A Segmentation-Based Method for Metal Artifact Reduction²

Hengyong Yu, PhD, Kai Zeng, MS, Deepak K. Bharkhada, MS, Ge Wang, PhD, Mark T. Madsen, PhD, Osama Saba, PhD, Bruno Policeni, MD, Matthew A. Howard, MD, Wendy R. K. Smoker, MS, MD, FACR

Rationale and Objectives. We propose a novel segmentation-based interpolation method to reduce the metal artifacts caused by surgical aneurysm clips.

Materials and Methods. Our method consists of five steps: coarse image reconstruction, metallic object segmentation, forward-projection, projection interpolation, and final image reconstruction. The major innovations are 2-fold. First, a state-of-the-art mean-shift technique in the computer vision field is used to improve the accuracy of the metallic object segmentation. Second, a feedback strategy is developed in the interpolation step to adjust the interpolated value based on the prior knowledge that the interpolated values should not be larger than the original ones. Physical phantom and real patient datasets are studied to evaluate the efficacy of our method.

Results. Compared to the state-of-the-art segmentation-based method designed previously, our method reduces the metal artifacts by 20–40% in terms of the standard deviation and provides more information for the assessment of soft tissues and osseous structures surrounding the surgical clips.

Conclusion. Mean shift technique and feedback strategy can help to improve the image quality in terms of reducing metal artifacts.

Key Words. Metal artifacts; mean shift filtering; image segmentation; projection interpolation; feedback.

In x-ray computed tomography (CT), the attenuation coefficient of high-density objects, such as surgical clips, metal prostheses, or dental amalgams, is much higher than that of soft tissues and osseous structures. Because the x-ray beam is highly attenuated by metals, an insufficient number of photons reach the detector, producing corrupted projection data. Consequently, images reconstructed by the traditional filtered back-projection (FBP) method are marred by starburst artifacts, often referred to as metal artifacts. These artifacts significantly degrade CT image quality and limit the usefulness of CT for many clinical applications because tissues in the plane of the metal appliance are severely obscured. Hence, there is an important need for methods that reduce metal artifacts.

The effects of metallic objects on x-ray scanning are 2-fold: beam hardening, due to the poly-energetic x-ray spectrum, and a poor signal-to-noise ratio from photon starvation. To suppress the metal artifacts, iterative reconstruction methods have been successfully applied that avoid the corrupted data. For example, Wang et al. used the maximum expectation maximization (EM) formula and algebraic reconstruction technique (ART) to iteratively deblur metallic artifact (1, 2). A key step in their algorithm is the introduction of a projection mask and the
computation of a 3D spatially varying relaxation factor that allows compensation for beam divergence and data incompleteness. However, this approach is computationally expensive and not practical for clinical imaging. Conventional FBP methods (3) are computationally efficient but produce image artifacts when complete and precise projection data are unavailable. Different linear and polynomial interpolation techniques have been developed for estimating the “missing” projection data (4–11). The major task of this method is to identify the corrupted segments in the sinogram and interpolate these data from noncorrupted neighbor projections. Because the first step of all the above methods requires segmenting of the metal parts from a coarse image reconstructed by FBP, segmentation is a key technique for metal artifact reduction.

Most researchers have adopted simple threshold methods to segment the metallic objects and then forward-projected them into the sinogram domain. These projections serve as a mask for interpolation and reconstruction (4–6, 10). The threshold-based methods may produce inaccurate segmentation of the metallic objects and, hence, the information from structures surrounding the metal may be lost. To improve the accuracy of this technique, Yazdi et al. (11) proposed to automatically detect and reconnect the edges of the metal projection in the original sinogram. The idea was to apply the interpolation scheme between the two corresponding projected edges belonging to the projection regions of the same object. While this method works well for metallic hip prostheses (11), it is not suitable for irregular or nonconvex objects. Recently, Bal and Spies (12) developed a novel method to reduce artifacts by in-painting missing information into the corrupted sinogram. The missing sinogram information was derived from a tissue-classes model extracted from the corrupted image, and a k-means clustering technique was adopted for image segmentation. However, this method strongly depends on the prior tissue model.

Here, we describe a novel segmentation-based interpolation method to suppress artifacts from metallic aneurysm clips. The major contributions include the use of a sophisticated mean-shift technique and a restriction on the interpolation values. The mean-shift technique is the state-of-the-art approach for feature space analysis in the computer vision field (13, 14). Because the mean shift technique is not sensitive to image noise and metal artifacts, it can provide high segmentation accuracy. Also, because the goal of the interpolation is to improve the projection data affected by the metallic objects, prior information is used to ensure that the interpolated values are consistent to the normal data, especially when multiple metallic objects are in the field of view. This is accomplished with a feedback strategy in the interpolation step to adjust the interpolated values automatically.

