CT metal artifact reduction using MR image patches

Nielsen, Jonathan Scharff; Edmund, Jens Morgenthaler; Van Leemput, Koen

Published in:
Proceedings of SPIE

Link to article, DOI:
10.1117/12.2293815

Publication date:
2018

Document Version
Publisher's PDF, also known as Version of record

Link back to DTU Orbit

Citation (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
CT metal artifact reduction using MR image patches

Jonathan S. Nielsen, Jens M. Edmund, Koen Van Leemput


Event: SPIE Medical Imaging, 2018, Houston, Texas, United States
CT Metal Artifact Reduction Using MR Image Patches

Jonathan S. Nielsen\textsuperscript{a,b}, Jens M. Edmund\textsuperscript{b,c} and Koen Van Leemput\textsuperscript{a,d}

\textsuperscript{a}Department of Applied Mathematics and Computer Science, Technical University of Denmark, Lyngby, Denmark.
\textsuperscript{b}Radiotherapy Research Unit, Department of Oncology, Gentofte and Herlev Hospital, Herlev, Denmark.
\textsuperscript{c}Niels Bohr Institute, University of Copenhagen, Copenhagen, Denmark.
\textsuperscript{d}Department of Radiology, Massachusetts General Hospital, Harvard Medical School, Boston MA, USA.

ABSTRACT

Metal implants give rise to metal artifacts in computed tomography (CT) images, which may lead to diagnostic errors and erroneous CT number estimates when the CT is used for radiation therapy planning. Methods for reducing metal artifacts by exploiting the anatomical information provided by coregistered magnetic resonance (MR) images are of great potential value, but remain technically challenging due to the poor contrast between bone and air on the MR image. In this paper, we present a novel MR-based algorithm for automatic CT metal artifact reduction (MAR), referred to as kerMAR. It combines kernel regression on known CT value/MR patch pairs in the uncorrupted patient volume with a forward model of the artifact corrupted values to estimate CT replacement values. In contrast to pseudo-CT generation that builds on multi-patient modelling, the algorithm requires no MR intensity normalisation or atlas registration. Image results for 7 head-and-neck radiation therapy patients with T1-weighted images acquired in the same fixation as the RT planning CT suggest a potential for more complete MAR close to the metal implants than the oMAR algorithm (Philips) used clinically. Our results further show improved performance in air and bone regions as compared to other MR-based MAR algorithms. In addition, we experimented with using kerMAR to define a prior for iterative reconstruction with the maximum likelihood transmission reconstruction algorithm, however with no apparent improvements.

Keywords: Computed Tomography, Metal Artifact Reduction, Bayesian modeling, Radiation Therapy

1. INTRODUCTION AND PURPOSE

CT images of patients with metal implants often suffer from severe streak and cupping artifacts, potentially leading to dosimetric errors in radiation therapy (RT) where the CT is used for patient specific electron density or mass stopping power estimation.\textsuperscript{1} Metal implants amplify the effects of beam hardening and noise, which are both major sources of artifacts in filtered back projection (FBP),\textsuperscript{2} the most widespread CT reconstruction algorithm. For this reason, a number of MAR algorithms have been proposed in the literature.\textsuperscript{3} The most straightforward type of method is image-based, aiming to directly replace CT values in corrupted regions of already reconstructed CT images, typically using segmentation or anatomical prior knowledge.\textsuperscript{3} A far more widespread method is sinogram (raw CT data) correction, which treats the x-ray measurements acquired through metal as missing data. The missing data are sometimes simply interpolated, sometimes estimated by forward projecting through a prior image generated by an image-based method.\textsuperscript{1,4,5} The clinically used oMAR algorithm (MAR for Orthopedic implants, Philips Healthcare), for instance, combines image space segmentation and projection completion to post-process artifact-corrupted images in an iterative algorithm.\textsuperscript{6} Such methods potentially introduce new artifacts as they impose prior information of limited quality on either the image or sinogram.\textsuperscript{7} A newer family of methods, model-based iterative reconstruction techniques (MBIR), attempt to include the underlying causes of the artifacts in the data acquisition model used for CT reconstruction. Although such techniques are slow and may sometimes require unavailable information such as the x-ray source spectrum and implant metal composition, incorporating prior knowledge of the expected CT reconstruction has been shown to help mitigate such issues.\textsuperscript{7-11}

