Remote, Depth-based Lung Function Assessment

Vahid Soleimani, *Student Member, IEEE,* Majid Mirmehdi, *Senior Member, IEEE,* Dima Damen, *Member, IEEE,* James Dodd, Sion Hannuna, Charles Sharp, Massimo Camplani, *Member, IEEE,* and Jason Viner

***Abstract*—*Objective:* We propose a remote, non-invasive approach to develop Pulmonary Function Testing (PFT) using a depth sensor. *Method:* After generating a point cloud from scene depth values, we construct a 3D model of the subject's chest. Then, by estimating the chest volume variation throughout a sequence, we generate volume-time and flow-time data for two prevalent spirometry tests: Forced Vital Capacity (FVC) and Slow Vital Capacity (SVC). *Tidal volume* and *main effort* sections of volume-time data are analysed and calibrated separately to remove the effects of a subject's torso motion. After automatic extraction of keypoints from the volume-time and flow-time curves, seven FVC (*FVC*, *FEV1*, *PEF*, $FEF_{25\%}$, $FEF_{50\%}$, $FEF_{75\%}$ and $FEF_{25-75\%}$) and four SVC measures (*VC*, *IC*, *TV* and *ERV*) are computed and then validated against measures from a spirometer. A dataset of 85 patients (529 sequences in total), attending respiratory outpatient service for spirometry, was collected and used to evaluate the proposed method. *Results:* High correlation for FVC and SVC measures on intra-test and intra-subject measures between the proposed method and the spirometer. *Conclusion*: Our proposed depth-based approach is able to remotely compute 11 clinical PFT measures, which gives highly accurate results when evaluated against a spirometer on a dataset comprising 85 patients. *Significance:* Experimental results computed over an unprecedented number of clinical patients confirm that chest surface motion is linearly related to the changes in volume of lungs, which establishes the potential towards for an accurate, low-cost and remote alternative to traditional cumbersome methods, like spirometry.**

***Index Terms*—Chest surface reconstruction, chest volume estimation, forced vital capacity (FVC), Kinect noise analysis, pulmonary function testing (PFT), spirometry, slow vital capacity (SVC).**

I. Introduction

PULMONARY Function Testing (PFT) is a vital component of clinical assessment in the investigation of respiratory diseases. This can be achieved by a variety of measures, including exercise testing, lung volume measurement, and dynamic breathing tests. Traditional measures of pulmonary function, such as spirometry [1] and whole body plethsmography [2] (which measures lung volumes and gas transfer) require patient co-operation and direct contact with the equipment. There are other measures of lung physiology which are even more invasive, such as arterial blood gas sampling (direct arterial sampling) and cardiopulmonary exercise testing (treadmill or exercise bike) [1]. Comparatively among these methods, spirometry is the most prevalent to assess lung function due to its portability, price, and accuracy for medical diagnosis.

V. Soleimani, M. Mirmehdi, D. Damen, S. Hannuna and M. Camplani are with the Department of Computer Science, University of Bristol, Bristol, BS8 1UB, UK. e-mail: vahid.soleimani@bristol.ac.uk, m.mirmehdi@bristol.ac.uk. J. Dodd, C. Sharp and J. Viner are with Academic Respiratory Unit, Southmead Hospital, Bristol, BS10 5NB, UK. email: james.dodd@bristol.ac.uk.

Manuscript received xx, 201x; revised xx, 201x.

To perform a spirometry test, patients are asked to breathe through a mouthpiece while a nose-clip is applied to prevent air leakage. The two primary clinical protocols undertaken with a spirometer are forced vital capacity (FVC) and slow vital capacity (SVC). The former comprises a maximal inspiration followed by a forced maximal expiration, and the latter a maximal inspiration followed by a slow, controlled, maximal expiration. Various clinical PFT measures, such as *FVC*, *FEV1*, *PEF*, $FEF_{25-75\%}$ (FVC measures) and *VC*, *IC*, *TV* and *ERV* (SVC measures) are calculated within a spirometry test [1], [3]. These PFT measures, and their combinations, are used in the diagnosis and assessment of *obstructive* lung diseases, e.g. Chronic Obstructive Pulmonary Disease (COPD) and Asthma, and *restrictive* lung diseases, e.g. lung fibrosis.

Although spirometry is an accurate and reliable clinical method, there are some disadvantages which limit its application. The spirometer is a particularly challenging device for certain clinical populations to perform with, such as the frail elderly, children, and cognitively impaired patients. It needs to be recalibrated at least every couple of days and a new mouthpiece and nasal clip are needed for each patient.

In this paper, we propose a novel depth-based method for remote lung function assessment by estimating and tracking the volume of the chest to compute clinically acquired FVC and SVC measures. For depth sensing, we use the Microsoft Kinect V2 RGB-D sensor [4] which is based on time-of-flight technology. The estimated measures are correlated against results obtained using a spirometer for 85 patients who attended a respiratory outpatient service for spirometry. In our previous work [5], we demonstrated that the Microsoft Kinect can be used to estimate chest volume and compute intra-test PFT measures. To the best of our knowledge, the only other work that remotely computes and reports PFT measures (just two, FVC and FEV1) is [6] which used the first generation, structured-light based, Kinect. Their study mainly focused on estimating passive airway resistance and was tested on 5 healthy subjects who were instructed to mark their inhalation and exhalation manually (using the computer mouse) during the test.

We extend our previous work in [5] by: (a) detailed analysis of volume-time data to automatically extract more reliable keypoints for calculating scaling factors and measures, (b) obtaining three more FVC measures i.e. $FEF_{25\%}$, $FEF_{50\%}$ and $FEF_{75\%}$, (c) performing comparative analysis of PFT measures obtained by the proposed method and spirometer, (d) investigating subjects' upper body motion during the test and its effects on volume-time data, (e) generalizing the intra-subject scaling factor, and (f) evaluating the proposed method on 85 actual patients (compared to 40 in [5]).

Our proposed system has been developed in response to

increasing clinical interest in contactless or remote techniques for respiratory assessment. It can be exploited for a wide range of potential applications, such as screening for respiratory diseases, home monitoring, and gating controls for radiological imaging techniques. The proposed system is easy to setup and does not require calibration on a daily basis. Due to remotely assessing the lungs, not only does it cut the costs (pneumatach and disposable accessories), but also it decreases infection risks caused by connecting to a pneumatach. Furthermore, our method requires no specialist training.

There are several recent studies that only estimate breathing rate, without performing PFT, using structured-light [6]–[18], time of flight cameras [5], [19], [20], video cameras [21], [22], and other remote sensors [23], [24]. These are briefly considered in Section II. In Section III, we present an overview and schematic of the proposed approach. Then in Section IV, we describe the Kinect noise analysis and filtering, 3D chest modelling, and volume estimation. This is followed in Section V with volume-time data keypoints computation and analysis. Extracting clinical PFT measures is presented in Section VI and our proposed method for scaling factor generalization is described in Section VII. The system configuration, the dataset, and the experimental results are presented in Section VIII. The paper is concluded in Section IX. A list of abbreviations used in the paper is provided in the Appendix.

II. Literature Review

Remote respiratory monitoring has recently become a potential solution and is attracting more researchers, especially since the availability of affordable depth sensors, such as the first generation Microsoft Kinect, and then later the Microsoft Kinect V2, which use structured-light and time-of-flight techniques respectively. While many works, referred to below, have investigated breathing rate, respiratory waveform estimation, and respiration resistance using depth sensors, we know of only our own earlier work [5] and Ostadabbas et al. [6] that have applied the Kinect to PFT measurement in particular. Further, we have many more subjects and a much wider range of PFT measures.

Structure-light approaches – Ostadabbas et al. [6] applied the first generation Kinect to compute two PFT measures (*FVC* and *FEV1*) for the estimation of airway resistance, defined as lung pressure divided by the airflow. Five healthy subjects were asked to blow through various numbers of straws (to induce varied airway resistance) while their lung volume was measured over time. They instructed subjects to press their back against the wall to restrict their body movement and use a wireless mouse to timestamp their inhalation and exhalation during the test. They reported on average 0.88 correlation between their method and spirometry for the *FEV1* measure.

Aoki et al. [7] proposed a non-contact respiration measurement technique, using the first generation Kinect, by extracting the volume of the thoracoabdominal region formulated on the skeleton joint positions available from the sensor. Respiration waveforms were generated by computing the changes to this volume. Their results were validated against an expiration gas analyser and flow meter and they reported 0.98 correlation between volume change (estimated by their method) and the air flow volume (measured by an expiratory gas analyser). Yu et al. [8] developed an elaborate calibration technique, along with a predefined chest wall mask, to approximately extract the subject's chest wall region and dimensions. The respiratory volume was estimated by using the computed length per pixel and depth information. Correlation of 0.96 was reported against a spirometer for estimating respiratory volume. Similar to [8], Seppanen et al. [9] used the first generation Kinect to estimate the respiration rate (of healthy subjects) by generating respiratory airflow waveform using several models from depth sensor data. The best coefficient of determination (R^2) between the spirometer signal and the estimated airflow signal was reported as 0.93. Benetazzo et al. [12] detected respiratory rates by applying a weighted averaging filter to the chest region pixels segmented by using the first generation Kinect skeleton's shoulder and torso joint positions. Their breathing rate results were evaluated against a spirometer, with an outcome of 0.98 correlation. Tahavori et al. [13] used a first generation Kinect placed above the participant's body who was required to be supine and obtained the average depth value of 16 regions of interest on the chest and abdomen over time to analyse their motion. After applying principal component analysis (PCA) to the average depth values of these regions, they demonstrated that the first principal component describes nearly 70% of the motion data variance in chest and abdomen surfaces. Other works of note that extract the respiratory rate using the structured-light based Kinect are [14]–[18].

In an example non-Kinect, yet structured-light approach, De Boer et al. [10] deployed two cameras as a stereo pair to capture a predefined light pattern projected onto the chest wall and estimate chest volume changes. The volume was defined as the enclosed space between the chest surface and the work bench, for which $R^2 = 0.91$ was reported when compared against a spirometer. The authors reported that their PFT measures correlated with the spirometer at $R^2 = 0.97$, but provided no further details.

Time-of-flight approaches – In [19], Ostadabbas et al. proposed a non-invasive, passive method, using the Microsoft Kinect V2 and a pulse oximeter, to assess the severity of airway obstruction as mild, moderate or severe. To estimate respiration airflow, 14 healthy subjects were asked to breathe through various straws to induce airway obstruction externally in a spontaneous breathing session while lying supine to minimize body movement effects. In a separate stage, they estimated breathing rate and tidal volume of 14 patients in a sitting position to classify their airway obstruction severity. In both parts, they asked each subject to perform some instructions, e.g. pressing the pulse oximeter buttons during the test. They reported 76.2% and 80% accuracy in detecting airways obstruction in healthy and ill subjects respectively.

Penne et al. [20] employed a time-of-flight camera and used the flat clinic bed in a calibration stage and fit a reference plane onto it. Then, with a test subject present, the two best-fitting planes were found for the chest and abdomen regions, and are used to compute the breathing signal. They compared their respiration signal against that obtained from an ANZAI belt

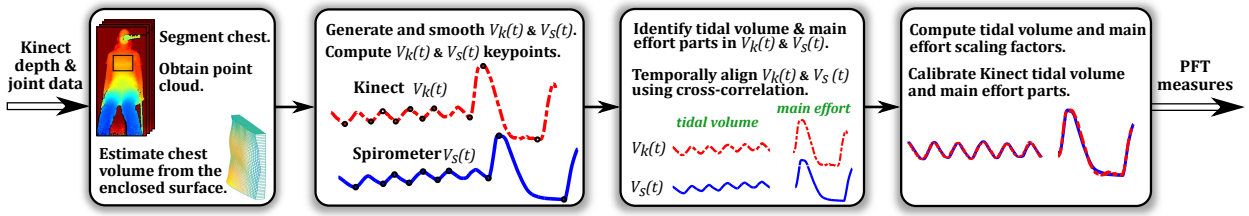


Fig. 1. A schematic of the proposed method.

for chest and abdomen regions and reported 0.85 and 0.91 correlation respectively.

