3D Camera and Pulse Oximeter for Respiratory Events Detection

Carmina Coronel, Christoph Wiesmeyr, Heinrich Garn, Bernhard Kohn, Markus Wimmer, Magdalena Mandl, Martin Glos, Thomas Penzel, *Senior Member, IEEE*, Gerhard Klösch, Andrijana Stefanic-Kejik, Marion Böck, Eugenijus Kaniusas, and Stefan Seidel

Abstract*—*Objective: **The purpose of this study was to derive a respiratory movement signal from a 3D time-of-flight camera and to investigate if it can be used in combination with SpO_2 to detect respiratory events comparable to polysomnography (PSG) based detection.** ***Methods:*** **We derived a respiratory signal from a 3D camera and developed a new algorithm that detects reduced respiratory movement and SpO_2 desaturation to score respiratory events. The method was tested on 61 patients' synchronized 3D video and PSG recordings. The predicted apnea-hypopnea index (AHI), calculated based on total sleep time, and predicted severity were compared to manual PSG annotations (manualPSG). Predicted AHI evaluation, measured by intraclass correlation (ICC), and severity classification were performed. Furthermore, the results were evaluated by 30-second epoch analysis, labelled either as respiratory event or normal breathing, wherein the accuracy, sensitivity, specificity and Cohen's kappa were calculated.** ***Results:*** **The predicted AHI scored an ICC $r = 0.94$ (0.90 – 0.96 at 95% confidence interval, $p < 0.001$) compared to manualPSG. Severity classification scored 80% accuracy, with no misclassification by more than one severity level. Based on 30-second epoch analysis, the method scored a Cohen's kappa $= 0.72$, accuracy $= 0.88$, sensitivity $= 0.80$, and specificity $= 0.91$.** ***Conclusion:*** **Our detection method using SpO_2 and 3D camera had excellent reliability and substantial agreement with PSG-based scoring.** ***Significance:*** **This method showed the potential to reliably detect respiratory events without airflow and respiratory belt sensors, sensors that can be uncomfortable to patients and susceptible to movement artefacts.**

***Index Terms*—Biomedical computing, detection algorithm, biomedical signal processing, biomedical monitoring.**

I. Introduction

SLEEP apnea is a disorder characterized by multiple episodes of apnea or hypopnea during sleep. As stated in the American Academy of Sleep Medicine (AASM) scoring manual, an apnea is defined as a reduction in airflow by 90% for at least 10 seconds, while hypopnea is an airflow reduction of at least 30%, with an associated decrease in blood oxygen saturation and/or arousal [1]. The gold standard for sleep diagnosis, including the detection of respiratory events such as apneas and hypopneas, is the overnight sleep recording called polysomnography (PSG), which is composed of multiple sensors simultaneously measuring different body signals. According to the AASM scoring manual, electrocardiogram, electroencephalogram (EEG), electrooculogram, leg and chin electromyograms, airflow, respiratory effort, and SpO_2 are mandatory signals in a PSG [1]. SpO_2 is a signal measured by a pulse oximeter and is the measure of blood oxygen saturation at the peripherals. Respiratory effort signals, preferably recorded at the thorax and abdomen by respiratory inductance plethysmography (RIP), are also included in the PSG [1].

The frequency of apnea and hypopnea events is a measure of severity of sleep apnea, and it is expressed as apnea-hypopnea index (AHI). The AHI is the number of apnea and hypopnea events per hour during sleep. An AHI between 5–15 is considered mild, 15?–30 is considered moderate and an AHI >30 is considered severe [2]. PSG systems typically provide scoring of respiratory events by their integrated software. However, the overall setup, with the attachment of numerous sensors and wires

Manuscript received October 21, 2019; revised January 30, 2020 and March 11, 2020; accepted March 25, 2020. Date of publication April 20, 2020; date of current version January 5, 2021. The project was supported by the Austrian Research Promotion Agency (FFG), project ID 859622. Thomas Penzel was partially supported by the RF Government under Grant 075-15-2019-1885. *(Corresponding author: Carmina Coronel.)*

Carmina Coronel, Christoph Wiesmeyr, Heinrich Garn, and Bernhard Kohn are with the AIT Austrian Institute of Technology GmbH, 1210 Vienna, Austria (e-mail: carmina.coronel@ait.ac.at; christoph.wiesmeyr@ait.ac.at; heinrich.garn@ait.ac.at; bernhard.kohn@ait.ac.at).

Markus Wimmer was with the Kepler University Hospital, Department of Neurology 2, 4020 Linz, Austria. He is now with the Salzkammergut Klinikum Vöklabruck, Oberösttereichische Gesundheitsholding, 4840 Vöcklabruck, Austria (e-mail: markus.wimmer@ooeg.at).

Magdalena Mandl is with the Kepler University Hospital, Department of Neurology 2, 4020 Linz, Austria (e-mail: magdalena.mandl@kepleruniklinikum.at).

Martin Glos is with the Advanced Sleep Research GmbH, Berlin, Germany and Interdisciplinary Sleep Medicine Center, Charité-Universitätsmedizin Berlin, 10117 Berlin, Germany (e-mail: martin.glos@charite.de).

Thomas Penzel is with the Advanced Sleep Research GmbH, Berlin, Germany and Interdisciplinary Sleep Medicine Center, Charité-Universittsmedizin Berlin, 10117 Berlin, Germany, and also with the Saratov State University, Saratov 410012, Russia (e-mail: thomas.penzel@charite.de).

Gerhard Klösch, Andrijana Stefanic-Kejik, Marion Böck, and Stefan Seidel are with the Medical University of Vienna, Department of Neurology, 1090 Vienna, Austria (e-mail: gerhard.kloesch@meduniwien.ac.at; andrijana.stefanic-kejik@meduniwien.ac.at; marion.boeck@meduniwien.ac.at; stefan.seidel@meduniwien.ac.at).

Eugenijus Kaniusas is with the Vienna University of Technology (TU Wien), Institute of Electrodynamics, Microwave and Circuit Engineering, 1040 Vienna, Austria (e-mail: eugenijus.kaniusas@tuwien.ac.at).