**METHODS**

**Algorithm Description**

CT scans were performed on a Siemens SOMATOM Sensation 16 scanner using helical scanning geometry. After the dataset was acquired, we applied the single-slice rebinning method (15) to convert the multislice dataset to a stack of fan-beam sinograms, each associated with one horizontal z-slice. Once the fan-beam sinograms were generated, metal artifact reduction and image reconstruction were performed for each z-slice in the 2D fan-beam geometry as illustrated in Figure 1. The fan-beam sinogram can be represented as $P_{\beta}(\gamma)$, where $\beta$ is an angle that indicates the position of the x-ray source, $\gamma$ is the fan-angle that refers to detector location, and the subscript $O$ represents the original dataset. For a full scan dataset from the Siemens SOMATOM Sensation 16 scanner, $\beta$ is discretized as $\beta_i$ with $i = 1, 2, \ldots, 1160$, and $\gamma$ is discretized as $\gamma_j$ with $j = 1, 2, \ldots, 672$. Our improved segmentation-based method for metal artifact reduction has steps that are similar to state-of-the-art methods (10). As shown in Figure 2, our method can be summarized into five steps for each fan-beam projection dataset.

**Step 1: Coarse image reconstruction**

For a specified region of interest (ROI) in the field of view (FOV), a coarse image $I_{O}(m,n)$ was reconstructed by the conventional full-scan FBP algorithm, where $m = 1, 2, \ldots, M$ and $n = 1, 2, \ldots, N$. Because the ROI is fixed, the
higher the $M$ and $N$, the smaller the pixel size. For the selection of $M$ and $N$, our general rule was to make each pixel represent a cubic region in the final 3D volume image.

**Step 2: Metallic object segmentation**

Using the state-of-the-art computer vision mean-shift technique (13, 14), we segmented the metallic objects from the coarse image $I_O$ and obtained a characteristic image $I_C$, defined as

$$I_C(m, n) = \begin{cases} 
1 & \text{if } I_O(m, n) \text{ belongs to metallic objects} \\
0 & \text{otherwise} 
\end{cases},$$

where the subscript “C” denotes a characteristic function. $I_C$ functions as an index for the metallic objects in the specified ROI. In the next subsection, we describe this procedure in detail.

**Step 3: Forward-projection**

In the same fan-beam geometry, we forward-projected the characteristic image $I_C$ into the sinogram domain and obtained $P_M(\beta, \gamma)$, which represents the intersectional length of the corresponding x-ray path with the metallic objects. We defined characteristic projection by

$$P_C(\beta, \gamma) = \begin{cases} 
1 & P_M(\beta, \gamma) > 0 \\
0 & \text{otherwise} 
\end{cases}. 
(2)$$

$P_C$ functions as an index to specify which pixel in the original sinogram had been corrupted by the high-intensity metallic objects.

**Step 4: Projection interpolation**

The corrupted projection data were deleted from the original projection $P_O$. From the feedback interpolation strategy (see later subsection for detail), we obtained a new interpolated projection $P_I$, where the subscript $I$ represents interpolation. A difference projection dataset $P_D$ was obtained by

$$P_D(\beta, \gamma) = P_O(\beta, \gamma) - P_I(\beta, \gamma). 
(3)$$

Obviously, $P_D$ was completely formed by the metallic objects.

**Step 5: Final image reconstruction**

A background image, $I_B$, and a metallic image, $I_M$, were reconstructed by the traditional FBP method from the interpolated projection, $P_I$ and difference projection $P_D$, respectively. The final image was composed by a scale scheme (16) as

$$I_F(m, n) = I_B(m, n) + \eta \times I_C(m, n) \times I_M(m, n),$$

where $\eta$ is the scale factor, ranging from 0.05 to 0.5, and $I_C$ functions as a mask to protect the background image from corruption by the metallic artifacts. To smooth the edge and preserve the structure of the metallic objects, $I_C$ is defined as the 2D convolution of the characteristic image $I_C$ and a normalized Gaussian kernel.

**Metallic Object Segmentation**

The mean shift technique was first proposed in 1975 by Fukunaga and Hostetler (17) and largely forgotten until Cheng’s work rekindled interest in it in 1995 (13). Based on the mean shift technique, in 2002, Comaniciu and Meer proposed a nonparametric method for analysis of a complex multimodal feature space and to delineate arbitrarily shaped clusters within it (14). The technique was successfully applied for discontinuity, preserving smoothing and image segmentation, and raised a hot topic in the computer vision field. Here, we use this technique to segment the metallic objects from the coarse image $I_O$.