In our preliminary work [5], we obtained several PFT measures of FVC and SVC tests remotely using Kinect V2 depth data by computing volume-time and flow-time curves of chest volume changes. We evaluated on 40 patients by comparing their computed measures to those obtained from a spirometer. These results are reproduced in the Section VIII of this paper.

RGB video camera approaches – Tan et al. [21] proposed a single video camera approach which used image subtraction to detect the motion of the chest and abdomen regions in subjects wearing a striped pattern shirt. After applying an averaging filter, the breathing signal was obtained from the number of moving pixels given a threshold. They evaluated their results against a stain gauge, a thermistor, and a flow monitoring system, but reported only subjective assessments. Frigola et al. [22] used optical flow to detect body movement to monitor inhalation and exhalation during sleep. Although they used an elastic cloth band as their groundtruth, comparative evaluation results were not reported.

Other sensor approaches – Other example methods of note to monitor respiratory rate are Scalise et al. [23] who used a Laser Doppler Vibrometer and Sato and Nakajima [24] who employed a stereo system with an infrared beam fibre grating projection. Further, there have been a number of marker-based (motion capture system) clinical works [25]–[28]. These approaches are expensive and require a complicated calibration process. They mainly focus on the existence of correlation between chest wall motion and actual lung volume changes.

III. Overview

Fig. 1 presents an overview of the proposed method. After identifying and segmenting the chest region in each depth frame of the sequence captured by the Kinect, the volume of the thoracic wall is estimated and the Kinect volume-time and flow-time curves are generated. Next, the Kinect volume-time curve is smoothed using a moving averaging filter and then keypoints are automatically computed for both the depth and spirometry measurements. After establishing linear scaling factors, needed to calibrate the curves from the depth sensor, PFT measures are computed on the depth sensor curve and their stability over multiple runs for the same subject is analysed and compared against the spirometer measures. We show that these scaling factors are subject-specific as they relate to the natural body motion of the subject while performing PFTs. Accordingly, by investigating subjects' trunk motion patterns, we generalize intra-subject scaling factors to compute intra-subject PFT measures.

IV. Chest Modelling and Volume-time Data Computation

Kinect V2 Sensor Noise – Kinect depth estimation suffers from measurement noise caused by the depth sensor technology. Since the Kinect V2 was released only recently, there is little public information on the nature and characteristics of its noise. We performed a planar noise analysis to find the optimal distance range between the sensor and the subject.

In this experiment, we estimated the sensor measurement error by placing the Kinect at various distances - from $60cm$ to $500cm$ at $20cm$ intervals - in front of a white wall under normal room temperature and lighting conditions, with the sensors optical axis approximately perpendicular to the wall. At each position, a sequence of 200 frames were recorded and $15K$ depth values were randomly sampled from a constant-size patch at the center of the sensor's viewpoint and the standard deviation was computed for them. Fig. 2 illustrates this standard deviation in mm plotted against the sensor distance to the wall. It shows a non-linear behaviour similar to the general ToF depth sensors [29]. Furthermore, a similar noise curve was reported by Breuer et al. [30]. Noise increases between 60 and $80cm$, and then drops to its minimum at $\sim 150cm$. Accordingly, we carried out all our experiments with the Kinect placed at $\sim 150cm$ from the subject. However, noise may vary under different environmental lighting and temperature conditions and also depends on the sensor temperature itself. These factors therefore require the optimal distance to be re-computed for the environment the device is to be used in.

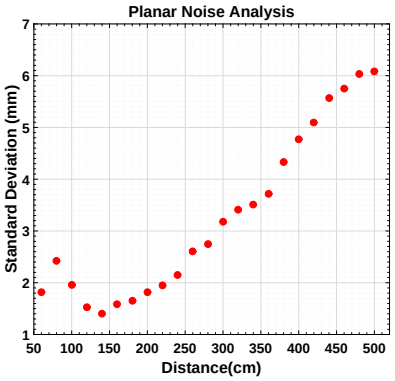


Fig. 2. Planar surface noise analysis within a distance range of $60 - 500cm$.

To filter noise in the measurements, an edge-preserving bilateral filter [31] was applied to each frame of our data:

$$BF[I]_p = \frac{1}{W_p} \sum_{q \in S} G_{\sigma_s}(\|p-q\|) G_{\sigma_r}(|I_p - I_q|) I_q, \quad (1)$$

where W_p is the normalization factor, G_{σ_s} is a spatial Gaussian kernel, G_{σ_r} is a range Gaussian kernel, p and q are the locations of central and neighbour pixels, $\|p-q\|$ is Euclidean distance between pixel locations p and q, and I is the image to be filtered. The range parameter σ_r of the bilateral filter was determined to be 1.5, which is approximately equal to the standard deviation of distance measuremnts obtained by the Kinect at the chosen distance of $\sim 150cm$. In particular, this value was selected, as in [32] according to the level of noise at this distance, to optimize the performance of the range component of the bilateral filter. For the spatial filter, we select $W_f = 13$ which guarantees a good trade-off between accuracy and processing speed, also reported by Camplani et al. in a similar filtering approach. Consequently, $\sigma_s = W_f/6$, such that the the significant part of the Gaussian kernel (up to $3\sigma_s$) is completely included within the selected window W_f [33].

Smoothed Volume-time Curve – The volume-time curve was obtained for each sequence by estimating the chest volume as a function of time. Smoothing of the volume-time curve, in one form or another, is routinely applied in all other works, for example in [19], [34]–[36]. Here, although the bilateral filter was applied to each frame of the depth sequence, the volume-time curve still remained considerably noisy (see Fig. 3) as the chest volume is estimated temporally in a very limited chest wall motion, i.e. $\pm 2.5cm$ approximately. Thus, we used a non-causal moving average filter, which is a low pass finite impulse response (FIR) filter [37], to eliminate high frequency noise of the Kinect volume-time curve, i.e.

$$V_{out}(t) = \frac{1}{N} \sum_{i=-(N-1)/2}^{(N-1)/2} V_{in}(t-i), \quad (2)$$

where $V_{in}(k)$ and $V_{out}(k)$ are the input and filtered volume-time curves respectively, and N is the averaging window size, which is computed as $N = 15$ based on the filter cut-off frequency of 1Hz [38]. The cut-off frequency was chosen according to the range of respiratory rates (frequency) for healthy adults at $12-20$ breaths/minute ($0.2-0.34$ Hz) [39], elderly at $16-25$ breaths/minute ($0.27-0.42$ Hz) [40], and those with severely pulmonary disorders at 36 breaths/minute (0.6 Hz) at most [41]. The computed range of respiratory rates for the 85 patients of our dataset, at $8-32$ breaths/minute ($0.13-0.53$ Hz), satisfies the chosen cut-off frequency of 1Hz.

3D Modelling of Thoracic Wall – After obtaining a point cloud representing the captured scene from the filtered depth images, a subject's chest area was segmented automatically using body joints estimated by Kinect software (SDK2.0), defined by *ShoulderRight*, *ShoulderLeft*, *SpineShoulder* and *SpineMid* joint positions. The chest wall surface was then reconstructed by applying a 2D Delaunay triangulation [42] on the point cloud (Fig. 4a).

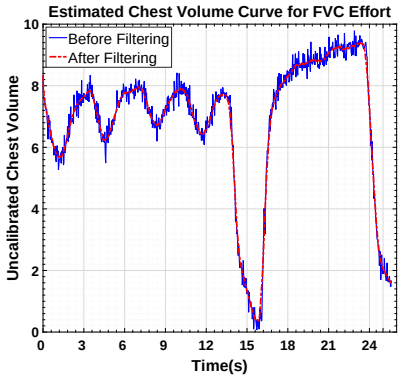
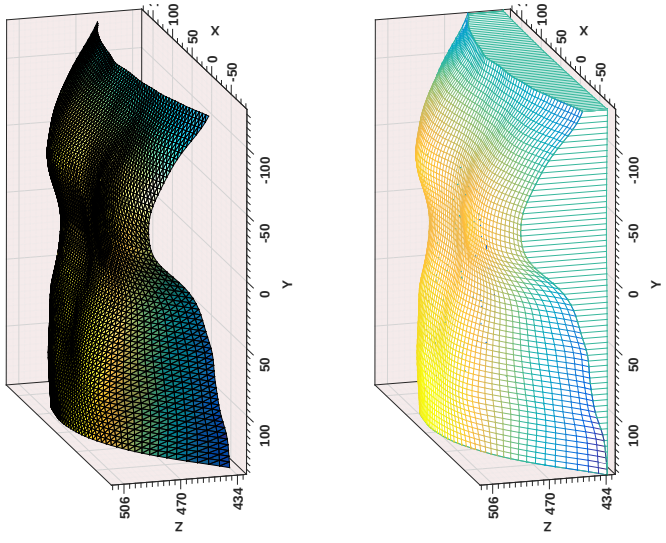


Fig. 3. Volume-time curve before and after applying moving averaging filter.

3D-chest-model based volume estimation – Given the 2.5D data, we proposed in [5] a method to approximate the chest volume by computing the volume between the model of the thoracic wall and a reference plane at a predefined distance from the camera. Our approach is sufficient to compute the volume-time curve $V(t)$ that models variations in the approximated volume, based on the assumption that body movements are minimal during PFT and can be ignored[1]. The reconstructed chest wall surface was then enclosed by surrounding lateral surfaces and a reference plane (Fig. 4b), and its volume was estimated using the Divergence Theorem. More information about our volume estimation can be found in [5].

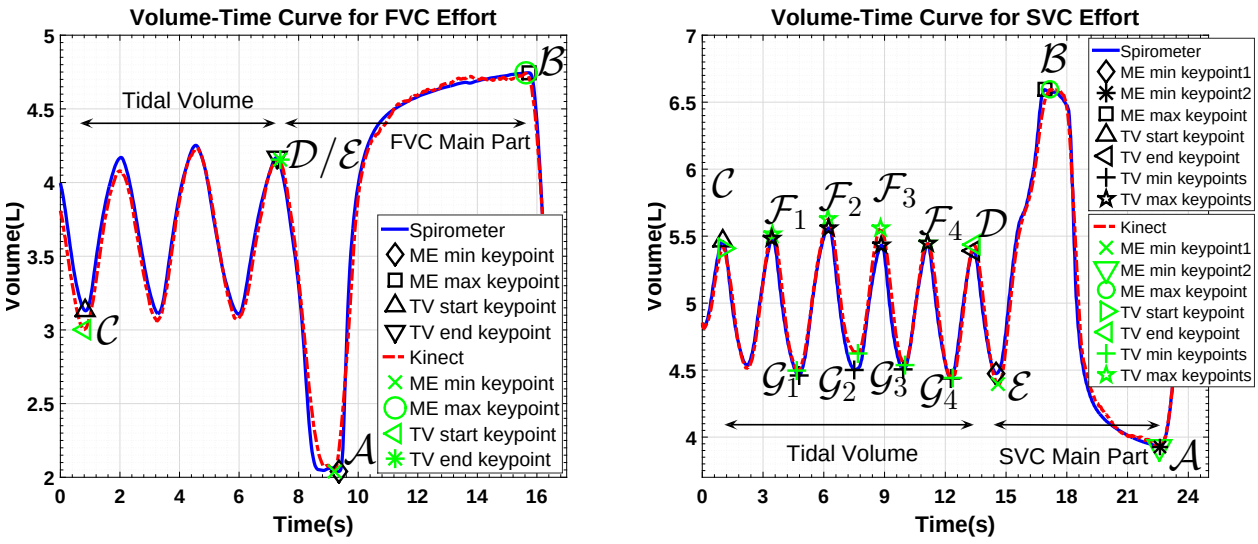
Chest-averaging based volume estimation – Similar to previous approaches [6], [12], [15], [17], [18], we also estimated the uncalibrated chest volume at time point t by computing the average distance of each pixel located in the chest region. Chest-averaging is simple and fast to compute.

