Motion artifact detection in four-dimensional computed tomography images

G Bouilhol¹,², M Ayadi², R Pinho², S Rit¹,² and D Sarrut¹,²
¹ University of Lyon, CREATIS; CNRS UMR 5220; Inserm U1044; INSA-Lyon; University Lyon 1, Villeurbanne, France
² Department of Radiation Oncology, Leon Berard Cancer Center, Lyon, France
E-mail: gauthier.bouilhol@gmail.com

Abstract. Motion artifacts appear in four-dimensional computed tomography (4DCT) images because of suboptimal acquisition parameters or patient breathing irregularities. Frequency of motion artifacts is high and they may introduce errors in radiation therapy treatment planning. Motion artifact detection can be useful for image quality assessment and 4D reconstruction improvement but manual detection in many images is a tedious process. We propose a novel method to evaluate the quality of 4DCT images by automatic detection of motion artifacts. The method was used to evaluate the impact of the optimization of acquisition parameters on image quality at our institute. 4DCT images of 114 lung cancer patients were analyzed. Acquisitions were performed with a rotation period of 0.5 seconds and a pitch of 0.1 (74 patients) or 0.081 (40 patients). A sensitivity of 0.70 and a specificity of 0.97 were observed. End-exhale phases were less prone to motion artifacts. In phases where motion speed is high, the number of detected artifacts was systematically reduced with a pitch of 0.081 instead of 0.1 and the mean reduction was 0.79. The increase of the number of patients with no artifact detected was statistically significant for the 10%, 70% and 80% respiratory phases, indicating a substantial image quality improvement.

1. Introduction
Motion artifacts appear in four-dimensional computed tomography (4DCT) images because of suboptimal acquisition parameters or patient breathing irregularities. Occurrence frequency of motion artifacts is high and they may introduce errors in radiation therapy treatment planning. Motion artifacts generally appear as discontinuities in voxel intensities perpendicularly to the scanning direction. They appear at the interface between groups of slices called stacks and composed of slices acquired during the same breathing cycle and attributed to the same respiratory phase.

Motion artifact detection can be useful for image quality assessment [1, 2, 3] and 4D reconstruction improvement [4, 5] but manual detection in many 4DCT images [6] is a tedious process and it is observer-dependent. Various algorithms have been developed to automatically detect motion artifacts. Discontinuities induced by motion artifacts can be detected by measuring the mean squared gray value difference between adjacent slices [4]. They can also be detected by measuring the difference between observed and predicted slices at the interface between stacks [5]. However, these methods have a limited efficiency because of their sensitivity to anatomic variations in the scanning direction. It is also possible to predict the
occurrence of artifacts directly from the respiratory signal but according to Han et al. [2] poor efficiency is associated with this method. For a given interface between two stacks, the method described by Han et al. consists of finding among the other reconstructed phases, a bridge stack containing common slices with each of the two considered stacks. A rigid spatial registration is then performed between each stack and the bridge stack. The registration result (contiguous, overlapping or separated stacks) then gives an indication of the presence or absence of an artifact at the interface of the two stacks. The advantage of this method relies on the ability to identify the artifact type (duplication or overlapping), but it has several drawbacks and limitations. First, access to the entire 4DCT dataset is required. In addition, as noted by the authors, the correct identification of bridge stacks is essential but it may not be possible in case of large irregular respiratory motions (leading to high magnitude artifacts). This method is not sensitive to anatomic changes in the scanning direction but it can be influenced by the deformation of structures between the respiratory phases. Finally, it can only be used for helical acquisitions since in cine mode the stacks contain the same slices whatever the phase so that it is impossible to find a stack with common sections with each of the two considered stacks.

Given the limitations of existing algorithms, a novel method for automatically detecting motion artifacts in 4DCT images has been proposed. The optimization of acquisition parameters at our institute led to a change in the pitch value used for treatment planning. The initial value of 0.1 has been changed to a value of 0.081. Our method was used to assess the image quality and thus quantify the impact of this optimization on the presence of motion artifacts in a large patient database.

2. Materials and methods

2.1. Artifact detection

The proposed method is based on Ehrhardt et al. [4]. It consists of computing the p-th power of the n-th order absolute difference of the mean gray value between adjacent slices:

\[
D(z) = |\Delta^n T(z)|^p
\]

where \(T(z)\) is the mean gray value in the slice \(z\). In a preliminary study, \(n = 3\) and \(p = 7\) were found to be suitable for thoracic acquisitions with the parameters used at our institute (see Patient database section). This brings out the strongest gray value discontinuities and lower the smallest ones such as those due to anatomical changes.

