A Reference Image Database Approach for NLM Filter-Regularized CT Reconstruction

Wei Xu and Klaus Mueller

***Abstract*— Low-dose CT is becoming more popular in recent years due to growing concerns on radiation exposure in medical scans. Regularized iterative CT reconstruction is a promising way to overcome the resulting noise and streak artifacts. We have recently described an approach that uses non-local means (NLM) filters in the regularization step. Traditional NLM filters evaluate reference neighborhoods in the current image for filtering. This can lead to poor results when the image is severely contaminated under very low low-dose conditions. We introduce a method that instead uses reference regions in artifact-free priors, that is, reconstructions that contain no artifacts but have similar anatomical and pathological structures. This requires (i) a comprehensive database containing a representative sample of possible anatomical, pathological, noise and streak images, and (ii) a suitable search+match strategy to locate the contaminated micro-feature and links it to its clean counterpart. We find that our method performs much better than traditional NLM under these adverse conditions, while being computationally efficient.**

I. Introduction

Low-dose CT has attracted more attention in recent years since it can reduce the radiation exposed to the scanned patients. To reconstruct a volume in low-dose CT, iterative methods are often applied to obtain a numerically optimized solution. However, severe noise or streak artifacts may still persist, making accurate diagnosis difficult. To suppress these artifacts, regularization is usually incorporated into the reconstruction process. One form of regularized CT reconstruction alternates the core reconstruction algorithm and the regularization step in each iteration step. In previous work [7], we have demonstrated the successful application of one neighborhood filter, the bilateral filter, as a regularization operator within an iterative algebraic reconstruction framework. In more recent work [6], we expanded on this idea by introducing (and comparing) a wider selection of neighborhood filters. Specifically, we found that the non-local means (NLM) filter [1] performs exceedingly better because it bases the filtering on redundant areas that exist in the image, using a patch-based similarity matching procedure to identify relevant pixels. In this context, a patch is a small adjacent area around a pixel.

NLM replaces the value of each pixel p_x with the weighted sum of the values of neighborhood pixels p_y, whose weights are determined by the patch similarity between p_x and p_y:

$$p'_x = \frac{\sum_{y \in W_x} \exp\left(-\sum_{t \in P} G_a(t)\left|p_{x+t} - p_{y+t}\right|^2 / h^2\right) \cdot p_y}{\sum_{y \in W_x} \exp\left(-\sum_{t \in P} G_a(t)\left|p_{x+t} - p_{y+t}\right|^2 / h^2\right)} \quad (1)$$

Here x and y are pixel locations, W_x is the neighborhood of x, P is the patch size of each pixel, t is the index within a patch, G_a is a Gaussian kernel and h is the parameter of NLM to control the overall smoothness of the filtering. Although NLM works well for artifacts suppression in modestly contaminated images, for more severe artifacts arising from very few projections or extremely low SNR, the regularization effect can still be quite limited. This is due to the fact that NLM depends on a high degree of redundancy inside an image and takes advantage of this redundancy to eliminate noise. However, when the matched neighborhood regions are heavily corrupted, the redundancy is too poor to be reliable.

In this paper, we modify the original NLM filter and introduce a *reference-based NLM (RNLM)* filter to take advantage of existing artifact-free patches that match the imaged original features. In this scheme, the reference images contain no artifacts but have similar anatomical and pathological structures than the image to be regularized. To find such reference regions, we propose (i) the construction of a comprehensive database containing all possible anatomical, pathological, noise and streak image regions, and (ii) a suitable search+match strategy that can locate the contaminated image regions in the database and link them to their clean counterparts. Our experiments indicate that the method allows even severe artifacts to be effectively reduced.

The paper is organized as follows: Section 2 presents related work and background, Section 3 describes our reference-based NLM algorithm and its workflow, Section 4 presents results, and Section 5 ends with conclusions.

II. Related Work and Background

Much work on reference-based filtering has appeared in computer graphics. Petschnigg *et al* used this type of approach in conjunction with a bilateral filtering module [3] for digital photography. Their work, called *Joint Bilateral Filter*, uses flash and non-flash image pairs to retain good aspects from both images for noise reduction, detail preservation and retention of ambient illumination. Another application is the colorization of grey-level images [4]. Here, a user provides the grey-level image to be colorized and a (color) reference image of a similar scene. To find the RGB color of a grey-level (target) pixel, the algorithm then searches the grey-level converted reference image for pixels with similar grey-level neighborhoods than the target pixel and then chooses the original color of the best matching reference pixel. In our proposed scheme, the reference images are a set of artifact-free medical images and the converted images used for patch matching which are these same images but degraded by some measurable artifact model. The artifact-free feature counterparts are then retrieved and used for

Wei Xu and Klaus Mueller are with the Computer Science Department, Stony Brook University (e-mail: {wxu, mueller}@cs.sunysb.edu).