Digital Object Identifier 10.1109/JBHI.2020.2984954

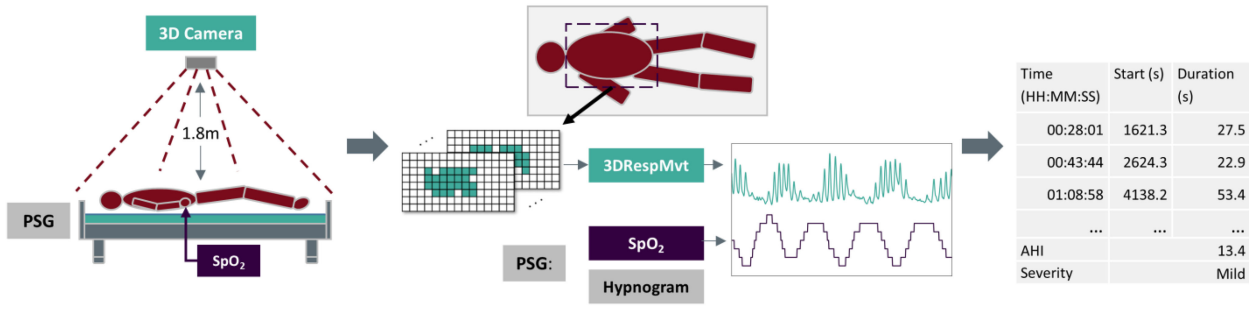

Fig. 1. Overview testing setup of the proposed method using 3D camera and SpO_2 to detect respiratory events.

to the body inhibits the patients' comfort level and may lead to further sleep loss.

Wireless, contactless, or simplified approaches have been considered in sleep medicine, especially for apnea detection. The use of doppler radars placed either under [3] or in the mattress [4] was previously studied. One novel approach previously introduced included a disposable strip called SleepStrip (S.L.P. Ltd, Tel-Aviv, Israel) [5]. And another was a wrist-worn device called WatchPAT (Itamar Medical Ltd, Caesarea Israel) that measures the peripheral arterial tone to detect apneas and hypopneas [6]. Sound-based approaches were also presented in previous studies, with the use of microphones measuring tracheal sounds [7]–[9]. Oxygen desaturation measured by SpO_2 was also introduced to detect respiratory events [10], [11], however oxygen desaturation occurs in the peripherals with a time delay, depending on the location of the pulse oximeter, the local blood perfusion, and due to other physiological reasons [12].

In this paper, we propose a method of detecting respiratory events by using SpO_2 measured by a pulse oximeter and respiratory movement derived from contactless 3D camera. The motivation of this approach is to limit the number of sensors attached to a patient, to promote better sleep quality and to lighten the load of sleep technicians during the overnight monitoring. Our method is also appropriate for patients with low tolerance against contact sensors. It has been reported that, globally, an estimated 936 million adults aged 30-69 years have mild to severe obstructive apnea [13]. Furthermore, individual waiting times for initial checkup and PSG appointment may vary from weeks to months [14]. Therefore, there is a great demand for PSG alternatives that facilitates testing of large number of patients for sleep apnea.

The approach presented in this paper uses the 3D time-of-flight (TOF) technology in the Kinect V2 (Microsoft, Redmond, VA, USA) camera to a derive respiratory movement signal. The 3D TOF camera measures the distance of objects from itself by the time it takes for light to be emitted, reflected by the objects, and received back by the camera. It has the advantage of being contactless, therefore patient mobility during sleep is not hindered. Moreover, the chances of having unusable signal due to improperly attached sensors, either due to incorrect placement or due to movements during sleep, are diminished. Therefore, we investigated if the derived respiratory movement signal from 3D camera and SpO_2 can reliably detect respiratory events with similar results to PSG-based detection.

We organized the paper as follows: Section II gives an overview of the proposed method, Section III provides information on the subjects included in the validation, Section IV presents the annotations used as reference for the validation, Section V provides the derivation of the respiratory movement signal from the 3D camera, Section VI explains the detection algorithm, Section VII describes the statistical methods used in validation, and Section VIII, IX, and X are the results, discussion and conclusion respectively.

II. The Proposed Method

Our proposed method utilizes the Kinect V2 camera to derive a respiratory movement signal. We call this signal 3DRespMvt. We chose Kinect V2 as it was reported to be applicable for body movement monitoring and appropriate for curved surfaces such as the chest region [15]. Furthermore, we validated the use of the Kinect V2 to derive a respiratory movement signal and found significant correlation with RIPsum (abdomen and thorax RIP) signal during respiratory events [16].

Using 3DRespMvt and SpO_2, we developed a novel algorithm to detect sleep respiratory events, such as apneas and hypopneas. An overview of the proposed method is illustrated in Fig. 1. The main tasks of the method are to detect respiratory events, calculate the AHI, and to estimate a severity level. Only the area of the upper body region was relevant in the derivation of 3DRespMvt. Active pixels in the upper body region relating to breathing were automatically selected and its depth information processed to obtain the 3DRespMvt. The decrease in 3DRespMvt and the desaturation in SpO_2 were detected and processed to score respiratory events and to estimate an AHI.

To evaluate our proposed method, we set up a simultaneous overnight recording of PSG and 3D camera. The 3D camera was placed approximately 1.8 m above the bed. The room was dimmed, and each patient was given a blanket. No special clothing for patients was required. The SpO_2 was measured as part of the PSG system by a pulse oximeter placed on the finger. The PSG software systems provided the hypnogram (sleep staging) and body positions (supine, left lateral, right lateral, prone). The total sleep time (TST) was derived from the hypnogram. The recording of the 3D camera was independent of the PSG. Therefore, a time synchronization was necessary for synchronization of the signals. To implement this, we incorporated a time signal sent by the recording computer into the PSG. Only the relevant

recording period where 3D camera and PSG were both recording were included in the analysis, due to the difference between the start and termination of the 3D camera and the PSG recordings. Therefore, the beginning and end sections of the recording where only one system was recording were removed from the analysis.