A promising source of such prior information is a coregistered MR scan that is acquired for e.g., tumor delineation in head-and-neck patients planned for RT. In such patients, the CT often suffers from metal artifacts due to dental implants.\textsuperscript{1} Since metal artifacts are often less pronounced and more localised on MR images, a coregistered MR scan can provide useful tissue information in areas where the CT is corrupted. Using MR for MAR is a relatively new idea that has mainly been investigated with an image-based approach, e.g. by finding replacement CT values in local windows guided by MR voxel intensity differences,\textsuperscript{12} or by creating a pseudo-CT (pCT) in which replacement CT values are assigned to discrete MR image segmentations.\textsuperscript{13} Because these
methods only use local MR intensity similarity to predict CT values, however, they typically produce errors in bone and air regions which both appear dark on MR images acquired with conventional sequences. We therefore propose in this paper a novel MR-based MAR algorithm that combines MR information in larger larger spatial neighborhoods with the local, corrupted CT values to make more accurate CT value predictions. Our approach is based on methods developed for pseudo-CT generation in MR-only radiotherapy and PET/MR attenuation correction applications. Here, CT values are predicted from image patches (clusters of neighboring voxels) in coregistered MR scans, using regression models learned from a database of matching MR patches and CT values. Our method combines such a regression model trained on CT value/MR patch pairs from the uncorrupted part of the patient, with a probabilistic forward model of the artifact-corrupted measurements to predict the true CT values using Bayesian inference.

We further provide an expectation maximisation (EM) algorithm for automatically choosing the regression model’s hyperparameters using Empirical Bayes estimation. This renders the method fully automatic and optimises the algorithm for each individual patient. Since the resulting method only requires data from the patient targeted for MAR, it avoids any inter-subject issues that arise during pCT generation such as the requirement for MR intensity normalisation or time-consuming atlas registration of the target patient. We will refer to our method as “kernel regression MAR” (kerMAR) in the remainder. It can be used both as-is or to generate a prior for subsequent model based iterative reconstruction. In this paper we will use the maximum likelihood transmission reconstruction (MLTR) algorithm for this, and refer to the resulting combination of kerMAR with MLTR as MLTR-k.

The artifact reduced images resulting from applying both algorithms are presented for 7 head-and-neck RT patients, and compared to those of the clinically used oMAR algorithm. To investigate the benefits of including the corrupted CT measurements for MAR, we additionally compare against pCT.

2. METHOD

2.1 Generative model

Consider a patient for which a medical CT and MR volume have been acquired and (assumedly) perfectly coregistered, associating each voxel pair with a unique index \( i \) from the set \( T \). In the presence of metal implants, the CT is corrupted by artifacts and so the CT measurements \( \{ t_i \}_{i \in T} \) may be incorrect; the problem of metal artifact reduction can then be viewed as estimating the true, unknown CT values \( \{ y_i \}, \forall i \in T \).

To achieve this goal, we establish the probabilistic relationship between \( y_i \) and \( t_i \) as well as the MR measurements. Contrary to CT, MR provides little contrast between bone and air, so the single voxel MR measurement is ambiguously related to \( y_i \) in such regions. We therefore extract larger spatial contexts from the MR image, using as our MR measurement for voxel \( i \) the patch \( m_i \) of size \( d = M^{1/3} \), an \( M \)-dimensional vector of MR intensity values from a \( d \times d \times d \) cuboidal window centered on the voxel. We then model the joint distribution of \( \{ m_i, y_i, t_i \}, \forall i \in T \), given hyperparameters \( \lambda = \{ \beta_m, \beta_y, \beta_t \} \):  

\[
p(\{ y_i, t_i, m_i \}| \lambda) = \prod_{i \in T} p(\{ y_i, t_i, m_i \}| \lambda) \tag{1}
\]

where  

\[
p(y_i, t_i, m_i| \lambda) = p(t_i| y_i, m_i, \lambda)p(y_i, m_i| \lambda).
\]

We learn \( p(y_i, m_i| \lambda) \) from samples. For this purpose we pick for each voxel \( i \in T \) a subset of indices \( A_i \) far from the metal implants with assumed uncorrupted CT values \( \forall n \in A_i : t_n = y_n \) and extract the CT value / MR patch pairs \( \{ y_n, m_n \}_{n \in A_i} \). \( p(y_i, m_i| \lambda) \) is then estimated using kernel density estimation on this dataset. Using Gaussian kernels with diagonal covariance matrices with components \( \beta_m^{-1}I_M \) and \( \beta_y^{-1} \), that thus factor into separate Gaussians, we get the kernel density estimate:  

\[
p(y_i, m_i| \lambda) \approx \frac{1}{|A_i|} \sum_{n \in A_i} \mathcal{N}(m_i|m_n, \beta_m^{-1}I_M)\mathcal{N}(y_i|y_n, \beta_y^{-1}). \tag{2}
\]
\( \mathcal{N}(\cdot | \nu, \Sigma) \) here denotes a Gaussian with mean \( \nu \) and covariance matrix \( \Sigma \), while \( \mathbf{I}_M \) is the identity matrix of size \( M \). 

