Published in final edited form as:

Phys Med Biol. 2016 April 21; 61(8): 3109–3126. doi:10.1088/0031-9155/61/8/3109.

Robust breathing signal extraction from cone beam CT projections based on adaptive and global optimization techniques

Ming Chao¹, Jie Wei², Tianfang Li³, Yading Yuan¹, Kenneth E Rosenzweig¹, and Yeh-Chi Lo¹

¹Department of Radiation Oncology, Mount Sinai Medical Center, New York, NY 10029, USA

²Department of Computer Science, City College of New York, New York, NY 10031, USA

³Department of Radiation Oncology, University of Pittsburgh Medical Center, Pittsburgh, PA 15232, USA

Abstract

We present a study of extracting respiratory signals from cone beam computed tomography (CBCT) projections within the framework of the Amsterdam Shroud (AS) technique. Acquired prior to the radiotherapy treatment, CBCT projections were preprocessed for contrast enhancement by converting the original intensity images to attenuation images with which the AS image was created. An adaptive robust z-normalization filtering was applied to further augment the weak oscillating structures locally. From the enhanced AS image, the respiratory signal was extracted using a two-step optimization approach to effectively reveal the large-scale regularity of the breathing signals. CBCT projection images from five patients acquired with the Varian Onboard Imager on the Clinac iX System Linear Accelerator (Varian Medical Systems, Palo Alto, CA) were employed to assess the proposed technique. Stable breathing signals can be reliably extracted using the proposed algorithm. Reference waveforms obtained using an air bellows belt (Philips Medical Systems, Cleveland, OH) were exported and compared to those with the AS based signals. The average errors for the enrolled patients between the estimated breath per minute (bpm) and the reference waveform bpm can be as low as -0.07 with the standard deviation 1.58. The new algorithm outperformed the original AS technique for all patients by 8.5% to 30%. The impact of gantry rotation on the breathing signal was assessed with data acquired with a Quasar phantom (Modus Medical Devices Inc., London, Canada) and found to be minimal on the signal frequency. The new technique developed in this work will provide a practical solution to rendering markerless breathing signal using the CBCT projections for thoracic and abdominal patients.

Keywords

breathing signal; cone beam CT projection; radiation therapy

1. Introduction

Respiratory motion plays an important role in radiation treatment of thoracic and abdominal cancer. Proper extraction of this motion is essential in ensuring precise target localization

and accurate delivery of radiation dosage. The respiratory signal, as a surrogate of the tumor position, is indispensable in simulation to generate respiratory correlated computed tomography (CT) or four-dimensional CT (4DCT) and treatment delivery (Ford *et al* 2002, Keall *et al* 2004, Gierga *et al* 2005, Keall *et al* 2006, Oppenheim and Schafer 2009). The common technique is to use external markers or abdominal belts, measuring the marker displacement or the abdominal pressure to produce respiratory signals. These methods entail additional equipment such as an infrared camera, a pressure sensor, or a spirometry and the corresponding signal processing devices, which requires extra effort to set up, calibrate, and maintain, but may not be necessary. Alternatively, implanting a radiopaque fiducial marker near the tumor can directly provide information on tumor motion. However, this involves an invasive procedure and may not be suitable for every patient. With the advent of the on board imager on the clinical linear accelerators, the cone beam CT (CBCT) can improve patient setup significantly; furthermore, the raw projection images can be utilized to extract breathing signals. An Amsterdam Shroud (AS) image based technique was developed as a marker free tracking option, which shows great potential in clinical application. The main idea is to form the breathing signal by concatenating each projection to display the distinct anatomy oscillations (Sonke *et al* 2005).

A number of investigations were carried out using the CBCT projections for respiratory motion extraction (Kavanagh *et al* 2009, Dhou *et al* 2013, Yan *et al* 2013, van Sornsens de Koste *et al* 2015). In all these methods the breathing signals can be reliably extracted from the projection images provided that there are sufficient distinct structures differentiating various anatomies. The projections are radiographic images, thus the contrast is inferior compared to tomographic images. Often the onboard imager calibration may drift, which further deteriorates the image quality (Yoo *et al* 2006). This can be corrected through imager calibration, using an air projection and a dark projection to compute the line integral. However, there might be fluctuations during these two projections. Therefore, frequent and regular acquisition of these images is needed to capture the time fluctuations. Another solution is to use a constant value, i.e. the air intensity, to correct the projection images. This technique converts an intensity image to an attenuation image with the air intensity correction (Chao *et al* 2015a). In some cases this strategy works well and significant improvements were observed. The disadvantage of this method is that it corrects the image globally. Thus, for local areas where the contrast is low, it fails to improve. Therefore, a mere calibration or a constant value correction may not improve the low contrast as needed for extracting breathing signals.

Different from the signal tracking techniques using fluoroscopic movies or external markers (Vedam *et al* 2003, Berbeco *et al* 2005), CBCT projection based techniques suffer from the adverse impact of the projection angular variation. At different gantry angles, the kV x-ray beam passes through varying thickness of the patient tissue, immobilization devices (if there is any), and/or the couch top, resulting in attenuation variations. For instance, when kV x-ray source faces downward at twelve o'clock position, the beam goes through less tissue to reach the diaphragm. When the x-ray source is rotated to three o'clock position, the beam will penetrate the thickest tissue before it reaches the contralateral diaphragm. The image contrast for diaphragm on these two typical views is vastly different. In Kavanagh *et al* (2009) the authors introduced a pixel summation plus a high pass filter method to rectify the

slow component caused by the gantry rotation. While it is easy to implement and yields good signals, the method strongly depends on the summation limits that need to be set manually. Whereas AS technique extracts breathing signals without these limits, attenuation variations in some views are so drastic that the oscillating structure almost disappears on these projection images. Hence restoration of the structure on the AS image is imperative if any robust signal can be reliably extracted thereof.

