

# 3D imaging system for respiratory monitoring in pediatric intensive care environment

*Haythem Rehouma<sup>a</sup>, Rita Noumeir<sup>a</sup>, Wassim Bouachir<sup>b</sup>, Philippe Jouve<sup>c</sup>, Sandrine Essouri<sup>c</sup>*

<sup>a</sup> *École de Technologie Supérieure, Montréal (QC), Canada*

<sup>b</sup> *LICEF Research Center, TÉLUQ University, Montréal, Canada*

<sup>c</sup> *CHU Sainte-Justine, Mother and Child University Hospital Center, Montréal, Canada*

---

## ABSTRACT

Assessment of respiratory activity in pediatric intensive care unit allows a comprehensive view of the patient's condition. This allows the identification of high-risk cases for prompt and appropriate medical treatment. Numerous research works on respiration monitoring have been conducted in recent years. However, most of them are unsuitable for clinical environment or require physical contact with the patient, which limits their efficiency. In this paper, we present a novel system for measuring the breathing pattern based on a computer vision method and contactless design. Our 3D imaging system is specifically designed for pediatric intensive care environment, which distinguishes it from the other imaging methods. Indeed, previous works are mostly limited to the use of conventional video acquisition devices, in addition to not considering the constraints imposed by intensive care environment. The proposed system uses depth information captured by two (Red Green Blue-Depth) RGB-D cameras at different view angles, by considering the intensive care unit constraints. Depth information is then exploited to reconstruct a 3D surface of a patient's torso with high temporal and spatial resolution and large spatial coverage. Our system captures the motion information for the top of the torso surface as well as for its both lateral sides. For each reconstruction, the volume is estimated through a recursive subdivision of the 3D space into cubic unit elements. The volume change is then calculated through a subtraction technique between successive reconstructions. We tested our system in the pediatric intensive care unit of the Sainte-Justine university hospital center, where it was compared to the gold standard method currently used in pediatric intensive care units. The performed experiments showed a very high accuracy and precision of the proposed imaging system in estimating respiratory rate and tidal volume.

*Keywords:* Intensive care unit, RGB-D camera, Point cloud, Surface reconstruction, Tidal volume, Respiratory monitoring.

---

## 1. Introduction

The Pediatric Intensive Care Unit (PICU) receives patients in critical condition, from newborns to 18 years old. In a PICU room, the patient is closely monitored on a 24h/24h basis to intervene in case of sudden worsening. Vital signs, including the respiratory rate (RR), are continuously measured to ensure patient's safety. An abnormal RR (too high or too low) can be a predictor of a potential physiological deterioration. Thus, the RR is used with other respiratory parameters for the assessment of the patient's condition. These respiratory parameters, including also tidal volume (Vt) and retraction signs of the breast are collected through clinical examination and several medical devices. Measuring them accurately is crucial in PICU. If the child is mechanically ventilated, RR and Vt are directly measured by the ventilator. With spontaneous breathing patients, RR is monitored using thoracic electrodes (plethysmography) that are sensitive to movements with a high rate of erroneous measurements, while Vt is not assessed by medical devices during spontaneous breathing. Lastly, retraction signs of the breast are only estimated by clinical examination, in a subjective way, even when caregivers use a scoring system (Essouri et al., 2016).

PICU plays a pivotal role in acute care hospitals, but services are sometimes affected by complex settings and unexpected urgent interventions. In this environment, rooms are too small for their functions, equipment, and the number of patients. It is therefore important to manage the space properly. Any occupied space should not cause interruptions, and caregivers need to provide the appropriate services with sufficiently free space around the patient. Due to these constraints, previous works on contactless RR and Vt assessment were limited to proofs of concept. They were conducted for a general use case, and are thus inappropriate for Pediatric Intensive Care Environment (PICE) (Benetazzo et al., 2014; Ostadabbas et al., 2014; Katashev et al., 2015; Harte et al., 2016; Sharp et al., 2017).

In this work, we present a contactless imaging system specifically designed to monitor critically-ill children in PICU. Our system allows early identification of evolving child conditions through a 3D imaging measurement method. Using two RGB-D sensors, we first align point clouds captured from different view angles in a common coordinate system. The aligned point cloud is segmented to reconstruct the surface of body regions involved by respiration. The volume is then calculated for each frame through a recursive subdivision of the 3D space into cubic unit elements. Finally, Vt and RR are accurately estimated by analyzing the changes in the calculated volumes.

Breathing monitoring systems can be classified into two main categories according to whether they are designed in a contact or a contactless approach (Al-Khalidi et al., 2011). With contact based methods, sensing instruments are attached to the body to measure breathing parameters such as airflow from

acoustic breathing signals (Corbishley and Rodríguez-Villegas, 2008), body movement (Kan and Chen, 2012), and respiratory CO<sub>2</sub> emission and oximetry probe SpO<sub>2</sub> (Dullenkopf et al., 2003).

Despite their considerable potential in hospital wards, contact-based methods can cause disturbance to the patient, especially for children. For this reason, non-contact methods were given particular interest (Al-Khalidi et al., 2011). Over the last few years, several works based on non-contact sensors have been proposed for healthcare applications, such as respiratory assessment and sleep apnea detection. In this context, Lee et al. (2015) have presented a system based a Microwave Doppler Radar for capturing different respiratory dynamics like slow, normal, and fast breathing. This method is based on measuring the change in the phase and frequency of the transmitted signal due to the movement of the chest/abdomen during the process of inhalation and exhalation. Detailed information on the breathing activity was extracted by decomposing the breathing cycle into inhalation, pause, and exhalation. Although their technique showed a good correlation with a conventional measurement device of respiratory rate (respiration belt), the authors did not report any volume measures. Moreover, the system was not tested in a clinical environment.

Other computer vision-based approaches have been proposed for respiratory assessment. For example, Bai et al. (2010) describe a temporal differencing method to continuously record the patient's breath and send the respiratory information over the Internet network to the hospital. An alarm signal is sent when the breathing activity stops for more than 10 seconds or if the respiratory rate is too low or fast. The proposed system detects moving objects by image processing, and finds the differences between two continuous 2D frames to estimate breathing change. Despite the use of two webcams providing a wide detection range, only breath rate can be calculated using temporal differencing. Indeed, the use of 2D imaging methods does not allow volume calculation because of the lack of the motion information in the 3D space.

In another study, researchers used optical flow variations between frames to visualize the apparent velocity field of the entire body motion, including breast movement due to respiration (Nakajima et al., 2001, 1997). Nevertheless, their method does not allow the quantitative evaluation of respiratory parameters.

Within the last years, computer vision approaches were moving towards investigation of depth data in many important problems, such as detection, tracking and recognition. This trend resulted from the emergence of new depth acquisition devices with high accuracy. Such devices have greatly simplified the task of human breath detection, which have given rise to new possibilities for respiratory parameters estimation. In this direction, Xia and Siochi (2012) proposed a contactless respiratory monitoring system using a structured-light (SL) scanner. The structured light technology consists of projecting a known infrared pattern onto a scene. Depth data are then calculated by quantifying the differences between the captured pattern and the reference pattern. Therefore, the sensor can handle any change in the scene illumination by only using the depth information. Once the images are captured, the average depth is calculated over a thoraco-abdominal zone. This zone was manually determined by placing a translation surface on the thorax in the center of the image. However, the study was a simple proof of concept of how to use raw depth data to visualize respiratory motion signals. It does not allow a concrete evaluation of respiration by determining breathing parameters, such as RR and V<sub>t</sub>.

Another approach to estimate the respiratory rate was proposed (Benetazzo et al., 2014). The authors used a RGB-D SL camera to develop an algorithm for respiratory rate monitoring. Their algorithm requires the user to be sitting in frontal position with an angular orientation not exceeding 25°. Though the temporal resolution of their method was good, the spatial coverage was very limited and thus it was impossible to estimate V<sub>t</sub> accurately.

Tahavori et al. (2014) placed an RGB-D SL camera above the body to calculate the average depth value of sixteen regions of interest on the chest and abdomen. They apply the principal component analysis to the depth information of these regions and demonstrate that only the first principal component describes nearly 70% of the motion data variance in chest and abdomen surfaces. Aoki et al. (2015) presented a breathing detection system using a SL camera as well. Thorax movements were detected by recording the depth information during bicycle pedaling exercise. However, these works have been shown to not be suited to the clinical environment and specifically for the PICU because they do not provide quantitative measure, in addition to their setting complexity.

Lately, Harte et al. (2016) developed a prototype for chest wall motion assessment using four SL cameras. The system aims to fit the clinical environment more than the existing respiration assessment methods. It performs a 3D reconstruction of the patient's torso to assess how the chest wall moves. The obtained results were compared with those of a spirometer. The system showed a good correlation with spirometry in static objects. Despite the method providing additional information about chest movement with a high spatial coverage of respiration zone, the system is inappropriate in a PICU settings. This is because the cameras were placed around the patient supposing him in a standing position. More precisely, the patient must be in a precise marked position in order to meet the system conditions. This means that patients have to adapt their posture to the system and not the opposite. This makes the system unable to continuously monitor the respiration of children. In addition, the used hardware (four devices, excess of wiring) is not suitable as well for the PICE.