Several different camera technologies have been proposed and studied to measure respiration. Thermal, infrared and depth cameras have been tested to monitor either respiration, movement, cardiac activity, sleep and wake phase [17]–[22]. Falie and Inchim's study was one of the first detailing the use of depth camera to monitor respiratory effort but comparison metrics with PSG were not reported [23]. Using the Kinect sensor to monitor apnea was tested on 5 subjects, where they utilized frame differencing to determine respiratory rate [24]. However, their experiment involved using mimicked apnea events from healthy subjects. Yang *et al.* classified and detected respiratory events by a Support Vector Machine using depth camera features and audio features [25]. In [26], the study used a 3D camera to derive a respiratory index. based on Kinect and reported a correlation of $r = 0.823$ in 21 obstructive sleep apnea (OSA) patients.

Our approach differs in several ways: Firstly, we derived a respiratory signal similar to respiratory effort signals in PSG, enabling visual validation and comparison. Secondly, we combined the use of depth camera and SpO_2 to develop a novel algorithm to detect respiratory events. And lastly, we were able to validate our detection results using overnight 3D video and PSG recordings of 61 patients with suspected sleep apnea.

III. Subjects

Sixty-one subjects with suspected obstructive sleep apnea underwent full 3D video and -PSG overnight recording for one night at the sleep laboratories of Kepler University Clinic (JKU), Linz or at the Advanced Sleep Research GmbH, Berlin (ASR). The mean age of the subjects was 56 years old ($\pm$12.8). The youngest was 25 years old and oldest was 80 years old. There were 41 men and 20 women. The mean body mass index (BMI) was 29.8 kg/m^2 ($\pm$5.8).

This study was approved by the ethical committees of the state of Upper Austria (B-130-17) and of the Charité – Universitätsmedizin Berlin (EA1/127/16). Written consents from the subjects were acquired prior to inclusion in the study.

The PSG systems used by the clinics were Somnoscreen Plus with Domino software (Somnomedics, Randersacker, Germany) for JKU and EMBLA N7000 system with RemLogic 3.4.1 software, (Embla Systems, Broomfield, CO, USA) for ASR.

IV. Respiratory Events Annotations

Each recording was manually scored (manualPSG_1A) by visual inspection according to the AASM scoring manual with the recommended rule 1A for hypopneas [1]. This annotation served as a reference for the evaluation. We also manually scored the recordings using the acceptable 1B rule for hypopneas (manualPSG_1B). The 1A rule states that a hypopnea event is a decrease of at least 30% in airflow with an associated arousal or

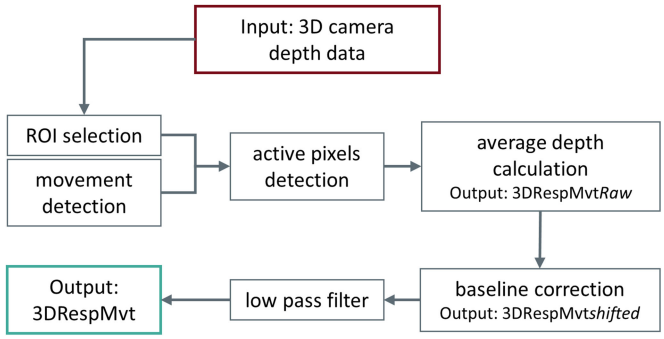

Fig. 2. The 3DRespMvt derivation stages.

associated oxygen desaturation of at least 3%, while 1B requires only an associated oxygen desaturation of at least 4% without arousals in addition to the airflow decrease [1].

For each annotation set, the apneas and hypopneas were pooled together and were referred to as respiratory events. Reference AHIs were calculated using the TST. Events occurring during wake periods were excluded from the computation. Furthermore, AHI during rapid eye movement (REM) sleep stages and during non-REM (NREM) were also calculated. Due to the difference in total recording time, only the events where 3D camera and PSG were both recording were included in the analysis.

Every PSG recording came with automatic scoring (autoPSG) of the respiratory events by the PSG software systems (ASR: RemLogic 3.4.1, JKU: Domino Software).

V. Respiratory Movement From 3D TOF Camera

The raw data of the 3D camera were processed to derive a 3DRespMvt. The main steps of the derivation are shown in Fig. 2. Prior to ROI selection, missing depth information were estimated by linearly interpolating between previous and succeeding frames. The dimension of a single image frame of the camera is 512 $\times$ 424 with a sampling frequency of 30 Hz.

As a frame includes both the full bed and the patient, a region of interest (ROI) was defined manually per recording to focus on the area of the upper body region, covering the chest and abdomen area. The dimension of the ROI differed for every recording. And since a patient was expected to move throughout the night, the ROI was placed strategically to cover most of the upper body. We set the depth information in the ROI as $d(t_m, i, j)$ where i and j are the pixel locations in the ROI in the two spatial axes. And t_m is the m-th temporal sample at a rate of 30 Hz.

A software for detection of body movements during sleep was applied, involving a convolution filter to detect changes in the signal. We used the same software that was introduced in [20], [22]. The recording was then divided into segments without movements and segments with movements. The segments without body movements were then processed to select active pixels relating to breathing. We denote the depth information in a segment as $d_s(t_m, i, j)$. The spectral distribution of each pixel in a segment was estimated by Fourier transform shown

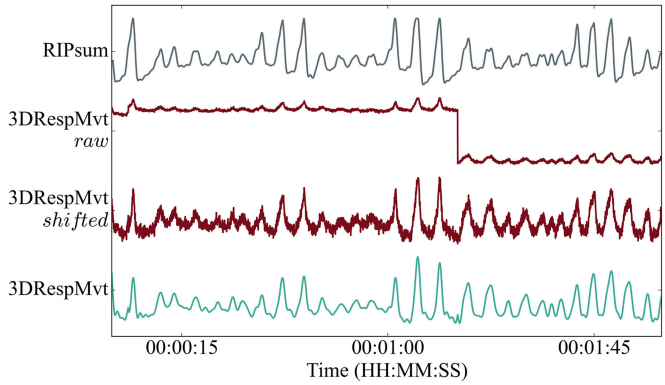

Fig. 3. The 3DRespMvt derivation stages and RIPsum = sum of RIP abdomen and thorax