In this work we present a study to extract breathing signals from CBCT projections with a modified AS technique. The objective of this paper is three-fold. First, we tried to enhance the image contrast of the CBCT projections; second, an adaptive robust z -normalization filtering was introduced to further boost the oscillating features on the AS image; and lastly, a two-step optimization method was devised to reliably and accurately extract the breathing signals from the enhanced AS image.

2. Materials and methods

2.1. Projection image preprocessing for contrast enhancement

Anatomy feature on these images plays a key role in obtaining the breathing signal, and is highly dependent on the image contrast. The raw projections are characterized with inferior contrast due to the fact that they are radiographic images. In order to obtain clear breathing signal, improvement on the contrast is necessary. Conversion from the raw projection images to attenuation images could enhance the contrast because the attenuation images represent the radiological line integral at each pixel. The intensity in the raw projections is given

$$I(E) = I_0(E) e^{-\int \mu(x, y, E) ds} \quad (1)$$

where $I(E)$ is the intensity on the image, $I_0(E)$ the air intensity, E the x-ray energy spectrum, $\mu(x, y, E)$ the attenuation coefficient at pixel location (x, y) . The exponential term is the line integral of attenuation along the x-ray transmission path. To convert the intensity images to attenuation images, we need to estimate the air intensity $I_0(E)$. There are several ways to estimate it, such as air scan or using the existing projections. In this work we chose the latter by performing a histogram based automatic searching of all the pixels for all projections. All pixel values are put in a histogram and then an upper bound of the histogram is searched iteratively until the largest number is found. This is the highest intensity value I_0 . A manual search of the air intensity could be done using an image viewer which displays the grayscale values, for example, ImageJ. Based on the air intensity, attenuation images are then obtained by simply taking a ratio for individual pixels to I_0 and a logarithm operation is applied:

$$\ln \left(\frac{I_0(E)}{I(E)} \right) = \int \mu(x, y, E) ds \quad (2)$$

Now each pixel value is a summation of the attenuation along the x-ray path in the anatomy. An attenuation image is generated. Figure 1 shows the comparison between the intensity

image and the attenuation image. The top plots are profiles corresponding to the dashed blue lines on two images: on the left (c) the raw projection and the right (d) the attenuation image. The contrast (defined as the difference between maximum and minimum normalized by maximum) on (a) is 0.021 while on (b) is 0.748, an improvement by more than one order of magnitude.

2.2. AS image technique and breathing signal extraction

As illustrated in figure 2, the AS image is generated using CBCT projections. It was originally introduced to extract the breathing signals to sort the respiratory phases (Sonke *et al* 2005). First a logarithmic transform is performed on each projection image followed by a derivative along the superior inferior (SI) direction to enhance the anatomical difference. Due to the preprocessing of the projection images, the logarithmic operation in this step may not be necessary. However, the derivative is indispensable to enhance the feature edge that is critical in forming AS images. Each row pixel on the resultant image is summed and averaged to form a one-column image. Subsequently the one-column images from all projections are concatenated horizontally to form a 2D image, which is the AS image. If significant anatomical change occurs during the course of imaging, clear wavy patterns should be visible on this image. The respiratory motion can be subsequently extracted out of the AS image with appropriate signal processing techniques.

The breathing signal extracted from an AS image is based on the matching between the adjacent vertical lines (columns) in the AS image. As a reference signal extraction technique, we follow the signal extraction algorithm developed in Yan *et al* (2013). For an AS image with m rows and n columns, one needs to find the breathing motion for column $i + 1$, $i = 1, \dots, n - 1$, denoted by C_{i+1} from the i th column C_i . C_i is first padded by l zeros on both ends ($l = 6$ in their original implementation), resulting in a new column vector of dimension $m + 2 * l$:

$$C'_i = [0, \dots, 0, C_i, 0, \dots, 0]^T \quad (3)$$

T is the transpose operator. The motion vector V_{i+1} for C_{i+1} relative to C_i is the moving value having the least square errors (LSE):

$$V_{i+1} = \underset{V \in [-l, l]}{\operatorname{argmin}} \sum_{j=1}^m (C_{i+1}[j] - C'_i[j + V + l])^2 \quad (4)$$

The breathing motion evaluated by using equation (4) is presented in figure 3 where figure 3(a) is the AS image and the extracted signal is depicted in figure 3(b). When the oscillating structure in the AS image is clean and strong, such as the left part of figure 3(a), the extracted breathing signal according to equation (4) is of high quality and stable. This is shown from the leading one-third of the extracted signal in figure 3(b). However, when the wavy pattern is not clear, for example, the middle part of the figure 3(a), no reliable

breathing signal can be extracted. We aim to improve both the AS image and the signal extraction algorithm.

2.3. Adaptive robust z-normalization filtering

The original projections undergo the intensity to attenuation conversion. Subsequently, the newly converted attenuation images are plugged in to create the AS image. Although the contrast is improved significantly, some projections due to the varying attenuation and possible radiation scatter could cause the oscillating structures barely visible on the AS image. As previously depicted in figure 3, one typical AS image, of neither the best nor the worst quality, after the intensity to attenuation conversion is presented. This is caused by the attenuation conversion that only takes the local behavioral properties into account: on the left one-third of the AS image, the (roughly) periodic motion due to breathing is clearly visible; while on the right two-thirds, especially the middle one-third, the signals are almost entirely buried in the very bright background. Without further processing, the subsequent signal retrieval will encounter serious troubles. This is reflected on the right panel of figure 3(b), the periodic signal disappears in the middle part. To combat this issue, the adaptive methodology should be called upon (Wei *et al* 2015). In the well-established color constancy topic in computer vision (Wei 2002), this issue have been investigated thoroughly.