In this work, we developed a non-contact breathing measurement system within the constraints imposed by the specific nature of PICU. The requirements of high spatial coverage, spatial and temporal resolutions are very important for estimating breathing patterns accurately. It can be seen from the previous works cited above that a lack of one of these three factors affects the performance of the overall assessment. In fact, the lack of spatial coverage or spatial resolution conducts to incomplete or partial information in space. In this case, only a part of respiratory information will be used which can lead to imprecise and inaccurate results, especially when estimating V<sub>t</sub>. Likewise, the lack of temporal resolution can provide misleading or incomplete information about the respiratory motion over time.

Our approach is thus based on considering these three important factors: spatial resolution, temporal resolution and spatial coverage of the respiration zone, while taking into account the constraints of the PICE environment.

## 2. Materials and methods

### 2.1. Camera system and sensing properties

An RGB-D sensor is able to capture three simultaneous streams: Color (RGB), Depth (D) and Infrared Radiation (IR). There have been multiple RGB-D cameras released by Intel and Microsoft over the last years. However, these devices presently work with a borderline level of acceptance of depth resolution. Many research works investigated the performance of RGB-D cameras such as Kinect v1 sensor (Khoshelham and Elberink, 2012; Mallick et al., 2014), Kinect v2 sensor (Lachat et al., 2015; Pöhlmann et al., 2016; Sarbolandi et al., 2015) and Intel RealSense RGB-D imaging systems (Keselman et al., 2017). The Kinect v2 sensor has been markedly used in many research works due its promising properties. Although the first version of Kinect (Kinect v1) was used in many applications such as tracking and recognition (Oikonomidis et al., 2011; Schwarz et al., 2012; Beserra Gomes et al., 2013; González-Ortega et al., 2014; Li et al., 2016), the Kinect v2 has greatly improved upon it and was used as well in many application such as mobile robot navigation (Bloesch and Rodriguez, 2015), human motion recognition (Lun and Zhao, 2015) and suicide attempts preventing (Bouachir et al., 2018). Table 1 presents the technical specifications of both Kinect v1 and v2. The Kinect v2 depth camera does not only present a significant higher color and depth resolution, but also a wider field of view in comparison to Kinect v1. Furthermore, it has been proven, in a comparative study that Kinect v2 surpasses Kinect v1 in terms of accuracy and precision (Gonzalez-Jorge et al., 2015). Accuracy is the ability of a measuring method or device to capture the true value of a measure while precision is the ability of a measurement to be consistently reproduced. The experimental study investigated the accuracy and precision of the two versions of Kinect cameras using a standard artefact based on five spheres and seven cubes. The tests were performed for two ranges (1m, 2m) and three inclination angles between each sensor and the artefact (45°, 90°, 135°). Table 2 summarizes the results of this analysis. It can be clearly seen that the accuracy is better for Kinect v2 at both 1 m and 2 m range. Precision as well is better for Kinect v2 mainly at the 2 m range.

**Table 1**

Comparative specifications and technical features of Microsoft Kinect v1 and v2 (Al-Naji et al., 2017; Bloesch et al., 2015)

Features	Kinect v1	Kinect v2
Dimensions (w x d x h) [mm]	180 x 35 x 25	249 x 66 x 67
Connection type	USB 2.0	USB 3.0
Depth sensor technology	SL	ToF
Frame rate	30 Hz	30 Hz
Minimum latency	102 ms	20-60 ms
Color camera resolution	640 x 480 px	1920 x 1080 px
Infrared camera resolution	320 x 240 px	512 x 424 px
RGB camera field of view	62° x 48.6°	84.1° x 53.8°
Depth image field of view	57° x 43°	70° x 60°
Operative measuring range	0.4 m – 3.5 m	0.5 m – 4.5 m
Skeleton joints number	20 joints	25 joints
Maximum skeletal tracking	2	6

**Table 2**

Precision and accuracy results for Kinect I and Kinect II (Gonzalez-Jorge et al., 2015)

Range	Accuracy (mm)		Precision (mm)	
	Kinect I	Kinect II	Kinect I	Kinect II
1 m	12	5	[2-6]	[2-6]
2m	25	5	< 12	< 8

Fig. 1 illustrates the Kinect v2 imaging system as well as the measurement technique. The camera includes three devices (RGB camera, an infrared emitter and an infrared camera) to acquire color, infrared and depth images of the scene. This is illustrated in Fig. 1.a. The color data arise from the RGB sensor, while infrared data and depth maps come from the depth sensor and have the same resolution. The color data have a very high-resolution of 1920 x 1080 px. The depth maps of inferior resolutions (512 x 424 px) are 2D images, where depth information is stored for each pixel. To estimate depth, the Kinect v2 uses the time-of-flight technique by measuring the round-trip time needed by a light pulse to travel from the sensor illuminator to the target object and back again (Fig. 1.b). The illuminator is a near-infrared laser diode emitting a modulated infrared signal to the object. The reflected light is collected by the sensor detector. A timing generator is used to synchronize the actions of the emitter and the sensor detector. The depth of each pixel is then calculated by Equation 1:

$$d = \frac{\Delta\phi}{4\pi f} \times c \quad (1)$$

where  $d$  is the distance to be measured (pixel depth),  $\Delta\phi$  is the phase shift between the emitted light and the reflected light,  $c$  is the speed of light ( $3 \times 10^8$  m/s) and  $f$  is the modulation frequency.

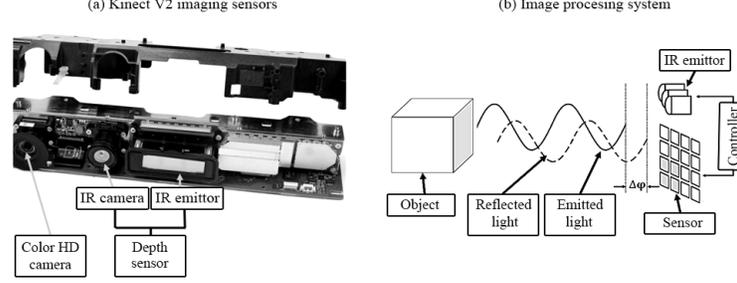


Fig. 1. Kinect v2 sensor system (a) components and (b) measurement technique (Foix and Aleny, 2011; Hansard et al., 2012).

In this study, two calibrated Kinect v2 sensors are used to capture the scene from two viewpoints simultaneously and automatically merge them. Once the views are aligned, a region of interest (ROI) is segmented. The ROI includes the body region surface involved in breathing from two angles of view, allowing a high coverage. The ROI surface is then reconstructed in order to calculate the volume at Frame  $t$ . Fig. 2 illustrates the process of the respiratory parameters calculation, starting from raw depth data acquisition and leading to the volume calculation at a given frame. Our system then calculates a volume time-curve from the calculated volumes in subsequent frames.  $V_t$  and  $RR$  are finally estimated from the volume-time curve.

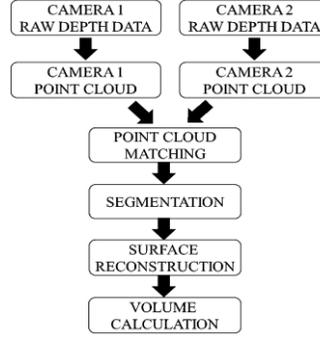


Fig. 2. Volume calculation workflow at frame  $t$ .

## 2.2. Automatic multi-view point-cloud matching

Point clouds are a set of points in the 3D space used to create a representation of a scanned physical object. Points in a point cloud are always situated on the external surfaces of the object. They are very useful for 3D modeling and remain the starting point in any 3D data processing application. A point cloud derives from raw data. Indeed, it is straightforwardly generated from depth data using the camera software development kit (SDK). In our approach, point clouds need to be available simultaneously from two different view angles to provide a high spatial coverage of the patient's torso. Accordingly, point clouds alignment in a same coordinate system is performed as a first step in our method. This is illustrated in Fig. 3, where a 2D marker is used to automatically align the point clouds. Our method assumes that the two cameras have a common view zone where the marker can be easily detected by both sensors. Each point cloud, covering a section of patient's torso, is thus aligned in the common coordinate system using the transformation matrix in Equation 2.

$$M = \begin{pmatrix} \cos \theta_y \cos \theta_z & \sin \theta_x \sin \theta_y \cos \theta_z + \cos \theta_x \sin \theta_z & -\cos \theta_x \sin \theta_y \cos \theta_z + \sin \theta_x \sin \theta_z & t_x \\ -\cos \theta_y \sin \theta_z & -\sin \theta_x \sin \theta_y \sin \theta_z + \cos \theta_x \cos \theta_z & \cos \theta_x \sin \theta_y \sin \theta_z + \sin \theta_x \cos \theta_z & t_y \\ \sin \theta_y & -\sin \theta_x \cos \theta_y & \cos \theta_x \cos \theta_y & t_z \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (2)$$

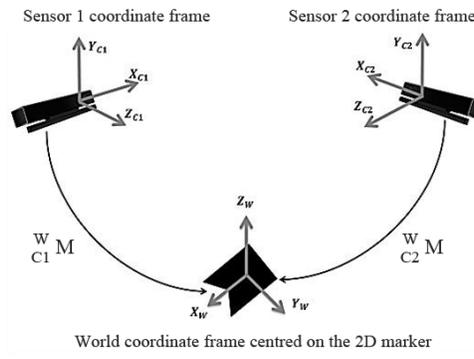


Fig. 3. Illustration of transformations between world coordinate frame  $(X_w, Y_w, Z_w)$  centred on the 2D marker and cameras coordinate frames  $(X_{c_i}, Y_{c_i}, Z_{c_i})$ ,  $i \in \{1,2\}$ .