Let $x_k$, $k = 1, 2, \ldots, K$, be a point in one finite set $\mathbf{X}$ (the “data” or “sample”) in the $d$-dimensional space $\mathbb{R}^d$. Let $G$
be a kernel and \( x \rightarrow (0, \infty) \) be a map \( W : x \rightarrow (0, \infty) \).

The sample mean with kernel \( G \) at \( x \) is defined as (13),

\[
M_s(x) = \frac{1}{\sum_{k=1}^{K} G(x_k - x) W(x_k) x_k}.
\] (5)

Let \( T \subseteq X \) be a finite subset. The evolution of \( T \) in the form of iterations \( t \leftarrow M_s(T) \) with \( M_s(T) = \{ M_s(t) : t \in T \} \) is called a mean shift procedure, and the difference \( M_s(t) - t \) is called mean shift. For each iteration, \( t \leftarrow M_s(t) \) is performed for all \( t \in T \) simultaneously. For each \( t \in T \), there is a sequence, \( t, M_s(t), M_s(M_s(t)), \ldots \), that is called the trajectory of \( t \). The weight function \( W(x_k) \) can be either fixed throughout the process or re-evaluated after each iteration. The algorithm halts when it reaches fixed points \( M_s(T) = T \).

For our application, the 2D gray-level image can be regarded as the 3D sampling points, that is, \( d = 3 \). The space of the 2D coordinate lattice is known as the spatial domain while the gray-level is called the range domain. The joint 3D spatial-range vector set of the coarse image \( I_o \) is used to form the finite set \( X \) and \( T \). Correspondingly, the size of the set is the total pixel number of \( I_o \), that is \( K = M \times N \). As pointed out by Comaniciu and Meer (14), an Epanechnikov kernel or a truncated normal kernel always provides satisfactory performance. Hence, we selected the multivariate Epanechnikov kernel, which is defined as the product of two radially symmetric kernels and the Euclidean metric allowing a single bandwidth parameter for each domain (14),

\[
G(x) = \frac{C}{h_i h_r} g\left(\frac{\|x'/h_i\|^2}{\|x'/h_r\|^2}\right),
\] (6)

with the common profile

\[
g(s) = \begin{cases} 1 & |s| < 1 \\ 0 & \text{otherwise} \end{cases}.
\] (7)

In Equation 6, \( x' \) is the spatial component, \( x' \) is the range component, \( h_i \) and \( h_r \) are the kernel bandwidth parameters, and \( C \) is the corresponding normalized constant. The weighting function is fixed throughout all iterations. To construct the weighting \( W \), we first computed the gradient image \( I_G \) of \( I_o \), which is defined as

\[
I_G(m, n) = \left| \nabla I_o(m, n) \right| = \sqrt{\left( \frac{\partial I_o(m, n)}{\partial m} \right)^2 + \left( \frac{\partial I_o(m, n)}{\partial n} \right)^2}.
\] (8)

Because the mean shift procedure preserves discontinuities, a small weight was assigned to a pixel in discontinuous regions, while a large weight was assigned to that in flat regions. Hence, the weighting function \( W \) was constructed as:

\[
W(m, n) = 1 - \gamma I_G(m, n)/\max(I_G),
\] (9)

where \( \max(I_G) \) is the maximum of the gradient image \( I_G \) and \( 0 \leq \gamma < 1 \) is a scale factor. As a result, all the necessary components have been determined for the mean shift technique.

Let \( x_k \) and \( z_k \) be the 3D input and output points in the joint spatial-range domain with \( k = 1, 2, \ldots, K \). Based on the work of Comaniciu and Meer (14), the mean shift smoothing filtering for each point can be performed as:

1. Initialize \( l = 1 \) and \( y_{k, l} = x_k \);
2. Compute \( y_{k, l+1} = M_s(y_{k, l}) \) according to Eq.(5) until converging to the final result \( y_{k, L} \);
3. Assign the output \( z_k = (x_k, y_{k, L}) \).

After the coarse image has been smoothed by preserving the discontinuity, the metallic object segmentation can be implemented by the grouping method proposed in subsection 4.2.1 in Comaniciu and Meer (14). Noting the fact that metallic objects have higher attenuation than human tissues, the simple threshold method can also be employed after mean shift smoothing filtering. For briefness, we used the latter in our study. To compensate for the blurring of the projection introduced by the focal spot of the x-ray source, a morphological dilation was performed to the above segmentation result. Finally, the characteristic image \( I_t \) was obtained. Figure 3 presents some typical interim images to illustrate the mean shift technique.

