



# Evaluation Methodology for Respiratory Signal Extraction from Clinical Cone-Beam CT (CBCT) using Data-Driven Methods

Adam Tan Mohd Amin<sup>1</sup>, Siti Salasiah Mokri<sup>1</sup>, Rozilawati Ahmad<sup>2</sup>, Fuad Ismail<sup>3</sup>, Ashrani Aizzuddin Abd Rahni<sup>1\*</sup>

<sup>1</sup>Faculty of Engineering and Built Environment,  
Universiti Kebangsaan Malaysia, Bangi, 43600, MALAYSIA

<sup>2</sup>Faculty of Health Science,  
Universiti Kebangsaan Malaysia, Jalan Raja Muda Abdul Aziz, Kuala Lumpur, 50300, MALAYSIA

<sup>3</sup>Universiti Kebangsaan Malaysia Medical Centre,  
Jalan Yaacob Latif, Bandar Tun Razak, Kuala Lumpur, 56000, MALAYSIA

\*Corresponding Author

DOI: <https://doi.org/10.30880/ijie.2021.13.05.001>

Received 26 March 2021; Accepted 30 April 2021; Available online 31 July 2021

**Abstract:** The absence of a ground truth for internal motion in clinical studies has always been a challenge to evaluate developed methods to extract respiratory motion especially during a 60-second cone-beam CT (CBCT) scan in Image-Guided Radiotherapy Treatment (IGRT). The unavailability of a gold standard has led this study to present a methodology to manually track respiratory motion on a clinically acquired CBCT projection data set over a 360° view angle. The tracked signal is then used as a reference to assess the performance of four data-driven methods in respiratory motion extraction, namely: the Amsterdam Shroud (AS), Local Principal Component Analysis (LPCA), Intensity Analysis (IA), and Fourier Transform (FT)-based methods that do not require additional equipment nor protocol to the existing treatment delivery. The assessment using this reference signal includes both quantitative and qualitative analysis. It is found out quantitatively that all four methods managed to extract respiratory signals that are highly correlated with the reference signal, with the LPCA method displaying the highest correlation coefficient value at 0.9108. Furthermore, the normalized root-mean-squared amplitude error of detected peaks and troughs within the signal from the LPCA method is also lowest at 1.6529 % compared to the other methods. This result is further supported by qualitative analysis via visual inspection of each extracted signal plotted with the reference signal on the same axes.

**Keywords:** Cone-beam CT, image-guided radiotherapy treatment, data-driven methods, respiratory motion, internal motion

## 1. Introduction

Image-Guided Radiation Therapy (IGRT) has played a crucial role in radiotherapy treatment delivery for cancer patients. This is especially true for treatments involving tumors located at sites such as lung and upper abdomen that are largely affected by involuntary internal motion. Motion such as that due to respiration are vividly visible from the lungs observed on projections acquired during a 1-minute Cone-Beam CT (CBCT) scan. A 3D image that is reconstructed

\*Corresponding author: [ashrani@ukm.edu.my](mailto:ashrani@ukm.edu.my)

2021 UTHM Publisher. All rights reserved.

[penerbit.uthm.edu.my/ojs/index.php/ijie](http://penerbit.uthm.edu.my/ojs/index.php/ijie)

without taking this motion into account is considered to be a compromised averaged image that actually contains motion artefacts, thus reducing the accuracy of the treatment. However, the unavailability of a ground truth that represents a patient's internal motion poses a challenge to clinically validate modern developed motion tracking algorithms [1].

To overcome this, many researchers have adopted various methods such as tracking the respiratory signal on patients via strapping a bellows belt [2] or utilizing surface guided radiotherapy (SGRT) using optical cameras [3], [4] to detect variations in the patients' chest and abdomen as a surrogate signal. Others also used manual tracking of visible moving features on raw projection data [4]-[8] thus producing a reference signal without using additional equipment or tampering with the existing treatment protocols.

A methodology to evaluate extracted respiratory signals from clinical CBCT projection data is presented in this paper. Since it is difficult to define a gold standard ground truth for respiratory motion, a reference signal is manually tracked and extracted as a viable alternative. Using this reference signal, the performance of different data driven methods to extract respiratory signals from raw CBCT projection data are able to be evaluated. Qualitatively, each extracted signal is assessed by plotting it with the manually tracked reference signal on the same axes. The signals are also assessed quantitatively via calculating the correlation coefficient between each extracted signal and the reference signals. The normalized root-mean-squared error values of the signals' peak and trough positions are also calculated to determine their accuracies.