To model \( p(t_i | y_i, \mathbf{m}_i, \lambda) \), we assume the artifacts add zero mean Gaussian noise to \( y_i \), making \( t_i \) independent of \( \mathbf{m}_i \) given \( y_i \). We let the variance of the artifact noise be position dependent, letting it vary with position as \( \beta^{-1}_t(x_i) = f(x_i) | \beta^{-1}_t | \), where \( \beta^{-1}_t \) is one of the hyperparameters of the model. \( f(x) \) is here a user-specified function with \( 0 \leq f(x_i) \leq 1 \) that scales the variance of the artifact noise and thus the credibility of the CT measurements, quantifying the observation that artifact corruption decays with distance to the metal implants. We define \( f(x) \) in section 3. Suppressing the position dependence of \( \beta_t \) in the notation, the measured CT value is thus modelled as:

\[
p(t_i | y_i, \mathbf{m}_i, \lambda) = \mathcal{N}(t_i | y_i, f(x_i) | \beta^{-1}_t |) = \mathcal{N}(t_i | y_i, \beta^{-1}_t),
\]

where the \( \mathbf{m}_i \) dependence disappeared from the right-hand side due to the conditional independence of \( t_i \) and \( \mathbf{m}_i \).

### 2.2 Inference of the uncorrupted CT values for kerMAR

Given the set of MR patches and CT values \( \{ \mathbf{m}_i, t_i \}, \forall i \in \mathcal{T} \), we wish to infer \( \{ y_i \} \). Using Bayes’ rule, we have that:

\[
p(\{ y_i \} | \{ \mathbf{m}_i, t_i \}, \lambda) = \frac{p(\{ \mathbf{m}_i, y_i, t_i \} | \lambda)}{p(\{ \mathbf{m}_i, t_i \} | \lambda)}
\]

with

\[
p(\{ \mathbf{m}_i, t_i \} | \lambda) = \prod_{i \in \mathcal{T}} p(\mathbf{m}_i, t_i | \lambda)
\]

and the marginal likelihood

\[
p(\mathbf{m}_i, t_i | \lambda) = \int_{y_i} p(\mathbf{m}_i, y_i, t_i | \lambda) dy_i
\]

\[
= \int_{y_i} \mathcal{N}(t_i | y_i, \beta_t^{-1}) \left[ \frac{1}{|\mathcal{A}_i|} \sum_{n \in \mathcal{A}_i} \mathcal{N}(y_i | y_n, \beta_y^{-1}) \mathcal{N}(\mathbf{m}_i | \mathbf{m}_n, \beta_m^{-1} \mathbf{I}_M) \right] dy_i
\]

\[
= \frac{1}{|\mathcal{A}_i|} \sum_{n \in \mathcal{A}_i} \left[ \int_{y_i} \mathcal{N}(t_i | y_i, \beta_t^{-1}) \mathcal{N}(y_i | y_n, \beta_y^{-1}) dy_i \right] \mathcal{N}(\mathbf{m}_i | \mathbf{m}_n, \beta_m^{-1} \mathbf{I}_M)
\]

\[
= \frac{1}{|\mathcal{A}_i|} \sum_{n \in \mathcal{A}_i} \mathcal{N}(t_i | y_i, \beta_t^{-1} + \beta_y^{-1}) \mathcal{N}(\mathbf{m}_i | \mathbf{m}_n, \beta_m^{-1} \mathbf{I}_M).
\]

where we in the last step recognised the integral as a convolution over Gaussians, leading to a new Gaussian with the variances added (see e.g. [21, p. 112]). Inserting this result along with eqn. (1) in eqn. (3) yields the final posterior:

\[
p(\{ y_i \} | \{ \mathbf{m}_i, t_i \}, \lambda) = \prod_{i \in \mathcal{T}} p(y_i | \mathbf{m}_i, t_i, \lambda)
\]