[1]This assumption is revisited in Sections V-B, V-C and VII



(a) (b)

Fig. 4. (a) Reconstructed chest surface, (b) Chest surface confined by reference plane and lateral sides. Images are reproduced from [5].



(a) (b)

Fig. 5. (a) Kinect and spirometer FVC volume-time curve and their corresponding keypoints, (b) Kinect and spirometer SVC volume-time curve and their corresponding keypoints.

We report results using both the 3D-chest-model based [5] and chest-averaging methods in Sections VIII-B and VIII-D.

V. Volume-time Data Keypoints and Analysis

All PFT tests start with a few cycles of normal breathing, called *tidal volume*, followed by the intended lung function test, called *main effort*. Since our Kinect volume-time data measures the chest volume in cubic metres (m^3) relative to an arbitrary plane, as opposed to the spirometer's air volume measure in litres, we need to linearly scale the y-axis in the volume-time curves (using computed scaling factors) to enable the correlation of computed measures. Note that this is not to imply that the Kinect truly measures *lung volume*: chest volume is a proxy for the amount of air within the lungs that we show is linearly related to air flow as measured by spirometry.

A. Keypoints Computation

Several keypoints were automatically computed from the volume-time curves to (a) identify *tidal volume* and the *main effort*, (b) establish scaling factors, and (c) compute PFT measures. Five keypoints are required for separating *tidal volume* and *main effort* in the FVC and SVC volume-time curve, $V(t)$, which are named as $\{\mathcal{C}, \mathcal{D}\}$ (beginning and end of *tidal volume*) and $\{\mathcal{E}, \mathcal{A}, \mathcal{B}\}$ (beginning to the end of *main effort*), as illustrated in Fig. 5.

In order to compute keypoints correctly, first we need to find the FVC and SVC volume-time curve extrema which identify respiratory cycles during the PFT test. Since the curve can be noisy (e.g. because of chest movement and coughing), local minima or maxima may be incorrectly selected. To avoid false local extrema, the difference between two consecutive turning points, which are introduced as local extrema, needs to be greater than a threshold γ. Considering V_{min} and V_{max} as the smallest and greatest estimated chest volume in a sequence (volume-time curve global minimum and maximum), $[V_{max} - V_{min}]$ indicates the maximum volume of exchanged air that occurs during *main effort*. A fraction of this exchanged volume is defined as γ to identify local extrema, i.e. $\gamma = \frac{1}{\rho}[V_{max} - V_{min}]$, where ρ is defined as the ratio of the greatest exhaled air during *main effort* (6.8) to the smallest exhaled air during *tidal volume* (0.35) among all sequences, which is $\rho = \sim 20$.

Note that SVC volume-time curve presents inhalation and exhalation in the opposite direction to the FVC volume-time curve. This means, while an increase in FVC volume-time curve corresponds to exhalation, it indicates inhalation in the SVC volume-time curve. This is similar to the volume-time curves obtained from the spirometer.

FVC keypoints – In FVC, keypoints $\mathcal{D}$ and $\mathcal{E}$ are coincident in $V(t)$. Since lungs always contain a residual air volume, the amount of exhaled air volume in deep expiration is greater than inhaled air in a deep inspiration. Hence, keypoints $\mathcal{A}$ and $\mathcal{B}$, indicating the beginning and end of deep expiration respectively, are more detectable than other points. They were extracted, timestamped $t_\mathcal{A}$ and $t_\mathcal{B}$ respectively, as a pair of consecutive minimum and maximum points with the largest change in volume between them during expiration, such that,

$$[t_\mathcal{A}, t_\mathcal{B}] = \underset{t_i^x, t_i^y}{\arg\min} \left\{ x(t_i^x) - y(t_i^y) \right\},\ 1 = i \ldots n$$
$$\forall t_i^x, t_i^y \ni t_i^y > t_1^x, \tag{3}$$

where X and Y are sets of volume-time curve extrema computed as minima and maxima, $x(.) \in X$ and $y(.) \in Y$, t_i^x and t_i^y are each minimum and maximum corresponding timestamps, and n is computed as

$$n = \min(|X|, |\bigcup_{i=1}^{|Y|} y(t_i^y)|),\ t_i^y > t_1^x. \tag{4}$$

The local maximum directly before $t_{\mathcal{A}}$ was selected as $\mathcal{E}$ (and thus $\mathcal{D}$). The first extremum of the curve was selected as $\mathcal{C}$.

In addition to the volume-time curve, we also used the flow-time curve to compute some FVC measures. The flow is defined as the rate of changing volume, i.e. $\dot{V}(t) = \frac{\partial V}{\partial t}$.

FVC peak flow and time zero – To compute some FVC test measures, such as *FEV1*, we also needed to compute the Peak Flow (PF) point and 'time zero' t_0 (Fig. 6). PF is the point at t_{PF} with the maximum air flow speed during *main effort* exhalation,

$$t_{PF} = \underset{t\in[t_{\mathcal{A}},t_{\mathcal{B}}]}{\arg\max}\{\frac{\partial}{\partial t}\big(V(t)\big)\}. \tag{5}$$

Since *FEV1* is a timed PFT measure, instead of keypoint $\mathcal{A}$ (timestamped $t_{\mathcal{A}}$), a starting 'time zero' t_0 keypoint is used for computing *FEV1* (Fig. 6). This is because keypoint $\mathcal{A}$ is affected by hesitant or delayed exhalation in the *main effort* manoeuvre leading to an incorrect and decreased *FEV1* value. After subtracting $V(t_{\mathcal{A}})$ from the estimated volume, t_0 is computed using the back-extrapolation approach [1],

$$t_0 = t_{PF} - \big[V(t_{PF}) - V(t_{\mathcal{A}})\big] \times \left[\frac{\partial}{\partial t}\big(V(t)\big)\Big|_{t=t_{PF}}\right]^{-1} \tag{6}$$

SVC keypoints – In the SVC test, we extracted $\{\mathcal{C},\mathcal{D}\}$ and $\{\mathcal{E},\mathcal{B},\mathcal{A}\}$ keypoints for partitioning the volume-time curve into the *tidal volume* and *main effort* respectively as shown in Fig. 5b. Similar to the FVC keypoints extraction method, to be able to find other keypoints, we first computed $\{\mathcal{B},\mathcal{A}\}$ timestamps as,

$$\begin{aligned}[t_{\mathcal{B}}, t_{\mathcal{A}}] =& \underset{t_i^y, t_i^x}{\arg\max}\big\{y(t_i^y) - x(t_i^x)\big\},\ i = 1\dots m \\ & \forall t_i^y, t_i^x \ni\ t_i^x > t_1^y,\end{aligned} \tag{7}$$

where notations are similar to (3) and m was computed as

$$m = \min(|\bigcup_{i=1}^{|X|} x(t_i^x)|, |Y|),\ t_i^x > t_1^y\ . \tag{8}$$

Here, in the volume-time curve $V(t)$, inhalation in SVC shows as exhalation in FVC. Thus, we still used the exhalation part of the *main effort*, which is more reliable, to extract $\mathcal{B}$ and $\mathcal{A}$, similar to the FVC test. Keypoint $\mathcal{E}$ marks the beginning of inhalation in *main effort* and is determined as the local minimum directly before $t_{\mathcal{B}}$. Like FVC, $\mathcal{C}$ is chosen as the first extremum of the curve, and $\mathcal{D}$ is the local maximum directly before $t_{\mathcal{B}}$. For computing SVC measures, four maxima and four minima keypoints ($\mathcal{F}_i$ and $\mathcal{G}_i$ in Fig. 5b) from the *tidal volume* part are also required.

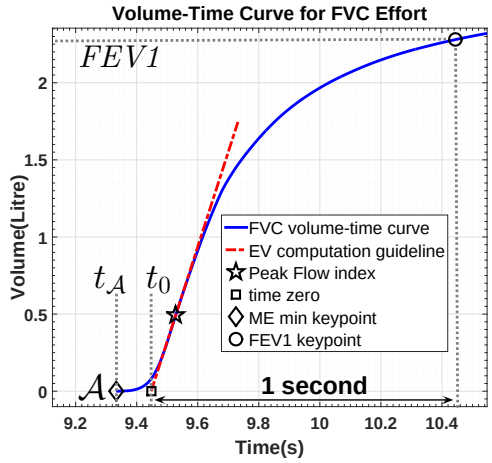


Fig. 6. 'time zero' and peak flow in FVC volume-time curve.

B. Tidal Volume Analysis and Calibration

To be able to extract PFT measures from the Kinect volume-time curve and compare them with those given by the spirometer, and thus evaluate our proposed method, we needed to (a) temporally align Kinect and spirometer volume-time curves and (b) compute scaling factors and (c) use them to calibrate the Kinect volume-time curve. We perform alignment and scaling separately for the *tidal volume* and *main effort* parts, to take into consideration any inevitable trunk movement when subjects take a deep inhalation, followed by a maximal exhalation.

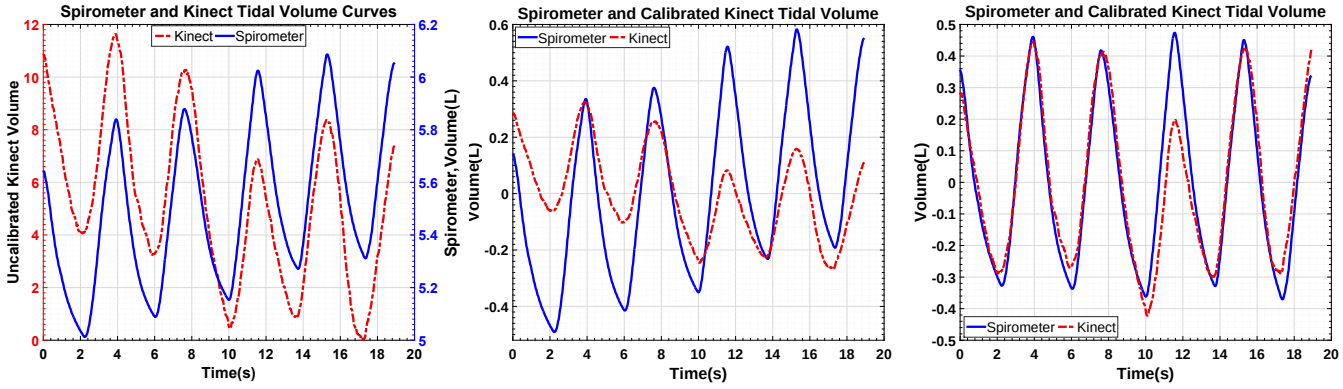
After selecting the *tidal volume* parts of the Kinect and spirometer volume-time curves using the $\mathcal{C}$ & $\mathcal{D}$ keypoints, we performed some pre-processing operations on these two sub-signals to allow them to be directly compared. The spirometer sub-signal was sampled at the Kinect sampling rate of 30Hz. Both signals are normalized to zero mean. Finally, the two sub-signals were synchronized by computing the optimal time delay using windowed cross-correlation,

$$\tau_{delay} = \underset{\tau}{\arg\max}\left(\sum_{-\infty}^{+\infty} V_k^*(t)V_s(t+\tau)\right), \tag{9}$$

where $V_k^*(t)$ and $V_s(t)$ denote the complex conjugate of Kinect normalized *tidal volume* and spirometer subsampled and normalized *tidal volume* curves respectively.

The *tidal volume* scaling factor can be computed using only a pair of consecutive minimum and maximum points [5], however this is not very reliable. We modelled it with a first degree polynomial, $\widehat{V}_s = \xi_{tv} \cdot \widehat{V}_k + \psi_{tv}$, where $\widehat{V}_s$ and $\widehat{V}_k$ are subsampled and aligned Kinect and spirometer *tidal volume* data, ψ_{tv} is the offset between the Kinect and spirometer *tidal volume* parts, and ξ_{tv} presents the *tidal volume* scaling factor. Since the Kinect and spirometer *tidal volume* parts were mean zero normalized, then $\psi_{tv} \approx 0$.