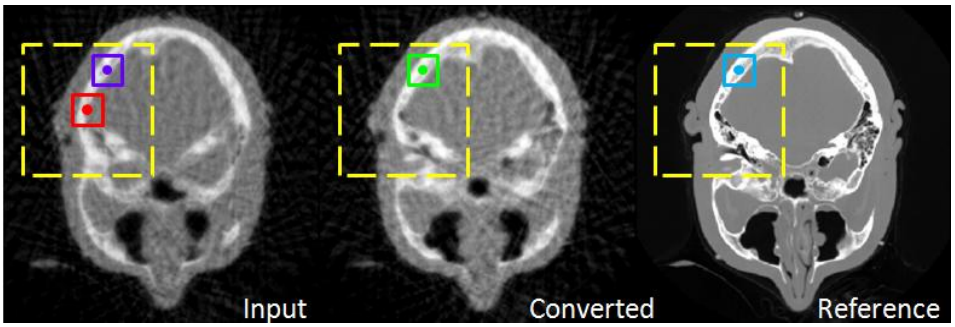


Fig. 1. Matching procedure (only one patch is shown in each image for ease of illustration). The red dot is the central (target) pixel of the yellow dashed window, which is the neighborhood area the filter operates on; the other dots and squares (purple, green, cyan) denote a pixel and its patch in the (original, converted reference, artifact-free reference) image.

regularization.

In the CT reconstruction literature, Yu et al. [9] proposed the PSRR method which replaces regions in a low-dose CT reconstruction with their embodiments in a normal-dose CT reconstruction, when unchanged. The NLM filter relaxes the need for registration as it (i) performs the registration on the fly via its patch-based search mechanism and (ii) utilizes features at a much smaller scale, exploiting the potentially large number of redundancies that might exist there. On the other hand, Kelm et al. [2] describe an approach that reconstructs volumes at two different thicknesses, from the same acquired projection data. They reconstruct the thicker slices at higher SNR, while the thinner slices have lower SNR. We extend this method to high-quality imagery not necessarily acquired simultaneously.

III. Methodology

Our method uses a set of artifact-free reference images R and their converted images C containing similar artifact types and distribution than the input image I. We call a reference image and its converted image a *reference image pair*. To obtain the reference image, a comprehensive population of common CT anatomical and pathological images is required. Further, to obtain the corresponding converted image a simulation must be performed to generate in the reference image every possible artifact type that may appear in the input image. This imagery then forms a comprehensive database D comprising both the clean reference images and every artifact-corrupted image.

A. The Reference-Based NLM (RNLM) Filter

Like the NLM, the RNLM works within a local neighborhood and computes the individual weight for each pixel inside this neighborhood as this pixel's contribution to the central pixel. However, there are two modifications (see Fig. 1). First, for NLM, the weight computation is based on the similarity between patches from the same input image I (comparing the red and the purple patches), while for RNLM, the central patch is still from input image I but the other patches are from the converted reference images C (comparing the red and the green patches). Second, the RNLM is a weighted sum of neighborhood pixels inside the reference image R (cyan dot in the reference image) and not from the input image I. Finally, the RNLM could have several reference image pairs to increase the effective patch-searching range. Formally:

$$I_x^{'} = \frac{\sum_{y\in W_x}\sum_{i\in RP}\exp\left(-\sum_{t\in P}G_a(t)\left|I_{x+t}-C(i)_{y+t}\right|^2/h^2\right)\cdot R(i)_y}{\sum_{y\in W_x}\sum_{i\in RP}\exp\left(-\sum_{t\in P}G_a(t)\left|I_{x+t}-C(i)_{y+t}\right|^2/h^2\right)} \quad (2)$$

where x and y are pixel positions, W_x is the neighborhood of x, P is the patch size of each pixel, RP is the set of reference image pairs, i and t are index of corresponding set, G_a is a Gaussian smoothing kernel and h is the parameter to control the overall smoothness of the filtering.

We choose the converted image instead of the clean reference image for patch matching since the latter would increase the patch difference, resulting in a much smaller weight after the exponential function is applied. By simulating the artifacts in the reference image we obtain a much more realistic comparison in the matching stage. Our results indicate that this strategy is well chosen.