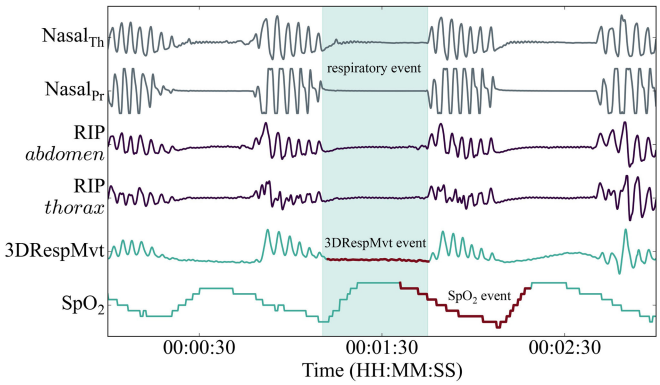

Fig. 4. SpO_2 event occurred 23 s after the onset of the 3DRespMvt (33 s in duration). Airflow sensors are for comparison purposes only. Nasal_{Th} = nasal thermistor, Nasal_{Pr} = nasal pressure cannula.

in (1), where $t_m \in \{0, \ldots, T\}$ denote the time indices in the particular segment.

$$D_s\,(k,\ i,j) = \sum_{t_m=0}^{T-1} d_s\,(t_m,i,j)\,e^{-j2\pi k t_m/T} \quad (i,j)\ \in \text{ROI}. \tag{1}$$

The pixels in the ROI which have higher energies in the frequency band between 10 to 40 breaths per minute were selected as active, where we defined A_s as the subset of the ROI which contained the active pixels' indices (i, j) of the segment s. The selected range of frequencies used was based on the respiratory rates between newborns and adults [27]. Afterwards, the 3DRespMvt*raw* was calculated from the mean depth of the active pixels (2), where N is the number of active pixels in A_s:

$$\text{3DRespMvt}raw\,(t_m) = \frac{1}{N}\sum_{i,j\in A_s} d\,(t_m,i,j) \tag{2}$$

For the segments with body movements, the active pixels of its previous segment were used to define A. As the set of active pixels changed from one segment to another, amplitude jumps in the signal occurred as seen in 3DRespMvt*raw* in Fig. 3. 3DRespMvt*raw* was corrected by subtracting the baseline curve as described in formula (3). The baseline, *bl*, was calculated per segment by polynomial curve fitting. The polynomial order increased logarithmically with the segment's length, with a maximum set at 10. This led to 3DRespMvt*shifted* signal, as shown in Fig. 3.

$$\text{3DRespMvt}shifted\ (t_m) = \text{3DRespMvt}raw\,(t_m) - bl\,(t_m) \tag{3}$$

After the shifting process, the signal was filtered using a low pass filter with a cutoff frequency of 1 Hz, wherein the chosen cutoff was based on reported respiratory rates [27], leading to the final signal, 3DRespMvt. In Fig. 4 the resulting 3DRespMvt is shown with other PSG respiratory related signals exhibiting movement during regular breathing and decrease in movement during respiratory events.

VI. Detection of Respiratory Events

The detection of respiratory events was performed by combining reduction events in 3DRespMvt and desaturation events in SpO_2. The algorithm first detected reductions in 3DRespMvt and labelled it as 3DRespMvt events. To this end an upper envelope curve was estimated by using spline interpolation on the positive peaks of 3DRespMvt. The algorithm then detected major peaks in the upper envelope curve. Afterwards, the algorithm tested segments in 3DRespMvt between the defined major peaks against an amplitude and duration threshold to label 3DRespMvt events. When the amplitude of the segment fell below 0.5 (threshold factor) of the previous peak with at least 10 seconds in duration it was marked as an event. The threshold factor was chosen based on the observed decrease in 3DRespMvt during respiratory events and on the reported average decrease in RIP signals [16], [28]. The duration threshold followed the minimum duration of events prescribed in the AASM scoring manual [1].

SpO_2 events were detected similarly by the algorithm. Major peaks in the SpO_2 were detected and were used to defined segments between the peaks. The segments were then tested for desaturation. In this study, we used 3% and 4% thresholds for SpO_2 events, chosen based on the AASM scoring manual [1].

The algorithm defined a respiratory event when a 3DRespMvt event was associated with a SpO_2 desaturation event occurring within 60 seconds after, wherein detected respiratory events have the same starting time and duration as its corresponding 3DRespMvt event. For example, in Fig. 4, the SpO_2 event occurred within 60 seconds after the start of the 3DRespMvt event, a respiratory event was detected with the same starting time and duration as the 3DRespMvt event.

The detection algorithm returned the time and duration of the detected events. It also returned the AHI based on TST. Similar to the AHI calculation of the reference annotations, events occurring during wake periods were excluded. Finally, a severity level was given based on the predicted AHI. In comparison to the reference annotations, REM-AHI and NREM-AHI were also computed by the algorithm. To be able to compare the results

with the reference annotations, only the events where 3D camera and PSG were both recording were included in the analysis.

VII. Statistical Analysis

The analysis was performed twice, using the 4% and 3% SpO_2 threshold. AHI and time-based evaluations were performed in reference to manualPSG_1A and manualPSG_1B. This was to compare performance of the proposed method to the two different AASM guidelines.

The estimated AHI values were evaluated using the intraclass correlation coefficient (ICC). The ICC r is a measure that not only evaluates correlation but also the level of agreement between two raters wherein the model used in this study is the two-way mixed effects using absolute agreement definition and single rater type [29]. The upper and lower bounds of ICC r at 95% confidence interval were reported in parenthesis. An ICC $r \geq 0.90$ denotes excellent agreement and correlation [29]. The evaluation by AHI was also performed for NREM-AHI and REM-AHI to evaluate the method for different sleep stages. For REM-AHI analysis, only 53 recordings were included because either no REM stage were observed in the recording, or no REM stage coincided with the relevant recording period. In AHI comparison, the median absolute difference, $|\Delta(\text{AHI})|$, was also calculated. Analysis by Bland-Altman was also performed to evaluate the range of agreement between the algorithm and the reference annotations. The analysis was reported with the limits of agreement at 95% confidence interval (± 1.96 standard deviation).

