BF and NLM algorithms, as well as specialized hardware to speed up computations. The resultant preferred parameters would need to be archived for future use, which presents considerable practical complexities, especially for image archival and retrieval systems.

References

[1] D. J. Brenner, E. J. Hall, "Computed tomography - an increasing source of radiation exposure," *N. Eng. J. Med*, vol. 357, pp. 2277-2284, 2007.

[2] T. C. Gerber *et al.*, Ionizing radiation in cardiac imaging: "A science advisory from the American Heart Association Committee on Cardiac Imaging of the Council on Clinical Cardiology and Committee on Cardiovascular Imaging and Intervention of the Council on Cardiovascular Radiology and Intervention," *Circulation*, vol. 119, pp. 1056-1065, 2009.

[3] W.A. Kalender, Y.Kyriakou, "Flat-detector computed tomography, " *Journal of European Radiology*, vol 17, pp.2767-2779, Nov 2007.

[4] T. G. Schmidt, R. Fahrig, N. J. Pelc, and E. G. Solomon, "An inverse geometry volumetric CT system with a large-area scanned source: feasibility study," *Medical Physics*, vol. 31, pp. 2623-2627,2004.

[5] C. H. McCollough, A. N. Primak, N. Braun, J. Kofler, L. Yu, and J. Christner, "Strategies for reducing radiation dose in CT," *Radiol Clin North Am*, vol. 47, pp. 27-40, 2009.

[6] J. Hsieh, *Computed Tomography: Principles, design, artifact and recent applications*. Bellingham, WA: SPIE Press, Ch. 3 ,2003.

[7] J. B. Thibault, K. D. Sauer, C. A. Bouman, and J. Hsieh, "A three-dimensional statistical approach to improved image quality for multislice helical CT," *Medical Physics*, vol. 34, pp. 4526-4544, 2007.

[8] J. Wang, T. Li, H. Lu, and Z. Liang. "Penalized weighted least –squares approach to sinogram noise reduction and image reconstruction for low

dose X-ray computed tomography," *IEEE Trans. Med. Imaging*, vol. 25, pp. 1272-1283, 2006.
[9] M. Kachelrieß, O. Watzke, W.A. Kalender. Generalized multi-dimensional adaptive filtering for conventional and spiral single-slice, multi-slice, and cone-beam CT. *Medical Physics*, vol. 28(4), pp. 475-490, 2001.
[10] L. Yu, A. Manduca, J. Trzasko, N. Khaylova, J. M. Kofler, C. H. McCollough, and J. G. Fletcher, "Sinogram smoothing with bilateral filtering for low-dose CT," *Progress in Biomedical Optics and Imaging – Proc. of the SPIE*, vol. 6913, article 691329, 2008.
[11] H. Lu, X. Li, L. Li, D. Chen, Y. Xing, J. Hsieh, Z. Liang. "Adaptive noise reduction toward low-dose computed tomography", in *Medical Imaging 2003: Physics of Medical Imaging. Proceedings of the SPIE,* M.J. Yafe and L. E. Antonuk, Eds., vol. 5030, pp. 759-766, 2003.
[12] A. Borsdorf, R. Raupach, T. Flohr, J. Hornegger. "Wavelet based noise reduction in CT-images using correlation analysis". *IEEE Trans. Med. Imaging,* vol 27, pp. 1685-1703, 2003.
[13] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," In *Proc. of the International Conference on Computer Vision*, Piscataway, NJ, pp. 839-846, 1998.
[14] A. Buades, B. Coll, and J. M. Morel, "A review of image denoising algorithms, with a new one," *SIAM Journal on Multiscale Model. Simul.*, Vol. 4, Issue 2, pp. 490-530, 2005.

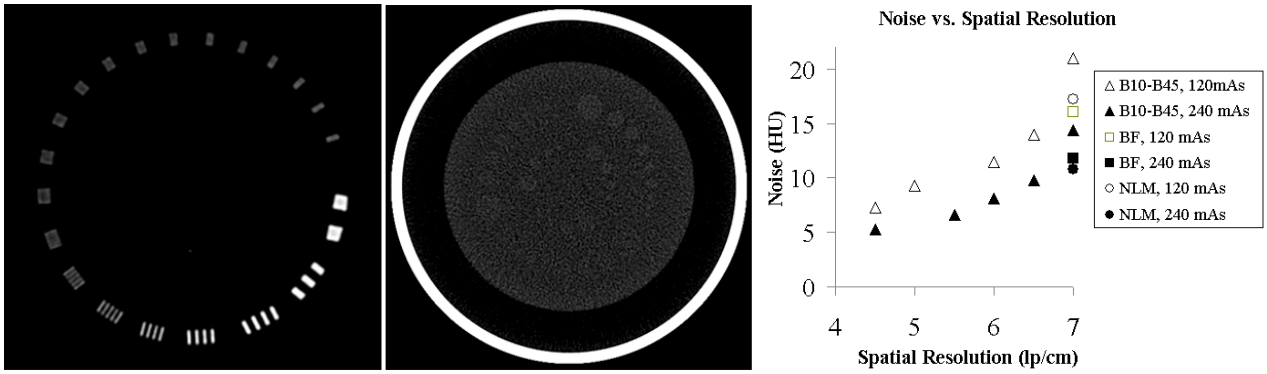

Fig. 1. Spatial resolution, low contrast detectability, and noise measurements using two CATPHAN modules scanned with full (240 mAs) and half dose (120 mAs). (a) Original full dose high contrast module. (b) Original full dose low contrast module. (c) Chart comparing noise vs. spatial resolution as determined after reconstruction with the B10-B45 kernels and further NLM/BF processing of the B45 reconstruction. These values were found for the 120mAs and 240 mAs scans of the high contrast module.

Fig. 2. Comparison of full dose (FD) and half dose (HD) abdomen slices (a) HD [B45] (b) HD, NLM processed (c) HD, BF processed (d) original FD.

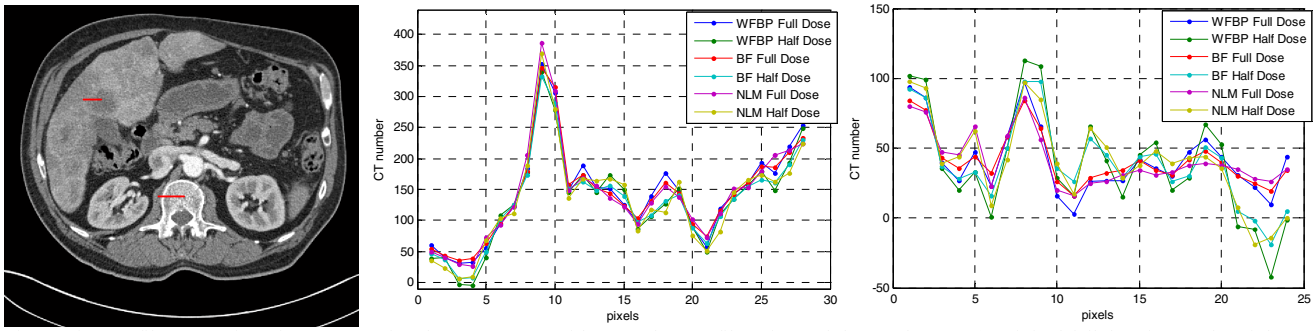

Fig. 3. Line profile comparison from conventional WFBP (B45) with BF and NLM filters in an abdomen dataset (a) Original full dose image (b) High contrast line profile (c) Low contrast line profile.