with

\[
p(y_i | \mathbf{m}_i, t_i, \lambda) = \frac{\mathcal{N}(t_i | y_i, \beta_t^{-1}) \left[ \frac{1}{|\mathcal{A}_i|} \sum_{n \in \mathcal{A}_i} \mathcal{N}(y_i | y_n, \beta_y^{-1}) \mathcal{N}(\mathbf{m}_i | \mathbf{m}_n, \beta_m^{-1} \mathbf{I}_M) \right]}{\sum_{n' \in \mathcal{A}_i} \mathcal{N}(t_i | y_{n'}, \beta_t^{-1} + \beta_y^{-1}) \mathcal{N}(\mathbf{m}_i | \mathbf{m}_{n'}, \beta_m^{-1} \mathbf{I}_M)}
\]

\[
= \frac{\sum_{n \in \mathcal{A}_i} \mathcal{N}(t_i | y_n, \beta_t^{-1} + \beta_y^{-1}) \mathcal{N}(y_i | y_n, \beta_y^{-1}) \mathcal{N}(\mathbf{m}_i | \mathbf{m}_n, \beta_m^{-1} \mathbf{I}_M)}{\sum_{n' \in \mathcal{A}_i} \mathcal{N}(t_i | y_{n'}, \beta_t^{-1} + \beta_y^{-1}) \mathcal{N}(\mathbf{m}_i | \mathbf{m}_{n'}, \beta_m^{-1} \mathbf{I}_M)}
\]

\[
= \sum_{n \in \mathcal{A}_i} v_{n'} \mathcal{N}(y_i | \mu_n, (\beta_y + \beta_t)^{-1}),
\]
where we have defined
\[ \mu_i^n = \frac{\beta_t}{\beta_t + \beta_y} t_i + \frac{\beta_y}{\beta_t + \beta_y} y_n \quad \text{and} \quad v_i^n = \frac{N(t_i|y_n, \beta_t^{-1} + \beta_y^{-1})N(m_i|m_n, \beta_m^{-1} I_M)}{\sum_{n' \in A_i} N(t_i|y_{n'}, \beta_t^{-1} + \beta_y^{-1})N(m_i|m_{n'}, \beta_m^{-1} I_M)}. \tag{6} \]

We now estimate the undistorted CT values \( \{y_i\} \) as the mean of the distribution \( p(\{y_i\}|\{m_i, t_i\}, \lambda) \), which yields
\[ \bar{y}_i = \int_{y_i} y_i p(y_i|m_i, t_i, \lambda) dy_i = \sum_{n \in A_i} v_i^n \mu_i^n, \quad \forall i. \tag{7} \]

We refer to this as the kerMAR estimate ("kernel regression MAR"). It is instructive to consider the following special cases for various \( \beta_t \):

- \( \beta_t \to 0 \) (CT measurement \( t_i \) fully corrupted):
  \[ \bar{y}_i = \sum_{n \in A_i} w_i^n y_n, \quad w_i^n = \frac{N(m_i|m_n, \beta_m^{-1} I_M)}{\sum_{n' \in A_i} N(m_i|m_{n'}, \beta_m^{-1} I_M)}. \tag{8} \]
  which corresponds to conventional kernel regression [21, 301-304]; the CT measurement \( t_i \) is completely discarded. This method has previously been used for MR-based pseudo-CT generation,\(^{14,15}\) and so we will in this paper refer to it as the pCT estimate.

- \( \beta_t \to \infty \) (CT measurement \( t_i \) not corrupted):
  \[ \bar{y}_i = t_i, \]
  where the CT measurement is used as-is and the MR measurement \( m_i \) is completely discarded.

- \( 0 < \beta_t < \infty \) (General case):

  For general hyperparameter settings, \( y_i \) is estimated according to eqn. (7) using a combination of the CT measurement \( t_i \) and the MR measurement \( m_i \).