Furthermore, 30-s epochs was also evaluated to measure inter-rater reliability using Cohen's kappa for a time-based evaluation. The epochs were categorized into respiratory event or normal breathing (1 or 0 respectively). When at least one second of respiratory event occurred in the epoch, the epoch was categorized as 1, to ensure that events occurring in two succeeding epochs were included. A Cohen's kappa >0.60 indicates substantial agreement between raters [30]. Accuracy, sensitivity, and specificity were also calculated. Calculations were performed for all epochs pooled together and per recording basis. In per recording bases, the median was calculated.

The severity classification was evaluated by accuracy. The categories of severity were normal ($\text{AHI} < 5$), mild ($5 \leq \text{AHI} < 15$), moderate $15 \leq \text{AHI} < 30$, and severe $\text{AHI} \geq 30$ [2]. Median $|\Delta(\text{AHI})|$ in misclassified recordings were also calculated.

Additionally, respiratory events annotated by manualPSG_1A were categorized according to the body position it occurred. The detection rate of the algorithm based on sleeping body positions such as supine, left lateral, right lateral, and prone was computed. A detected respiratory event overlapping with a manually annotated respiratory event was considered as detected.

VIII. Results

The detection was performed twice, using two oxygen desaturation thresholds: 3% and 4%. Results for both are reported in the AHI analysis and 30-s analysis. Severity classification and sleeping body position analysis were only performed using the 3% threshold.

TABLE I
AHI Evaluation of the 3DRespMovt – SpO_2 Detection Compared to Manual Annotations

		manualPSG_1A		manualPSG_1B	
SpO_2 threshold		3%	4%	3%	4%
AHI					
ICC r		0.94	0.86	0.91	0.97
	LB-UB[1]	0.90-0.96	0.51-0.94	0.51-0.97	0.95-0.98
	p	<0.001	< 0.001	= 0.001	< 0.001
median($\|\Delta(\text{AHI})\|$)		3.9	7.4	9.2	2.4
NREM-AHI					
ICC r		0.94	0.86	0.92	0.97
	LB-UB	0.90-0.96	0.51-0.94	0.58-0.97	0.96-0.98
	p	<0.001	< 0.001	< 0.001	< 0.001
median($\|\Delta(\text{AHI})\|$)		3.4	7.3	8.4	2.3
REM-AHI[2]					
ICC r		0.90	0.82	0.89	0.94
	LB-UB	0.84-0.94	0.56-0.92	0.69-0.95	0.89-0.96
	p	<0.001	< 0.001	< 0.001	< 0.001
median($\|\Delta(\text{AHI})\|$)		6.5	7.6	9.8	3.5

[1] LB-UB: lower bound – upper bound at 95% confidence interval

[2] Evaluation performed on 53 recordings only due to availability of REM sleep stage.

Using 3% threshold, the predicted AHI when compared to manualPSG_1A, resulted in ICC $r = 0.94$ with a median $|\Delta(\text{AHI})| = 3.9$, as seen in Table I. Moreover, the predicted NREM-AHI scored an ICC $r = 0.94$, while the predicted REM-AHI scored an ICC $r = 0.90$ when compared to manualPSG_1A. REM-AHI evaluation was performed on only 53 recordings due to the availability of the REM sleep stage. In Fig. 5 the Bland-Altman plot of predicted AHI against the manualPSG_1A are shown, where the mean AHI is 1.6. All except for two were within the limits of agreement.

Additionally, the algorithm was tested using the 4% threshold. This resulted in ICC $r = 0.86$ when evaluated against manualPSG_1A. The results were also tested against manualPSG_1B with an ICC $r = 0.97$ and median $|\Delta(\text{AHI})| = 2.4$. Predicted NREM-AHI scored an ICC $r = 0.97$ and the predicted REM-AHI scored an ICC $r = 0.94$ when compared to manualPSG_1B. The complete AHI evaluation by ICC is shown in Table I. The Bland-Altman plot shows a narrow limit of agreement with a mean of -2.5, with 2 recordings outside the limits of agreement between the predicted AHI and manualPSG_1B. For comparison purposes, the autoPSG AHI resulted in ICC $r = 0.94$ (0.74 – 0.98, $p < 0.001$) when compared to manualPSG_1A, and ICC $r = 0.95$ (0.89 – 0.97, $p < 0.001$) when compared to manualPSG_1B.

The detection (using 3% threshold) evaluated by 30-s analysis against manualPSG_A1, resulted in Cohen's kappa = 0.72, accuracy = 0.88, specificity = 91, and sensitivity = 0.80, when all epochs were pooled together. The desaturation using 4% threshold performed better when compared to manualPSG_1B with Cohen's kappa = 0.79, accuracy = 0.92, specificity = 95,