We note that in our scheme RP also includes the input image, but only when the denominator of (2) is very close to zero. This occurs when the neighborhood size is not big enough to contain similar patches than the central patch, which is quite rare. In that case, RNLM degenerates to normal NLM.

In (2), the pixel and neighborhood positions are directly taken from the input image. Therefore, image alignment or registration should be performed in advance. For now, we have employed image slices that were adjacent to the target image and so this step has been omitted. In fact, when the neighborhood is big enough, a rough alignment is sufficient to perform the patch matching. Future work will extend our current (proof-of-concept) framework in this direction.

B. The Database

First, suitable reference image pairs must be identified by matching the input image in the database D. A typical database is shown in Fig. 2. In this paper, we consider only two artifact types commonly appearing in low-dose CT – streak artifacts due to too few projections in a full viewing range (180˚) and noise due to low-dose projections. Therefore, the corresponding database is three-dimensional with each of anatomic feature, streak and noise taking up one axis. More artifact types could also be added according to the specific configuration of the data generation.

In this database, along the anatomic axis only ideal images which are artifact-free and generated with a full number (180) of projections are listed (green strip in Fig. 2). This axis contains every related anatomical image on hand. These *ideal items* are

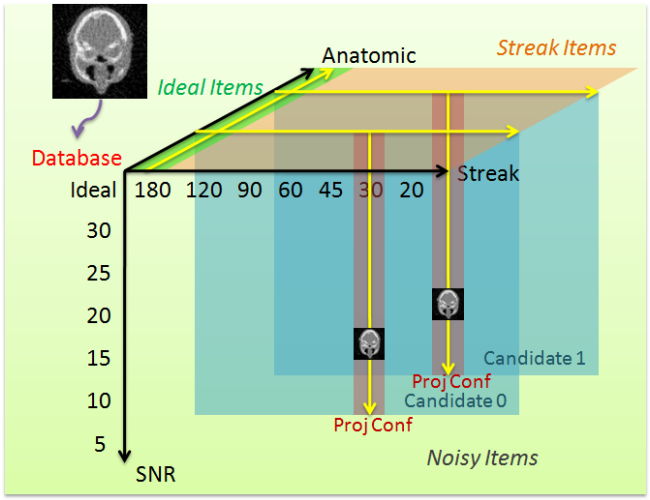


Fig. 2. A three dimensional database *D* for matching the input image: The anatomic axis (green strip) contains only ideal images which are artifact-free, generated with the full number (180) of projections and is used as reference images; the anatomic-streak plane (orange area) contains only streak images which are noise-free, and along the streak axis the number of projections used for reconstruction decreases so that the streak artifacts increase; the remaining space beside the orange plane contains different anatomic images with every combination of artifacts; along the SNR axis, the noise level increases.

the reference images used in the RNLM filter. Along the streak axis the number of projections decreases from 180 to 20, while along the SNR axis the noise level increases from noise-free (ideal) to SNR 5. The anatomic-streak plane (orange plane in Fig. 2) contains only noise-free images (we call *streak items*), while the remaining space contains *noisy items*, one for each ideal and streak item. Note that streak (noise) items with no streaks (noise) are items that apply in the regularization when only one of the two respective artifacts is present.

C. Database Matching

In our current work, we first find an appropriate reference image pair and then perform the regularization for the entirety of the target image. Given the (contaminated) input image, our reference image pair must satisfy two constraints: (i) the reference image must contain similar anatomical features than the input image and (ii) its converted counterpart must have similar artifact types than the input image. Finding such a matched reference image pair is equivalent to do high-dimensional space matching. The associated computations could be very complicated and time consuming. However, because of the special configuration of the database space, which is generated from items along the anatomy axis, it is sufficient to perform the search + match operation only along the yellow arrows in Fig. 2. We devise the search routine as shown in Fig. 3 to effectively reduce the computational complexity of the matching in terms of the number of loops from $O(A \cdot S \cdot N)$ to $O(A) + O(S) + O(N)$ to find each reference image pair, where A, S and N stand for the number of anatomic (we have not considered pathologies in the current work), streak and noise configurations, respectively. Thus, along each search direction (yellow arrow in Fig. 2) the candidate items only differ from each other in one aspect – the structural difference for ideal items, the streak difference for streak items and the noise level difference for noisy items. This allows for efficient similarity comparison.