### 2.3 Empirical Bayes hyperparameter estimation

The results generated by kerMAR depend directly on the settings of the hyperparameters \( \lambda = \{\beta_t, \beta_y, \beta_m\} \). We seek to learn these hyperparameters automatically from our data using empirical Bayes estimation, finding the set of hyperparameters that best explain the data \( \{t_i, m_i\}, \forall i \in T \), by maximising their marginal likelihood (eqn. (4)):
\[ \lambda^* = \arg \max \lambda \log(p(\{m_i, t_i\}|\lambda)). \]

To simplify this optimisation problem, we make the approximation during hyperparameter estimation that all data belong to either an uncorrupted set \( T_u \) or a fully corrupted set \( T_c \), with respectively \( \beta_t \to \infty \) and \( \beta_t = \beta_t^* \). These subsets are defined by thresholding \( f(x_i) \) such that \( T_u = \{i \in T|f(x_i) \leq 0.5\} \) and \( T_c = \{i \in T|f(x_i) > 0.5\} \).

For the optimisation, we use an expectation maximisation algorithm (EM)\(^{22,23}\) which iteratively performs two steps: The E-step, where for all \( i \in T \) are assigned probabilities to the \( n \in A_i \) regression points based on the current hyperparameter estimates by calculating the weights \( v_{i}^{n} \) (eqn. (6)); and the M-step where the hyperparameters are updated accordingly:

\[ \beta_y^{-1} \leftarrow \frac{1}{|T_u|} \sum_{i \in T_u} \sum_{n \in A_i} v_{i}^{n} (t_i - y_n)^2, \]

\[ \beta_t^{* -1} \leftarrow \frac{1}{|T_c|} \sum_{i \in T_c} \sum_{n \in A_i} v_{i}^{n} (t_i - y_n)^2 - \beta_y^{-1} \quad \text{and} \]

\[ \beta_m^{-1} \leftarrow \frac{1}{|T|} \sum_{i \in T} \sum_{n \in A_i} v_{i}^{n} (m_i - m_n)^T (m_i - m_n). \]

The algorithm is initialised by a guess at \( \lambda \). We chose \( \lambda = \{0, 0, 0\} \), corresponding to flat initial weights.
2.4 Maximum likelihood transmission reconstruction using kerMAR as prior (MLTR-k)

In addition to using the kerMAR algorithm alone, we also consider using it to define an image prior distribution for MBIR using the MLTR algorithm. Specifically, we use the kerMAR estimates of the voxels, \( \{\tilde{y}_i\}_{i \in \mathcal{T}} \), to define the following Gaussian prior distribution on the CT image \( \{z_i\}_{i \in \mathcal{T}} \):

\[
p(\{z_i\}_{i \in \mathcal{T}}) = \prod_{i \in \mathcal{T}} \mathcal{N}(z_i|\tilde{y}_i, \kappa^{-1}),
\]

where \( z_i \) is the (unknown) CT value in voxel \( i \), \( \kappa \) the precision of the prior and \( \tilde{y}_i \) is the kerMAR estimate obtained as explained earlier.

The MLTR\textsuperscript{20,24} algorithm is a gradient-based optimisation algorithm that iteratively maximises the Poisson likelihood of the x-ray intensity measurements \( \{A_j\}_{j \in \mathcal{S}} \), related to the sinogram by an exponential transform.\textsuperscript{2} Starting from an initial estimate of \( \{z_i\}, \forall i \in \mathcal{T} \), MLTR iteratively improves the image estimate via an additive step. Including the image prior alters this step, which, in terms of the system matrix\textsuperscript{2} \( \mathbf{L} \) with entries \( l_{j,i} \),\textsuperscript{24} becomes:

\[
\begin{aligned}
Z_i &\leftarrow Z_i + \frac{\sum_{j \in \mathcal{S}} l_{j,i} \left[ C e^{-\sum_{i \in \mathcal{T}} l_{j,i} z_i} - A_i \right] + 2\kappa (\tilde{y}_i - Z_i)}{\sum_{j \in \mathcal{S}} l_{j,i} [\alpha_j C e^{-\sum_{i \in \mathcal{T}} l_{j,i} z_i} + 2\kappa], \quad \forall i \in \mathcal{T} \text{ with } \alpha_j = \sum_{i \in \mathcal{T}} l_{j,i}.}
\end{aligned}
\]

This update step is performed in parallel for all voxels \( \{Z_i\}_{i \in \mathcal{T}} \).

Ideally, \( C \) should here be the emitted x-ray intensity; in practice, we did not have access to this information, and so we used the Noise Equivalent Count (NEC) scaling coefficient, which scaled the exponentially transformed sinogram\textsuperscript{2} such that the measurements became approximately Poisson distributed.