## 2. Methodology

A set of projections from a CBCT scan with Half-Fan scan operating mode [9] of a curative lung cancer patient based on the low-dose thorax protocol is used. The projections were acquired using the Varian On-Board Imager (OBI, Varian Medical Systems, Palo Alto, CA), at the Oncology and Radiotherapy Department, Universiti Kebangsaan Malaysia (UKM) Medical Center. Approval by the UKM Research Ethics Committee with the ethics code: NN-2017-118 was obtained prior to patient data acquisition. A total of 635 2D projections were recorded by the OBI between angles approximately  $90^\circ$  to  $-270^\circ$  in a  $360^\circ$  anti-clockwise rotation around the patient over a 1-minute time span. A summary of the parameters used in acquiring the projection data is shown in Table 1 below.

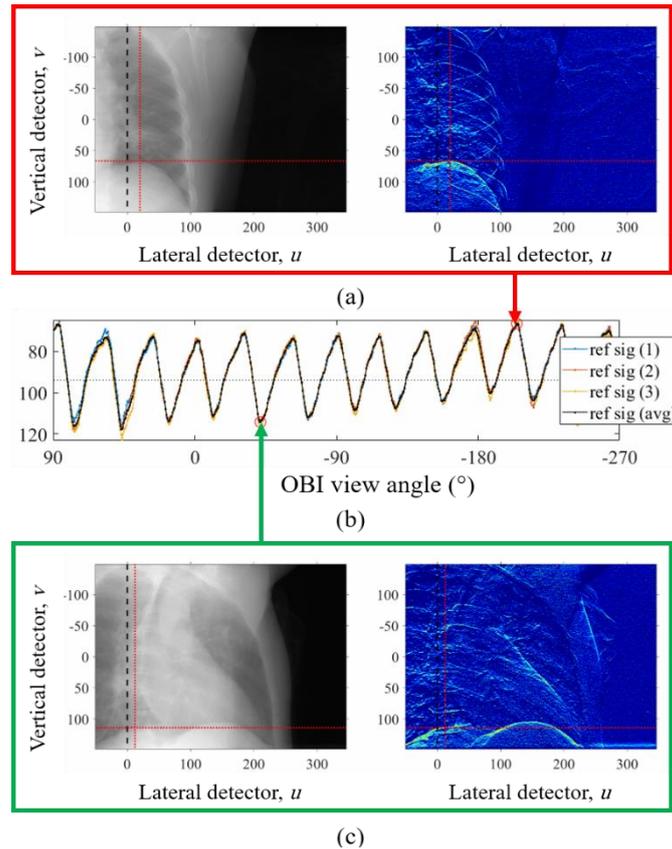
**Table 1 - Varian On-Board Imager (OBI) parameters for low-dose thorax protocol**

OBI Parameters	Values
Operating mode	Half-Fan
Projection views, $N$	635
Acquisition period, $T$ (s)	60
OBI start angle, $\theta_i$ ( $^\circ$ )	89.52
OBI stop angle, $\theta_e$ ( $^\circ$ )	-269.34
Detectors lateral, $u$ (pixels)	1024
Detectors vertical, $v$ (pixels)	768
Detector lateral size (mm)	397.31
Detector vertical size (mm)	297.98
Detector lateral offset (mm)	148
CTNC number range	273 to 284

### 2.1 Reference Respiratory Signal

The raw projection data are first normalized thus converting the intensity values to a more precise attenuation coefficient number, based on the CT Norm Chamber (CTNC) numbers [10] that is a unique feature in most Varian systems. Each projection is then edge enhanced and visualized in a 'jet' colormap scheme, where the coordinates on the 2D projection view ( $u, v$ ) corresponding to the apex one of the most prominent hemidiaphragm of the lung is identified and recorded. A straight horizontal line representing the longitudinal,  $v$ -coordinate and a vertical line representing the vertical,  $u$ -coordinate of the apex is then plotted on the projection, where the step in identifying the same hemidiaphragm apex is repeated on all of the available projections over the  $360^\circ$  view. An illustration of this process is depicted in Fig. 1. The coordinates for the apex of the same hemidiaphragm when the gantry rotates are shown for extreme conditions during breathing, i.e. (a) end-expiration at view angle,  $\theta = -204.6^\circ$  with identified ( $u, v$ ) coordinates (183, 67), and (c) end-inspiration at view angle,  $\theta = -41.8^\circ$  with identified coordinates: (163, 114). Only the longitudinal,  $v$ -coordinate is considered as the final reference respiratory signal, thus reflecting the lung volume variation in the Superior-Inferior (SI) direction.