Denote the AS image by I , the mean value for each pixel is evaluated by convolving I with a Gaussian filter G of a relatively large size s : in our work $s = 10$.

$$I_G = I * G \quad (5)$$

Besides the adaptive evaluation of mean values for each pixel, it is found that the spread statistics at each pixel should also be adaptively re-adjusted.

$$I_\sigma = \text{stdfilt}(I, M) \quad (6)$$

where **stdfilt** is the standard deviation filter (in MATLAB syntax) to compute the standard deviation (σ) for each pixel p in the circular mask M of size 10 with p as the center. The new image I_N is then computed as:

$$I_N = (I - I_G) / (I_\sigma + c) + I_{\text{all}} \quad (7)$$

where $/$ is the element-wise division; c is a small constant value to avoid over-emphasizing when a certain I_σ value is too small; I_{all} is the mean value of the original AS image only to ensure the non-negativity of the resulting I_N . In this work, after trial and error, c is chosen to be $\mu(I_\sigma)$, the mean value of all standard deviations across the AS image. This series of adaptive processing based on local mean value and standard deviations is referred to as *adaptive robust z-normalization*. without constant c equation (7) is the actual z-

normalization procedure. Only the mean and the standard deviation are the adaptively readjusted values for each pixel; whereas the small constant c in equation (7) is utilized to avoid possible overflow when the local standard deviation is too small, making the transform dictated by equation (7) *robust*. The image shown in figure 3 is thus filtered, and the result is demonstrated in figure 4 where the local brightness and contrast are put in better quality: the originally missing detail in the middle part is now effectively restored.

2.4. Global regularity of breathing signals based on spectral density analysis

In AS image regions with considerable noise, such as the rightmost one-third in figure 4(a), the contrast enhanced by the adaptive robust z -normalization described in section 2.3 alone is inadequate to produce reliable breathing signals: there are a multitude of incomplete breathing cycles on the rightmost one-third shown in figure 3(b).

There are two problems with the motion estimation adopted by the approach based on equation (4):

1. The use of LSE, or the L_2 norm, in the minimization of correspondence errors. According to robust statistics (Hoaglin *et al* 2000), the L_1 norm is more desirable than the L_2 norm since the latter overly amplifies the negative impacts of outliers.
2. Probably more importantly, in this breathing signal extraction approach, to estimate movement for column $i + 1$, only its difference from column i is taken account of. In principle, this *local* correspondence search method is extremely vulnerable to noises within these small regions, which cannot be combated by any image enhancement techniques alone. This is a more serious problem for this AS breathing signal extraction procedure.

To effectively tackle the matching trouble in local regions, the global properties or the regularities of breathing signals existing in larger scales should be first examined and exploited (Wei and Li 2014). Toward that end, as done by most signal processing schemes (Oppenheim and Schaffer 2009), the power spectral density (PSD) of breathing signals extracted by the AS procedure, equation (4), gives rise to a valuable indicator of the global property of the breathing signal. The Welch's PSD estimation method is among the most popularly used due to its robust nature (Oppenheim and Schaffer 2009). Welch's method proceeds by first partitioning the entire sequence into several overlapping segments—in our work the entire sequence is put into eight overlapping segments, with 50% overlapping size between every two adjacent segments. Each segment is windowed by a Hamming window. The *periodogram* of the discrete Fourier transform of each windowed segment is next evaluated. Finally the periodograms for all segments are averaged to produce the Welch PSD. This averaging periodogram can effectively combat the afore-mentioned local matching problems encountered by the AS locally matching scheme. Therefore, Welch's PSD estimation method can effectively reveal the large-scale regularity of the breathing curve with the local errors significantly mitigated. In figure 5, the PSDs evaluated by Welch's method for four breathing signals are tabulated as illustrations.

After evaluating the PSD using Welch's method, the PSDs for all breathing signals share common behavioral properties. Four typical PSD curves randomly selected from our data sets are tabulated in figure 5. It can be observed that the PSD curves of all breathing signals consist of two components: the left one at peak $\sim 0.1025\text{Hz}$ ($=2.5 * 0.041\text{ Hz}$) with significant lower frequency; while the right one at peak $\sim 0.328\text{ Hz}$ ($=8 * 0.041\text{ Hz}$) with considerably higher frequency. The lower component is likely caused by the global position and angular shift of the estimated signals; while the higher component, around the range of human respiratory frequencies, is due to the periodic respiratory movements of the patients. This two-peak phenomenon in the Welch PSD is present consistently for all the breathing signals extracted by the AS procedure. Therefore, although in many local regions the breathing signals extracted by the AS procedure is unreliable, yet in the large scales, the local noises can be effectively suppressed. The two components in the low and relatively high spectrum account for most energies and thus capture the global shape of the actual breathing signals.

2.5. Two-step breathing signal extraction algorithm

The global or large-scale breathing signal regularity as expanded on in section 2.4 provides the valuable basis to achieve effective noise reduction in the frequency domain: starting from the initial breathing signal V extracted by the AS procedure, equation (4), the global breathing curve can be effectively represented by two groups of Fourier coefficients for V , one in the very low frequency regions, corresponding to the left component; the other in the spectral domain determined by the possible human breathing frequencies. In the spectral domain F_V of V :