3. EXPERIMENTS

**Image material and processing** We considered CTs and T1-weighted MRs from 7 head and neck RT patients. The CTs were acquired on a Philips Brilliance Big Bore helical CT scanner at a KVP of 120kEV and tube currents from 272-433mA. They were reconstructed in 512x512 2D slices by the scanner software using FBP at resolutions of (1.2 × 1.2 × 2.0mm). We will refer to these images as FBPs in the remainder. The MRs were acquired with a Philips Panorama 1.0T HFO scanner at resolutions of (0.5 × 0.5 × 5.5mm) except patient 7 with (0.5 × 0.5 × 6.5mm). For patients 2-6 the MR repetition- and echo times were \( (TR/TE) = (520.2ms/10ms) \), for patient 1,7 \( (TR/TE) = (572.2ms/10ms) \). The MRs were rigidly coregistered to the CTs using mutual information coregistration\textsuperscript{25,26} and resampled to the CT resolution. oMAR reconstructions were made by the scanner software and available alongside the FBPs.

The unit of the CT measurements was displaced Hounsfield Units [2, p. 475] (HU + 1024), so the minimum CT measurement was 24. The kerMAR/pCT estimates and their hyperparameters were accordingly calculated in displaced HU and afterwards subtracted by 1024 to yield HU. The unit of the MR measurements had no particular physical meaning other than depending on \( TR \) and \( TE \).

**kerMAR implementation** We used cubic patches with dimensions in units of voxel size of 7 × 7 × 7, or 8.4 × 8.4 × 14mm, chosen in preliminary investigations to provide better results than smaller patches while being computationally favourable. To evaluate the kerMAR and pCT estimates we used the FBPs and T1w MRs described in the previous paragraph.

The FBPs were reconstructed in axial 2D slices, only some of which were reconstructed using any metal projection data, which led us to expect a sharp boundary in the superior-inferior direction between slices containing corrupted CT values and those that did not. To approximately find these potentially corrupted slices, we noted that the intra-slice maximum HU value increased abruptly to the scanner cut-off value of 4095 (\( HU = 3071 \)) in a subset of consecutive slices. Assuming these to be corrupted, 24 slices centered on them were chosen as the voxel set \( \mathcal{T} \), excluding air voxels outside the body outline using a watershed segmentation.
Next we segmented the metal voxels by thresholding using Otsu’s method on an FBP calculated by the ASTRA toolbox\textsuperscript{27} that allowed higher CT values than the clinical FBP and thus for better distinction between high intensity streaks and the implant metal. We excluded these metal voxels from the kerMAR/pCT artifact reduction and used them to define the variance scaling function $f(x)$ as $f(x) = 1 + \tanh(-s(x)^2/500)$, where $s(x)$ is the Euclidean distance from position $x$ to the nearest metal voxel in mm. The characteristic squared distance of $500\text{mm}^2$ was chosen by experimentation.

We next found $|A_i| = 500$ kerMAR regression points for each $i \in T$ among the metal free axial slices. We ideally wanted $A_i$ to contain the points with most similar MR patches, i.e. smallest $(m_i - m_n)^T(m_i - m_n)$. Since finding these by direct calculation was too time consuming we used the approximate search algorithm “Fast PatchMatch” presented by Ta et. al. in.\textsuperscript{28} To further increase speed, we only used PatchMatch for $i \in T_c$, selecting instead based on the CT distance $(t_i - y_n)^2$ for $i \in T_n$ where the CT was less corrupted. We specifically used a K-means clustering with 6 clusters to classify the voxels, picking the 500 closest points from the assigned class.

Having determined $A_i$, the hyperparameter estimates were then calculated with 3 iterations of the EM-algorithm, at which point the hyperparameters changed by less than 0.1%. Using these hyperparameter estimates, the kerMAR (eqn. (7)) and pCT (kerMAR with $\beta_t = 0$ applied to $T_c$, eqn. (8)) estimates were finally calculated.

**MLTR-k implementation and raw data preprocessing** The CT sinograms were interpolated from helical to planar sinograms, exponentially transformed to transmissions and NEC scaled\textsuperscript{2} to make the data approximately Poisson distributed for use as intensity measurements for MLTR. The NEC scaling coefficient $C$ was calculated for each patient on an air scan as $\sim 3500$, varying little between patients.