1. for each *ideal item*
   Compute similarity value *VI* with input image *I*
   Insert into the ordered list according to *VI*
   Pick top *M* items as reference images *R*
2. for each reference image *R*
   for each *streak item* of *R*
   Compute similarity value *VS* with input image *I*
   Update current largest value
   Pick the streak item *S* with the largest *VS*
3. for each selected streak item *S*
   for each *noisy item* of *S*
   Compute similarity value *VN* with input image *I*
   Update current largest value
   Pick the noisy item N_o with the largest *VN*
4. Return *M* pairs of (R, N_o)

Fig. 3. Pseudocode for database matching.

1. for each noisy item of a selected streak item *S*
   Subtract the noisy item from *S* to get image *D*
   for each line of *D*
   Compute the absolute mean of the block area around that line
   A perturbation magnitude spectrum *P* is obtained
2. Subtract the input image from *S* to get image *ID*
   for each line of *ID*
   Compute the absolute mean of the block area around that line
   Obtain the perturbation magnitude *IP*
3. for each line of *IP*
   if *IP* value is within spectrum *P*
   Pick the noise level with the closest value
   else
   Set the line as ineffective
4. Pick the top 10% of the effective lines according to the closest value and return the most often selected noise level

Fig. 4. Pseudocode for the similarity metric for noisy items.

D. Similarity Metric

The choice of similarity comparison metric varies with image type. For ideal item comparison, since the items are artifact-free, regardless of the type of input image, the structural differences dominate the similarity evaluation. Therefore, any common image evaluation metric (RMS, CC, etc.) works well. Similarly, for the comparison of streak items, which are noise-free and structurally similar to the input image, the streak level difference dominates the evaluation.

However, using this same metric will make the comparison fail to find the similar noise level items. We expect a metric to evaluate the noise level as the magnitude of perturbation around the ideal image profile, while common metrics would boost such difference and end up suggesting noise-free items. To obtain the appropriate perturbation magnitude description, since we know the noise-free items of these noisy candidates, subtracting the noisy from the noise-free items will reveal the level of noise contained in each noisy item. So we perform the comparison as Fig. 4. For step 2, although the input image and the streak item are anatomically quite similar to one another, the subtraction not only reveals the noise level but also inevitably keeps the effect from structural difference. To eliminate such unwanted side

effects, we discard (in step 3) the line out of the range of the magnitude spectrum as one dominated by structural differences, and right after that (in step 4), we pick only the top 10% suggested solutions.

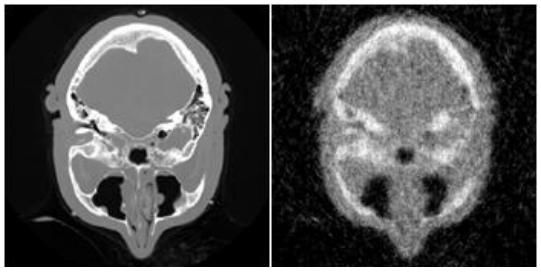


Fig. 5. One of two reference image pairs matched from the database (left: perfect, right: target artifact-matched low-dose reconstructions.

IV. Results

We tested our algorithm using the NIH Visible Human brain (size 256^3). To generate the database *D*, we first picked representative image slices from the volume to represent various anatomical items. Then after simulating the projections of each slice, we chose different numbers of projections (180,120, 90, 60, 45, 30, 20) and noise levels (noise-free, SNR 30, 25, 20, 15, 10, 5) in the projections to generate a complete database of corrupted and uncorrupted CT reconstructions using the OS-SIRT 5 reconstruction algorithm [5]. A new slice was chosen from the brain volume, 30 X-ray projections were simulated, noise with SNR 12 was added, and an OS-SIRT 5 reconstruction was performed. This (target) image is shown in Fig. 7b, along with its non-corrupted counterpart in Fig. 7a.

Using the database matching we obtained 2 reference image pairs (SNR 10 with 30 projections and SNR 10 with 20 projections, respectively). One of these pairs is shown in Fig. 5. To determine the noise level of these, the magnitude spectrum was evaluated (plotted in Fig. 6). Using the two reference image pairs, we regularized the target using traditional target image-based NLM (Fig. 7c), RNLM using the perfect reference images for matching (Fi. 7d), and RNLM using the target artifact-matched reference images for matching (Fig. 7e). All NLM use a 7×7 window and optimized parameter settings. We observe that RNLM is clearly superior to NLM, and that matching the target artifacts brings additional significant improvements (see for example, the locations pointed to by the three arrows). Each result is also evaluated by a perceptual quality metric E-CC (edge-based correlation coefficient) [8] in Fig. 7. In terms of computational performance, RNLM is similar to NLM, and with GPU acceleration its runtime is around 0.1s for a megapixel image. There is constant overhead for the indexing, while the database matching is a preprocessing step.