To account for unwanted errors while the coordinates are identified, the steps are repeated 3 times for each set of projection data where an average signal is obtained, as shown in Fig. 1(b). This average signal is thus used as the reference respiratory signal.



**Fig. 1 - Manual respiratory signal tracking on raw and edge-enhanced projection images, the latter viewed in ‘jet’ colormap scheme: (a) end-expiration phase, (b) the extracted respiratory signals, and (c) end-inspiration phase**

## 2.2 Data Driven Methods

Four data-driven methods to directly extract respiratory signal from the acquired projection data are compared. A description of each method is given as follows:

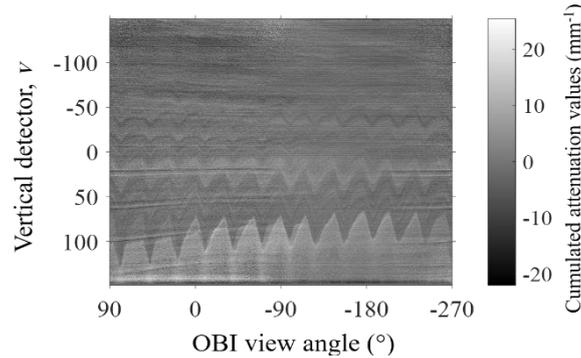
### 2.2.1 Amsterdam Shroud (AS) Method [2]

Each edge enhanced 2D projection view is summed laterally, generating a column vector that consists the vertical intensities that corresponds to the Superior-Inferior (SI) condition of that particular projection view. Thus, when all of the vertical column vectors for all views are concatenated into a 2D image, the observed pattern represents the SI variation across the entire 360° projection views. The 2D image is known as the Amsterdam Shroud (AS) image, as shown Fig. 2. As suggested by literature, an improved version of the original AS method [11] can be achieved by improving the image quality of this 2D AS image, i.e. enhancing the edges features, such as incorporating an adaptive  $z$ -normalization filter. The AS method then works to extract the 1D signal that is contained within the 2D AS image. This includes comparing consecutive column vectors and identifying the amount of pixel shift that is present using the L2-minimization criterion. However, this method requires an additional band-pass filtering step since apart from the desired respiratory signal, the SI variation from the concatenated vector columns also includes other motions such as internal cardiac motion, and anisotropic intensity-attenuation variability on the projection data due to different angle views during gantry rotation. The band-pass filter is set within the range of 0.20 Hz to 0.33 Hz that corresponds to a typical person’s respiration rate at 12-20 respirations per minute [12].

### 2.2.2 Local Principal Component Analysis (LPCA) Method [8]

In the original implementation of LPCA, a foreground AS image is first generated by removing the background image elements via a total variation,  $TV/L^1$  model from the 2D AS image. Then, PCA is employed locally via sliding a

window throughout the foreground-AS image with an adequate size of  $w = 55^\circ$ . The window is slid sequentially for each projection view across the OBI view angles, in which the principal component eigenvectors for consecutive sliding windows are compared at each step. The most correlated pairs of eigenvectors are evaluated where the corresponding principal component coefficient is kept as the eventual extracted respiratory signal.