In fact, each of the two cameras infers its relative position from the detected marker, which represents the world coordinate system. This presumes the estimation of two matrices  ${}^W_{c_1}M$  and  ${}^W_{c_2}M$  from the cameras coordinate systems to the world coordinate system. In Equation 2, the transformation matrix has six variables  $(\theta_x, \theta_y, \theta_z, t_x, t_y, t_z)$ . It can be expressed as combinations of three parameters coming from 3D translation  $(t_x, t_y, t_z)$  and three other parameters coming from 3D rotation  $(\theta_x, \theta_y, \theta_z)$ . By calculating the rotation  $R$  and the translation  $t$ , we can find the transformation matrix. To find the optimal transformation, we used the Procrustes analysis recognized for its effectiveness to resolve these types of problems (Goodall, 1991; Rohlf and Slice, 1990; Dryden et al., 2016). Procrustes analysis is the process of superimposing one collection of marker configurations on another by translating, scaling, and rotating them, so that the distances between corresponding points in each configuration are minimized. The Procrustes distance (Rohlf, 1999) is based on a least-square fit method and requires two aligned shapes with one-to-one point correspondence.

The process of superimposing a marker on another is divided into five main steps: marker detection, finding centroids, marker scaling, finding rotation and translation, and finally Procrustes distance computing. The first step uses only color data to detect the marker with a simple thresholding applied on the input images. The number of vertices of the detected area is compared to the number of vertices of the known shape to eliminate false results. If many images are detected, a subpixel precision processing technique is applied to refine the marker vertex locations. The second step uses the geometric model of the marker and computes its center of mass, so that the target marker can be placed over the reference configuration. In the third step, differences in size between configurations are removed by rescaling each configuration. Then, the differences in orientation is achieved by rotating one configuration (the target) around its centroid until it shows minimal offset in location of its landmarks relative to the other configuration (the reference). To transform a detected shape by the camera  $X_1 = (x_{11}, x_{21}, \dots, x_{n1})^T$  to an already known shape  $X_2 = (x_{12}, x_{22}, \dots, x_{n2})^T$ , we use Equation 3, where  $R$  is the rotation and  $t$  is the translation.

$$X_2 = R \times X_1 + t \quad (3)$$

To compute the Procrustes distance between the target and reference structures, we apply the Equation 4.

$$P_d^2 = \sum_{j=1}^n \left[ (x_{j1} - x_{j2})^2 + (y_{j1} - y_{j2})^2 \right] \quad (4)$$

These steps are repeated in order to minimize  $P_d^2$  and subsequently compute the optimal alignment. This is also illustrated in Fig. 4.

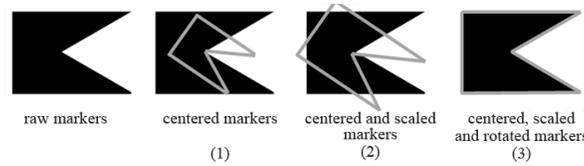


Fig. 4. Procrustes superimposition method - (1) Translation to a common origin, (2) scaling to unit centroid size, and finally (3) rotation to minimize the sum of squared distances between the corresponding shapes.

### 2.3. Region of interest extraction

The extraction is performed using Cloud Compare (Girardeau-Montaut, 2015) and Point Cloud Library (Rusu and Cousins, 2011). By including classical computer vision functions and algorithms, Cloud Compare allows 3D data processing and visualization. The contributor community is growing and expanding its applications in many research and industry fields. As such, Cloud Compare is continuously updated and becoming a standard tool in 3D data processing. Cloud Compare uses the Point Cloud Library as a third-party library to provide a set of additional computer vision algorithms, such as 3D data filtering, projections, feature estimation, etc... Point Cloud Library is a C++ library containing various algorithms to process all forms of point cloud data. This includes color data, depth data, point clouds, mesh data, noisy data and even reconstructed models. Point Cloud Library also includes numerous filters for data cleaning. These filters can process the data based on the position of the points in addition to other parameters. For example, some Point Cloud Library filters can be used to drop any points with an intensity value below a certain threshold. In our work, the 3D vision libraries are used for extracting the region of interest, as well as for cleaning the point cloud.

Once point cloud matching is performed, we first extract a rectangular cuboidal region of interest (ROI) including the thoraco-abdominal region using Cloud Compare. The clouds are selected at once and then aligned together. Our camera system is positioned in a manner to ensure the inclusion of the thoracic-abdominal area in the extracted region. It should be noted that we are not performing a precise segmentation of the thoracic-abdominal region by finding its boundaries. Instead, we perform a coarse segmentation by extracting a rectangular cuboid including the thoracic-abdominal region. Since our method for volume calculation is based on a subtraction technique, a precise segmentation of the ROI is not needed and only the moving volume due the chest contraction and expansion between subsequent frames is retained. The rest of volume is removed by the subtraction operation. Moreover, our coarse extraction technique allows a significant decrease of the computation time. The extracted 3D point cloud may contain noise that appears as clusters of neighboring points. We remove this noise using the Statistical outlier removal filter of the Point Cloud Library. This filter allows removing points that do not statistically fit with the rest of the data. The principle is to calculate the mean distance from each point to all its neighbors. The distribution is assumed to be Gaussian with a mean and standard deviation. Then, a threshold value is computed based on the mean and standard deviation of all distances. The filter finally keeps points whose mean distance is below the threshold value.

#### 2.4. Surface reconstruction

Because of the presence of holes and surface discontinuity, the point cloud information is not sufficient to calculate the volume. An intermediary mesh with closed gaps then needs to be generated. Using meshes simplifies surface reconstruction significantly. Thus, our surface reconstruction scheme follows three essential phases. Once the surface is scanned and the point cloud is calculated, a minimum spanning tree propagation technique (Pettie and Ramachandran, 2002) is applied in order to compute and orient normals. In our case, this technique allows to close the reconstructed surface. Its main principle consists in constructing a graph over the point cloud for all the vertexes through the k-nearest neighbors of each point. Then, the orientation of the vertex with the highest z value is calculated. Afterward, the correction of the direction of the entire vertex is conducted across the graph. Finally, the surface is reconstructed using Poisson surface reconstruction (Hoppe, 2008; Kazhdan et al., 2013; Berger et al., 2014), which takes as input a group of points with oriented normals and calculates a closed volume. By acting on a set of 3D points with oriented normals, the method solves for an approximate indicator function of the inferred solid, whose gradient best matches the input normals. The indicator function is zero everywhere except close to the surface. Note that all surfaces are closed by considering a reference plane at a well-defined distance from the subject's back and the lateral chest wall.

#### 2.5. Extraction of respiratory signal

The volume of the reconstructed surface is calculated using Cloud Compare (Girardeau-Montaut, 2015). Our method relies on the octree 3D structure representation. Based on a hierarchical tree structure, an octree partitions the 3-D space (Meagher, 1980). Starting from a root node in the form of single large cube, the octree is recursively subdivided into eight equal sized sub-cubes. This subdivision process continues until a predefined maximal depth is reached or if the regions are empty. The final volume is computed for each frame by multiplying the number of octrees by the unit size.

As a result, we compute a 1D signal where the frequency is the respiratory rate. On the other hand, the change in the signal amplitude is the key to estimate the tidal volume  $V_t$ . Note that the position of the reference plane is not important in  $V_t$  estimation, as even if the real volume of the thoracic-abdominal area is not accurate, this does not affect the accuracy of volume difference between frames with our subtraction method. The ROI volume is calculated at each frame to estimate a surrogate of patient's real volume-time curve. After detecting relevant peaks and minima of the curve, the tidal volume is deduced by subtracting volume values corresponding to consecutive extrema points. On the other hand, the respiratory rate is calculated from the volume-time curve by simply counting the number of peaks in a minute. In fact, each cycle has only one peak corresponding to the end of an inspiration.

To improve the accuracy of our method, we compute the average duration of a respiratory cycle ( $D$ ) using Equation 5:

$$D = \frac{\sum_{i=1}^{N_p-1} d_i}{N_p-1} \quad (5)$$

where  $N_p$  is the number of peaks of the volume-time curve in a minute and  $d_i$  is the temporal distance between peaks  $i$  and  $i + 1$ .