We chose the MLTR-k precision $\kappa = 5 \cdot 10^6$ by experimentation, finding its exact value within an order of magnitude non-critical for the results. We used the clinical FBP as initialisation, stopping the algorithm after 300 iterations at which point the voxel averaged step was smaller than $\sim 10^{-10}$ compared to the initial $\sim 10^{-7}$. The MLTR iterative step eqn. (9) was calculated using the GPU accelerated primitives of the ASTRA toolbox\textsuperscript{27} to evaluate the forward projections $\sum_{i \in T} l_{j,i}(\cdot)$ and back projections $\sum_{j \in S} l_{i,j}(\cdot)$.

Since MLTR reconstructs the image in attenuation coefficients [2, p. 475], the kerMAR estimate $\tilde{y}_i$ and the FBP used for initialisation were linearly transformed from displaced HU to attenuation coefficients for use in MLTR-k. We found thus transform by linearly fitting the FBP to an MLTR reconstruction with 2000 iterations calculated on a set of axial slices far from the metal implants. We then applied it to the kerMAR and FBP before using them as prior and initialisation respectively, and applied the inverse to the final MLTR-k reconstruction.

### 4. RESULTS

We calculated three sets of artifact-reduced axial CT images for the 7 head-and-neck RT patients: 1) kerMAR, 2) MLTR-k using kerMAR as prior and 3) pCT (kerMAR with $\beta_t = 0$ applied to $T_c$), along with the uncorrected FBP and MLTRs (MLTR-k with $\kappa = 0$). For comparison to the clinical practice we additionally show the oMARs.

Some representative axial images are shown in figure 1 with arrows pointing to regions of interest. The blue arrows point to highly corrupted regions close to metal implants; the yellow arrows to bone/air regions where the MR-based kerMAR and pCT algorithms are error prone; the green arrows to large discrepancies on the pCT for two patients with a poorly registered MR due to its low longitudinal resolution; and the red arrows to a case where MLTR-k potentially show improvement over kerMAR.

The empirical Bayes estimated kerMAR precisions are of similar magnitude over the patients, with means ± standard deviation of $\langle\beta_t^*\rangle = (0.27 \pm 0.06) \cdot 10^{-5}$, $\langle\beta_y\rangle = (4.3 \pm 0.9) \cdot 10^{-5}$ and $\langle\beta_m\rangle = (0.0097 \pm 0.009) \cdot 10^{-5}$.
Figure 1: Representative axial images of the 7 patients. Left-Right: The uncorrected FBP; the kerMAR algorithm; the commercial, clinically used oMAR algorithm; kerMAR with $\beta_t = 0$ applied to the corrupted set of voxels $T_c$ defined during hyperparameter estimation (pCT); the MLTR algorithm initialised with the FBP and with the kerMAR as prior (MLTR-k); the MLTR algorithm without a kerMAR prior. The hyperparameters for patients 1-7 were: $\beta_t^* = (0.13, 0.31, 0.30, 0.28, 0.27, 0.35) \cdot 10^{-5}$, $\beta_y = (3.8, 6.5, 5.0, 4.2, 3.5, 4.2, 2.9) \cdot 10^{-5}$, $\beta_m = (0.0048, 0.0084, 0.0061, 0.0080, 0.0034, 0.032, 0.0053) \cdot 10^{-5}$. 
5. DISCUSSION

In this paper, we presented a novel MR-based metal artifact reduction algorithm (kerMAR) that used patches drawn from a coregistered MR and the corrupted CT measurements to predict the uncorrupted CT values. It automatically optimised its hyperparameters on each individual patient and only required data from the patient itself, thus requiring no MR intensity normalisation or database registration in contrast to pCT generation algorithms. We additionally experimented with using it as a prior for MBIR using the MLTR algorithm (MLTR-k).

Comparison to existing MR-based MARs  The main challenge when using MR images from conventional sequences for CT value prediction is the difficulty of disambiguating air and bone. Our kerMAR algorithm first addresses this issue by employing larger spatial contexts in the form of image patches for this prediction rather than single voxel intensities. As evident on the pCT results (fig. 1 column 4), however, we found the use of MR patches only to be insufficient. kerMAR therefore further included the corrupted measurements in the prediction to help resolve this disambiguation issue. As seen by the yellow arrows, this led to far better bone/air disambiguation.