**Fig. 2 - Amsterdam Shroud (AS) image**

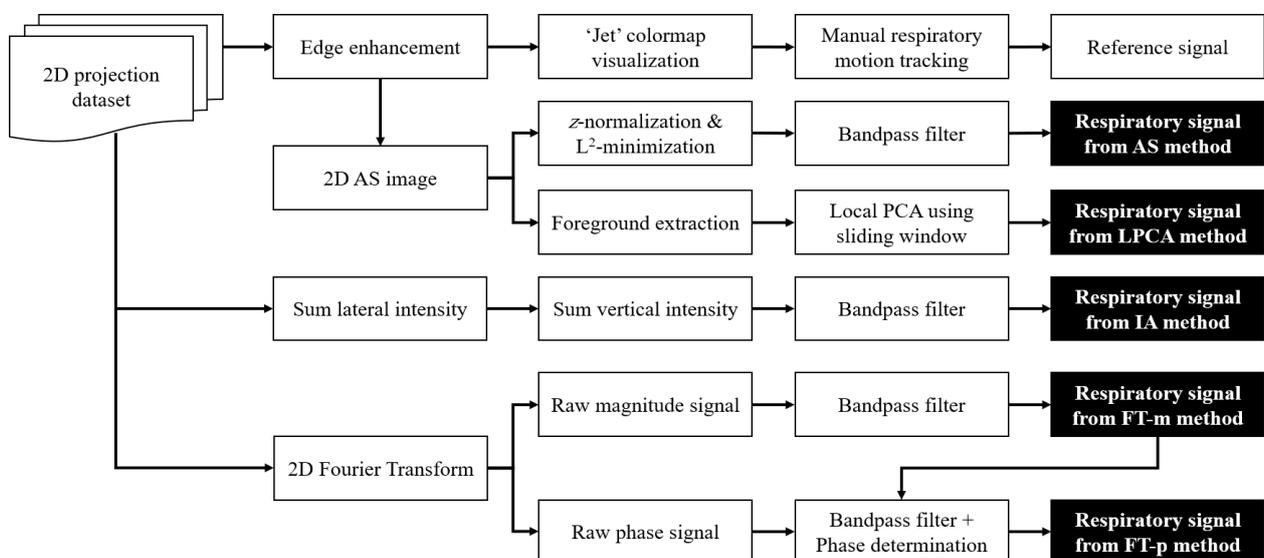
### 2.2.3 Intensity Analysis (IA) Method [13]

This method provides an alternative to the 1D signal extraction step from the AS method. Instead of just summing along the lateral direction of each projection view, it extends the summation along the vertical direction of the AS image. Thus, the summation output of both lateral and vertical direction is a representation of intensity-attenuation pixel information for each projection view. These values when plotted across all of the view angles generates a 1D signal. However, a similar band-pass filter is also required to distinguish the desired respiratory signal from all other signals it contains.

### 2.2.4 Fourier Transform (FT)-based Methods [7]

In this method the Fourier Transform (FT) of each 2D projection is found. Both the magnitude (FT-m) and phase (FT-p) of the FT can be used to extract a respiratory signal. In the FT-magnitude (FT-m) method, the absolute values of the FT at the origin (0, 0) in Fourier space are kept, producing a signal similar to the IA method that is essentially the intensity-attenuation pixel or magnitude information for each projection view. On the other hand, the FT-phase (FT-p) method depends on the basic theory that any physical variation that occurs geometrically in Cartesian space would result to a phase shift in Fourier space. Thus, the first phase values (0, 1) in Fourier space that corresponds to the vertical SI variation in all of the projection data, are kept and plotted. Both approaches require a bandpass filter to discern the desired respiratory signal.

A summary of the four aforementioned data-driven methods is shown in Fig. 3.



**Fig. 3 - Flowchart of the methodology used to extract the reference respiratory signal and implementation of four data driven methods**

### 2.3 Extracted Signal Evaluation

Using the manually tracked reference respiratory signal described earlier, all of the signals extracted from the AS, LPCA, IA, and FT-based data driven methods are respectively evaluated using quantitative and qualitative assessments.

Quantitatively, the correlation between each of the extracted signals,  $s$  and the reference signal,  $r$  are evaluated based on the Pearson linear correlation coefficient,  $\rho$  described by (1), where  $n$  is the projection view,  $n = 1, 2 \dots N$ . This is to gauge the overall correlation of each extracted signal with the reference, in which higher correlation is represented by values that are closer to 1.

$$\rho = \frac{\sum (s_n - \bar{s})(r_n - \bar{r})}{\sqrt{\sum (s_n - \bar{s})^2 \sum (r_n - \bar{r})^2}} \quad (1)$$

Intuitively, a qualitative assessment is done by visualizing the extracted signals with respect to the reference signal on the same plot axes respectively. The extreme normalized signal amplitude values:  $-1$  and  $1$ , each represents the deepest end-inspiration and shallowest end-expiration conditions during breathing.