We then deduce the respiratory rate RR using Equation 6:

$$RR(\text{respirations/minute}) = \frac{60 \text{ seconds}}{D} \quad (6)$$

The tidal volume is the volume of air inhaled or exhaled from a person's lungs in a cycle. For more accuracy, the final tidal volume in a cycle is calculated as the average value of inspiratory and expiratory volumes. The tidal volume per minute is thus the average of all tidal volumes during a minute as shown in Equation 7:

$$V_t = \frac{\sum_{i=1}^{N_p} tv_i}{N_p} \quad (7)$$

where  $tv_i$  is the tidal volume of the cycle  $i$ .

## 2.6. Experiments

Our experiments have been conducted in an intensive care room at Sainte-Justine Hospital in Montreal. To simulate the breathing activity, we used a baby mannequin, designed according to neonatal anatomical and physiological characteristics and with an artificial test lung for infants (MAQUET Medical Systems, 1 Liter Test Lung 190). The lung is branched to a mechanical ventilator (Servo i, Maquet Inc, Sweden). The ventilator is a bedside machine used to push a volume of air into the lungs. The pushed volume is usually adjusted by caregivers according to the baby's weight and condition.

Two cameras can be disposed according the different schemes shown in Fig. 5. Considering the limited space in a PICU, the cameras can be placed on two of the four legs of the bed. Since the knowledge of lateral surface motion is important for a complete torso reconstruction, the mannequin's lateral sides should be covered by the field of view of the two cameras together. In Fig. 5, we illustrate all possible combinations. Only the four first configurations are admissible (a, b, c and d), as the other configurations do not allow coverage to both lateral sides. These first four positions were tested experimentally and only positions (c) and (d) were retained as shown in Fig. 5. In fact, the depth sensor is placed on the left side of the camera as illustrated in Fig. 1. On this basis, depth views are not symmetrical. In configuration (a), the camera placed at the right of the patient (camera 1) allows to have a good point cloud of the right lateral side of the torso whereas the left camera (camera 2) does not cover the left side of the torso due to the position of the infrared sensor. In configurations (c) and (d), both cameras allow good point clouds of both lateral sides. The sensors are finally positioned in the top right and the bottom left of the bed (configuration (c)), both in the direction of 45° and at a distance of 1m to the crib mattress. This positioning offers a high spatial coverage since we cover the top and lateral sides of the baby.

For system calibration, the 2D marker is placed on the bed in such a way to be in a common field of view of the two cameras. The cameras infer their relative positions from the detected marker. We then remove the marker and place the baby mannequin in the bed.

In order to evaluate the performance of our method, the ventilator is used as gold standard. In PICE and for health professional decision-makers, the ventilator is considered as the most reliable method to provide accurate and precise quantitative measures for RR and Vt. Thus, ventilator measures are recorded in parallel to our experiments and are considered as ground-truth data.

We simulate spontaneous breathing of a patient with different volumes. Note that our mannequin lung supports volumes from 10 mL to 1L. Therefore, we used the same mannequin to test different volumes for all ages. Two primary modes were used to push the air into the artificial lungs: the neonatal and the adult mode. The air volumes for neonatal mode are respectively: 10 ml, 20 ml, 30 ml, 40 ml, 50 ml and 100 ml. For adult mode, the volumes are respectively: 150 ml, 200 ml, 250 ml, 300 ml, 350 ml, 400 ml, 450 ml and 500 ml. Vt and RR are computed with our method. The results are then compared with the ventilator reference values. The complete settings for each experiment are listed in Table 3.

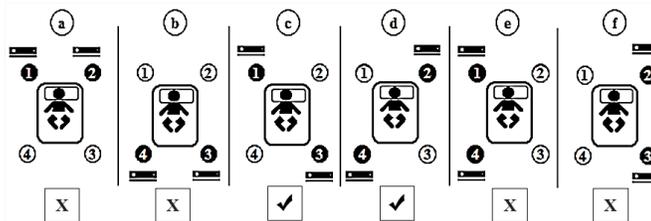


Fig. 5. Cameras placement: possible combinations. The cameras are placed in the (a) bed top, (b) bed bottom, (c) top right and bottom left, (d) top left and bottom right, (e) bed right side, (f) bed left side. Positions c and d can be accepted. Positions a, b, e and f are rejected.

**Table 3**

Reference ventilator settings for each experiment, with Vt: tidal volume, RR: respiratory rate and IT: Inspiration Time.

	EXPERIMENT NUMBER	Vt (ml)	RR (resp./min)	IT (seconds)
<i>Adult mode</i>	1	500	20	0.9
	2	450	20	0.9
	3	400	20	0.9
	4	350	20	0.9
	5	300	20	0.9
	6	250	20	0.9
	7	200	30	0.9
	8	150	30	0.9
	9	100	30	0.9
<i>Pediatric mode</i>	10	50	40	0.5
	11	40	40	0.5
	12	30	40	0.5
	13	20	40	0.5
	14	10	50	0.5

To verify the applicability of the proposed method on a real patient, a second test was conducted by measuring the breathing pattern of a mechanically ventilated infant. This test involved a 4 months and 20 days old female, weighed 6.6 kg. The patient was sleeping and requiring the ventilator support for breathing. The test was performed in a PICU room of Sainte-Justine Hospital, one of the largest pediatric health centers in Canada. This experiment was conducted with approval from the Research Ethics Board (REB) of the hospital. Fig. 6 shows the PICU room setup, where Kinect sensors were placed to

accommodate the patient and the already existing medical equipment. In case of emergency, the cameras can be easily and quickly detached from the bed allowing the urgent transport of the patient. This configuration was checked and validated by the equipment inspection team of the Hospital.

Note that the breathing activity can be controlled totally or partially by the mechanical ventilator. For example, the ventilator performed the entire breathing activity in the first test with the mannequin. In the second test (with a real patient), the ventilator is doing the preponderance of the breathing work, while the patient is partially contributing in the respiration. The ventilator settings are set to  $Vt = 40$  mL and  $RR = 20$  respirations/minute. The final  $Vt$  and  $RR$  values displayed by the ventilator are not only controlled by the ventilator, but also by the patient's breathing effort.



Fig. 6. System setup and Kinect sensors locations during the experiment on a critically ill patient in the PICU of Sainte-Justine Hospital.

## 2.7. Cloud-to-cloud distance

The common Euclidean metric  $\mathcal{L}_2$  is adopted to calculate distance between clouds. We consider  $S_1$  and  $S_2$  the external surfaces respectively in the initial and final state (before and after being inflated with air), as indicated in Fig. 7. Point clouds of the surface  $S_2$  are regarded as “target” points  $q = (q_x, q_y, q_z)$ , whereas the point clouds of the surface  $S_1$  are considered as points  $p = (p_x, p_y, p_z)$  in the “initial position”. The distance between  $p$  and  $q$  is calculated using  $\mathcal{L}_2$  in the space  $\mathbb{R}^3$ . The aim is to find corresponding 3D points before and after surface displacement from  $S_1$  to  $S_2$ . Consider that we have  $M$  source points cloud  $p_i$  on the surface  $S_1$ . Points  $p_i, i \in \{1..M\}$  from  $S_1$  are projected on  $S_2$  using the normal vector at each source point. The projected points are noted  $q'_i, i \in \{1..M\}$ . To find a corresponding destination point in  $S_2$ , the nearest neighbor is selected in  $q_i, i \in \{1..M\}$ . Then, the displacement distance is computed for each pair in the cloud using Equation 8, where  $p$  represents the “initial” point in  $S_1$  surface and  $q$  is the “target” point in  $S_2$  surface.

$$\|p - q\| = \sqrt{(p_x - q_x)^2 + (p_y - q_y)^2 + (p_z - q_z)^2} \quad (8)$$

The maximum displacement is selected for each cloud. For each experiment, these steps are repeated over each pair of  $N_p$  clouds. To compute the maximum displacement  $\Delta d$  in each experiment, we used Equation 9, where the maximum displacement is first calculated over one pair of point cloud and then calculated over the  $N_p$  clouds of each experiment.

$$\Delta d = \max_{j \in [1, N_p]} \left( \max_{i \in [1, M]} \left( \|p_i^j - q_i^j\| \right) \right) \quad (9)$$

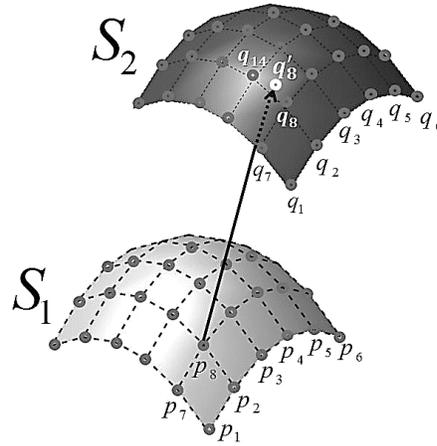


Fig. 7. Finding corresponding “pairs” of 3D points between surfaces before and after respiratory displacement of the test lung surface. The surface  $S_1$  represents the minimum displacement and the surface  $S_2$  represents the maximum displacement after being inflated with air volume corresponding to the experiment settings.  $q'_8$  is the projection of  $p_8$  on  $S_2$  and  $q_8$  is the closest neighbor to  $q'_8$ . Therefore, the corresponding point of  $p_8$  is  $q_8$ .

