

Received May 26, 2020, accepted June 7, 2020, date of publication June 10, 2020, date of current version June 23, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3001343

# Automated Detection of Movements During Sleep Using a 3D Time-of-Flight Camera: Design and Experimental Evaluation

MARKUS GALL<sup>1</sup>, HEINRICH GARN<sup>1</sup>, (Senior Member, IEEE), BERNHARD KOHN<sup>1</sup>,  
KATHARINA BAJIC<sup>1</sup>, CARMINA CORONEL<sup>1</sup>, STEFAN SEIDEL<sup>2</sup>, MAGDALENA MANDL<sup>3</sup>,  
AND EUGENIJUS KANIUSAS<sup>4</sup>

<sup>1</sup>New Sensor Technologies, Austrian Institute of Technology (AIT) GmbH, 1210 Vienna, Austria

<sup>2</sup>Department of Neurology, Medical University of Vienna, 1090 Vienna, Austria

<sup>3</sup>Department of Neurology II, Kepler University Hospital Linz, 4020 Linz, Austria

<sup>4</sup>Institute of Electrodynamics, Microwave and Circuit Engineering, Vienna University of Technology, 1040 Vienna, Austria

Corresponding author: Markus Gall (markus.gall@ait.ac.at)

M. Gall reports grants from The Austrian Research Promotion Agency (FFG, nr. 860159) during the conduct of the study.

**ABSTRACT** Analyses of sleep-related movement disorders have gained importance due to an increase in life expectancy. The present approaches for measuring movements are based on electromyography or accelerometry and provide only local or specific results from muscles/limbs to which sensors have been attached. The motivation of this work was to investigate the detection of a more complete spectrum of sleep-related movements using a three-dimensional (3D) camera instead of the current conventional methods. In contrast to most of the previously published literature, this method allows for the detection of movements even when patients are covered with a blanket. This is the first work to evaluate movement detection with a clinical dataset and replicate the clinical environment in a laboratory setup. The laboratory setup allowed for the characterization of detectable movements through the determination of speed and amplitude limits. We used the Kinect One time-of-flight sensor to record 3D videos. Movements were quantified based on the temporal depth change in these 3D videos. A computer-controlled lifting table allowed for the controlled simulation of movements. Our algorithm detected movements with amplitude values  $>3.0$  mm and velocity values  $>3.5$  mm/s with an F1 score  $\geq 95\%$ . The shortest reliably detected duration of movement was 350 ms. In an ethically approved clinical study including 44 patients, 93.1% of electromyography-detected leg movements were also found in 3D. A significant correlation ( $\rho = 0.86$ ) was found between movements detected by the 3D system and polysomnography. The 3D system detected 31.2% more movements than electromyography. In addition to obtaining a broader spectrum of movements not limited to local and muscle/limb-specific movements, the usage of a contactless 3D camera simplifies the recording setup and preserves natural sleeping behavior. The presented 3D system may become useful for diagnostic purposes during sleep studies.

**INDEX TERMS** Computer vision, infrared image sensor, medical information system, motion detection, patient monitoring.

## I. INTRODUCTION

Reliable and sensitive detection of movements during sleep is relevant for clinics to correctly estimate the severity of sleep-related movement disorders such as periodic limb movements (PLM) or REM-sleep behavior disorder (RBD).

The associate editor coordinating the review of this manuscript and approving it for publication was Huazhu Fu<sup>1</sup>.

PLM and RBD are becoming increasingly common due to the increased prevalence in the aging population [1].

This study is the first to evaluate a method for detecting and quantifying movements occurring during sleep in an experimental and clinical setup using a three-dimensional (3D) camera. Our method allows for the patients/test-objects to be covered with a blanket. We characterize detectable movements by determining the speed and amplitude limits.

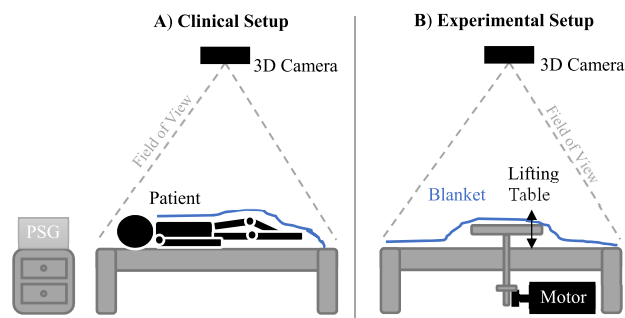
The presented 3D system is also evaluated based on data from clinical routines. This work is based on our previous work by Garn *et al.* [2], in which periodic leg movements detected by the 3D system were compared to periodic leg movements annotated by PSG using a dataset of 10 patients.