Additionally, the peaks and troughs of the signals are detected based on a deflection point detection algorithm [14], in which the normalized root-mean-squared percentage error,  $e$  of the detected points between the extracted signals,  $ds$  and reference signal,  $dr$  respectively are also evaluated based on (2), where  $m = 1, 2 \dots M$  corresponds to the detected peaks and troughs for each signal. The errors are assessed based on the extracted signal's phase,  $e_p$  and normalized amplitude displacement error,  $e_a$ . Both metrics reflect the accuracies of the extracted signals since one of the objectives of respiratory motion tracking is for the projection data to be sorted according to either phase and/or amplitude binning prior to 4D retrospective reconstruction, in which the occurrence accuracy of peaks and troughs are essential. Percent errors closer to 0% indicate a higher accuracy in the tracked peaks and troughs of an extracted signal.

$$e = \frac{1}{\max(dr) - \min(dr)} \sqrt{\frac{\sum (dr_m - ds_m)^2}{M}} \quad (2)$$

### 3. Results and Discussion

The evaluated correlation coefficient values are as shown in Fig. 4. Generally all of the extracted respiratory signals using the data-driven methods display a high correlation with respect to the manually tracked reference signal, with the LPCA method being the highest. This indicates that preliminarily, all of the data-driven methods described in this study are able to extract respiratory signal from the acquired actual patient 2D projection data. Nonetheless, a detailed observation on the qualitative assessment of the extracted signals shown in Fig. 5 provides more insight on the actual performance of these methods.

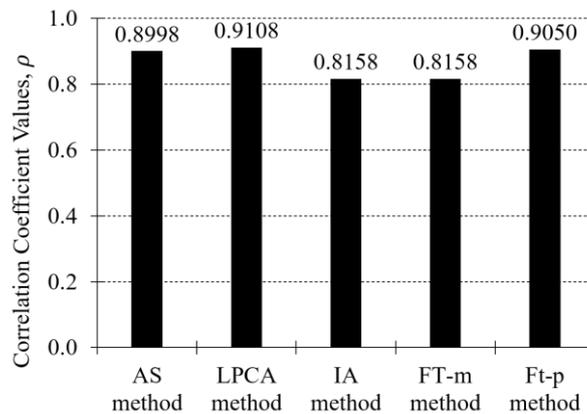
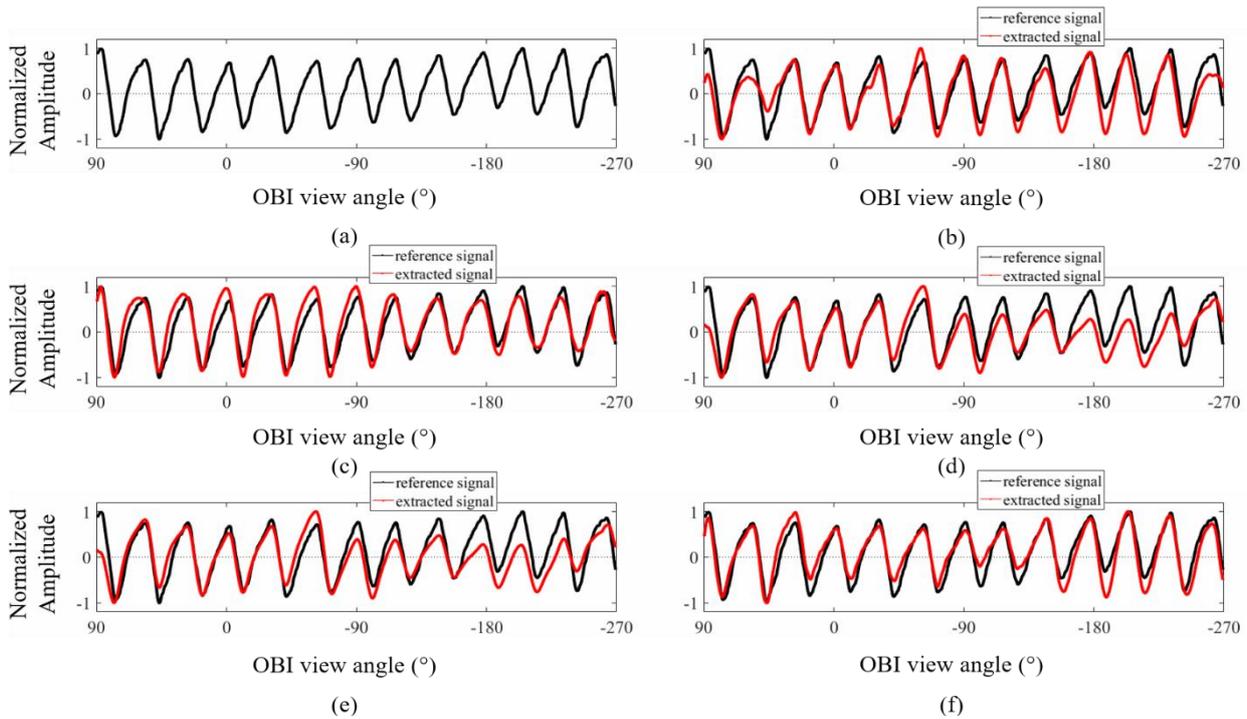