We make the following three comments on the implementation of the mean shift filtering. First, the constant \( C \) in Equation 6 can be omitted since it has been canceled in Equation 5. For the same reason, \( W \) can also be defined as \( W(m, n) = \max(I_G) - \lambda I_G(m, n) \). Second, many techniques can be used to reduce the computational time.
required for the mean shift technique (14). Third, the final segmentation is not sensitive to the selection of the threshold because the mean shift filtering sharpens the edges of the metallic objects.

**Feedback-Based Interpolation Strategy**

After the characteristic projection function \( P_C \) has been determined in the third step of the proposed method, we delete the corrupted data and interpolate them via linear or polynomial interpolation in the fourth step. It is well known that a metallic object corrupts a strip-like region in the projection domain. When there is only one metallic object in the field of view, it is reasonable to interpolate the missing corrupted data. However, when there are multiple metallic objects inside the FOV, the reliability of the interpolated value is lower, especially when it is near the overlap region of multiple strips. To improve the interpolation accuracy, we propose a feedback strategy for the interpolation based on prior information.

Assume that the original projection data are \( P_O(\beta, \gamma) \) and interpolated value is \( P_I(\beta, \gamma) \). Because the interpolation is performed only for the region in which the projections have been corrupted by the high-density metallic objects, the difference projection \( P_D(\beta, \gamma) \) should be not be smaller than zero. That is, the interpolated value \( P_I(\beta, \gamma) \) should not be larger than \( P_O(\beta, \gamma) \). This fact provides the prior information in our feedback strategy. The feedback strategy can be described by the following procedure:

1. Perform the interpolation for the missing corrupted dataset indexed by \( P_C \);
2. Compare the original and interpolated values and set \( P_C(\beta, \gamma) = 0 \) if \( P_I(\beta, \gamma) > P_O(\beta, \gamma) \);

![Figure 3. Illustration of metallic object segmentation using the mean shift technique.](image-url)

(a) The original coarse image, (b) weighting function, (c) image after mean shift filtering and (d) final metallic object characteristic function.
3. Repeat (i) and (ii) until all the interpolated values are no larger than the original ones.

In our study, the linear interpolation was adopted since the strips of metallic clips in the projection domain are very thin. For more complex applications, our feedback scheme can be used directly without any modification. Figure 4 shows two representative interpolated profiles with and without the feedback strategy.

RESULTS

Algorithm Implementation

We set up a platform to reconstruct and display images from data collected along a helical scan. First, the sinogram datasets were rebinned from spiral cone-beam geometry into fan-beam geometry. Then, images were reconstructed using a conventional fan-beam FBP algorithm. Between these two steps, we inserted some additional processing steps proposed in this paper to reduce the metal artifacts. All the reconstructed images were converted into the standard Hounsfield units (HU) by assuming the attenuation coefficient of water was 0.019/mm. The proposed method was tested by both a physical clip phantom and clinical patient datasets. The corresponding parameters were summarized in Table 1.

As a benchmark, the segmentation-based interpolation method proposed by Wei et al. (10) was also implemented. There are three major differences between Wei et al.’s method and ours. First, Wei et al. used an average filtering plus the threshold method to segment the metallic objects, while we adopted the mean shift filtering plus threshold method and morphological dilation. Second, Wei et al. used the polynomial interpolation for the missing values, while we used linear interpolation along with the feedback scheme. Third, Wei et al. generated the final image with the mask \( I_C \) (see Equation 4), while we substituted \( I_C \) with \( I_C / H \) to smooth the edge and preserve the structure of the metallic objects. For comparison, both the threshold for metallic segmentation and the scale factor \( \eta \) in the last step were set to be the same values for both methods.

Clip Phantom Experiment

The phantom was a water-filled plastic cylinder with an inside diameter of 21.6 cm and a length of 18.6 cm with a wall thickness of 3.2 mm. An 8-mm plastic rod was located in the center of the phantom, and five different aneurysm clips were attached with rubber bands to the rod. The axial separation between clips was at least 2 cm. The phantom was scanned with clinical brain CTA parameters: 120 kVp and 496 mAs. Figure 5 shows a typical phantom slice reconstructed by the proposed method.
and Wei et al.’s segmentation-based interpolation method (10). From the local magnified images, it is observed that the bright artifacts around the clip are better suppressed by our method compared to Wei et al.’s. This results from the more accurate metallic object segmentation, provided by the mean shift filtering, and the feedback strategy that provides improved correction of the interpolated values. Note the intensities of artifacts surrounding the metal vary significantly more than elsewhere, and the tissues are of major diagnostic interest, a standard deviation (SD) is utilized to quantify the improvement of the proposed method.
where $D_C$ is a characteristic function indicating the surrounding region of the metallic objects and $T$ is the ideal average gray level (HU) around the metallic objects. Let $I^C$ denote the 2D convolution of the characteristic image $I$ and a specific kernel function, $D_C$ can be obtained by subtracting $I^C$ from the characteristic function of $I$. Because the phantom was filled with water, we set $T = 0$ HU and select $D_C$ as the last row in Figure 5. The SDs from our clip phantom images in Figure 5 are listed in Table 2.