$$F_V = \text{FFT}(V) \quad (8)$$

$\text{FFT}(V)$ is the fast discrete Fourier transform of V . As discussed before (see figure 5), the energy of F_V is mostly concentrated on the low-frequency component around 0.1 Hz and the respiratory frequency around 0.3 Hz. To take advantage of this behavioral property of F_V , one can use a few low frequency coefficients around 0.1 Hz together with several higher frequency coefficients around 0.3 Hz to represent F_V . In this work, after intensive statistical studies over the data sets available to us, one can use three coefficients $F_{\text{Low}3}$ of the largest magnitude whose frequencies are less than 0.15 Hz and larger than 0 to approximate the contribution of the low-frequency component; whereas one can use the three coefficients $F_{\text{Res}3}$ of the largest magnitude whose frequencies are larger than 0.15 Hz to represent the contribution of the respiratory motion. Although there are only six numbers in these two sets, while in total there are around 640 elements in F_V , based on the statistical study conducted over all data sets, these $<1\%$ coefficients can capture more than 90% of the total energy of F_V . The new signal V' is reconstructed from these two sets:

$$V' = \text{IFFT}(F_{\text{Low}3} \cup F_{\text{Res}3}) \quad (9)$$

IFFT is the inverse fast discrete Fourier transform. Due to Parseval's theorem, V' can load more than 90% of the energy of the original V as well.

In figure 6(a), the V' reconstructed by equation (9) for the AS-reconstructed signal V shown in figure 3(b) is depicted. The energy loaded by these six dominant coefficients is 91%. In figure 6(b), another V' for a better V is shown, here 94% of the energy is captured. From these two plots and indeed from all the available data sets, it is consistently observed that although the amplitudes between V and its corresponding V' may differ, the large-scale regularities of V are captured well by V' . Consequently the local noises present in the rightmost part of V (blue) in figure 6(a) has a better chance to be extracted using the V' signal.

As shown in figure 6, V' , as a spectrally smoothed signal of V , can only capture the global regularities of V . It cannot serve as a viable proxy for the real breathing signal because it is still the result of the estimation based on equation (4), which is extremely vulnerable to local noises. The global regularity exhibited in V' , however, can be exploited to steer the search for better quality breathing signals. Toward that end, a second respiratory motion search is invoked under the guidance of V' . Since the actual amplitudes of V' s are unreliable, a directional vector D is first formulated using V' :

$$D = [0, \text{sgn}(\text{diff}(V'))], \quad (10)$$

where $\text{sgn}()$ is the sign function, i.e. $\text{sgn}(x \geq 0) = 1$; $\text{sgn}(x < 0) = -1$; while diff is the first-order difference function: $\text{diff}(V') = [V'(2) - V'(1), V'(3) - V'(2), \dots, V'(n) - V'(n-1)]$. The leading 0 is padded to ensure the length of D is the same as n , the column number of the AS image. The breathing signal B is then estimated using the optimization procedure below, with notations similar to those of equation (4).

$$B_{i+1} = \underset{B \in D(i+1) * [0, l]}{\text{argmin}} \sum_{j=s}^e |C_{i+1}[j] - C'_i[j+B+l]| \quad (11)$$

There are three major differences between equations (11) and (4):

1. Instead of indiscriminately searching from $[-l, l]$, the optimization process is *guided or constrained* by the corresponding element $D(i+1)$: if $D(i+1)$ is one, namely, according to equation (10), the motion dictated by V' is positive, then the candidates are among $[0, l]$; whereas if $D(i+1)$ is -1 , the searching space is limited to $[-l, 0]$.
2. Instead of using all rows in each column, i.e. j ranges in $[1, m]$, here j takes its value in $[s, e]$, where $s \geq 1$ and $e \leq m$. Thus columns C'_i s can be either all the rows in the AS image or a portion of it. Note that the row direction is along the SI view of the projection images. This choice was made to reflect the fact that the longitudinal field of view is limited in CBCT scans. The diaphragm, while

preferably chosen, may not always be in the field of view. It is our goal to ensure that our approach can still work even if only a portion of the lung information, i.e. $s > 1$ and $e < m$, is available. Therefore, in our algorithm and the test results reported in the sequel, not all the rows of the AS image (see figures 3(a) and 4(a)) are used, but only a certain p percentage of them is employed to emulate the clinical situation. $p\%$ row means that $\frac{e-s}{m} = \frac{p}{100}$.

3. In view of the non-robustness of the L_2 error norm, in equation(11), the absolute difference is used instead, whence the serious noises present in the AS images cannot make too much negative impact. In our algorithm, the first optimization step is the call to equation (4) with the L_2 norm used thereof switched to L_1 to ensure the robustness and the corresponding change of actual rows used in the summation. This revised equation (4) is rewritten as a new equation below to facilitate notations similar to equation (11).

$$V_{i+1} = \operatorname{argmin}_{V \in [-l, l]} \sum_{j=s}^e |C[j] - C'_i[j+V+l]| \quad (12)$$

The new two-step breathing signal extraction algorithm from an input z -normalized AS image I is summarized below:

1. *First local search step:* call equation (12) to estimate initial breathing signal V ;
2. *Large-scale regularity evaluation:* from V call equations (8)–(10) to obtain the directional vector D ;
3. *Second constrained search step:* using D and call equation (11) to arrive at the final breathing signal B .

Figure 7 illustrates the breathing signals extracted using the two-step algorithm (figure 7(a)) and L_2 norm (figure 7(b)) based on the same patient data as in figures 3(a) and 4(a). The average breath per minute (bpm) evaluated from (figure 7(a)) is 26.24, compared to the reference mean bpm 26 that was obtained with the air bellows belt (see section 2.7 for details of the air bellows belt). The error is merely 0.24. As a contrast, if the L_2 error norm is used, as done by the reference AS procedure, the extracted breathing signal is illustrated in figure 7(b), the corresponding bpm is 24.73, with an error 1.27, farther away from that extracted by L_1 . This is typical for all the data sets that we tested on: the performance delivered by L_2 norm is always a little worse than that by use of L_1 error norm.