Fig. 4 - Correlation coefficient values,  $\rho$  of the respiratory signals extracted using the data-driven methods



**Fig. 5 - Extracted respiratory signals compared with (a) reference signal using (b) Amsterdam Shroud (AS) method, (c) Local Principal Component Analysis (LPCA) method, (d) Intensity Analysis (IA) method, (e) Fourier Transform-magnitude (FT-m) method, and (f) Fourier Transform-phase (FT-p) method**

As described in the Methodology section, the signals extracted from IA and FT-m methods should be similar, and both methods do yield the same results: both with the same correlation coefficient value at 0.8158, and also having the same pattern as shown in Fig. 5(d) and (e). However, during the OBI view angles ranging approximately between  $-90^\circ$  to  $-270^\circ$ , i.e. the second half of the  $360^\circ$  gantry rotation, the amplitudes are significantly inaccurate although the phase of the extracted signals correlate with the reference signal. Since both methods rely heavily on the total amount of intensity-attenuation values that is contained by each 2D projection image data, any distorted variation of total intensity values would affect the results. Here, since the Half-Fan (HF) operating mode is implemented for a thoracic/abdomen protocol, a significant distinction between the first and second half of the  $360^\circ$  gantry rotation would be the visibility of the beating heart. It is known that the anatomical location of the heart is closer to the left-lateral position of the patient. Therefore, due to the lateral shift of the detector in HF mode to achieve a larger field-of-view, the heart is only visible during the first half of the rotation.

On the other hand, for the extracted respiratory signal using AS method, although having a higher correlation coefficient value at 0.8998, it can be observed in Fig. 5(b) that there exists potential phase inaccuracies and false peaks during the first half of the OBI view angles ranging between  $90^\circ$  to  $-90^\circ$ . This may be due to the nature of the AS method that is significantly dependent on the quality of the oscillation features in the AS image shown in Fig. 2. During this period, the visibility of the beating heart actually affects the clarity of the amplitude variation of the diaphragm. The heart obscures the view of the diaphragm and other respiratory motion-affected organs hence exacerbating the performance to extract the respiratory signal.

Apart from that, another disadvantage of the AS method is its vulnerability of having to use a bandpass filter to distinguish the desired respiratory signal from other motions that is contained within the AS image. In fact, this situation is also apparent on the other IA and FT-based methods where the bandpass filter is needed to decouple apparent motions from different sources (respiration and cardiac) and angular variation mentioned earlier. To avoid false extremes: both peaks and troughs, the parameters in the bandpass filter can be optimized to achieve a better signal performance of course with the expense robustness of the algorithm.

Although the FT-p method displayed favorable results: with a correlation coefficient of 0.9050, and minimal amplitude discrepancies when compared to the reference signal as shown in Fig. 5(f), this method requires an additional step in determining the phase direction of the signal. The FT-p method utilizes the direction of the acquired signal by FT-m method, as shown in Fig. 3, since the latter method - similar to the IA method, is not susceptible to signal direction confusion.

Analyzing the problem at hand, one might be intrigued to eventually use Principal Component Analysis (PCA) to address the issue of distinguishing different motion signals from one another, since that is the main concept behind its application. The LPCA method used here managed to display superior performance in terms of highest correlation coefficient value at 0.9108 and is considered to be the best extracted signal with respect to the reference signal shown in Fig. 5(d).