**Patient Study**

Under the approval of the institutional review board committee, the University of Iowa, we obtained the raw data from a clinical study at our institution. The patient was a 59-year-old woman with a known history of hypertension. She presented with a “worst headache of life” and right hemibody numbness. When she was first seen in the emergency department, her Glasgow Coma Scale score was 10 and the World Federation of Neurologic Surgeons score was 4. A head CT scan was obtained showing diffuse subarachnoid hemorrhage in the basal cisterns and Sylvian fissures, left greater than right. CT angiography demonstrated a left middle cerebral artery aneurysm. She was taken to the operation room and the aneurysm was clipped. She had numerous head CT scans after surgery for assessment of increased intracranial pressure to rule out rebleeding and hydrocephalus. Figure 6 presents the reconstructed images of a representative CT slice. Comparing the local magnification of the image reconstructed by our method and the state-of-the-art method (10), our method provides better quality with fewer artifacts around the metallic clip. Because the metallic clip was surrounded by soft tissues, we set $T = 30$ HU and selected the characteristic function as in the bottom row in Figure 6 for quantitative measurement. The quantified results are listed in Table 2, which shows that our method reduces the artifacts by 43% of the SD compared to the state-of-the-art method (10). This enabled us to visualize soft tissue structures around the metallic clips and improve clinical diagnostic accuracy.

**DISCUSSION**

In the proposed metal artifact reduction method, there are several parameters that needed to be specified empirically. General speaking, the larger the bandwidth parameters $h_r$ and $h_s$, the sharper is the discontinuity of the metallic objects. The selection of $h_s$ should be larger than the size of a pixel. On the other hand, metallic objects will be erased if their sizes are in the same order as $h_s$. Four different techniques can be considered to optimize the bandwidth parameters (14). The local adaptive solutions can also be used to solve the difficulties generated by the narrow peaks and the tails of the underlying density (18). Because the mean shift procedure smoothes the whole image while preserving the discontinuity, the selection of a threshold to segment metallic objects is more flexible than the conventional methods (4–6, 10). The general rule is that the threshold should be smaller than the metallic gray level and larger than the maximum gray level of other human tissues.

As an initial study, our method was implemented in MatLab on a regular PC (2.8 GHz CPU). The total computational time was about 10 minutes to process one 512×512 slice. The forward-projection step accounted for most of this time. Currently, a ray-driven method was adopted for the forward-projection of the characteristic function $I^C$ into the projection domain. In the future, we plan to use a voxel-driven method for the forward-projection and only project the metallic object parts of the characteristic function $I^C$. Furthermore, we may transplant the codes into C++. Hence, the total time will be reduced in an order of magnitude, making the final time less than one minute. If it is necessary, parallel-computing techniques can be used for real-time or near real-time performance.

**CONCLUSION**

We have proposed a clinically feasible approach for surgical aneurysm clip artifact reduction, which reduced metal artifacts by 20–40% in terms of the standard devia-
tions with the soft tissues and osseous structures surrounding the metallic clips. Compared to the conventional segmentation-based methods, our method has two distinguishing features: mean shift filtering for metallic object segmentation and feedback-based interpolation for projection data consistency. The method may be also applied in other applications such as dealing with metallic prostheses and dental amalgams. It is hypothesized that the imaging performance in the 3D case will be similar to that in the 2D case. Further efforts will be made to improve effective metal artifact reduction techniques and bring them into clinical arenas.

**Figure 6.** Representative images of the patient head reconstructed using different methods. The left column image were directly reconstructed from the fan-beam sinogram without any correction. The middle column images were corrected using the conventional segmentation-based interpolation method. The right column images were corrected using our proposed method. The middle row images are the local magnification of the top row images. The bottom row images are the characteristic functions for quantitative measurement of the middle row images. The significant differences between (e) and (f) are indicated by circles. The display windows are [-100,300] HU.
REFERENCES