To improve the sensitivity to artifacts and further limit the sensitivity to anatomical changes, images were first automatically cropped in the left-right and antero-posterior directions to fit the lung boundaries and the detection was performed separately for each side (right and left) of the cropped region. This prevents the false detection of artifacts at the top of the diaphragmatic dome or at the apex of the lungs when these structures are located at the same height on both sides. Separation of right and left also improves the sensitivity to artifacts clearly appearing in only one of the two sides. In addition, only the voxels within the limits of the patient are considered to avoid the influence of objects outside of the patient (e.g. immobilization systems). This improvement step relies on the segmentation of the lungs and the patient boundaries. It could be automatically performed using previously described methods [7, 8].

Furthermore, the detection was performed using two grayscale windows excluding bone structures. Voxels of intensity between -2000 and 220 HU (lung window) were considered between the superior-inferior limits of the lungs, and voxels of intensity between -180 and 220 HU (soft tissue window) below the lungs. This way, sudden changes in voxel intensity due to the bone structures or gas in the abdomen are ignored.

Artifacts are then detected and located by applying two thresholds, \(s_1\) (lung window) and \(s_2\)
(soft tissue window), above which the values of $D$ correspond to discontinuities considered as artifacts.

Finally, a minimum distance of 3 pixels (in the scanning direction) between two artifacts is introduced in order to limit the detection of more than one artifact within an area smaller than a stack. The principle of this artifact detection method is shown in Figure 1.

![Figure 1](image_url)

**Figure 1.** Principle of the artifact detection method. The red area represents the result of the automatic segmentation of the lungs used to crop the image. Blue spikes indicate where artifacts occur.

### 2.2. Threshold adjustment

The adjustment of the detection thresholds, $s_1$ and $s_2$, is automatic and specific to each image. Based on a heuristic, the adjustment method takes into account that the number of detected artifacts not only depends on the thresholds and the number of artifacts in the image, but also on the patient’s morphology and the scanned length. Discontinuities appear at the interface between structures of different densities. Consequently, the variation of the average intensity $I(z)$ between two slices is not the same for thin or obese patients (contrast variation, especially in the abdomen), or for patients with many or few dense structures in the lungs. It is therefore possible to adjust the detection thresholds based on the number of detected artifacts. The first step is to determine the numbers of artifacts detected with the two grayscale windows, $h_1(s)$ and $h_2(s)$ respectively, for thresholds ranging from 1 to $10^{2p}$ on a logarithmic scale. The choice of $s_1$ and $s_2$ is then carried out in different ways for the two windows. Low and high threshold limits are introduced to improve the detection for patients of unusual morphology.

For the lung window, the threshold $s_1^{\text{max}}$ corresponding to the largest number of detected artifacts is determined. If multiple values match, the smallest one is selected. The decreasing cumulative number of artifacts, $H_1(s)$, is then calculated for $s$ ranging from $10^{2p}$ to $s_1^{\text{max}}$. The line passing through $H_1(s_1^{\text{max}})$ and the previous value, $H_1(s_1^{\text{max}+})$, is then determined and its intersection with the x-axis was found to be a good estimate of $s_1$. If $s_1$ is smaller than a minimum value $s_1^{\text{inf}}$ or if no artefact is detected above an upper limit $s_1^{\text{sup}}$, then no artifact is detected. Figure 2 illustrates the $s_1$ adjustment method.

For the soft tissue window, the threshold $s_2^{\text{max}}$ corresponding to the largest number of detected artifacts is determined. If multiple values match, the largest one is selected and $s_2 = s_2^{\text{max}}$. If $s_2$ is smaller than a minimum value $s_2^{\text{inf}}$ then $s_2 = s_2^{\text{inf}}$, and if no artifact is detected above an upper limit $s_2^{\text{sup}}$, an intermediate value $s_2^{\text{int}}$ is chosen as $s_2$. Figure 3 illustrates the $s_2$ adjustment method.

We can see that when the threshold value decreases, the number of detected artifacts increases to a maximum and then decreases (see Figures 3 and 4). This is due to the minimum distance...
of 3 pixels imposed between two consecutive detected artifacts. Indeed, for low threshold values, the number of sets of three consecutive slices for which the value of $D$ is smaller than the threshold is low. The determination of optimal values of $s_{\text{inf}}^1$, $s_{\text{sup}}^1$, $s_{\text{inf}}^2$, $s_{\text{int}}^2$ and $s_{\text{sup}}^2$ was performed using two leave-one-out cross-validations, one for each window. To do so, artifacts were visually identified in 27 respiratory phases randomly selected from the 4DCT scans of 25 patients in the database. For each image, the analysis consisted of assessing the sensitivity $Se$ and specificity $Sp$ for the 26 other images and for a large number of values $s_{\text{inf}}^1$, $s_{\text{sup}}^1$, $s_{\text{inf}}^2$, $s_{\text{int}}^2$ and $s_{\text{sup}}^2$. To measure the specificity, the number of true negatives was determined assuming a stack size equal to the collimation width. It corresponds to the maximum stack size $[2]$ and thus a low estimate of the specificity. The product $Se \times Sp$ was calculated to determine the optimal values for each of the 27 combinations.
2.3. Patient database