However, in many cases, this approach is insufficient to deal with an incremental or decremental trend in the data that can appear in one or both of the Kinect and the spirometer data. Fig. 7a shows example Kinect and spirometer *tidal volume* curves each plotted on a different scale, with the left y-axis for the uncalibrated Kinect volume and the right y-axis for the spirometer volume (L). Both curves exhibit such a trend which makes the extraction of a correct scaling factor (or an alignment process) a cantankerous task (see Fig. 7b). This trend might occur because of one or more reasons: the use of a nasal Oxygen mask by patients during the test (which affects only the spirometer data), lung hyperinflation, or the subject's body movements.



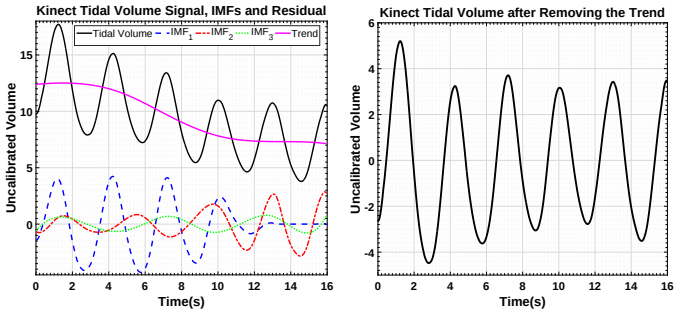
(a) (b) (c)

Fig. 7. (a) Existing trends in spirometer and Kinect *tidal volume* curves, (b) Incorrect Kinect Calibrated *tidal volume* because of existing trend, (c) Correct Kinect calibrated *tidal volume* after removing the trend.

A simple approach to modelling the trend to help eliminate it would be a linear regression model. However, we found this to be insufficient due to the non-linear nature of the trend, thus we applied Empirical Mode Decomposition (EMD) [43] to estimate the trend more accurately. EMD is an adaptive method to decompose a non-linear and non-stationary signal in the time domain into its individual components (Intrinsic Mode Functions or IMFs) and a residual r, from which no more IMFs can be extracted and can be said to represent the signal's trend (10):

$$s(t) = \sum_{j=1}^{l} IMF_j(t) + r(t). \tag{10}$$

Figs. 8a and 8b present the first three IMFs and the residual of a *tidal volume* curve (where the residual displays the signal trend), and the modified *tidal volume* curve after applying EMD. Fig. 7c shows the Kinect and spirometer *tidal volume* curves with their trend estimated and removed by EMD, and the Kinect curve has been calibrated using the correct *tidal volume* scaling factor. Note we used the modified *tidal volume* curves to compute scaling factors only and other analysis were performed on the original Kinect and spirometer data.



(a) (b)

Fig. 8. (a) Original Kinect *tidal volume* curve, *IMF*s and the residual signal, (b) *Tidal volume* curve after removing the trend.

C. Main Effort Analysis and Calibration

As stated in Section V-B, the Kinect and spirometer volume-time curves were aligned only using their *tidal volume* sections to avoid errors arising from the subjects upper body movement during *main effort*. Then, the *main effort* scaling factor (ξ_{me}) was obtained by solving $\widehat{V}_s = \xi_{me} \cdot \widehat{V}_k + \psi_{me}$, using only the $\mathcal{A}$ & $\mathcal{B}$ keypoints on each signal as they are less affected by motion artefacts and thus more reliable. Unlike in the *tidal volume* calibration process where ψ_{tv} was zero, ψ_{me} here correlates with body movement and appears as an offset along the y-axis. However, in scenarios where subjects are stationary during the whole test (e.g. see Fig. 5b), then $\psi_{me} \approx 0$, and there is no offset between the *tidal volume* and *main effort* parts.

We calibrated the *tidal volume* and *main effort* parts individually and generated two calibrated Kinect volume-time curves. For the first (*tidal volume* calibrated), the whole Kinect volume-time curve is scaled by multiplying by the *tidal volume* scaling factor ξ_{tv}, as computed in Section V-B. Then, it was vertically aligned with the spirometer tidal volume part by making both the Kinect and spirometer tidal volume part zero-mean, as shown in Fig. 9a. For the second (*main effort* calibrated), the whole Kinect volume-time curve was scaled by multiplying the *main effort* scaling factor ξ_{me}, computed in this section, and vertically aligned with the spirometer *tidal volume* part by adding the *main effort* offset ψ_{me} to all Kinect volume-time data, as shown in Fig. 9b.

VI. Computation of Clinical PFT Measures

FVC measures – Within a FVC spirometry test, several clinical measures are provided by the spirometer software. In addition to these numerical measures, there are two common 'qualitative' presentations of lung function test, i.e. volume-time curve and flow-volume loop (Figs. 10a and 10b), that pulmonologists often use these graphs to visually diagnose problems in the patient's breathing function.

The 7 most significant FVC measures that we compute using the proposed Kinect FVC volume-time and flow-time data are: (i) *FVC* as the maximum amount of air in litres blown out after a maximal inhalation, determined as the volume change

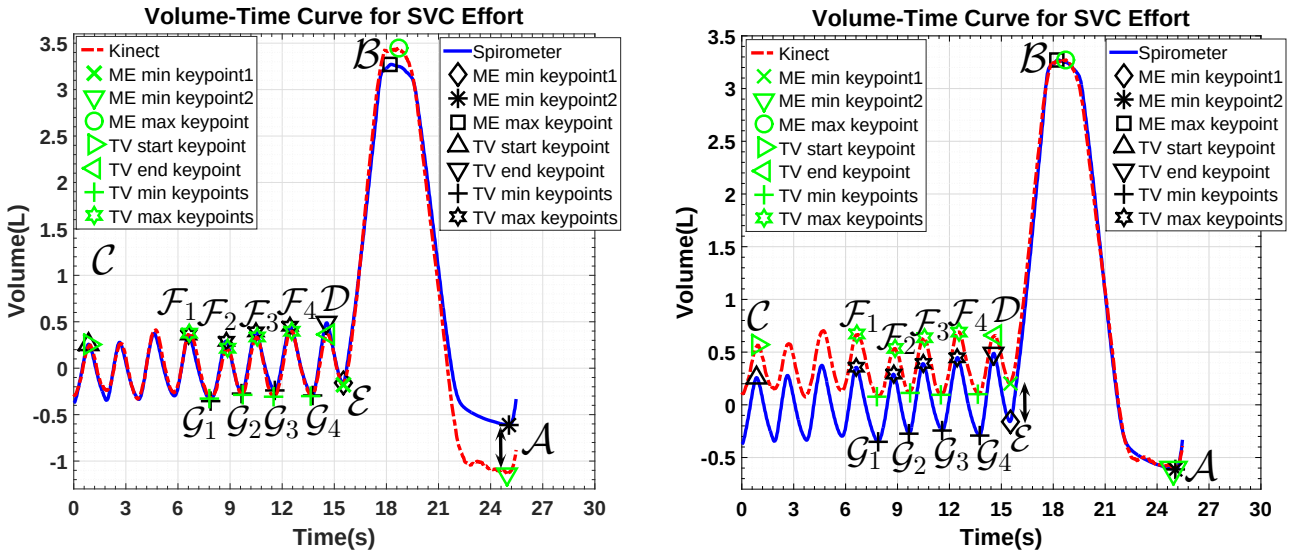


Fig. 9. (a) Volume-time curve calibrated using spirometer *tidal volume* part, (b) Volume-time curve calibrated using spirometer *main effort* part.

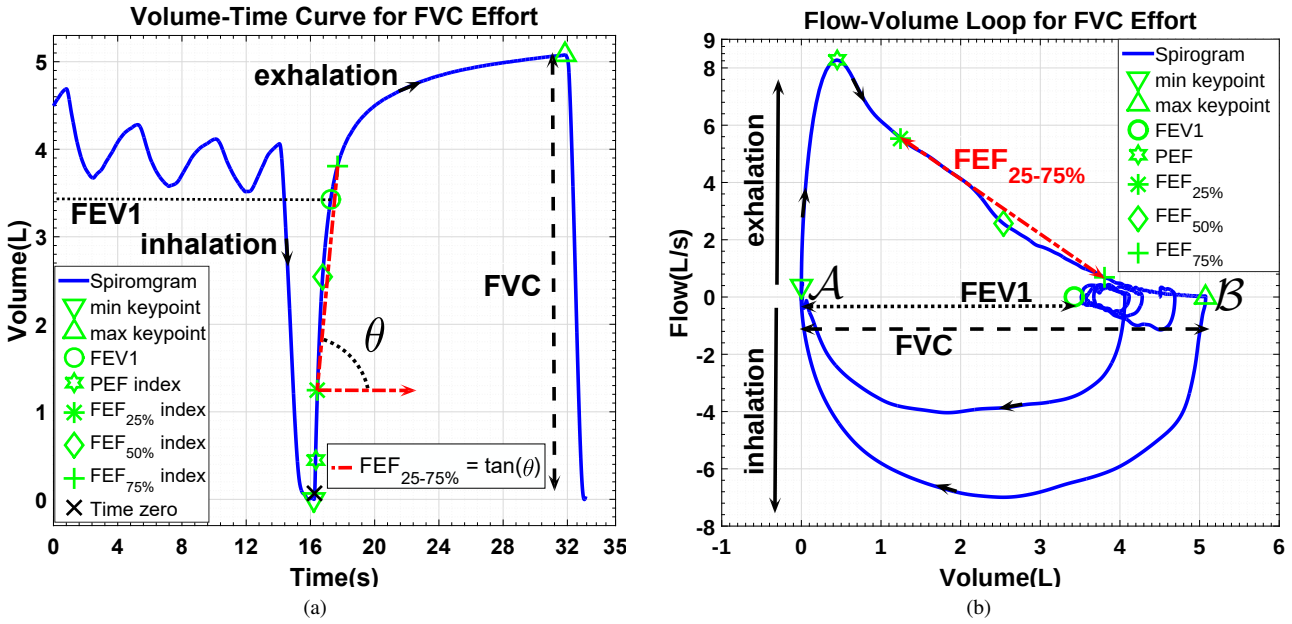


Fig. 10. (a)-(b) FVC measures on volume-time curve and flow-volume loop extracted from our dataset.

between keypoints $\mathcal{A}$ & $\mathcal{B}$, i.e. $FVC = \big[V(t_{\mathcal{B}}) - V(t_{\mathcal{A}})\big]$, (ii) *FEV1* (Forced Expiratory Volume) as the volume of air forcibly expired in 1 second starting from 'time zero' (6), i.e. $FEV1 = \big[V(t_0 + 1) - V(t_0)\big]$, (iii) *PEF* (Peak Expiratory Flow) as the maximum speed of exhaled air, i.e. $PEF = \dot{V}(t_{PF})$, (iv) $FEF_{25\%}$ (Forced Expiratory Flow as flow of exhaled air at 25% of *FVC*, i.e. $FEF_{25\%} = \dot{V}(t_{0.25FVC})$, (v) $FEF_{50\%}$ as flow of exhaled air at 50% of *FVC*, i.e. $FEF_{50\%} = \dot{V}(t_{0.5FVC})$, (vi) $FEF_{75\%}$ as flow of exhaled air at 75% of *FVC*, i.e. $FEF_{75\%} = \dot{V}(t_{0.75FVC})$, and (vii) $FEF_{25-75\%}$ as the mean forced expiratory flow between 25% and 75% of the *FVC*, computed as $FEF_{25-75\%} = \frac{0.75FVC - 0.25FVC}{t(FEF_{25\%}) - t(FEF_{75\%})}$.