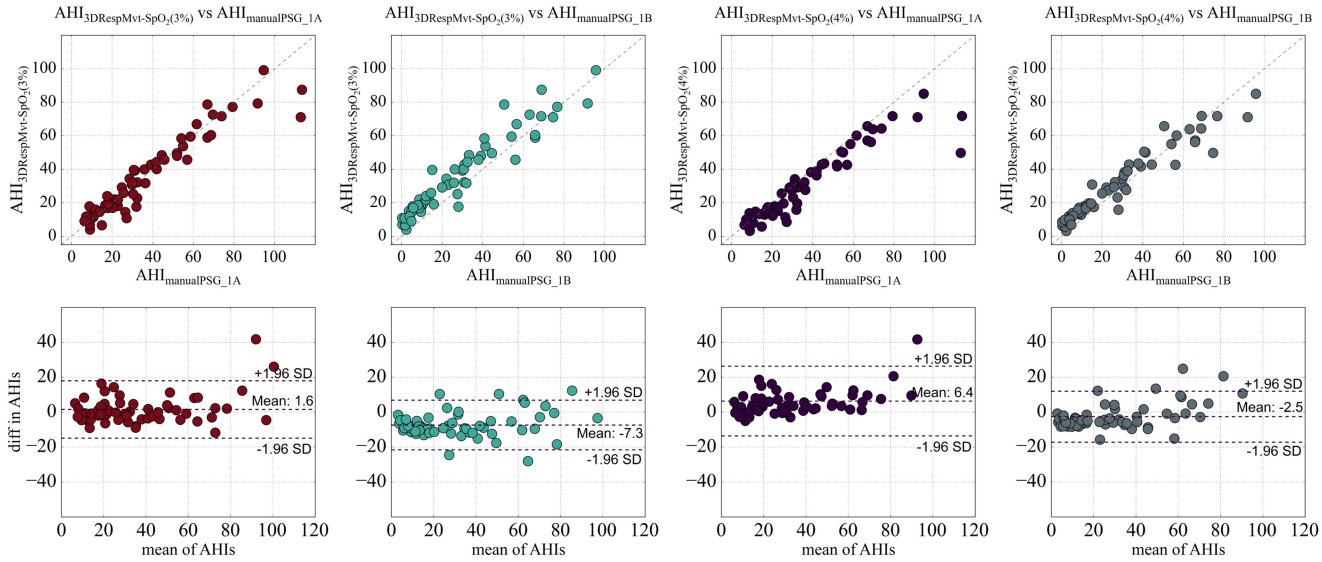

Fig. 5. The predicted AHI by 3DRespMvt-SpO$_2$ algorithm versus manualPSG_1A and manualPSG_1B AHI. Scatterplots on the top half and Bland-Altman plots on the bottom half. AHI: # of events / hour. ±1.96 SD: standard deviation at 95% confidence interval, diff in AHI: difference in AHI.

TABLE II
EVALUATION BY 30-S EPOCH OF THE 3DRESPMOVT – SPO$_2$ DETECTION VERSUS MANUAL ANNOTATIONS

	manualPSG_1A		manualPSG_1B	
SpO$_2$ threshold	3%	4%	3%	4%
(all epochs[1]/ median[2])				
Cohen's kappa	0.72 /0.65	0.70 /0.61	0.69 /0.60	0.79 /0.73
Accuracy	0.88 /0.90	0.88 /0.89	0.87 /0.89	0.92 /0.93
Sensitivity	0.80 /0.82	0.70 /0.69	0.87 /0.89	0.84 /0.86
Specificity	0.91/ 0.92	0.96/ 0.96	0.87 /0.88	0.95 /0.95

[1] The Cohen's kappa for all 30 s epochs (n = 53,938, 1: 16756, 0: 37182) pooled together
[2] The median Cohen's kappa per recording analysis.

TABLE III
SEVERITY CLASSIFICATION BY 3DRESPMVT – SPO$_2$ (3%) DETECTION COMPARED TO MANUALPSG_1A

		Predicted			
	n = 61	<5	mild	moderate	severe
Reference	<5	0	0	0	0
	mild	1	8	3	0
	moderate	0	2	14	3
	severe	0	0	3	27

and sensitivity = 0.84. The median results per recording analysis are given in Table II.

Severity classification (using 3% threshold) resulted in 80% (n = 49) accuracy. The median $|\Delta(\text{AHI})|$ for the misclassified recordings was 5.46. As presented in Table III, 6 recordings were overestimated while the other 6 were underestimated. No underestimation or overestimation by more than 1 severity level occurred.

TABLE IV
DETECTION RATE PER BODY POSITION BY 3DRESPMVT – SPO$_2$ (3%). EVALUATION BASED ON MANUALPSG_1A

body position	# events	# detected events	detection rate
supine	7572	6123	0.81
left lateral	1186	876	0.74
right lateral	2437	1765	0.72
prone	96	61	0.64

The detection rate (detected based on 3% threshold) per body position, as shown in Table IV, are 81%, 74%, and 72% (compared to manualPSG_1A) in the supine, left lateral, and right lateral positions respectively. 64% of the 96 respiratory events during a prone position were detected.

IX. DISCUSSION

Our proposed system using 3D TOF camera and SpO$_2$ (3% and 4% threshold) performed well in detecting respiratory events with Cohen's kappa >0.60 and ICC >0.86 when evaluated against the manualPSG_1A. Compared to manualPSG_1B, the AHI evaluation resulted in Cohen's kappa >0.60 and excellent ICC r >0.90. Our method using the 3% threshold performed better when compared to manualPSG_1A, and when the 4% threshold was used, it performed better when compared to manualPSG_1B. These results and the Bland-Altman plots have shown that by changing the SpO$_2$ thresholds, the method adapted to the two scoring guidelines set by AASM. Furthermore, the interrater-reliability as measured by the Cohen's kappa has shown that not only are the predicted AHIs in agreement with the references, but also by time-based analysis, indicating that the events were found in the correct time and duration. Furthermore, evaluation between NREM-AHI and AHI respectively

also resulted in ICC $r > 0.80$, showing that the proposed method worked well in different sleep stages.

In addition, the results by severity classification (using 3% threshold) resulted in 80% accuracy with only 6 recordings underestimated. Given that the average $|\Delta(\mathrm{AHI})|$ was only 5.46 for the misclassified recordings, the predicted AHIs were not far from the severity level cutoff value.

The 3D camera is a contactless device therefore efficiency in all sleeping body positions is important. Our proposed method resulted in good detection rate >0.70 for supine, left lateral, and right lateral positions. The prone position was only observed in 3 patients with a detection rate of 0.61.

In Fig. 5, two outliers can be observed in the Bland-Altman plot between the predicted AHI (using 3% threshold) and manualPSG_1A. The outlier with >40 AHI difference has a high number of respiratory events in manualPSG_1A that were associated with arousals but without SpO_2 desaturation. This is one of the limitations of the proposed method, as the EEG was not included to determine arousals, the method couldn't detect any of these events. Further studies are needed to address this limitation with patients with a higher number of arousal-associated respiratory events. The other outlier was also underestimated as the threshold used for 3DRespMvt was too restrictive. A change of threshold factor to 0.7 from 0.5 would have made the predicted AHI closer to the reference. This leads us to one point of improvement in our method: an algorithm to select the appropriate threshold for 3DRespMvt event detection per patient based on physiological factors and body positions.