2.6. Impact of gantry rotation on the breathing signal

In contrast to the breathing signal tracking approaches using external sources such as the real-time position management (RPM) system (Varian Medical Systems, Palo Alto, CA) or the fluoroscopic tracking where both the patient and the imaging devices remain still, CBCT projections are acquired at various gantry angles. As briefly mentioned in the Introduction, varying gantry angles incur anisotropic attenuation on the projection images. Thus, the contrast of structures of interest deteriorates, which compromises the breathing signal extraction. To quantitatively assess the effect of the gantry rotation on the breathing signal

extraction, we scanned a Quasar phantom (Modus Medical Devices Inc., London, Canada) with programmed SI moving amplitude of 2.0 cm and a cycle of 4.0 s. The AS obtained from the phantom data and the extracted signal is shown in figure 8. The regular wavy pattern is conspicuously present on the AS image (figure 8(a)). As depicted in figure 8(b), the extracted signal (blue curve) faithfully reproduced the programmed sine wave (green curve). The relative error of the bpm is as small as 0.0049. This experiment demonstrates that the impact of gantry rotation is minimal on the extracted breathing signal cycle/frequency, in spite of the harmful signal diminishment from attenuation variations, which the first part of our effort is devoted to reestablish.

2.7. Case studies and evaluation

CBCT projection image sets from five patients (four lungs and one liver) were used in assessing the performance of the proposed algorithm. All the projections were acquired with the Varian Onboard Imager on the Clinac iX System Linear Accelerator (Varian Medical Systems, Palo Alto, CA). The imager has a geometric dimension of 40×30 cm with 1024×768 pixels in transverse and longitudinal directions correspondingly. All image sets have 640–645 projection views, acquired with a full circle of 360 degrees using a half fan CBCT acquisition mode. Reference waveforms of these patients were acquired with the air bellows belt from Philips (Philips Medical Systems, Cleveland, OH). The air bellows belt is an elastic belt positioned around the patient's abdomen, expanding and contracting with respiratory motion. It contains a pressure transducer, converting the pressure waveform to a voltage signal, which is then digitized and transmitted to the CT scanner system. The pressure waveform is equivalent to the breathing amplitude that can be utilized for comparison with the extracted signals. Normalized waveform difference between the reference and the computation could be theoretically used as a quantitative measure for the algorithm evaluation. However, owing to the fact that the waveforms are sensitive to perturbations, phase shifts often happen between the bellows' based signals and the computed signals since they were acquired at different time, which prevents meaningful comparison. Thus, the breathing frequency or bpm is compared to that from the breathing signals extracted from CBCT projections. The breathing frequency is chosen to be the quantitative indicator because it is a physiological strain index unique for every patient and does not change much over a short period of time, in our case a few weeks, during the radiation therapy course. The Quasar phantom experiment described in section 2.6 serves as a quantitative evaluation of the proposed algorithm since the ground truth is known in this situation.

3. Results

We applied this algorithm to five data sets with the reference waveforms, i.e. breathing signals obtained with the air bells belt. With eight choices of starting location s and ending location e as used in equations (11) and (12), the statistics of the estimated bpm's against the reference signals for all five patients are tabulated in figure 9.

In figure 9, for row size around 55% to 65%, the bpm calculated with the proposed algorithm, shown as the blue circles in the middle of each group, is exceedingly close to the

mean bpm of the reference signals denoted by the red circles. Whereas if the region is chosen relatively smaller, like 40% or 45%, the discrepancy from the reference mean bpm is larger, especially in the first two patients' data. This attributes to the fact that these regions happen to be in the upper lungs where respiratory motion is not so pronounced. When the chosen region is large, say, 65% to 75%, breathing motion can be detected reliably despite that for patient 3 the bpm fluctuates a bit. This might be caused by the unclear motion in this region. At any rate, all the extracted bpm, even with regions as small as 40% of the original size and as large as 75%, the extracted bpm are still well within the breath rate range of the reference signals. One need to know that the same patient could undergo different breathing patterns during two imaging periods that are often apart by one to two weeks, thus the extracted signal bpm agree with the reference ones fairly well. As a comparison, the bpm estimated by the original AS technique are also shown in each group as circles filled by green color. Due to the presence of more noises in the original AS method, the estimated bpm tend to be farther away from the known mean bpm: some are larger (patients 1, 2, and 4), while others are smaller (patients 3 and 5). In patients 1, 3, 4, the estimated bpm are close to or slightly outside the known upper or lower bounds. For patients 2 and 5, the bpm given by the original AS technique are indeed safely within the known bounds, but still clearly worse than those by the proposed algorithm with different row choices from 40% to 75%.

In figure 10, five breathing signals extracted by the proposed algorithm for each of the five patients with row size 60% are illustrated. The grand average of errors between the estimated bpm and the measured bpm is -0.07 with standard deviation 1.58. Since the average bpm is >14 , this average error is well below 1%. In table 1, the grand average and standard deviations of these errors for all eight different row size choices, namely, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, are presented. Here again the best results are attained in the middle, by 55% and 60% row size.

To further showcase the performance of extracted signals against the reference ones, we compare our project-extracted signals with the reference respiratory signals measured by belt system as the proxy of ground truths. To minimize the possible contributions of the signal magnitudes, both types of respiratory signals are first normalized within each respiration cycle, namely, setting 1 at end-inspiration and 0 at end-expiration. The resultant pair of reference and estimated respiration signals for patient 1 are overlapped and illustrated in figure 11(a). Due to the phase difference, although the bpm estimated by the proposed method are persistently close to those offered by belt bellows, as reported in figure 9 and table 1, the crude difference between this pair of signals cannot yield meaningful dissimilarity. As done in signal processing (Oppenheim and Schaffer 2009), the comparison should be conducted in the spectral domain.