FVC, *FEV1* and $FEF_{25-75\%}$ are illustrated in Figs. 10a and 10b and *PEF*, $FEF_{25\%}$, $FEF_{50\%}$ and $FEF_{75\%}$ measures are marked on flow-volume loop in Fig. 10b. Note that since the last four measures are computed using volume-time and flow-time data, only their corresponding locations are marked as 'index' on volume-time curve in Fig. 10a (using their timestamps).

SVC measures – Within an SVC test, four clinical measures are provided by the spirometer software, and only one 'qualitative' presentation of lung function, i.e. the volume-time curve (Fig. 11), which we compute on Kinect volume-time data: (i) *VC* (Vital Capacity) as the volume change between full inspiration and complete expiration between keypoints $\mathcal{B}$ & $\mathcal{A}$, i.e. $VC = \big[V(t_{\mathcal{B}}) - V(t_{\mathcal{A}})\big]$, (ii) *IC* (Inspiratory Capacity)

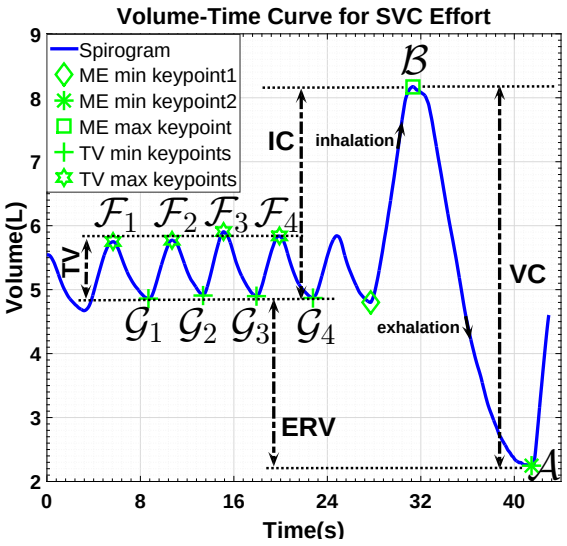


Fig. 11. SVC measures on volume-time curve extracted from our dataset.

as the volume change between taking a slow, full inspiration and the passive end-tidal expiration, i.e. difference of volume at keypoint $\mathcal{B}$ and the average volume at group keypoints $\mathcal{G}$ within the *tidal volume* section,

$$IC = V(t_{\mathcal{B}}) - \frac{1}{4}\sum_{i=1}^{4} V(t_{\mathcal{G}i}), \tag{11}$$

(iii) *TV* (Tidal Volume) as the volume of air inspired and expired at rest condition, i.e. the average volume difference between group keypoints $\mathcal{F}$ & $\mathcal{G}$,

$$TV = \frac{1}{4}\sum_{i=1}^{4} \big[V(t_{\mathcal{F}i}) - V(t_{\mathcal{G}i})\big], \tag{12}$$

and (iv) *ERV* (Expiratory Reserve Volume) as the volume change between passive end-tidal expiration and complete expiration, i.e. difference of the average volume at group keypoints $\mathcal{G}$ within the *tidal volume* section and volume at keypoint $\mathcal{A}$,

$$ERV = \frac{1}{4}\sum_{i=1}^{4} V(t_{\mathcal{G}i}) - V(t_{\mathcal{A}}), \tag{13}$$

Note that, based on spirometry experiment protocols [1], each FVC and SVC test should be repeated several times (at least three) to ensure consistency.

VII. Scaling Factor Generalization

So far we have shown that we can compute PFT measures from the Kinect volume-time and flow-time curves which have been calibrated by applying scaling factors computed using the corresponding spirometer volume-time curve. We refer to this as an 'intra-test' procedure. However, we need to remove this dependency, so we can compute PFT measures for a new trial[2] using only Kinect volume-time data - i.e. a more practical 'intra-subject' procedure.

[2]A trial refers to each performance of the FVC/SVC test by each subject.

As the change in the distance of the Kinect to a subject's thoracic wall is directly related to the change in their lung volume, our scaling factors are specific to each subject. In theory, this relationship should remain unchanged for a subject who performs a test several times (even on different days) with the same system configuration. However, in practice, this is only true for the *tidal volume* scaling factors, but not for the *main effort* scaling factor due to the subject's trunk motion. Since there is no significant movement during *tidal volume*, it should be possible to detect body movement during *main effort* by comparing scaling factors ξ_{tv} and ξ_{me}. However, even when ξ_{tv} and ξ_{me} are very similar (i.e. $\xi_{tv}/\xi_{me} \approx 1$), which implies there is no torso motion, the Kinect volume-time curve might still be affected by body movements. This can be categorized in two ways: (a) backward motion at the beginning of deep inhalation (between $\mathcal{E}$ and $\mathcal{A}$ keypoints) for FVC and SVC tests, and (b) forward lean at the beginning or middle of the deep and fast exhalation (after $\mathcal{A}$ in both tests), and then a move back at the end of exhalation such that it compensates the first forward lean - which might be also accompanied by the motion pattern in (a) as well. Figs. 12a and 12b present two examples of volume-time curves related to categories (a) and (b) and their scaling factors. The effects of similar motion artifacts on chest volume estimation, have also been reported in Yu et al. [8], Ostadabbas et al. [19], and Soleimani et al. [5], previously.

The similarity of the motion patterns of trunk movements across different trials of a subject allows us to estimate the best matching scaling factors for calibrating the Kinect volume-time curve of a new trial. This means that unless there is unexpected body movement, we can train our system to learn the *tidal volume* and *main effort* scaling factors for each subject, which enables us to compute PFT measures directly from the Kinect volume-time curve without using spirometer data when testing.

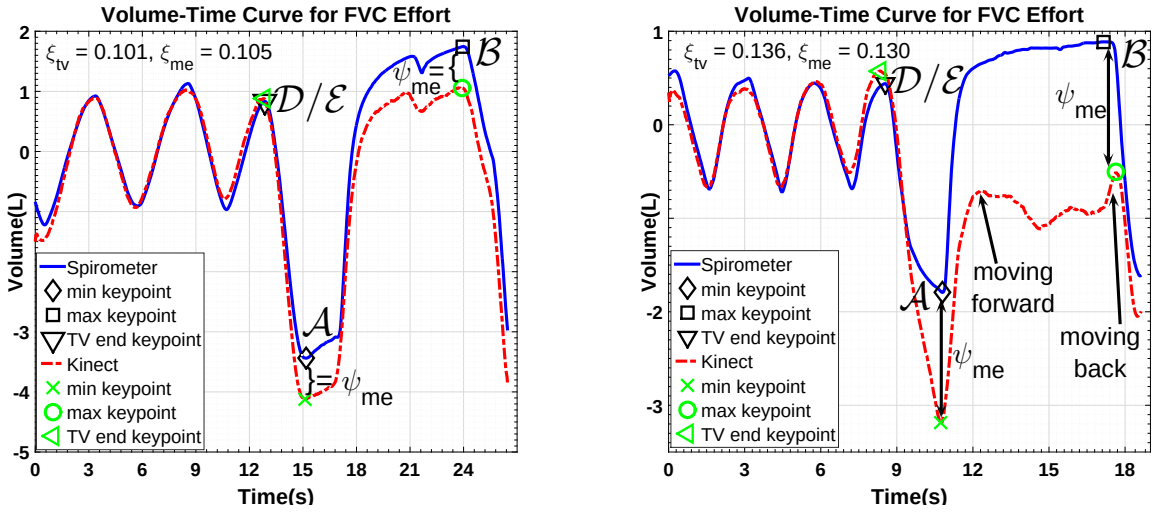
Training phase – We used training data, provided as pairs of corresponding Kinect and spirometer volume-time curves from training trials, to compute training *tidal volume* scaling factors $\{\xi_{tv}^{\ell}\}_{\ell=1}^{n_{tv}}$ and training *main effort* scaling factors and offsets $\{(\xi_{me}^{\ell}, \psi_{me}^{\ell})\}_{\ell=1}^{n_{me}}$, as explained in Sections V-B and V-C. n_{tv} and n_{me} are number of *tidal volume* and *main effort* training trials.

Testing phase – We calibrated the Kinect volume-time curve of a test trial, by applying the best matching scaling factors and offsets learned from the training phase. Our analysis showed that because the spirometer volume-time curve is always correct, then similar Kinect volume-time curves can be calibrated using similar scaling factors and offsets. Thus, to calibrate the test Kinect volume-time curve, we found the best matching scaling factors and offsets from the training phase using the curve similarity measures

$$F_{tv} = \frac{1}{4}\sum_{i=1}^{4} \Big[V_k(t_{\mathcal{F}i}) - V_k(t_{\mathcal{G}i})\Big], \tag{14}$$

$$F_{me} = \Big[V_k(t_{\mathcal{B}}) - V_k(t_{\mathcal{A}})\Big], \tag{15}$$

where $V_k(t)$ is the original Kinect volume-time curve, and



(a) (b)

Fig. 12. (a)-(b) Different types of torso motion which affect *main effort* even while ξ_{tv} & ξ_{me} values (mentioned in top of the figures) are very close to each other.

$t_{\mathcal{A}}, t_{\mathcal{B}}, t_{\mathcal{F}_i}$ and $t_{\mathcal{G}_i}$ are automatically computed keypoint timestamps, as introduced in Section V-A.

For the FVC test, the estimated *main effort* scaling factor ξ'_{me} was computed as

$$\xi'_{me} = \xi^k_{me} \ni k = \underset{j\in[1..n_{FS}]}{\arg\min} \left\{ \left| F^{test}_{me} - F^{j}_{me} \right| \right\}, \tag{16}$$

where F^{test}_{me} denotes the *main effort* curve similarity measure extracted from the test Kinect volume-time curve in (15), F^{j}_{me} is the same measure for the jth training Kinect volume-time curve, j denotes different trials, $\{\xi^{\ell}_{me}\}^{n_{FS}}_{\ell=1}$ states the training *main effort* scaling factors, and n_{FS} is the total number of training FVC and SVC trials for this subject. Since Vital Capacity, $\left|V_s(t_{\mathcal{A}}) - V_s(t_{\mathcal{B}})\right|$, is equal for FVC and SVC tests (notwithstanding the reproducibility measurement error), we also used training SVC trials to estimate the best matching scaling factors for the FVC test trial. As no measure is computed from the *tidal volume* section in FVC tests, F_{tv} was not extracted and therefore, ξ'_{tv} was not computed.

Similarly, for the SVC test, the estimated *tidal volume* scaling factor ξ'_{tv} and the estimated *main effort* scaling factor and offset (ξ'_{me}, ψ'_{me}), were computed as

$$\xi'_{tv} = \xi^k_{tv} \ \ni \ k = \underset{j\in[1..n_S]}{\arg\min} \left\{ \left| F^{test}_{tv} - F^{j}_{tv} \right| \right\}, \tag{17}$$

$$(\xi'_{me}, \psi'_{me}) = (\xi^k_{me}, \psi^k_{me}) \ni k = \underset{j\in[1..n_{FS}]}{\arg\min} \left\{ \left| F^{test}_{me} - F^{j}_{me} \right| \right\}, \tag{18}$$

where n_S is the total number of only SVC training trials. We do not use the *tidal volume* section of the FVC volume-time curves in the estimation of ξ'_{tv} because the *tidal volume* breathing cycles are too short in FVC tests and are not reliable for computing F_{tv} and consequently the *tidal volume* scaling factor. Note that in all FVC and SVC tests, $\psi'_{tv} \approx 0$.

After calibrating the Kinect volume-time curve of the *test* trial using the estimated *tidal volume* and *main effort* scaling factors and offsets, PFT clinical measures were computed from the calibrated Kinect volume-time curve.

The proposed scaling factor generalization was evaluated using leave-one-out cross-validation, which repeatedly takes one trial as the *test* and the rest as the training data. Leave-one-out is a more suitable approach, instead of k-fold cross-validation or other conventional validation methods, due to the limited number of FVC and SVC trials for each subject.