Other points of improvement and interest for future work are (i) automatic selection of ROI based on body movement and position, (ii) investigation of different distances between the camera and bed, as the Kinect V2 has a reported efficient range up to 4.5m only [31], (iii) adaption of a wireless pulse oximetry to measure SpO_2, to further promote an all wireless detection process, and (iv) algorithm to classify apnea types into obstructive, central, and mixed. We have already visually tested 3DRespMvt capability for classification of apneas in [16] with 80% accuracy.

The results showed that the detections were comparable to the detected events scored by the PSG software systems used in the clinics. The autoPSG scored an ICC $r = 0.94$ compared to manual annotations. The results were also comparable to other novel approaches, such as the use of Doppler radars under the mattress that resulted in 96% and 90% sensitivity for AHI >15 and >30 respectively with specificity of 100% and 79% [3]. Another novel device is the Sonomat with a reported correlation of 0.89 compared to PSG [4]. Testing SleepStrip reported a correlation of 0.71 between the device's results and PSG results [5]. Validation studies comparing the WatchPAT calculated AHI to PSG's AHI reported an intra-class correlation coefficient (ICC) of 0.93 [32] and Spearman's $r = 0.802$ [33].

In comparison to previous studies that used 3D camera to detect respiratory events, we provided a larger number of subjects with recordings taken from sleep clinics. We were able to evaluate our proposed methods with high correlation and agreement in patients with suspected sleep apneas. In [23] a signal was derived from the upper body region, separated in twelve zones, using a 3D camera. In comparison to our work, we employed an algorithm to select only active pixels related to breathing. Unfortunately, comparison between the detection can't be made as [23] didn't report any comparison metrics with PSG recordings. The authors of [24] tested their method on 5 healthy subjects between the ranges of 1-5 years old where the apnea events were mimicked. They reported a Pearson's $r^2 = 0.98$, tested with or without blankets, with or without lights. In this study, the recordings were performed in a dimmed room and patients were provided blankets. Our setup followed the standard PSG overnight recording except for the addition of the 3D camera. Furthermore, the respiratory events in our study were not mimicked events but rather real events of adult patients observed in a sleep laboratory. Yang *et al.* [25] tested their method of respiratory events classification and detection by 3D camera and audio feature on only 4 patients and reported an error rate of 0.4%. In addition, a respiratory index based on Kinect was developed by [26] where they reported a correlation of 0.823 with AHI in 21 OSA patients. In this study, we were able to develop an algorithm with significant results tested on 61 patients. Furthermore, we utilized a novel combination of SpO_2 and the respiratory movement signal derived from 3D camera for the detection of respiratory events and for the estimation of AHI.

In addition to the results, we see practical applications for our approach: (i) as backup for airflow and/or RIP signals in case of artefacts and thereby saving the recording for assessment, or (ii) intentional replacement of the standard sensors for better sleep quality of patients and reduction of workload of sleep laboratory technicians, or (iii) as a simplified home or mobile testing apparatus for sleep apnea detection.

The overall setup proposed here has a patent pending in the Austrian Patent Office [34].

X. Conclusion

In this paper, we propose a novel detection method based on (1) a respiratory movement signal, derived from the thorax region of a 3D camera depth data and (2) SpO_2, measured by pulse oximetry. The proposed method aims to limit the number of sensors attached to a patient and to not hinder sleep quality of patients. Furthermore, having fewer contact sensors than full PSG helps to lighten the monitoring work of sleep technicians.

Our detection method outputs the starting time and duration of respiratory events and estimates the AHI. We tested our method on synchronized PSG and 3D camera recordings. We compared the detected events to manually annotated events based on standard PSG signals. Our method, without using nasal airflow and respiratory belt sensors, has shown high correlation and agreement with PSG-based detection. Therefore, the proposed method can be used as an alternative for full PSG recording for automatic sleep respiratory event detection and AHI estimation. This can help increase the number of patients screened for sleep apnea syndrome.

REFERENCES

[1] R. B. Berry, C. L. Albertario, and S. M. Hardin, "For the American academy of sleep medicine. The AASM manual for the scoring of sleep and associated events: Rules, terminology, and technical specifications," version 2.5. Darien, IL: *American Academy of Sleep Medicine*, 2018.

[2] American Academy of Sleep Medicine Task Force, "Sleep–Related Breathing Disorders in Adults: Recommendations for Syndrome Definition and Measurement Techniques in Clinical Research," *Sleep*, vol. 22, no. 5, pp. 667–689, Aug. 1999.

[3] M. Kagawa, H. Tojima, and T. Matsui, "Non-contact diagnostic system for sleep apnea–hypopnea syndrome based on amplitude and phase analysis of thoracic and abdominal Doppler radars," *Med. Biological Eng. Comput.*, vol. 54, no. 5, pp. 789–798, Aug. 2015.

[4] M. B. Norman *et al.*, "Validation of the Sonomat: A contactless monitoring system used for the diagnosis of sleep disordered breathing," *Sleep*, vol. 37, no. 9, pp. 1477–1487, Sep. 2014..

[5] E. Dinç, *et al.*, "Reliability of SleepStrip as a screening test in obstructive sleep apnea patients," *European Archives of Oto-Rhino-Laryngol.*, vol. 271, no. 10, pp. 2813–2818, Oct. 2014.

[6] N. A. Collop, *et al.*, "Obstructive sleep apnea devices for out-of-center (OOC) testing: technology evaluation," *J. Clin. Sleep Medicine*, vol. 7, no. 5, pp. 531–548, Oct. 2011.

[7] T. Penzel and A. Sabil, "The use of tracheal sounds for the diagnosis of sleep apnoea," *Breathe*, vol. 13, no. 2, pp. e37–e45, Jun. 2017.