In Fig. 7, the source point  $p_8$  on the surface  $S_1$  (before displacement) is projected on the surface  $S_2$  (after displacement) using the normal vector in  $p_8$ . As can be seen, the nearest neighbors of the projected point  $q'_8$  are  $q_8$  and  $q_{14}$ . Since  $q_8$  is closest to  $q'_8$  than  $q_{14}$  as  $\|q'_8 - q_{14}\| > \|q'_8 - q_8\|$ , it will be selected as the corresponding point of  $p_8$ . Finally, the depth displacement distance is computed for the pair  $(p_8, q_8)$  by calculating  $\|p_8 - q_8\|$ .

The maximum displacement  $\Delta d$  is computed for different combinations of ventilator Vt settings. Point clouds corresponding to experiments listed in Table 3 are used in order to determine  $\Delta d$ .

## 2.8. Accuracy and precision evaluation

In this section, we analyze the accuracy and precision of our method in estimating Vt and RR. The quantitative values resulting from our 3D system are compared to the reference values measured by the ventilator ("ground truth" data) using the root mean square deviation (RMSD), the relative error (RE) and the relative standard deviation (RSD) metrics. The root mean square deviation is used to evaluate accuracy. RMSD is calculated as:

$$RMSD = \sqrt{\frac{\sum_{i=1}^n (x_i - \hat{x}_i)^2}{n}} \quad (10)$$

where  $x$  is the reference value (measured by the ventilator),  $\hat{x}_i$  is the estimated value in cycle  $I$  (measured by our 3D system) and  $n$  is the number of experiments.

The relative error has also been used to determine the accuracy of the quantitative values measured by our 3D system. To calculate RE, we use Equation 11:

$$RE = 100 \times \left| \frac{\hat{x} - \bar{x}^T}{\bar{x}^T} \right| \quad (11)$$

where  $\bar{x}^T$  is the mean of reference values and  $\hat{x}$  is the mean of estimated values over a number of observations.

Precision is studied through the variation of Vt values for one minute. The amount of measure repetition is determined by the number of cycles in one minute. To evaluate precision, we used the relative standard deviation (RSD) calculated as:

$$RSD = 100 \times \frac{SD}{|\bar{x}|} \quad (12)$$

where  $\bar{x}$  is the sample mean value and SD is its standard deviation.

## 3. Results and Discussion

The results obtained from the setup in Table 3 are illustrated in Fig. 8 and Fig. 9. Fig. 8 shows the results of the first nine experiments corresponding to adult mode, whereas Fig. 9 shows the results of five experiments corresponding to neonatal mode. These curves present a good regularity, as the calculated signal is periodic.

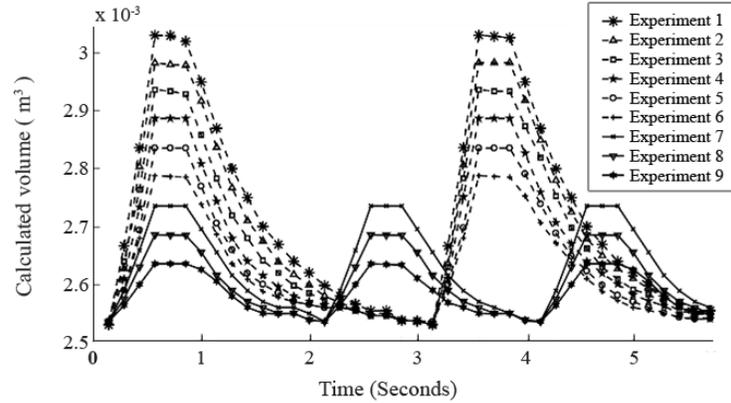


Fig. 8. Volume-time curves resulting from our 3D imaging system during volume-controlled ventilation in adult mode. These volumes were introduced by the mechanical ventilator into the artificial test lung to evaluate our method accuracy with big volumes. The displayed curves illustrate volume changes over time. The corresponding ventilator settings are listed in Table 3 from experiment 1 to experiment 9. For each of the experiments, it can be seen that volume values present high and low peaks. The highest peaks correspond to the end of air inhalation, while the lowest peaks correspond to the end of air exhalation. After reaching the highest peaks, the signal values slightly decrease over a short period of time. This slight decrease corresponds to relaxation time separating the inspiration and the expiration. The relaxation time takes approximately 0.4 seconds.

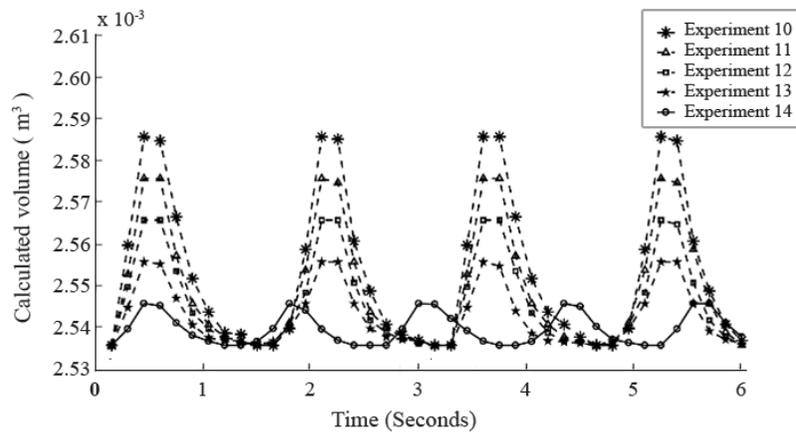


Fig. 9. Volume-time curves resulting from our 3D imaging system during volume-controlled ventilation in neonatal mode. These volumes were introduced by the mechanical ventilator into the artificial test lung to evaluate our method accuracy with small volumes. The displayed curves illustrate volume changes over time. The corresponding ventilator settings are listed in Table 3 from experiment 10 to experiment 14.

In volume-controlled ventilation, the air volumes are delivered with the same quantity at regular intervals by the ventilator. Curves for experiments 1 to 6 have equidistant peaks, with a time distance of about 3 seconds. Similarly, curves corresponding to experiments 7 to 9 and experiments 10 to 13 have also equidistant peaks, (2.14 seconds and 1.42 seconds respectively). This is expected, as the mentioned experiments present the same RR ventilator settings. Since RR values are increasing from experiments 1 to 14, the period is expected to decrease. For a clearer illustration, we show the calculated first five cycles in experiment 1 in Fig. 10. The resulting signal from our 3D imaging system is periodic and regular, and the waveform of one period seems to be the same between the 5 cycles. Moreover, peaks are shown very clearly. An approximation of the respiratory rate can be visually calculated by a simple peak analysis technique. A more precise method to calculate the respiratory rate is to involve the distance between consecutive peaks. This method was detailed in section 3.5. Fig. 11 illustrates the second cycle of Fig. 10 (zoomed in), to show how tidal volume is calculated by our system for a single respiration.

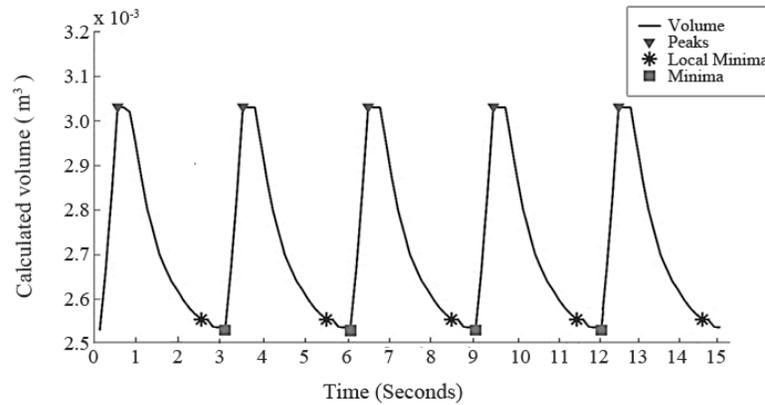


Fig. 10. Volume-time waveform resulting from our 3D imaging system for only the first experiment (ventilator settings:  $V_t = 500$  ml, RR = 20 respirations/minute and IT = 0.9 seconds). The figure illustrates 5 cycles from which breathing characteristics are exhibited over the 105 frames (15 seconds). These characteristics uniquely identify the breathing patterns delivered by the mechanical ventilator. Peaks positions are shown with triangles markers. They represent the end of inspiratory rise time. An asterisk marker represents the end of expiration, while a rectangle represents the start of inspiration.