4DCT images of 114 lung cancer patients were analyzed with the proposed method. Each patient dataset consisted of 10 respiratory phases reconstructed with a 3 mm slice thickness. Acquisitions were performed with a Brilliance Big Bore CT scanner (Philips) in helical mode using 400 or 800 mAs/slice, a collimation width of $16 \times 1.5$ mm and a rotation period of 0.5 seconds. For 74 patients, a pitch of 0.1 was used and for 40 patients, the pitch was 0.081. Given a maximum acquisition time of 120 seconds, the maximum scanned length is about 45 cm with a pitch of 0.081, which is suitable for the acquisition of the entire lung.

3. Results

3.1. Threshold adjustment

Optimal values $s_{\text{inf}}^1 = 7 \times 10^9$ and $s_{\text{sup}}^1 = 3 \times 10^{12}$ were obtained for each of the 27 combinations and $s_{\text{inf}}^2 = 3 \times 10^5$, $s_{\text{int}}^2 = 1 \times 10^7$ and $s_{\text{sup}}^2 = 5 \times 10^7$ were obtained for 22 of the 27 combinations. With these values, a sensitivity of 0.73 and a specificity of 0.97 were observed. Hence the corresponding product $Se \times Sp$ was 0.70.

3.2. Patient database

Automatic detection of motion artifacts in the 4DCT images of 114 patients showed that 92% of patients had at least one artifact in one of the phases with a pitch of 0.1, and 85% with a pitch of 0.081. In addition, the average number of respiratory phases affected by artifacts was 6.2 with a pitch of 0.1, and 5.2 with a pitch of 0.081. The 40%, 50% and 60% respiratory phases were the least affected by artifacts. These phases are indeed the most stable since they are located near end-exhale. In phases where motion speed is high (10%, 20%, 30%, 70%, 80% and 90%), the number of detected artifacts was systematically reduced with a pitch of 0.081 instead of 0.1 and the mean reduction was 0.79 (SD 0.23). In the other phases where motion speed is low, the mean reduction was 0.09 (SD 0.17). The reduction of the number of detected artifacts induced an increase of the number of patients with no artifact detected. This increase was statistically significant for the 10% ($p = 0.041$), 70% ($p = 0.004$) and 80% ($p = 0.025$) respiratory phases (see Figure 4).

\begin{figure}
\centering
\includegraphics[width=0.7\textwidth]{artifact_detection.png}
\caption{Percentage of patients with no artifact detected according to the respiratory phase for a pitch of (light blue) 0.1 and (dark blue) 0.081. In green are presented the $p$-values and those reflecting a significant difference are underlined.}
\end{figure}

4. Discussion

A motion artifact detection method has been proposed. The basic principle is similar to the method described by Ehrhardt et al. [4] but several improvements are proposed by our method. The $Se \times Sp$ score of 0.70 obtained in this study is comparable to the results obtained by Han.
et al. [2] who reported a mean $Se \times Sp$ value of 0.72 (5 patients) but our method does not require access to the entire 4DCT dataset and it is not limited to helical acquisitions. Despite a slightly lower sensitivity (0.73 vs. 0.87), the advantage of the method proposed in this paper is its high specificity (0.97 vs. 0.82). Moreover, false negative results corresponded to low magnitude artifacts. This method requires automatic segmentation of the lungs but only basic segmentation is needed since it is used to determine the lung bounding box. Motion artifacts in 4DCT images could be automatically detected in a large database. Results compare favorably to visual detection reported in literature (90% of patients have at least one artifact [6]).

5. Conclusion

4DCT image quality was improved with a pitch of 0.081 instead of 0.1. In respiratory phases with high motion speed, the increase of the number of patients with no artifact detected was statistically significant. This increase can be considered as clinically relevant since motion artifacts may impact the treatment planning quality. Breathing irregularities remain the main cause of motion artifacts in 4DCT images. Reducing the pitch allows to improve the robustness with respect to low frequency respiration cycles [2]. Methods such as patient training, data sorting optimization or data interpolation with deformable registration can also help to reduce frequency and magnitude of motion artifacts.

References