VIII. Experimental Results and Discussions

A. System Configuration and Data Acquisition

In each acquisition, the subject was asked to sit-up straight on a chair without armrests, facing the Kinect placed at a distance of 1.5m away from the subject and at a height of 0.6m (Fig. 13). This distance was chosen based on our study in Section IV. The subject was asked to put on a reasonably tight T-shirt to help improve the tracking accuracy of chest motion. Although putting subjects in supine position would have restricted their body movement during the Kinect test, we preferred to perform the test in the sitting position to simulate the spirometry setup. Moreover, it was difficult for fragile COPD patients to accomplish the *main effort* part of the test correctly in supine position.

The instruments used in our experiments were the Kinect V2 Microsoft depth sensor and the 'HDpft 1000 High Definition' spirometer, which provides raw volume-time and flow-time data at 200Hz for FVC and 50Hz for SVC. For validating the proposed method we compared our results with measures taken from the spirometer software.

Following ethical approval, we collected 529 Kinect and spirometry sequences on 85 patients attending respiratory clinic at Southmead Hospital in Bristol with a range of lung pathologies as they underwent their routine spirometry tests. The collection spanned several months between March and July of 2015. For each subject at least three FVC and three SVC efforts were recorded. The 36 male and 49 female

TABLE I
INTRA-TEST CORRELATION COEFFICIENT, MEAN AND STANDARD DEVIATION OF L_2 ERROR, AND RATIO OF EACH MEASURE'S L_2 ERROR TO THE MEAN VALUE OF THAT MEASURE FOR FVC MEASURES.

	3D-chest-model				Chest-averaging				Soleimani et al. [5]			
	λ_v	μ_v	σ_v	Ω_v	λ_m	μ_m	σ_m	Ω_m	λ_p	μ_p	σ_p	Ω_p
FVC	0.999	0.006	0.041	0.002	0.999	0.005	0.039	0.002	0.999	0.029	0.049	0.009
FEV1	0.929	0.285	0.241	0.137	0.940	0.266	0.217	0.127	0.947	0.284	0.220	0.127
PEF	0.756	1.685	1.284	0.490	0.774	1.618	1.259	0.464	0.805	2.008	1.325	0.613
FEF25%	0.701	1.696	1.282	0.597	0.719	1.650	1.246	0.572	–	–	–	–
FEF50%	0.687	0.931	0.916	0.375	0.729	0.877	0.830	0.340	–	–	–	–
FEF75%	0.577	0.576	0.637	0.559	0.595	0.528	0.576	0.554	–	–	–	–
FEF25-75%	0.719	0.757	0.676	0.414	0.728	0.737	0.665	0.409	0.790	0.642	0.539	0.333

TABLE II
INTRA-TEST CORRELATION COEFFICIENT, MEAN AND STANDARD DEVIATION OF L_2 ERROR, AND RATIO OF EACH MEASURE'S L_2 ERROR TO THE MEAN VALUE OF THAT MEASURE FOR SVC MEASURES.

	3D-chest-model				Chest-averaging				Soleimani et al. [5]			
	λ_v	μ_v	σ_v	Ω_v	λ_m	μ_m	σ_m	Ω_m	λ_p	μ_p	σ_p	Ω_p
VC	0.999	0.011	0.043	0.004	0.999	0.011	0.045	0.004	0.999	0.009	0.039	0.003
IC	0.998	0.045	0.040	0.019	0.998	0.043	0.040	0.019	0.997	0.048	0.047	0.020
TV	0.973	0.066	0.066	0.072	0.976	0.059	0.065	0.065	0.962	0.074	0.088	0.087
ERV	0.991	0.049	0.048	0.105	0.992	0.046	0.046	0.098	0.994	0.046	0.045	0.091

patients were aged between 24-83 years old (mean of 61.7), height of between 147.9-191.2cm (mean of 166.2cm), weight of 19.1-146.8kg (mean of 77.9kg), and BMI of between 6.9-45.7kg/cm^2 (mean of 28.1).

B. Intra-test Results

Tables I and II report the 3D-chest-model and the chest-averaging correlation coefficients (λ_v & λ_m) between the Kinect and the spirometer for all FVC and SVC test measures, along with the mean (μ_v & μ_m) and standard deviation (σ_v & σ_m) of the L_2 error for all 85 subjects (529 sequences). For each measure, we also report the ratio of the mean of the L_2 error to the mean value of that measure (Ω_v & Ω_m). These tables also present our previous results from [5] on 40 subjects (247 sequences). We note that the quality of the data for the first 40 subjects was very similar to the next 45 subjects (we verified this by observing the similarity of the correlation results for the two sets). This was expected as all the data were captured under similar conditions in the same clinic.

(a) (b)

Fig. 13. (a) The proposed setup showing the Kinect in front and the spirometer mouthpiece (when used) on the subject, (b) A typical Kinect depth image.

TABLE III
INTRA-SUBJECT CORRELATION COEFFICIENT, MEAN AND STANDARD DEVIATION OF L_2 ERROR, AND RATIO OF EACH MEASURE'S L_2 ERROR TO THE MEAN VALUE OF THAT MEASURE FOR FVC MEASURES.

	3D-chest-model				Chest-averaging			
	λ'_v	μ'_v	σ'_v	Ω'_v	λ'_m	μ'_m	σ'_m	Ω'_m
FVC	0.968	0.213	0.215	0.074	0.975	0.200	0.186	0.071
FEV1	0.906	0.332	0.280	0.163	0.927	0.299	0.243	0.146
PEF	0.753	1.756	1.301	0.523	0.769	1.717	1.286	0.508
FEF25%	0.703	1.757	1.272	0.633	0.715	1.735	1.254	0.621
FEF50%	0.682	0.933	0.910	0.385	0.715	0.882	0.822	0.354
FEF75%	0.585	0.570	0.606	0.564	0.603	0.509	0.540	0.553
FEF25-75%	0.717	0.758	0.662	0.425	0.721	0.727	0.670	0.417

TABLE IV
INTRA-SUBJECT CORRELATION COEFFICIENT, MEAN AND STANDARD DEVIATION OF L_2 ERROR, AND RATIO OF EACH MEASURE'S L_2 ERROR TO THE MEAN VALUE OF THAT MEASURE FOR SVC MEASURES.

	3D-chest-model				Chest-averaging			
	λ'_v	μ'_v	σ'_v	Ω'_v	λ'_m	μ'_m	σ'_m	Ω'_m
VC	0.956	0.237	0.239	0.084	0.963	0.214	0.248	0.075
IC	0.915	0.269	0.269	0.116	0.919	0.279	0.271	0.119
TV	0.888	0.118	0.137	0.129	0.924	0.098	0.110	0.107
ERV	0.737	0.297	0.310	0.592	0.750	0.280	0.300	0.561

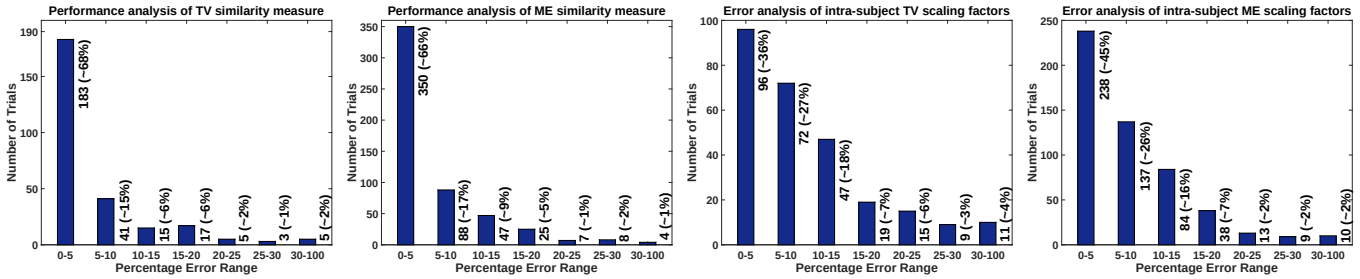
The results show that the Kinect and the spirometer correlate well for the FEV1 measure in the FVC tests and across all the SVC measures. The correlation amongst the other FVC measures is less strong due to the potential issues described later in subsection VIII-E. The results from both volume estimation methods are very close, with those from the chest-averaging based method just edging ahead. This confirms that the 3D-chest-model volume estimation method, with its greater space requirements and time complexity, does not necessarily obtain better results than the simple and fast chest-averaging approach. The *FVC* and *VC* results (gray background rows) are highly correlated due to the rescaling of the y-axis in the volume-time curves using their respective keypoints ($\mathcal{A}$ & $\mathcal{B}$).

In comparison to our previous work [5], where we performed only intra-tests for 40 patients, the proposed method achieved extremely similar, if not better, results. For example, we obtained reduced mean error (in μ_m) for all measures except *VC* and $FEF_{25-75\%}$ and improvement in *TV* measure correlation coefficient (λ_v & λ_m) and mean error (μ_v & μ_m) - across 85 patients including the same 40 from [5].

C. Intra-subject Results

Generalizing the scaling factor to compute intra-subject FVC and SVC measures is one of the major extensions in this study compared to our previous work [5]. Tables III and IV present the correlation coefficients (λ'_v & λ'_m), and the mean (μ'_v & μ'_m) and standard deviation (σ'_v & σ'_m) of L_2 error for FVC and SVC computed measures for all 85 subjects. It also reports the ratio of mean of the L_2 error to the mean value of that measure (Ω'_v & Ω'_m). Similar to the intra-test results, the chest-averaging based method provides slightly better results.

The FVC test results in Table III, λ'_v & λ'_m indicate strong correlation of the *FVC* and *FEV1* measures against the spirom-



(a) 269 Trials (b) 529 Trials (c) 269 Trials (d) 529 Trials

Fig. 14. (a) Performance analysis of intra-subject *tidal volume* similarity measure, i.e. F_{tv} (14), (b) Performance analysis of intra-subject *main effort* similarity measure, i.e. F_{me} (15), (c) Intra-subject *tidal volume* scaling factors error analysis, (d) Intra-subject *main effort* scaling factors error analysis.

eter, with the other five measures correlating reasonably well at a minimum of 0.603 for $FEF_{75\%}$ in the Chest-averaging model. Furthermore, good correlation can be seen between the intra-subject and intra-test FVC measures (Tables I and III).

The SVC results λ'_v & λ'_m in Table IV also show strong correlation against the spirometer for *VC*, *IC*, and *TV* measures and good correlation for *ERV*. However, the differences between intra-subject mean (μ'_v & μ'_m) and standard deviation (σ'_v & σ'_m) of errors (Table IV) and their intra-test counterparts (μ_v & μ_m and σ_v & σ_m from Table II) are higher than these differences in FVC test. This is because SVC requires two scaling factors for the *tidal volume* and *main effort* parts of the curve, in addition to estimating the offset ψ'_{me}.

D. Statistical Analysis of Intra-subject Scaling Factors

The *Tidal volume* and *main effort* test trials are calibrated using intra-subject scaling factors ξ'_{tv} & ξ'_{me}, which are chosen from the training sets $\{\xi^{\ell}_{tv}\}_{\ell=1}^{n_S}$ & $\{\xi^{\ell}_{me}\}_{\ell=1}^{n_{FS}}$, respectively, using (16), (17), and (18) based on the similarity measures in (14) and (15). The performance of the similarity measures, in terms of choosing the best intra-subject scaling factors from the training set, is evaluated by computing the normalised L_2 error:

$$E^c_{\xi'_{tv}} = \frac{\sqrt{(\xi'_{tv} - \xi^c_{tv})^2}}{\xi^c_{tv}}, \tag{19}$$

$$E^c_{\xi'_{me}} = \frac{\sqrt{(\xi'_{me} - \xi^c_{me})^2}}{\xi^c_{me}}, \tag{20}$$

where ξ^c_{tv} & ξ^c_{me} are the closest scaling factors in the training set to the original scaling factors of the test trial ξ^o_{tv} & ξ^o_{me}. The original scaling factors were computed using the corresponding spirometer data as explained in Sections V-B and V-C.