As discussed in section 2.4 and illustrated by figure 5, the higher component (the two-peak phenomenon as presented section 2.4) around unit 10 ($0.41 \text{ Hz} = 24.6 \text{ bpm}$) in the PSD estimate by Welch's method is caused by respiration, the area due to the projection-measured within that region which differs (orange area in figures 11(b) and (d)) from those (blue area in figures 11(b) and (d)) owing to belt-measure signals, the supposedly ground truth, indicates the estimation error in the PSD domain. We thus use the ratio of the orange

area over the blue area as the dissimilarity measure of these two signals: the smaller this ratio is, the better the corresponding signal matching. In figures 11(a) and (b), the comparison between these two signals for patient 1 is depicted, where the dissimilarity measure is small at 0.06. To showcase an entirely different situation, in figures 11(c) and (d), the comparison between the belt-measured signal and the original AS technique based method is illustrated, where the dissimilarity measure is 0.71. The dissimilarity measures of the proposed algorithm for different row choices are summarized in figure 12(a). As a comparison, in figure 12(b), the dissimilarity measures due to the original AS technique for the five data sets are shown side by side with the dissimilarity measures for the proposed algorithm with eight different row choices (40–75%). It is observed that the proposed method outperforms the original AS technique for all five patients the most for patient 5 by around 30% and the least for patient 2 by 8.5%, which is similar to the observation reported in figure 9.

4. Discussions

AS technique was introduced as an alternative solution to phase sorting for respiratory correlated CBCT. It is powerful but its success depends largely on the projection image contrast, particularly the visibility of the oscillating anatomy on the projections. In reality, this is not a trivial issue partially owing to the projection angular variation. In Kavanagh *et al* (2009) they had a detailed description on how the gantry angle affecting the attenuation on the projection images. Other investigations using various techniques, such as feature extraction, optical flow models, and Fourier analysis, to extract breathing signals from CBCT projections similarly rely on the image contrast (Kavanagh *et al* 2009, Vergalasova *et al* 2012, Dhou *et al* 2013, Yan *et al* 2013, Park *et al* 2014, van Sornsens de Koste *et al* 2015). However, to the authors' best knowledge, few of them have showed the work on the improvement of the projection image contrast. In this study, we strived to enhance the image contrast, particularly improving the contrast for the oscillating structures. As a first attempt, conversion from the intensity image to attenuation image, i.e. re-normalization of the original projections based on the air intensity was carried out. Significant improvement was observed. However, for regions where x-ray was extensively attenuated after passing through the mobile structures, e.g. when gantry rotated to the contralateral side, this global normalization failed to yield satisfactory results. To tackle this issue, an adaptive robust *z*-normalization filtering was introduced as a local contrast enhancement on the AS image. As demonstrated in the results, reasonable image contrast was achieved such that continuous breathing signals could be rendered in this used-to-be problematic temporal zone.

How to reliably extract breathing signals without too much noise is of utmost importance. In previous work based on the AS technique (Sonke *et al* 2005, Yan *et al* 2013), a single step that minimized the mean square errors between two adjacent lines was employed to extract the signals. As shown in results, while this method could generate acceptable breathing signals, essential improvement could and should be made to more reliably and robustly extract breathing motion. The proposed adaptive robust *z*-normalization alone cannot combat this problem, either. However, although there are many local errors in the AS signal extraction approach, a global analysis using Welch's PSD estimation method reveals the global or large-scale regularity of the signal, with peaks in two different spectra, as

evidenced in figure 5. With the signals estimated by the AS technique as the starting point, we approximate this signal using two groups of coefficients in these two spectra and use the trends in this approximation to constrain a second optimization step between two adjacent lines, thus the global regularity of the signal is taken account of for the purpose of breathing signal extraction. Empirical studies conducted demonstrated promising performance in a consistent manner by use of this new two-step optimization algorithm based on global spectral density analysis. A comparison between results using the original AS and the proposed technique as illustrated in figures 9 and 12 demonstrated that significant improvement was accomplished.

Breathing signal extraction using CBCT projections including the AS technique often requires that distinctive structures especially the diaphragm be in the field of view to attain acceptable signals. However, for a large amount of thoracic cancer patients the tumor targets locate on the upper lobes of the lungs. As such, the extraction of a diaphragm based breathing signal would encounter a problem owing to the limited longitudinal field of view of the current Clinic mounted on board imager system. Additionally, owing to the half fan image acquisition mode, some projections might lose the oscillating structures significantly. In such challenging cases our proposed algorithm, as demonstrated in figure 9, would still function well. For the eight selected zones, although the bpms of the extracted signal show apparent variations, the standard deviations for all zones are reasonably consistent. All bpms fall in the range of the reference bpms from waveforms using the air bellows belt. This result indicates that our AS-based breathing signal extraction algorithm, unlike other similar methods, can yield satisfactory breathing signal in most thoracic regions.

Breathing signals extracted by the proposed algorithm largely agreed with the bellows belt based signals, as shown in the results. However, clinically of more relevance is that breathing signal, being a surrogate of the tumor motion, should correctly indicate the tumor's actual motion so that respiration correlated CT can be properly reconstructed. Otherwise, the subsequent treatment planning and the treatment outcome would be jeopardized. Studies have been carried out to predict tumor locations using external marker based respiratory signals and the discrepancies could be more than five mm (Ozhasoglu and Murphy 2002, Keall *et al* 2006, Ionascu *et al* 2007). CBCT projection based signals, on the other hand, are extracted directly from the patient's anatomy change over scanning time. Thus it is inherently correlated with the tumor motion. In principle, one could directly detect the tumor motion from the CBCT projections, e.g. as reported in (Yang *et al* 2012). However, owing to the inferior image contrast on the projections, sophisticated image processing is needed and considerable uncertainties still exist. Another option is to calculate the tumor motion using the reconstructed CBCT, but that requires the phase sorting first before the temporally correlated CBCT is available. Tumor motion can also be retrospectively obtained from the same patient's 4D fan beam CT when the patient was simulated for dosimetric calculation and planning (Hallman *et al* 2012, Chao *et al* 2015b). The downside of this choice is that 4DCT is often reconstructed with external marker based signal, but at least it provides the best motion information one could have. This study is underway and a research reporting our progress is forthcoming.