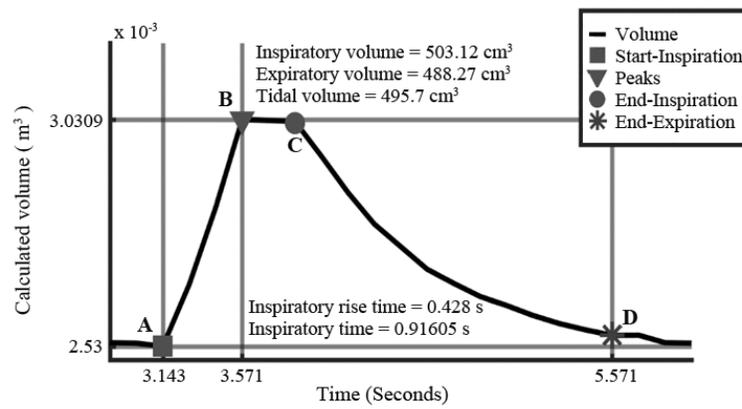


Fig. 11. Analysis of only one complete cycle obtained by our 3D imaging system. This corresponds to cycle 2 of experiment 1 extracted and zoomed from Fig. 10. The reference points A, B, C and D allow calculating inspiratory rise time (A–B), inspiratory time (A–C), and expiratory time (C–D). Point A represents the start of inspiration, point B is a peak position, point C represents the end of inspiration and the start of expiration and finally, point D is the end of expiration. For this cycle, the inspiratory rise time is 0.4286 seconds and the inspiratory time is 0.91605 seconds.

**Table 4**

Comparison between the values provided by the ventilator (gold standard) and those computed by our imaging system over one minute. *Vt*: Tidal volume, *RR*: respiratory rate, *RE*: Relative error and *RSD*: Relative standard deviation. For our method column, we present the mean value of all cycles over one minute. This value is constant in 'ventilator' columns as the ventilator delivers a constant *Vt* with a constant *RR* in each experiment.

EXPERIMENT NUMBER		Vt (ml)				RR (resp./min)			
		Ventilator	Our method	RE (%)	RSD (%)	Ventilator	Our method	RE (%)	RSD (%)
<i>Adult mode</i>	1	500	498,64	0,27	0,79	20	19,1973	4,01	5,68
	2	450	453,36	0,75	0,289	20	19,8329	0,83	4,39
	3	400	401,47	0,37	0,71	20	18,9734	5,13	5,43
	4	350	348,53	0,42	0,71	20	19,3495	3,25	1,76
	5	300	298,91	0,36	0,89	20	20,4946	2,47	4,59
	6	250	250,09	0,04	1,34	20	19,9994	0,003	1,25
	7	200	200,37	0,18	0,96	30	29,3494	2,16	3,305
	8	150	149,63	0,24	2,17	30	29,2494	2,5	3,285
	9	100	101,37	1,37	3,14	30	29,4945	1,68	3,46
<i>Pediatric mode</i>	10	50	48,63	2,74	6,09	40	39,9548	0,11	2,35
	11	40	44,37	10,92	7,32	40	39,3861	1,53	2,56
	12	30	27,63	7,9	9,05	40	39,3804	1,55	2,23
	13	20	21,05	5,25	10,40	40	39,6108	0,97	1,69
	14	10	7,95	20,5	36,23	50	49,4554	1,09	1,60

The validation on both neonate and adult's volumes is reported in Tables 4 through 8. Experiments 1 to 9 report the validation on adult's volumes, while experiments 10 to 14 report the validation on neonate's volumes. The mean values between cycles over one minute are shown in Table 4. The resulting RMSD between measured and reference values shows an error of  $\pm 1.94$  ml and  $\pm 0.59$  respirations per minute for respectively tidal volume and respiratory rate. These small RMSD value indicate that the estimation of our method is very close to the reference values given by the ventilator. Furthermore, the differences of *Vt* mean values given by the ventilator and our method are in the scale of a few milliliters while maximum difference between the two methods does not exceed 4.37 ml (in experiment 11). On the other hand, the differences of *RR* mean values given by the 2 methods are in the order of one respiration per minute. Table 4 also shows the relative error values between experiments for *Vt* and *RR*. *RE* values for *Vt* are small for large volumes and larger for small volumes. For example, the *RE* value for *Vt* in experiment 1 is 0,27, while in experiment 14, the *RE* value of *Vt* is 20,5. It can be concluded that our method is more accurate for large volumes. However, our system achieves approximately the same accuracy for small and big respiratory rate values.

We also analyzed the *RSD* values in Table 4 in order to study the precision of our system. Note that that each respiratory cycle in our volume-time curve is considered here as an observation. We can see again that *RSD* values for *Vt* are small for big volumes and larger for small volumes (e.g. *RSD* is 0,79 in experiment 1 and 36,23 in experiment 14). Our method is consequently more precise for large volumes. On another level, our method presents approximately the same precision for small and big *RR* values.

In order to further assess the precision of our 3D imaging system, we endorse the obtained results based on repetitive testing. Each experiment is replicated 5 times. Mean values, relative error and relative standard deviation are calculated for each set of observations. Tidal volume observations are shown in Table 5. This table shows the results of each repetition, as well as the average result over all the repetitions. Table 6 shows the calculated respiratory rates for each experiment, as well as the average result. We notice a small RMSD between measured and reference values for both *Vt* ( $\pm 8,9432$  ml) and *RR* ( $\pm 1.3654$  respirations per minute), respectively calculated from Table 5 and Table 6. These results mean that the estimated values by our method are very close to the reference values given by the ventilator. Additionally, the tidal volume RMSD is in the scale of few milliliters, while the respiratory rate RMSD is always in the order of one respiration per minute.

Table 7 shows the relative error and the relative standard deviation of calculated tidal volumes. RE and RSD values are better for larger volumes than smaller volumes. In experiment 14, the RE and RSD values of  $V_t$  are respectively 46,54% and 40,52%. Our method presents almost better accuracy and precision for larger volumes.

Table 8 summarizes RE and RSD results for respiratory rate. In total, we have 30 observations for  $RR = 20$ , 15 observations for  $RR = 30$ , 20 observations for  $RR = 40$  and 5 observations for  $RR = 50$ . To sum up, the proposed system is able to ensure approximately the same accuracy and precision for both high and low respiratory rates.

To explain the performance difference between small and large volumes in Table 7, the camera depth accuracy is compared to the distance between the test lung's initial external surface (when it is empty) and final external surfaces (after being filled with air). As shown in Fig. 12,  $\Delta d$  variation is very small and almost constant when the volumes are below 50 mL. However,  $\Delta d$  variations are linear and proportional to  $V_t$  for volumes between 100 mL and 500 mL. Let  $p$  be the camera's depth accuracy ( $p = 2$  mm for the Kinect v2 sensor). The gray area in Fig. 12 presents a depth displacement inferior to depth resolution ( $\Delta d < p$ ). This means that the hardware configuration limit is reached in this region, causing relatively inaccurate depth values. Seen from this perspective, the obtained results for neonates' volumes are quite reasonable.

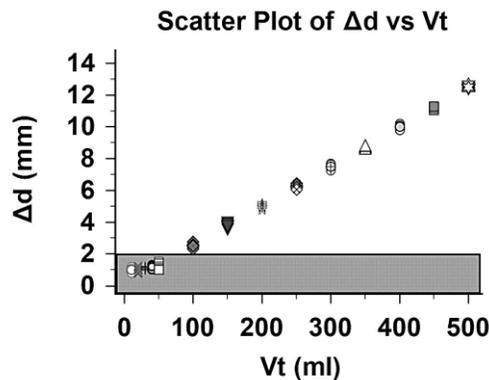


Fig. 12. Scatter plot of  $\Delta d$  versus  $V_t$  and depth accuracy limit for Kinect v2 hardware. The gray zone presents the area where the camera cannot generate depth values accurately. Thus, this region presents the hardware limit.

**Table 5**

$V_t$  Reference ventilator values and our method mean  $V_t$  value estimation over five observations for each experiment.

	EXPERIMENT NUMBER	Ventilator	#1	#2	#3	#4	#5	Average $V_t$ (ml)
<i>Adult mode</i>	1	500	498,64	485,88	502,5	494,33	502	496,67
	2	450	453,36	447,23	448,5	450	446,66	449,15
	3	400	401,47	398,66	397	395,33	395,66	397,63
	4	350	348,53	347,60	345,66	348,48	349,66	347,99
	5	300	298,91	294,43	298,43	295,96	296,43	296,83
	6	250	250,09	251,89	251	249,5	248,28	250,15
	7	200	200,37	199	199,66	199,71	199,5	199,65
	8	150	149,63	145	149	148	140	146,32
	9	100	101,37	98,75	99,5	98,33	99,44	99,48
<i>Pediatric mode</i>	10	50	48,63	44,55	49,03	45,73	46,26	46,84
	11	40	44,37	35,77	36,65	32,87	34,95	36,92
	12	30	27,63	22,57	24,17	25,13	26,08	25,12
	13	20	21,05	15,44	14,28	15	15,28	16,21
	14	10	7,95	4,96	2,36	4,57	6,88	5,34

**Table 6**

$RR$  Reference ventilator values and our method mean  $RR$  value estimation over a number of observations for each experiment. This number is different for each set of experiments (30 observations for  $RR = 20$ , 15 observations for  $RR = 30$ , 20 observations for  $RR = 40$  and 5 observations for  $RR = 50$ ).