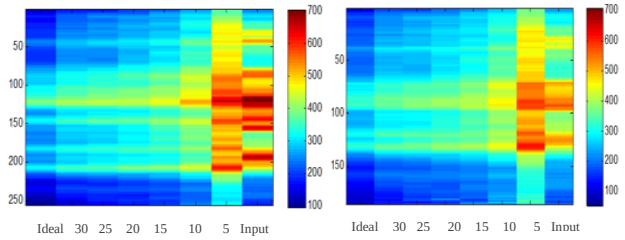


Fig. 6. The spectrum for mean magnitude of noise perturbation for the selected streak item under all noise levels: the left shows all lines of the images while the right shows only effective lines; from column 1 to 7 the noise level increases, while column 8 shows the mean magnitude of noise perturbation for the input image.

V. Conclusions and Future Work

We have extended NLM-filtering for regularization in an iterative CT reconstruction framework to reference image-based NLM filtering, where the reference image is another artifact-free image with similar anatomical features than the target. Future work will generalize this procedure to a pixel- or region-based matching approach, involving more than two reference-image pairs for regularization. This affects mostly the anatomy and pathology dimension. For the latter, we also plan to enrich the database with specific pathologies. Finally, we also plan to incorporate rough registration for better window placement.

References

[1] A. Buades, B. Coll, J.Morel, "A review of image denoising algorithms with a new one," Multi-Scale Modeling Simulation, 4(2): 490-530, 2005.

[2] Z. Kelm, D. Blezek, B. Bartholmai, B. Erickson, " Optimizing non-local means for denoising low dose CT," IEEE ISBI 662-665, 2009.

[3] G. Petchnigg, M. Agrawala, et al, "Digital photography with flash and no-flash image pairs," ACM Trans. Graph. 23, 3, 664–672, 2004.

[4] T. Welsh, M. Ashikhmin, K. Mueller, "Transferring color to greyscale images," ACM Trans. on Graphics, 21(3):277-280, 2002.

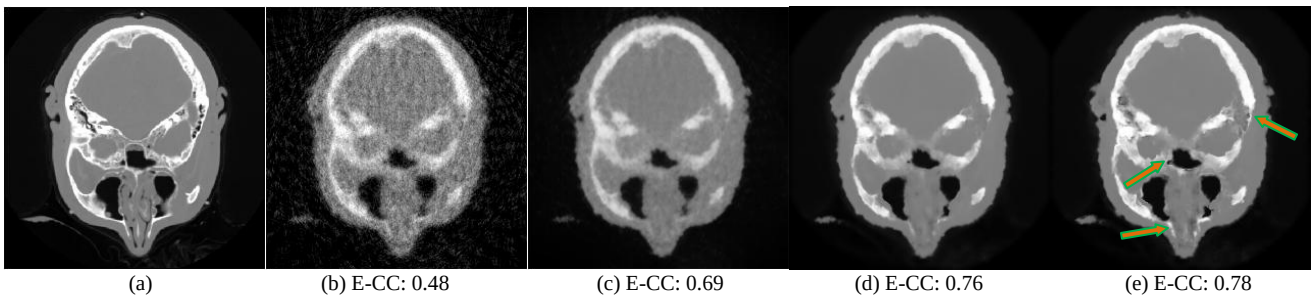
[5] F. Xu, W. Xu, D. Agard, K. Mueller et al, "On the Efficiency of Iterative Ordered Subset Reconstruction Algorithms for Acceleration on GPUs", Computer Methods and Programs in Biomedicine 98(3):261-270, 2010.

[6] W. Xu and K. Mueller, "Evaluating popular non-linear image processing filters for their use in regularized iterative CT", IEEE MIC, 2010.

[7] W. Xu and K. Mueller, "A Performance-Driven Study of Regularization Methods for GPU-Accelerated Iterative CT," HPIR, pp.20-23, 2009.

[8] W. Xu and K. Mueller, "Learning Efficient Parameter Settings for Iterative CT Reconstruction Algorithms," Fully3D, pp 251-254, 2009.

[9] H. Yu, S. Zhao, E. Hoffman, G. Wang, "Ultra-low dose lung CT perfusion regularized by a previous scan," Acad. Radiol. 16:363-373, 2009.



(a) (b) E-CC: 0.48 (c) E-CC: 0.69 (d) E-CC: 0.76 (e) E-CC: 0.78

Fig. 7. Target images: (a) perfect, (b) low-dose reconstruction; Regularization results: (c) traditional NLM, (d) reference image-based using the perfect reference images for matching (e) reference image-based using the target artifact-matched reference images for matching.