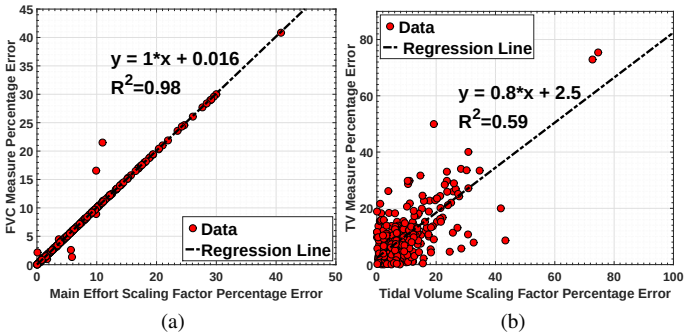
Figs. 14a and 14b report the distribution of these errors for all *tidal volume* and *main effort* trials, respectively, in the range $0-30\%$ at 5% interval and then in the entire $30-100\%$ range. As can be seen, $\sim$83% of *tidal volume* scaling factors and $\sim$83% of *main effort* scaling factors are within an error of less than 10%. Only $\sim$2% of *tidal volume* scaling factors and $\sim$1% of *main effort* scaling factors have errors of greater than 30%.

Further, for each test trial, to compare the estimated intra-subject *tidal volume* and *main effort* scaling factors ξ'_{tv} & ξ'_{me} to the original scaling factors ξ^o_{tv} & ξ^o_{me}, their normalised L_2 error is computed similar to (19) and (20). As seen in Figs. 14c and 14d, which present the distribution of errors for all *tidal volume* and *main effort* trials, $\sim$81% of *tidal volume* scaling factors and $\sim$87% of *main effort* trials have an error of less than 15%. Only $\sim$4% of *tidal volume* scaling factors and $\sim$2% of *main effort* scaling factors have an error of greater than 30%.

We also analysed the correlation between the *tidal volume* and *main effort* scaling factor normalized L_2 errors $E^o_{\xi'_{tv}}$ & $E^o_{\xi'_{me}}$, and error of FVC and SVC computed measures. Figs. 15a and 15b, present this correlation for *FVC* and *TV* measures. As can be seen, there is a high correlation between the *FVC* measure error and the *main effort* scaling factor error across all trials. This correlation is less strong for the *TV* measure error and *tidal volume* scaling factor error. The reason for this is, *tidal volume* scaling factors are computed using all data points of *tidal volume* part of volume-time curve, and *TV* measure itself is computed using group keypoints $\mathcal{F}$ & $\mathcal{G}$ (12). However, *FVC* measure and *main effort* scaling factors are both computed using the same keypoints $\mathcal{A}$ & $\mathcal{B}$. Thus, they are better correlated (15a) than the *TV* measure error and the *tidal volume* scaling factor error (15b).

E. Measurement Stability

It is important to note that even spirometer readings differ between multiple consecutive trials for the same subject, thus



(a) (b)

Fig. 15. (a) Correlation between the *FVC* measure error and the *main effort* intra-subject scaling factor error $E^o_{\xi'_{me}}$, (b) Correlation between the *TV* measure error and the *tidal volume* intra-subject scaling factor error $E^o_{\xi'_{tv}}$.

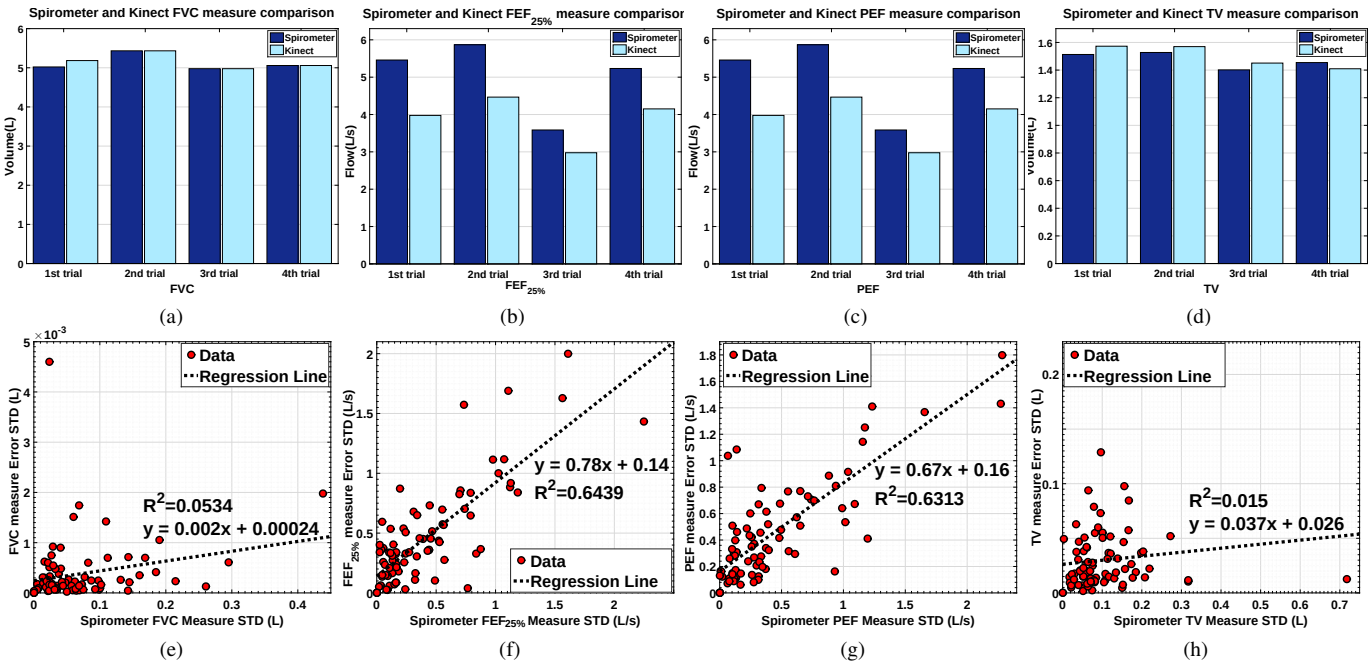


Fig. 16. (a)-(d) Comparison of spirometer provided PFT measures (*FVC*, $FEF_{25\%}$, *PEF* and *TV*) to the computed ones from the Kinect for 4 trials of a sample patient, (e)-(h) Regression of standard deviation of spirometer provided measures (*FVC*, $FEF_{25\%}$, *PEF* and *TV*) in all trials for each patient (85 subjects in total) and standard deviation of those measures error computed by the proposed method.

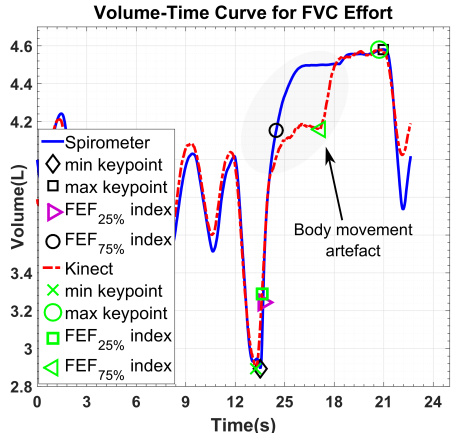


Fig. 17. An example of how body movement can affect the computation of the $FEF_{25-75\%}$ measure.

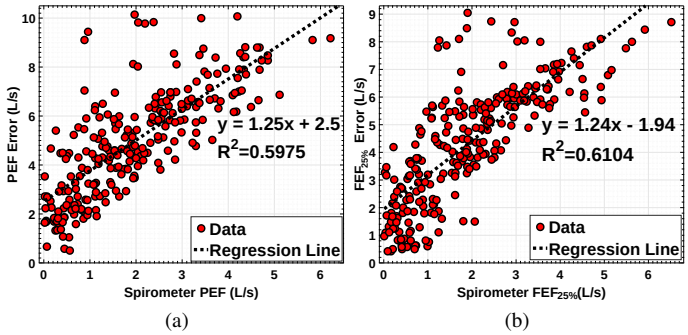


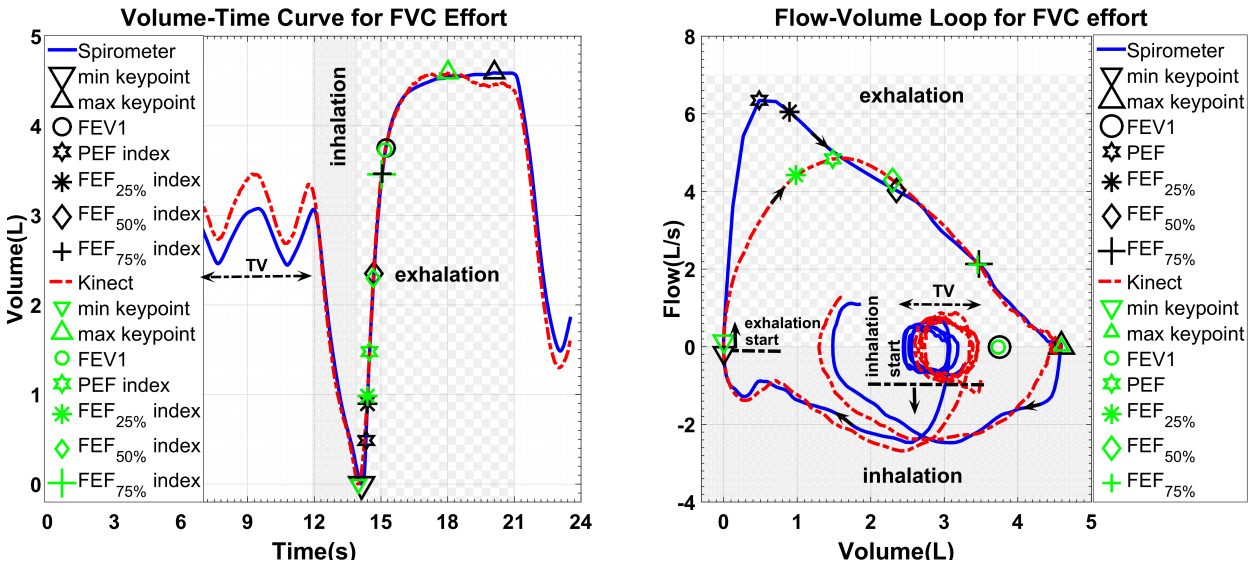
Fig. 18. The proposed method's error increases due to a subject's inevitable trunk movement while they blow harder and faster into the spirometer to achieve higher (better) *PEF* and $FEF_{25\%}$ measures.

requiring at least three trials with similar readings before a clinician considers the results. This is illustrated in Figs. 16(a)-16(d) which present some examples measures (*FVC*, $FEF_{25\%}$, *PEF*, and *TV*), provided by the spirometer and the proposed method for one subject from four consecutive trials.

To find out the correlation between spirometry reproducibility and the proposed method's error in computation of measures, we obtained the standard deviation of each measure and its corresponding error in all repeated trials for each patient. Figs. 16(e)-16(h) show the computed correlation for *FVC*, $FEF_{25\%}$, *PEF*, and *TV* measures. These results indicate that when the measures provided by the spirometer are less consistent, the error between measures obtained by the proposed method and the spirometer increases.

The subject's body movement during a test is a primary reason for poor correlation and this is more evident in *main effort* measures. A specific example of how body movement (due to expiration pressure) can affect the $FEF_{25-75\%}$ measure is in Fig. 17, where the estimation of $0.75FVC$ is sometimes compromised. In another observation, illustrated in Fig. 18, we found that as the $FEF_{25\%}$ and *PEF* readings from the spirometer increases, our proposed method's error also increases. To the best of our knowledge, this happens as subjects try to attain better lung function measures by blowing faster into the spirometer which inevitably results in more trunk movement.