	EXPERIMENT NUMBER	Ventilator	#1	#2	#3	#4	#5	Average $RR$ (resp./min)
<i>Adult mode</i>	1	20	19,19	19,33	19,76	19,5	19,33	19,5617
	2	20	19,83	19,71	19,23	19,60	19,87	
	3	20	18,97	19,70	19,79	19,77	19,08	
	4	20	19,35	19,79	19,60	19,71	19,66	
	5	20	20,49	19,60	19,5	19,61	19,33	
	6	20	19,99	19,33	19,78	18,86	19,5	
<i>Pediatric mode</i>	7	30	29,35	29,60	29,05	29,60	29,33	29,4690
	8	30	29,25	29,66	29,03	29,79	29,71	
	9	30	29,49	29,5	29,35	29,75	29,60	
	10	40	39,95	39,66	39,77	38,18	39,33	
	11	40	39,38	39	39,5	39,16	38,87	
	12	40	39,38	39,71	39,66	39,77	38,87	
	13	40	39,61	39,71	39,60	39,5	39,24	
	14	50	49,45	49,79	49,66	49,34	49,17	

**Table 7**

Vt accuracy and precision results of our method. RE is the relative error between Vt observations mean value and Vt ventilator value. RSD is the relative standard deviation of Vt observations mean value. For each experiment, five observations are made to calculate RE and RSD.

	EXPERIMENT NUMBER	Average value	RE (%)	RSD (%)
<i>Adult mode</i>	1	496,67	0,66	1,38
	2	449,15	0,19	0,60
	3	397,63	0,59	0,63
	4	347,99	0,57	0,43
	5	296,83	1,06	0,62
	6	250,15	0,06	0,55
	7	199,65	0,17	0,25
	8	146,32	2,45	2,70
	9	99,48	0,52	1,17
<i>Pediatric mode</i>	10	46,84	6,32	4,10
	11	36,92	07,69	11,89
	12	25,12	16,26	7,60
	13	16,21	18,93	16,90
	14	5,3462	46,54	40,52

**Table 8**

RR accuracy and precision results of our method. RE is the relative error between RR observations mean value and RR ventilator value. RSD is the relative standard deviation of RR observations mean value. Multiple observations were used to calculate RE and RSD and the number of observations is different for each set of experiments (30 observations for RR = 20, 15 observations for RR = 30, 20 observations for RR = 40 and 5 observations for RR = 50).

	EXPERIMENT NUMBER	Average value	RE (%)	RSD (%)
<i>Adult mode</i>	1			
	2			
	3	19,5617	2,19	1,68
	4			
	5			
	6			
	7			
	8	29,4690	1,77	0,84
	9			
<i>Pediatric mode</i>	10			
	11	39,3945	1,51	1,06
	12			
	13			
	14	49,5043	0,99	0,49

The results of the second test conducted on a mechanically ventilated patient are presented in Fig. 13, Table 9 and Table 10. Fig. 13 shows both reference ventilator signal and the reconstructed signal. Our method results are well correlated with the ventilator measurements ( $\rho = 0.97$ ). Table 9 shows the calculated tidal volumes and respiratory rates average for each experiment. We notice that RR and Vt values are changing between the different observations. This is due to the patient's respiratory effort, which is varying with time. To assess the precision of our system, we calculate the relative error and the relative standard deviation, as shown in Table 10. The RE and RSD values of RR are respectively 3,2516% and 9,8727%, which shows good accuracy and precision in estimating the respiratory rate. The RE and RSD values of Vt are respectively 9,169% and 12,3274%. When comparing with experiments conducted on our mannequin, the nearest experiment is experiment 11 (see Tables 5 and 7). In fact, the average of the real patient volume is 38,9398 mL (ventilator), which is near to the tested Vt in experiment 11 (40 mL). The RE and RSD values are respectively 7.69% and 11.89% for the mannequin against 9,169% and 12,3274% for the real patient. Although results on the mannequin are slightly better, we can conclude that our method presents a high accuracy and a good precision in estimating the tidal volume of a real patient. Moreover, this experiment confirms the applicability of our method in the PICU environment, where space arrangement including the configuration of cables and cameras was optimized to facilitate the system deployment.

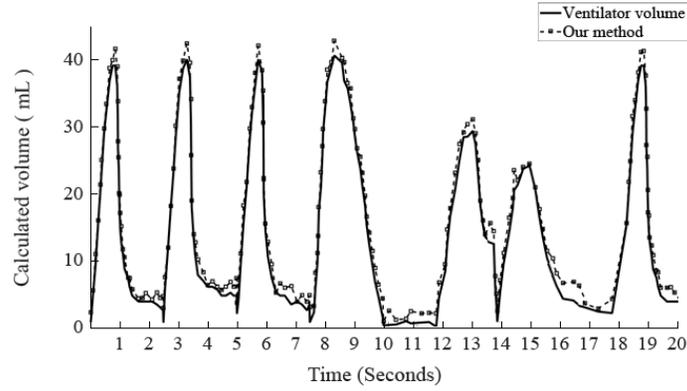


Fig. 13. Results of the second test conducted on the mechanically ventilated patient. The y-scale of our method results (dotted line, squares) was adjusted to the ventilator set of data (solid line) in order to make results fit into one graph and to study the applicability of our proposed method on a real patient.

**Table 9**

Results of the second test conducted on a mechanically ventilated patient. Mean ventilator values are compared to our method mean values with five one-minute observations for  $V_t$  and  $RR$ . The ventilator input parameters set by the nurse are:  $V_t = 40$  ml,  $RR = 20$  resp./minute. The final  $V_t$  and  $RR$  values displayed by the ventilator are not only controlled by the ventilator, but also by the patient's breathing effort.

	#1		#2		#3		#4		#5	
	Ventilator	Our method								
$V_t$	38,193	34.596	42,498	41.793	36,149	31.482	34,113	30.346	43,746	39.015
$RR$	20,5	19.715	25	23,971	23,5	23,346	25,5	24,489	28	26,966

**Table 10**

Accuracy and precision results for the test conducted on a mechanically ventilated patient. Mean ventilator values are compared to our method mean values for  $V_t$  and  $RR$ . RE is the relative error between our method and the ventilator values, while RSD is the relative standard deviation of the observations.

	Ventilator average	Our method average	RE (%)	RSD (%)
$V_t$	38,9398	35,4464	9,169	12,3274
$RR$	24,5	23,697	3,2516	9,8727

## 4. Conclusions

We proposed a 3D imaging system for respiratory monitoring in pediatric intensive care environment. The system uses two depth cameras providing a high spatial coverage of body regions involved in the respiration, notably the top of the torso surface and its lateral sides. Our system first aligns point clouds in a common coordinate system. Then, it performs a Poisson surface reconstruction of a segmented region of interest. The volume is computed in each frame using an octree subdivision technique of the 3D space. This allows the construction of a volume-time curve as a 1D signal that we use to compute respiratory parameters.

Our system is a very promising support tool intended to assist caregivers to monitor respiration in PICU environment. The major innovation of our monitoring system is the ability to provide a quantitative measure of tidal volume, and respiratory rate for spontaneous breathing patients in a PICU room. Note that these parameters are not assessed by medical devices during spontaneous breathing, as they can only be indicated by the ventilator when the patient is mechanically ventilated. The proposed system has been evaluated experimentally. The quantitative measures were compared to the ground truth data of the mechanical ventilator. The system was able to accurately and precisely track both tidal volume changes and respiratory rate.

Our future work will be articulated around two directions. Firstly, we plan to develop a GPU-based platform for real-time acquisition, processing and data analysis, for a more practical level. Secondly, we will extend the system ability by characterizing the retraction signs of the breast, an important parameter in PICU, whose actual assessment is based on the clinician's observation. We thus aim to develop an objective method for monitoring retraction signs of the breast.

## Acknowledgements

This work was supported by grants from the Natural Science and Engineering Research Council of Canada (NSERC) and Medteq.