The few previous MR-based MARs in the literature that we are aware of\(^\text{12, 13}\) suffer from this disambiguation issue. The algorithm introduced by Anderla et al.\(^\text{12}\) works by looking in a 5x5x5 voxel window on the MR around each corrupted voxel (classified using Otsu’s thresholding method), finding the voxel with smallest MR intensity difference to the window center and assigning its CT value to the voxel center on the CT. While this effectively constrains the search space to the very local 5x5x5 window and thus improves the bone/air disambiguation, it limits the potential accuracy of the algorithm and leads to failure in heavily corrupted regions.\(^\text{12}\) The algorithm presented by Delso et al.\(^\text{13}\) is more promising in this regard. It relies on threshold based implant segmentation and water/fat classification on a 3D MR dataset to create a pCT, replacing only soft tissue regions with database values empirically corrected on the patient data. This approach allows for improved handling of heavily corrupted soft tissue regions, but cannot address corrupted high or low intensity regions, in contrast to our proposed kerMAR algorithm. Additionally, coregistration errors between the CT and MR can lead to serious misestimations using the method, as reported by the authors. The green arrows on fig. 1 point to a case where we encountered such coregistration errors, thus leading to a poor pCT, but where the inclusion of the corrupted measurement in kerMAR led to good results.

Comparison of kerMAR to oMAR  The commercial oMAR algorithm (Philips Healthcare) has been found in a few other studies to improve dosimetric accuracy when used for photon RT dose planning\(^\text{1, 29, 30}\) and generally improving the image quality.\(^\text{30, 31}\) In this study, we found comparable visual improvement to oMAR with our kerMAR algorithm in terms of artifact reduction, with the following potential improvement as pointed to by the blue arrows on fig. 1: Our MR based kerMAR appears to better suppress high intensity streaks close to the implants than oMAR.

Kidoh et al. found in\(^\text{31}\) a statistically significant tendency with oMAR to introduce additional artifacts in the form of image blur and unnatural features. In kerMAR, the inclusion of the corrupted measurement in the prediction seems to act as a barrier against such tendencies, albeit an imperfect one. In particular, some erosion of thin low and high intensity areas is visible on the kerMAR when the MR and CT were imperfectly coregistered due to the low axial resolution of the MR. A higher resolution MR that provides better coregistration could potentially increase the accuracy of kerMAR.

Benefit of kerMAR as a prior in MLTR-k  In addition to using kerMAR as an image based MAR algorithm, we used it to define an image prior for MBIR with the MLTR algorithm (MLTR-k). Since such prior modelling may be used to integrate prior known features of the image in the reconstruction,\(^\text{2, 7, 8, 24}\) our hypothesis was that using kerMAR as a prior would alter the likelihood function such that the reconstructed image contained fewer artifacts.

We observe this on fig. 1 near the red arrows where MLTR-k provided streak suppression over MTLR and FBP.
while disagreeing with its prior, kerMAR, on the dental CT values. In terms of image quality, however, MLTR-k did not lead to improvement over kerMAR as the MLTR-k results in general are of lower visual quality and only improved slightly upon the MLTR and FBP.

**Clinical feasibility** The oMAR algorithm runs on the order of minutes per patient,\(^{31}\) which is suitable for clinical use. On our system (Dell Precision M3100 Laptop, CPU: Intel Core i7-4712HQ @ 2.3GHz, RAM: 16Gb) and our largely unoptimised Python implementation, kerMAR takes around 10-30min. per patient, the most time consuming part by far being the search for the regression point sets in the corrupted volume using Fast PatchMatch. Considering the results of the algorithm by Ta et al.,\(^{28}\) optimised code on a dedicated system should take on the order of tens of seconds per patient.

KerMAR only uses data from a single patient and in principle works independently of the specific MR sequence employed, since all parameters are picked on a per-patient basis, as mentioned earlier. kerMAR thus excludes issues of MR intensity normalisation\(^ {17–19}\) and inter-patient registration issues, adding to its clinical attractiveness; the main requirement for its clinical implementation is an infrastructure for acquisition in similar patient orientations, coregistration and joint storage of the MR and CT volumes, which is clinical routine in RT clinics that uses both MR and CT for RT planning.

6. CONCLUSION
We presented a novel, clinically feasible MR-based algorithm for automatic CT metal artifact reduction (MAR), referred to as kerMAR. It requires no MR intensity normalisation or atlas registration. Image results for 7 head-and-neck RT patients suggest a potential for better suppression of high intensity streaks near the metal implants in the oral cavity as compared to the oMAR algorithm (Philips) used clinically. Our results further suggest improved performance in air and bone regions as compared to existing MR-based MAR algorithms. Using kerMAR as a prior for MLTR (MLTR-k) did not provide apparent improvement.

**REFERENCES**