[8] A. Yadollahi, E. Giannouli, and Z. Moussavi, "Sleep apnea monitoring and diagnosis based on pulse oximetery and tracheal sound signals," *Med. Biological Eng. Comput.*, vol. 48, no. 11, pp. 1087–1097, Aug. 2010.

[9] M. Glos, *et al.*, "Tracheal sound analysis for detection of sleep disordered breathing," *Somnologie*, vol. 23, no. 2, pp. 80–85, Apr. 2019.

[10] D. W. Jung, *et al.*, "Real-time automatic apneic event detection using nocturnal pulse oximetry," *IEEE Trans. Biomed. Eng.*, vol. 65, no. 3, pp. 706–712, Mar. 2018.

[11] D. S. Morillo and N. Gross, "Probabilistic neural network approach for the detection of SAHS from overnight pulse oximetry," *Med. Biological Eng. Comput.*, vol. 51, no. 3, pp. 305–315, Nov. 2012.

[12] E. Kaniusas, *Biomedical Signals and Sensorsin I*. Heidelberg: Springer Berlin, 2012, pp. 239–240.

[13] V. Benjafield, *et al.*, "Estimation of the global prevalence and burden of obstructive sleep apnoea: A literature-based analysis," *The Lancet Respiratory Medicine*, vol. 7, no. 8, pp. 687–698, Jul. 2019.

[14] W. Flemons, N. Douglas, S. Kuna, D. Rodenstein, and J. Wheatley, "Access to diagnosis and treatment of patients with suspected sleep apnea," *Amer. J. Respiratory and Critical Care Medicine*, vol. 169, no. 6, pp. 668–672, Mar. 2004.

[15] S. T. L. Pöhlmann, E. Harkness, C. Taylor, and S. Astley, "Evaluation of kinect 3d sensor for healthcare imaging," *J. Med. Biological Eng.*, vol. 36, no. 6, pp. 857–870, Dec. 2016.

[16] C. Coronel, *et al.*, "Measurement of respiratory effort in sleep by 3D camera and respiratory inductance plethysmography," *Somnologie*, vol. 23, no. 2, pp. 86–92, May 2019.

[17] G. Matar, J. Lina, J. Carrier, and G. Kaddoum, "Unobtrusive sleep monitoring using cardiac, breathing and movements activities: An exhaustive review," *IEEE Access*, vol. 6, pp. 45129–45152, Aug. 2018.

[18] A. Procházka, M. Schätz, O. Vyšata,and M. Vališ, "Microsoft kinect visual and depth sensors for breathing and heart rate analysis," *Sensors*, vol. 16, no. 7, pp. 996, Jun. 2016.

[19] A. Procházka, *et al.*, "Extraction of breathing features using MS Kinect for sleep stage detection," *Signal, Image and Video Process.*, vol. 10, no. 7, pp. 1279–1286, May 2016.

[20] H. Garn, *et al.*, "0678 Contactless 3D Detection Of Leg Movements In Sleep," *Sleep*, vol. 41, no. 1, pp. A251–A251, Apr. 2018.

[21] M. Hu, *et al.*, "Combination of near-infrared and thermal imaging techniques for the remote and simultaneous measurements of breathing and heart rates under sleep situation," *PLOS ONE*, vol. 13, no. 1, Art. no. e0190466, Jan. 2018.

[22] M. Gall, *et al.*, "A novel approach to assess sleep-related rhythmic movement disorder in children using automatic 3D analysis," *Frontiers in Psychiatry*, vol. 10, Oct. 2019.

[23] D. Falie and M. Ichim, "Sleep monitoring and sleep apnea event detection using a 3D camera," *8th Int. Conf. Commun.*, Jul. 2010.

[24] A. Al-Naji, K. Gibson, S. Lee, and J. Chahl, "Real Time Apnoea Monitoring of Children Using the Microsoft Kinect Sensor: A Pilot Study," *Sensors*, vol. 17, no. 2, pp. 286, Feb. 2017.

[25] C.Yang, G.Cheung, V.Stankovic, K.Chan and N.Ono, "Sleep apnea detection via depth video and audio feature learning," *IEEE Trans. Multimedia*, vol. 19, no. 4, pp. 822-86-835, Nov. 2017.

[26] C. Veauthier, *et al.*, "Contactless recording of sleep apnea and periodic leg movements by nocturnal 3-D-video and subsequent visual perceptive computing," *Scientific Rep.s*, vol. 9, no. 1, Nov. 2019.

[27] W. Lindh, M. Pooler, C. Tamparo, and B. Dahl, *Delmar's Comprehensive Medical Assisting*, 4th ed. Clifton Park, NY: Delmar Cengage Learning, 2010, pp. 267.

[28] G. A. Gould, *et al.*, "The sleep hypopnea syndrome," *Amer. Review of Respiratory Disease*, vol. 137, no. 4, pp. 895–898, Apr. 1988.

[29] T. Koo and M. Li, "A guideline of selecting and reporting intraclass correlation coefficients for reliability research," *J. Chiropractic Medicine*, vol. 15, no. 2, pp. 155–163, Jun. 2016.

[30] J. Landis and G. Koch, "The measurement of observer agreement for categorical data," *Biometrics*, vol. 33, no. 1, pp. 159, Mar. 1977.

[31] E. Lachat, H. Macher, T. Landes, and P. Grussenmeyer, "Assessment and calibration of a RGB–D camera (kinect v2 sensor) towards a potential use for close-range 3D modeling," *Remote Sensing*, vol. 7, no. 10, pp. 13070–13097, Oct. 2015.

[32] D. Hong, *et al.*, "0492 Comparing three home sleep apnea testing devices to polysomnography: Simultaneous and multi-night assessments," *Sleep*, vol. 40, no. 1, pp. A183–A184, Apr. 2017.

[33] E. Körkuyu, *et al.*, "The efficacy of watch PAT in obstructive sleep apnea syndrome diagnosis," *European Archives of Oto-Rhino-Laryngol.*, vol. 272, no. 1, pp. 111–116, May 2014.

[34] AIT Austrian Institute of Technology, "Verfahren zur detektion von atemaussetzern," *Austrian Patent*, A 50757/2018.