The deflection point detection algorithm [14] determined that there is a total of  $M = 25$  peaks and troughs from the respiratory signal within the 1-minute period. An example of the algorithm detecting the peaks and troughs of both the extracted signal using LPCA method along with the manually extracted reference signal as shown in Fig. 6. Two metrics were evaluated to further support the current results, namely the normalized root-mean-squared percentage error for phase,  $e_p$  and amplitude,  $e_a$  differences as described in the Methodology section earlier. Fig. 6 displays the means in which both  $e_p$  and  $e_a$  are measured ( $m = 3$  and  $m = 24$  respectively, where  $m = 1, 2 \dots M$ ) from the OBI view angle ( $^\circ$ )- and normalized amplitude-axes. Locating the occurrences of these maxima and minima points is essential to the next step in dynamic reconstruction, where the projection views could be sorted into bins either via phase- and/or amplitude-binning [15]. The phase and amplitude NRMSE for all methods are as shown in Table 2.

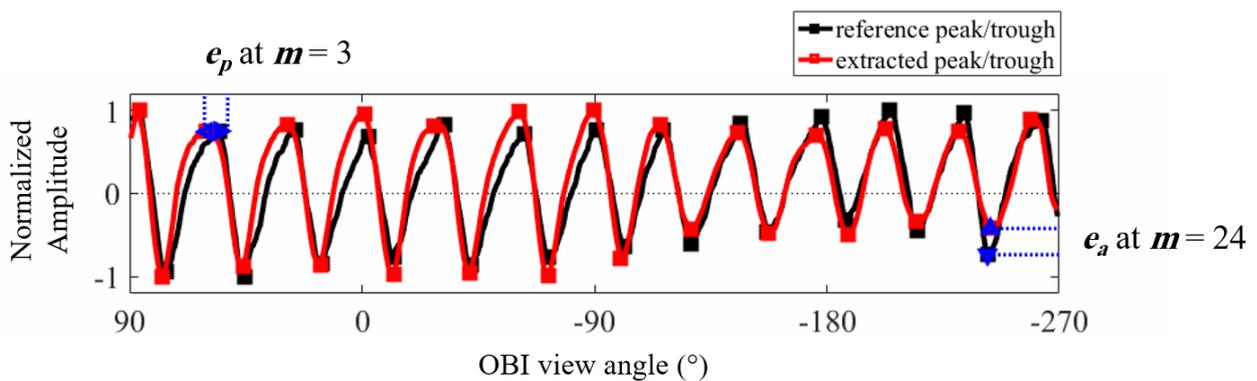


Fig. 6 - Identified peaks and troughs of the extracted signal using LPCA method with the reference signal

Table 2 - % Normalized Root Mean Squared Error (NRMSE) values of the detected peaks and troughs occurrences between the extracted respiratory and reference signals

% NRMSE	AS	LPCA	IA	FT-m	FT-p
Phase, $e_p$ (%)	0.0801	0.1134	0.0462	0.0462	0.0514
Amplitude, $e_a$ (%)	2.9529	1.6529	3.5813	3.5813	2.3170

Based on the values shown in Table 2, it can be seen that the LPCA method displayed the least amplitude error,  $e_a$  at only 1.6529% compared to the other data-driven methods. However, the phase error,  $e_p$  for the LPCA method is the highest at 0.1134%. Hence the best extraction method if phase-binning is to be used as the sorting method in 4D reconstruction is either the IA (or FT-m both of which are actually the same) method. Thus, the LPCA method would still be preferred if a decision is to be made by also optimally considering the performance of all of the data-driven methods based on the correlation assessment. The robustness of using the LPCA method is apparent, since it does not require the additional step of filtering the extracted signal.

#### 4. Conclusion

The difficulty of defining a data driven gold standard ground truth for internal motion has posed a challenge to clinically validate developed methods and algorithms since to the knowledge of the authors there is no such standard to measure the respiratory signal. In this study, a methodology to manually track a reference respiratory signal from an acquired lung cancer patient CBCT projection data over a 360° view angle is described. The reference signal is then used to compare and evaluate four main data driven methods. All methods are able to extract the desired respiratory signal, with the LPCA method yielding the highest correlation value with the reference signal at 0.9108, and lowest normalized root-mean-squared amplitude error at 1.6529% thus indicating its robustness.