5. Conclusion

Anatomy feature plays a key role in yielding breathing signals from the CBCT projection images. AS technique has proved to be an efficient tool but its application is limited by the low image contrast and the signal extraction algorithms. In this work we endeavored to tackle these obstacles by introducing innovative algorithms (Wei *et al* 2014) into the signal extraction procedure. The adaptive image filter facilitated the contrast enhancement significantly and the two-step extraction method provided a robust algorithm to extract less noisy breathing signals. The new method will offer a practical solution to obtaining markerless breathing signal and help better control breathing motion in radiation therapy.

Acknowledgments

Ming Chao and Jie Wei contributed equally to this work. Wei's work is supported in part by NSF grant number IIP-0968922 and NIH grant number U54CA132378-8. We would like to thank Dr Sebastien Brousmiche of University in Louvain-la-Neuve, Belgium and Dr Hao Yan of UT Southwestern Medical Center at Dallas, TX for discussions and help at the early stage of this work. The authors would like to thank anonymous reviewers for the valuable suggestions that significantly improved the quality of the manuscript.

References

- Berbeco RI, Mostafavi H, Sharp GC, Jiang SB. Towards fluoroscopic respiratory gating for lung tumours without radiopaque markers. *Phys Med Biol*. 2005; 50:4481–90. [PubMed: 16177484]
- Chao M, Brousmiche S, Yuan Y, Rosenzweig K, Lo Y. SU-D-207-01: markerless respiratory motion tracking with contrast enhanced thoracic cone beam CT projections. *Med Phys*. 2015a; 42:3218.
- Chao M, Yuan Y, Sheu RD, Wang K, Rosenzweig KE, Lo YC. A feasibility study of tumor motion estimate with regional deformable registration method for 4-dimensional radiation therapy of lung cancer. *Technol Cancer Res Treat*. 2015b
- Dhou S, Motai Y, Hugo GD. Local intensity feature tracking and motion modeling for respiratory signal extraction in cone beam CT projections. *IEEE Trans Biomed Eng*. 2013; 60:332–42. [PubMed: 23193225]
- Ford EC, Mageras GS, Yorke E, Rosenzweig KE, Wagman R, Ling CC. Evaluation of respiratory movement during gated radiotherapy using film and electronic portal imaging. *Int J Radiat Oncol Biol Phys*. 2002; 52:522–31. [PubMed: 11872300]
- Gierga DP, Brewer J, Sharp GC, Betke M, Willett CG, Chen GT. The correlation between internal and external markers for abdominal tumors: implications for respiratory gating. *Int J Radiat Oncol Biol Phys*. 2005; 61:1551–8. [PubMed: 15817361]
- Hallman JL, Mori S, Sharp GC, Lu HM, Hong TS, Chen GT. A four-dimensional computed tomography analysis of multiorgan abdominal motion. *Int J Radiat Oncol Biol Phys*. 2012; 83:435–41. [PubMed: 22197238]
- Hoaglin, DC., Mosteller, F., Tukey, JW. *Understanding Robust and Exploratory Data Analysis*. New York: Wiley; 2000.
- Ionascu D, Jiang SB, Nishioka S, Shirato H, Berbeco RI. Internal–external correlation investigations of respiratory induced motion of lung tumors. *Med Phys*. 2007; 34:3893–903. [PubMed: 17985635]
- Kavanagh A, Evans PM, Hansen VN, Webb S. Obtaining breathing patterns from any sequential thoracic x-ray image set. *Phys Med Biol*. 2009; 54:4879–88. [PubMed: 19636080]
- Keall PJ, et al. The management of respiratory motion in radiation oncology report of AAPM Task Group 76. *Med Phys*. 2006; 33:3874–900. [PubMed: 17089851]
- Keall PJ, Starkschall G, Shukla H, Forster KM, Ortiz V, Stevens CW, Vedam SS, George R, Guerrero T, Mohan R. Acquiring 4D thoracic CT scans using a multislice helical method. *Phys Med Biol*. 2004; 49:2053–67. [PubMed: 15214541]
- Oppenheim, A., Schaffer, R. *Discrete-Time Signal Processing*. Englewood Cliffs, NJ: Prentice Hall; 2009.