PEF and $FEF_{25\%}$ are more affected by the patient's trunk translation because (a) they are calculated using flow data which is the first derivative of the volume over time and so is sensitive to displacements, and (b) PFT and $FEF_{25\%}$ are located at the beginning of *main effort* section (Fig. 19), which



(a) (b)

Fig. 19. Although the spirometer and the Kinect volume-time curves are very similar in *main effort* in (a), the corresponding flow-volume loop is different at the beginning of exhalation in (b).

is more affected by the movement. Even subtle movements caused by leaning forward, due to forcible expiration, affects keypoint positions of these measures. In Fig. 19a, although the *main effort* parts of the curves match very well, their flow-volume loop is considerably different in Fig. 19b between the start of exhalation and the location of the $FEF_{50\%}$ point.

IX. Conclusion and Future Work

We proposed a remote, non-invasive depth-based approach for Pulmonary Function Testing. The proposed system generates Kinect-based volume-time and flow-time curves, and by locating several keypoints automatically, we computed several FVC and SVC measures which we compared against a spirometer, and evaluated their reproducibility. We analysed the subject's trunk motion pattern to generalize scaling factors to be able to compute intra-subject PFT measures for each subject, without having to use a spirometer to calibrate against for each trial. We validated our system in a clinical environment with 85 actual patients and achieved high intra-test and intra-subject correlation against the spirometer.

This work is a considerable step forward in the development of remote non-contact monitoring of patients with respiratory disease. This 'real world' clinical data, collected from a large group of patients with a wide range of lung function is unique. We are able to accurately obtain respiratory measures remotely which has potential clinical applications for monitoring of patients in the home, gating (timing) of thoracic imaging and synchronisation with ventilatory support. In summary, in this work we have taken a vital step towards the aim of applying the Kinect as an independent surrogate for spirometry by only needing the spirometer one time for each patient to obtain a personalised scaling factor.

In our future work, we plan to use two Kinects to decouple body motion and chest motion to increase the accuracy of our PFT measures. We also plan to use machine learning techniques to generalize the scaling factors by introducing parameters such as height, weight and age in the estimation and remove the need for subject specific spirometry.

Appendix

Table V presents list of abbreviations, corresponding terms, and their brief description.

Acknowledgement

The authors would like to thank the University of Bristol Alumni Foundation for funding this research and the patients at Southmead Hospital in Bristol who participated in it. The authors would also like to thank SPHERE[3], an EPSRC Interdisciplinary Research Centre, for providing the opportunity of collaboration between engineering and clinical researchers. Data from this study are unavailable for sharing due to insufficient consent from study participants.

References

[1] M. Miller et al., "Standardisation of spirometry." *The European Respiratory Journal*, vol. 26, no. 2, pp. 319–38, 2005.
[2] C. P. Criée et al., "Body plethysmography–its principles and clinical use." *Respiratory medicine*, vol. 105, no. 7, pp. 959–971, Jul. 2006.
[3] R. Pierce, "Spirometry: an essential clinical measurement." *Australian family physician*, vol. 34, no. 7, pp. 535–9, jul 2005.
[4] Microsoft, "Kinect for Windows," https://dev.windows.com/en-us/kinect, [Accessed: Jan. 2016].
[5] V. Soleimani et al., "Remote pulmonary function testing using a depth sensor," in *IEEE BioCAS*, Oct 2015, pp. 1–4.
[6] S. Ostadabbas et al, "A passive quantitative measurement of airway resistance using depth data." in *IEEE-EMBS*, 2014, pp. 5743–5747.
[7] H. Aoki et al., "Non-contact Respiration Measurement Using Structured Light 3-D Sensor," in *SICE*, 2012, pp. 614–618.
[8] M.-C. Yu et al., "Noncontact respiratory measurement of volume change using depth camera," in *IEEE-EMBS*, 2012, pp. 2371–2374.

[3] *Sensor Platform for HEalthcare in a Residential Environment.*

TABLE V
LIST OF ABBREVIATIONS AND THEIR CORRESPONDING TERM AND DESCRIPTION.

Abbr.	Term	Description
BMI	Body Mass Index	A measure of body fat based on weight and height.
COPD	Chronic Obstructive Pulmonary disease	Long-term respiratory conditions characterized by airway obstruction.
EMD	Empirical Mode Decomposition	An adaptive method to decompose a signal into its individual components.
ERV	Expiratory Reserve Volume (SVC measure)	Volume change between passive end-tidal expiration and complete expiration.
$FEF_{25\%}$	Forced Expiratory Flow 25% (FVC measure)	Flow of exhaled air at 25% of *FVC*.
$FEF_{50\%}$	Forced Expiratory Flow 50% (FVC measure)	Flow of exhaled air at 50% of *FVC*.
$FEF_{75\%}$	Forced Expiratory Flow 75% (FVC measure)	Flow of exhaled air at 75% of *FVC*.
$FEF_{25\text{-}75\%}$	Forced Expiratory Flow 25-75% (FVC measure)	Mean forced expiratory flow between 25% and 75% of *FVC*.
FEV1	Forced Expiratory Volume (FVC measure)	Volume of air forcibly expired at the 1st second.
FVC	Forced Vital Capacity (FVC measure)	Maximum amount of air in litres blown out after a maximal inhalation.
FVC	Forced Vital Capacity (test name)	Lung function tests based on forced blowing manoeuver.
IC	Inspiratory Capacity (SVC measure)	Volume change between full inspiration and a passive end-tidal expiration.
IMF	Intrinsic Mode Functions	Decomposed individual components obtained by applying EMD.
PEF	Peak Expiratory Flow (FVC measure)	Maximum speed of exhaled air in FVC test
PCA	Principle Component Analysis	A statistical procedure for high-dimensional data analysis.
PF	Peak Flow	Maximum speed of air flow in FVC test.
PFT	Pulmonary Function Testing	Respiratory tests for assessing lung function.
SVC	Slow Vital Capacity (test name)	Lung function tests based on slow blowing manoeuver.
SDK	Software Development Kit	Set of software tools which allows the creation of applications.
TV	Tidal Volume (SVC measure)	Volume of air inspired and expired at rest condition.
VC	Vital Capacity (SVC measure)	Volume change between full inspiration and complete expiration.

[9] T. M. Seppanen et al., "Accurate measurement of respiratory airflow waveforms using depth data," in *Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE*, Aug 2015, pp. 7857–7860.

[10] W. H. de Boer et al., "SLP: A Zero-Contact Non-Invasive Method for Pulmonary Function Testing," *BMVC 2010*, pp. 85.1–85.12.

[11] E. Bernal et al., "Non contact monitoring of respiratory function via depth sensing," in *IEEE-EMBS*, 2014, pp. 104–105.

[12] F. Benetazzo et al., "Respiratory rate detection algorithm based on RGB-D camera: theoretical background and experimental results," *Healthcare Technology Letters*, vol. 1, no. 3, pp. 81–86, Sep. 2014.

[13] F. Tahavori et al., "Marker-less Respiratory Motion Modeling Using the Microsoft Kinect for Windows," in *SPIE Medical Imaging*, 2014.

[14] P. Noonan et al., "Accurate Markerless Respiratory Tracking For Gated Whole Body PET Using the Microsoft Kinect," in *IEEE Nuclear Science Symp. and Medical Imaging*, 2012, pp. 3973–3974.

[15] J. Xia and R. A. Siochi, "A real-time respiratory motion monitoring system using KINECT: proof of concept," *Medical Physics*, vol. 39, no. 5, pp. 2682–2685, 2012.

[16] M. Martinez and R. Stiefelhagen, "Breath Rate Monitoring During Sleep using Near-IR Imagery and PCA," in *ICPR*, 2012, pp. 3472–3475.

[17] N. Bernacchia et al., "Non contact measurement of heart and respiration rates based on Kinect," in *MeMeA*, 2014, pp. 1–5.

[18] N. Burba et al., "Unobtrusive measurement of subtle nonverbal behaviors with the Microsoft Kinect," in *IEEE Virtual Reality Short Papers and Posters*, 2012, pp. 1–4.

[19] S. Ostadabbas et al., "A vision-based respiration monitoring system for passive airway resistance estimation," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 9, pp. 1904–1913, 2016.

[20] J. Penne et al., "Robust real-time 3D respiratory motion detection using time-of-flight cameras," *Computer Assisted Radiology and Surgery*, vol. 3, no. 5, pp. 427–431, 2008.

[21] K. S. Tan et al., "Real-time vision based respiration monitoring system," in *Int. Sym. on Comm. Sys. Net. and Dig. Sig. Proc.*, 2010, pp. 770–774.

[22] M. Frigola et al., "Vision Based Respiratory Monitoring System," in *Conf. on Control and Automation*, 2002, pp. 1–5.

[23] L. Scalise et al., "Optical method for measurement of respiration rate," in *Int. Workshopsh on Medical Measurement and App.*, 2010, pp. 19–22.

[24] I. Sato and M. Nakajima, "Non-contact Breath Motion Monitoring System in Full Automation," in *IEEE-EMBS*, 2005, pp. 3448–3451.

[25] G. Ferrigno et al., "Three-dimensional optical analysis of chest wall motion," *J Appl Physiol*, vol. 77, no. 3, pp. 1224–1231, 1994.

[26] S. J. Cala et al., "Chest wall and lung volume estimation by optical reflectance motion analysis," *J. Appl Physiol*, vol. 81, no. 6, pp. 2680–2689, 1996.

[27] R. L. Dellaca et al., "Estimation of end-expiratory lung volume variations by optoelectronic plethysmography," *Critical care medicine*, vol. 29, no. 9, pp. 1807–1811, 2001.

[28] A. Aliverti and A. Pedotti, "Opto-electronic plethysmography," in *Mechanics of Breathing*. Springer Milan, 2002, pp. 47–59.

[29] L. Marvin et al., "Time-of-flight sensor calibration for accurate range sensing," *Computer Vision and Image Understanding*, vol. 114, no. 12, pp. 1318 – 1328, 2010.

[30] T. Breuer et al., "Low-cost commodity depth sensor comparison and accuracy analysis," in *SPIE Security+ Defence*, 2014.

[31] S. Paris et al., "Bilateral filtering: Theory and applications," *Foundations and Trends in Computer Graphics and Vision*, vol. 4, no. 1, pp. 1–73, 2008.

[32] M. Camplani et al., "Depth-color fusion strategy for 3-d scene modeling with kinect," *IEEE Transactions on Cybernetics*, vol. 43, no. 6, pp. 1560–1571, 2013.

[33] D. A. Forsyth and J. Ponce, *Computer Vision: A Modern Approach*. Prentice Hall Professional Technical Reference, 2002.

[34] J. Fei and I. Pavlidis, "Analysis of breathing air flow patterns in thermal imaging," in *EMBS 2006*, pp. 946–952.

[35] R. Marani et al., "A new system for continuous monitoring of breathing and kinetic activity," *Journal of Sensors*, vol. 2010, 2010.

[36] C. Park and B. Lee, "Real-time estimation of respiratory rate from a photoplethysmogram using an adaptive lattice notch filter," *Biomedical engineering online*, vol. 13, no. 1, 2014.

[37] S. W. Smith, *Digital signal processing: a practical guide for engineers and scientists*. Newnes, 2003.

[38] S. K. Mitra and Y. Kuo, *Digital signal processing: a computer-based approach*, 4th ed. McGraw-Hill New York, 2006, vol. 2.

[39] W. F. Ganong and K. E. Barrett, *Review of medical physiology*, 24th ed. Appleton & Lange Norwalk, CT, 1995.

[40] J. McFadden et al., "Raised respiratory rate in elderly patients: a valuable physical sign." *Br Med J (Clin Res Ed)*, vol. 284, no. 6316, pp. 626–627, 1982.

[41] M. A. Cretikos et al., "Respiratory rate: the neglected vital sign," *Medical Journal of Australia*, vol. 188, no. 11, p. 657, 2008.

[42] M. de Berg et al., *Computational Geometry: Algorithms and Applications*, 3rd ed. Springer-Verlag, 2008.

[43] N. E. Huang et al., "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Math., Phys. and Eng. Sciences*, vol. 454, pp. 903–995, 1998.