## References

- Al-Khalidi, F. Q., Saatchi, R., Burke, D., Elphick, H., & Tan, S., 2011. Respiration rate monitoring methods: A review. *Pediatric pulmonology*, 46(6), 523-529.
- Al-Naji, A., Gibson, K., Lee, S.-H., Chahl, J., 2017. Real Time Apnoea Monitoring of Children Using the Microsoft Kinect Sensor: A Pilot Study. *Sensors (Basel)* 17 (2), 286.
- Aoki, H., Nakamura, H., Fumoto, K., Nakahara, K., Teraoka, M., 2015. Basic study on non-contact respiration measurement during exercise tolerance test by using kinect sensor. In: *System Integration (SII), 2015 IEEE/SICE International Symposium on*. IEEE, pp. 217–222.
- Bai, Y., Li, W., Chen, Y., 2010. Design and implementation of an embedded monitor system for detection of a patient's breath by double webcams. In: *Medical Measurements and Applications Proceedings (MeMeA), 2010 IEEE International Workshop on*. IEEE, pp. 171–176.
- Benetazzo, F., Freddi, A., Monteriù, A., Longhi, S., 2014. Respiratory rate detection algorithm based on RGB-D camera: Theoretical background and experimental results. *Healthc. Technol. Lett.* 1 (3), 81–86.
- Berger, M., Tagliasacchi, A., Seversky, L., Alliez, P., Levine, J., Sharf, A., Silva, C., 2014. State of the art in surface reconstruction from point clouds. In: *EUROGRAPHICS star reports*. Vol. 1. pp. 161–185.
- Beserra Gomes, R., Ferreira da Silva, B. M., Rocha, L. K., Aroca, R. V., Velho, L. C. P. R., Gonçalves, L. M. G., 2013. Efficient 3D object recognition using foveated point clouds. *Comput. Graph.* 37 (5), 496–508.
- Bouachir, W., Gouiaa, R., Li, B., & Noumeir, R., 2018. Intelligent video surveillance for real-time detection of suicide attempts. *Pattern Recognition Letters*, 110, 1-7.
- Corbishley, P., Rodríguez-Villegas, E., 2008. Breathing detection: Towards a miniaturized, wearable, battery-operated monitoring system. *IEEE Trans. Biomed. Eng.*, 55(1), 196-204.
- Dryden, I. L., Mardia, K. V., 2016. *Statistical shape analysis: With applications in r*, 125–173.
- Dullenkopf, A., Bernardo, S. D., Berger, F., Fasnacht, M., Gerber, A. C., & Weiss, M., 2003. Evaluation of a new combined SpO<sub>2</sub>/PtcCO<sub>2</sub> sensor in anaesthetized paediatric patients. *Pediatric Anesthesia*, 13(9), 777-784.
- Essouri, S., Baudin, F., Chevret, L., Vincent, M., Emeriaud, G., Jouvot, P., 2017. Variability of Care in Infants with Severe Bronchiolitis: Less-Invasive Respiratory Management Leads to Similar Outcomes. *The Journal of Pediatrics* 188, 156–162.e1.
- Fankhauser, P., Bloesch, M., Rodriguez, D., Kaestner, R., Hutter, M., Siegwart, R., 2015. Kinect v2 for mobile robot navigation: Evaluation and modeling. In: *Advanced Robotics (ICAR), 2015 International Conference on*. IEEE, pp. 388–394.
- Foix, S., Alenya, G., & Torras, C., 2011. Lock-in time-of-flight (ToF) cameras: A survey. *IEEE Sensors Journal*, 11(9), 1917-1926.
- Girardeau-Montaut, D., 2015. *Cloud Compare—3D Point Cloud and Mesh Processing Software*. Open Source Project.
- Gonzalez-Jorge, H., Rodríguez-González, P., Martínez-Sánchez, J., González-Aguilera, D., Arias, P., Gesto, M., & Díaz-Vilariño, L., 2015. Metrological comparison between Kinect I and Kinect II sensors. *Meas. J. Int. Meas. Confed.* 70, 21–26.
- González-Ortega, D., Díaz-Pernas, F. J., Martínez-Zarzuela, M., & Antón-Rodríguez, M., 2014. A Kinect-based system for cognitive rehabilitation exercises monitoring. *Computer Methods and Programs in Biomedicine* 113 (2), 620–631.
- Goodall, C., 1991. Procrustes methods in the statistical analysis of shape. *Journal of the Royal Statistical Society. Series B (Methodological)*, 285–339.
- Hansard, M., Lee, S., Choi, O., & Horaud, R. P., 2012. *Time-of-flight cameras: principles, methods and applications*. Springer Science & Business Media.
- Harte, J. M., Golby, C. K., Acosta, J., Nash, E. F., Kiraci, E., Williams, M. A., Naidu, B., et al., 2016. Chest wall motion analysis in healthy volunteers and adults with cystic fibrosis using a novel Kinect-based motion tracking system. *Medical & Biological Engineering & Computing* 54 (11), 1631–1640.
- Hoppe, H., 2008. Poisson surface reconstruction and its applications. In: *Proceedings of the 2008 ACM symposium on Solid and physical modeling*. ACM, pp. 10–10.
- Kan, Y. C., & Chen, C. K., 2012. A wearable inertial sensor node for body motion analysis. *IEEE Sensors Journal*, 12(3), 651-657.
- Katashev, A., Romberg, K., Danielsson, A., Saraste, H., 2015. Application of 3d scanner for estimation of chest movement in scoliotic patients. In: *16th Nordic-Baltic Conference on Biomedical Engineering*. Springer, pp. 63–66.
- Kazhdan, M., Hoppe, H., 2013. Screened poisson surface reconstruction. *ACM Transactions on Graphics (TOG)* 32 (3), 29.
- Keselman, L., Woodfill, J. I., Grunnet-Jepsen, A., Bhowmik, A., 2017. Intel realsense stereoscopic depth cameras. *arXiv preprint arXiv:1705.05548*.
- Khoshelham, K., Elberink, S. O., 2012. Accuracy and resolution of Kinect depth data for indoor mapping applications. *Sensors (Basel)* 12 (2), 1437–1454.

- Lachat, E., Macher, H., Landes, T., Grussenmeyer, P., 2015. Assessment and calibration of a RGB-D camera (Kinect v2 Sensor) towards a potential use for close-range 3D modeling. *Remote Sensing* 7 (10), 13070–13097.
- Lee, Y. S., Pathirana, P. N., Evans, R. J., & Steinfurt, C. L., 2015. Noncontact detection and analysis of respiratory function using microwave Doppler radar. *Journal of Sensors*.
- Li, B. Y. L., Xue, M., Mian, A., Liu, W., Krishna, A., 2016. Robust RGB-D face recognition using Kinect sensor. *Neurocomputing* 214, 93–108.
- Lun, R., Zhao, W., 2015. A survey of applications and human motion recognition with microsoft kinect. *International Journal of Pattern Recognition and Artificial Intelligence* 29 (05), 1555008.
- Mallick, T., Das, P. P., Majumdar, A. K., 2014. Characterizations of noise in Kinect depth images: A review. *IEEE Sensors Journal* 14 (6), 1731–1740.
- Meagher, D. J. R., 1980. Octree encoding: A new technique for the representation, manipulation and display of arbitrary 3-d objects by computer. *Electrical and Systems Engineering Department Rensselaer Polytechnic Institute Image Processing Laboratory*.
- Nakajim, K., Matsumoto, Y., Tamura, T., Nakajim, K., Matsumoto, Y., Tamura, T., 2001. Development of real-time image sequence analysis for evaluating posture change and respiratory rate of a subject in bed. *Physiological Measurement* 22 (3), N21–N28.
- Nakajima, K., Osa, A., Miike, H., 1997. A method for measuring respiration and physical activity in bed by optical flow analysis. In: *Engineering in Medicine and Biology Society, 1997. Proceedings of the 19th Annual International Conference of the IEEE*. Vol. 5. IEEE, pp. 2054–2057.
- Oikonomidis, I., Kyriazis, N., Argyros, A., 2011. Efficient model-based 3d tracking of hand articulations using kinect. In: *BmVC*. Vol. 1. p. 3.
- Ostadabbas, S., Bulach, C., Ku, D., Anderson, L., Ghovanloo, M., 2014. A passive quantitative measurement of airway resistance using depth data. In: *Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE*. IEEE, pp. 5743–5747.
- Pettie, S., Ramachandran, V., 2002. An optimal minimum spanning tree algorithm. *Journal of the Association for Computing Machinery* 49 (1), 16–34.
- Pöhlmann, S. T. L., Harkness, E. F., Taylor, C. J., Astley, S. M., 2016. Evaluation of Kinect 3D Sensor for Healthcare Imaging. *Journal of Medical and Biological Engineering* 36 (6), 857–870.
- Rohlf, F. J., 1999. Shape Statistics: Procrustes Superimpositions and Tangent Spaces. *Journal of Classification* 16 (2), 197–223.
- Rohlf, F. J., Slice, D., 1990. Extensions of the Procrustes method for the optimal superimposition of landmarks. *Systematic Biology* 39 (1), 40–59.
- Rusu, R. B., Cousins, S., 2011. 3D is here: Point cloud library (pcl). In: *Robotics and automation (ICRA), 2011 IEEE International Conference on*. Vol. 1–4. IEEE, pp. 1–4.
- Sarbolandi, H., Lefloch, D., Kolb, A., 2015. Kinect range sensing: Structured-light versus Time-of-Flight Kinect. *Computer Vision and Image Understanding* 139, 1–20.
- Schwarz, L. A., Mkhitarian, A., Mateus, D., Navab, N., 2012. Human skeleton tracking from depth data using geodesic distances and optical flow. *Image and Vision Computing* 30 (3), 217–226.
- Sharp, C., Soleimani, V., Hannuna, S., Camplani, M., Damen, D., Viner, J., Dodd, J. W., et al., 2017. Toward respiratory assessment using depth measurements from a time-of-flight sensor. *Frontiers in Physiology* 8, 65.
- Tahavori, F., Alnowami, M., Wells, K., 2014. Marker-less Respiratory Motion Modeling Using the Microsoft Kinect for Windows. *SPIE Med. Imaging* 44, 1–10.
- Xia, J., Siochi, R. A., 2012. A real-time respiratory motion monitoring system using KINECT: Proof of concept. *Medical Physics* 39 (5), 2682–2685.