- Ozhasoglu C, Murphy MJ. Issues in respiratory motion compensation during external-beam radiotherapy. *Int J Radiat Oncol Biol Phys*. 2002; 52:1389–99. [PubMed: 11955754]
- Park JC, Kim JS, Park SH, Webster MJ, Lee S, Song WY, Han Y. Four dimensional digital tomosynthesis using on-board imager for the verification of respiratory motion. *PLoS ONE*. 2014; 9:e115795. [PubMed: 25541710]
- Sonke JJ, Zijp L, Remeijer P, van Herk M. Respiratory correlated cone beam CT. *Med Phys*. 2005; 32:1176–86. [PubMed: 15895601]
- van Sornsen de Koste JR, Dahele M, Mostafavi H, Sloutsky A, Senan S, Slotman BJ, Verbakel WF. Markerless tracking of small lung tumors for stereotactic radiotherapy. *Med Phys*. 2015; 42:1640–52. [PubMed: 25832054]
- Vedam SS, Kini VR, Keall PJ, Ramakrishnan V, Mostafavi H, Mohan R. Quantifying the predictability of diaphragm motion during respiration with a noninvasive external marker. *Med Phys*. 2003; 30:505–13. [PubMed: 12722802]
- Vergalasova I, Cai J, Yin FF. A novel technique for markerless, self-sorted 4D-CBCT: feasibility study. *Med Phys*. 2012; 39:1442–51. [PubMed: 22380377]
- Wei J. Color object indexing and retrieval in digital libraries. *IEEE Trans Image Process*. 2002; 11:912–22. [PubMed: 18244685]
- Wei J, Cai B, Zhang L, Fu BM. Automatic classification and quantification of cell adhesion locations on the endothelium. *Ann Biomed Eng*. 2015; 43:1803–14. [PubMed: 25549777]
- Wei J, Li G. Automated lung segmentation and image quality assessment for clinical 3D/4D computed tomography. *IEEE J Transl Eng Health Med*. 2014; 2:1800110. [PubMed: 25621194]
- Wei J, Yuan A, Li G. An automatic toolkit for efficient and robust analysis of 4D respiratory motion (best in physics). *Med Phys*. 2014; 41:473.
- Yan H, Wang X, Yin W, Pan T, Ahmad M, Mou X, Cervino L, Jia X, Jiang SB. Extracting respiratory signals from thoracic cone beam CT projections. *Phys Med Biol*. 2013; 58:1447–64. [PubMed: 23399757]
- Yang Y, Zhong Z, Guo X, Wang J, Anderson J, Solberg T, Mao W. A novel markerless technique to evaluate daily lung tumor motion based on conventional cone-beam CT projection data. *Int J Radiat Oncol Biol Phys*. 2012; 82:e749–56. [PubMed: 22330989]
- Yoo S, Kim GY, Hammoud R, Elder E, Pawlicki T, Guan H, Fox T, Luxton G, Yin FF, Munro P. A quality assurance program for the on-board imagers. *Med Phys*. 2006; 33:4431–47. [PubMed: 17153422]

Figure 1. The top plots (a) and (b) are the gray scale value profiles along the dashed lines on projection images before the intensity to attenuation conversion (c), and after conversion (d), respectively.

Figure 2. Illustration of the steps used to generate the AS image: (a) original projection image; (b) logarithmic transform and superior-inferior derivative to enhance features; (c) horizontally summed pixels; (d) concatenation of all projections to form a 2D AS image (cropped to a smaller region showing the wavy pattern).

Figure 3.
Low quality of the 2D AS image (a) and the extracted signal (b).

Figure 4.

The AS image shown in figure 3 is processed by the proposed adaptive robust z -normalization step. The resultant image and the corresponding extracted breathing signal based on the original AS procedure are shown on the left (a) and right (b).

Figure 5.

The PSD estimated by Welch's method, the densities after position 50 are of extremely small magnitudes and are thus discarded for better visualization of important details in the regions of interest. One unit on the x -axis is 0.041 Hz.

Figure 6.

V (blue) and V' (green) of two AS-estimated breathing signals for two patients. (a) is for the patient's signal shown in figure 3(b) and (b) is for another patient. The quality of the signals shown in (b) is slightly better.

Figure 7.

(a) Breathing signals extracted by the proposed two-step algorithm, estimated bpm is 26.24.

(b) Extracted breathing signals when L_2 error norm is used in equations (12) and (11), estimated bpm is 24.73. The mean bpm from the air bellows belt is 26.

Figure 8.

(a) The AS image from the Quasar phantom with predefined motion along SI moving amplitude of 2.0 cm and motion cycle of 4.0 s. (b) The extracted signal (blue) overlapped with the known programmed sine wave (green), the relative error of the extracted bpm is merely 0.0049.

Figure 9.

The reference bpm for five patient data sets as numbered in the horizontal axis: red circle—mean bpm, red bars—breath rate range. The average bpm estimated by the proposed algorithm for eight different row sizes: the blue circles from left to right within each of the five data sets: 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%. The bpm computed by the original method for all five data sets are shown as circles filled by green color in each group.

Figure 10.

Breathing signals extracted by the proposed algorithm with 60% row size with difference choices of (s, e) (shown in the first row) as used in equations (12) and (11) for each column. Each row corresponds to one patient, D1–D5 are the patients 1–5 in figure 9.

Figure 11.

Normalized respiratory signals measured by the proposal algorithm (orange) and belt bellows (blue). (a) The overlapped signals in the spatial domain. (b) The corresponding power spectral density (PSD) estimated by Welch's method. The dissimilarity of the projection-based signal over the belt-measure one is measured by the ratio of the difference higher component, the orange area in (b), over the area of the belt-measured higher component in the PSD, the blue area in (b), which is owing to the respiratory motion only, see figure 5 for the two-peak phenomenon and corresponding discussions in section 2.4. Panels (a) and (b) are the pair for patient 1 using the proposed algorithm with small similarity (0.06), while panels (c) and (d) are the pair for patient 5 using the original AS technique, where the dissimilarity, the orange area over the blue area, is far larger (0.71).

Figure 12.

(a) Dissimilarity measures for the proposed algorithm for all five data sets, with different row choice. (b) Dissimilarity measures by original AS technique in blue color and the proposed algorithm with eight different row choices (40%–75%) in red color for all five data sets.

Table 1

Grand means and standard deviations of the bpmms of the breathing signals extracted by the proposed algorithm from the measured bpmms. Each column corresponds to a difference choice of row size.

	40%	45%	50%	55%	60%	65%	70%	75%
Mean	-1.40	0.08	-0.33	-0.44	-0.07	-0.49	-0.99	-0.91
Std	4.08	4.63	4.95	3.03	3.46	3.69	2.75	3.12