#### Acknowledgement

This research is funded by the Ministry of Education Malaysia and Universiti Kebangsaan Malaysia (UKM) under the grant number: FRGS/1/2019/TK04/UKM/02/5 and GGP-2019-004.

## References

- [1] Vergalasova, I., & Cai, J. (2020). A modern review of the uncertainties in volumetric imaging of respiratory-induced target motion in lung radiotherapy, *Med. Phys.*, 47, 10, e988 - e1008.
- [2] Chao, M., Wei, J., Li, T., Yuan, Y., Rosenzweig, K. E., & Lo, Y. C. (2016). Robust breathing signal extraction from cone beam CT projections based on adaptive and global optimization techniques, *Phys. Med. Biol.*, 61, 8, 3109-3126.
- [3] Akintonde, A., McClelland, J., Grimes, H., Moinuddin, S., Sharma, R. A., Rit, S., & Thielemans, K. (2017). Data Driven Cone Beam CT Motion Management for Radiotherapy Application, 2017 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), 1-4.
- [4] Tsai, P., Yan, G. Liu, C., Hung, Y., Kahler, D., L., Park, J., Potter, N., Li, J. G., & Lu, B. (2020) Tumor phase recognition using cone-beam computed tomography projections and external surrogate information, *Med. Phys.*, 47, 10, 5077-5089.
- [5] Wei, J., & Chao, M. (2018). A constrained linear regression optimization algorithm for diaphragm motion tracking with cone beam CT projections, *Phys. Medica*, 46, 7-15.
- [6] Zhang, L., Zhang, Y., Zhang, Y. Harris, W. B., Yin, F., Cai, J. & Ren, L. (2017). Markerless four-dimensional-cone beam computed tomography projection-phase sorting using prior knowledge and patient motion modeling: A feasibility study, *Cancer Transl. Med.*, 3, 6, 185-193, 2017.
- [7] Vergalasova, I., Cai, J., Giles, W., Segars, W. P., & Yin, F. F. (2013). Evaluation of the effect of respiratory and anatomical variables on a Fourier technique for markerless, self-sorted 4D-CBCT, *Phys. Med. Biol.*, 58, 20, 7239-7259.
- [8] Yan, H., Wang, X., Yin, W., Pan, T., Ahmad, M., Mou, X., Cerviño, L., Jia, X., & Jiang, S. B. (2013). Extracting respiratory signals from thoracic cone beam CT projections, *Phys. Med. Biol.*, 58, 5, 1447-1464.
- [9] Mohd Amin, A. T., Abd. Rahni, A. A., Mokri, S. S. & Ahmad, R. (2017). Modeling the Varian On-Board Imager (OBI): Cone-beam CT (CBCT) operating modes, 2017 IEEE Int. Conf. Signal Image Process. Appl., 117-122.
- [10] Cropp, R. J. (2011). Implementation of respiratory-correlated cone-beam CT on Varian linac systems," University of British Columbia.
- [11] Zijp, L., Sonke, J., & van Herk, M., (2004). Extraction of the respiratory signal from sequential thorax Cone-Beam X-ray images," *Int. Conf. Use Comput. Radiat. Ther.*, 507-509.
- [12] Lindh, W. Q., Pooler, M., Tamparo, C. D., Dahl, B. M. & Morris, J. (2013). *Delmar's comprehensive medical assisting: administrative and clinical competencies* (5th ed.). Cengage Learning.
- [13] Kavanagh, A., Evans, P. M., Hansen, V. N., & Webb, S. (2009). Obtaining breathing patterns from any sequential thoracic x-ray image set," *Phys. Med. Biol.*, vol. 54, no. 16, 4879-4888, 2009.
- [14] Samir, M., Golkar, E., & Abd. Rahni, A. A. (2015). Comparison between the Kinect™ V1 and Kinect™ V2 for Respiratory Motion Tracking, 2015 IEEE Int. Conf. Signal Image Process. Appl., 150-155.
- [15] O'Brien, R. T., Cooper, B. J., Kipritidis, J., Shieh, C. C., & Keall, P. J., (2014). Respiratory motion guided four dimensional cone beam computed tomography: Encompassing irregular breathing, *Phys. Med. Biol.*, 59, 3, 579-595.