The clinical state-of-the-art technology to assess movements during sleep is polysomnography (PSG) [3]. In a PSG scan, movements are detected using electromyography (EMG), where electrodes placed on muscles, such as the tibialis anterior muscles, measure their activity [4]. The signals obtained are then either manually annotated [5] or processed using automatic algorithms [6], [7]. Attached sensors and connecting cables are inconvenient for the patient and disturb his or her natural sleep behavior [8]. In addition, the complex setup, high cost, and limited availability in sleep laboratories of PSG limit its utility [5]. Moreover, sensors tend to detach upon contact; hence, data quality is frequently diminished by artifacts [9]. Questions have been raised as to whether it is sufficient to monitor only the tibialis muscles [4]. EMG detects only the activity of muscles where electrodes are applied, making it difficult to obtain the other movements when monitoring only one muscle. Actigraphy has been successfully introduced as an alternative in the medical domain [10]–[14]. Activity localization is, however, restricted since only specific limb movements to which the sensors are attached are provided. Contactless solutions would be most desirable.

Today, most sleep laboratories apply video PSG using 2D near-infrared cameras. The contactless approach is comfortable for the patient [15] and preserves natural sleep behavior. Infrared videos of sleep environments can be automatically analyzed by image processing software [16]–[19]. However, these works evaluated their systems only with a small number of healthy volunteers and/or did not allow for the usage of a blanket. To the best of our knowledge, no work using 2D infrared video on a representative sample of data obtained from a clinical routine in which participants were allowed to use a blanket to detect movements during sleep has been conducted.

2D infrared images allow only scene analysis in two dimensions. This limitation has been addressed by methods based on 3D videos that even allow for the detection of movements along the camera's viewing direction.

Numerous studies using 3D sensors, such as Microsoft Kinect One (Microsoft corp., Redmond, Washington, USA), in healthcare applications have been published [20]–[23]. The way 3D technology is presently used in the sleep environment was proposed by Grimm *et al.* [24], who suggested a method to estimate four classes of body postures. Lee *et al.* [25] and Yu *et al.* [26] estimated body postures and derived full body movements from posture changes using 3D cameras. However, both works relied on patients sleeping without a cover. The performance is still unknown when patients use a blanket. A method for using 3D videos to detect respiratory events was recently shown in our work by Coronel *et al.* [27] and by Yang *et al.* [28]. Most recently, a system using the



**FIGURE 1.** A) Clinical setup in which a patient lies in bed with the Kinect TOF camera mounted on the ceiling above his or her head at a distance of 1.8 m. The camera obtains the depth and infrared images of the patient lying in bed covered with a blanket. This setup allows the simultaneous application of synchronized PSG recordings. B) In the experimental setup, the patient is replaced by a lifting table actuated by a computer-controlled stepper motor. The lifting table is also covered with a blanket.

Microsoft Kinect One time-of-flight (TOF) sensor to record sleep apnea, periodic leg movements and detect sleep was published [29]. A significant moderate correlation between 3D-detected periodic leg movements and those obtained by PSG was found.

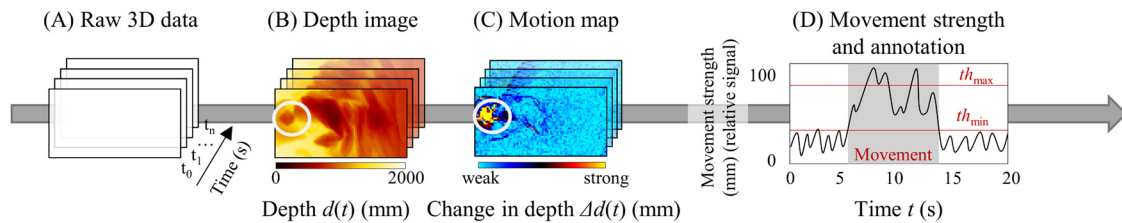
To the best of our knowledge, no literature quantified limits of detection in terms of amplitude and speed of vertical movements. With the 3D system presented in this work, we aim to simplify the recording setup, preserve natural sleep behavior, and improve results by detecting a more thorough movement spectrum than the currently used muscle/limb-specific methods.

This paper first describes the design of the 3D system for the detection of movements of a patient lying in bed. Next, in an experimental setup the 3D system was evaluated based on movements executed by a computer controlled lifting table to characterize speed and amplitude limits of detectable movements. Finally, we describe the clinical study where PSG and 3D depth videos were recorded simultaneously and time-synchronized. Using the collected data, the 3D system was evaluated based on ground truth EMG leg movement annotations.

## II. DESIGN OF THE 3D SYSTEM TO DETECT MOVEMENTS

### A. EXPERIMENTAL AND CLINICAL RECORDING SETUP

The recording setup used at clinical sites consisted of a Kinect One v2 time-of-flight camera (Microsoft corp., Redmond, Washington, USA) [30], [31] mounted on the ceiling above a patient lying in bed (Fig. 1A). A single camera recorded 3D depth and 2D infrared videos simultaneously. The distance between the surface of the bed and the sensor was approximately 1.8 meters. Pöhlmann *et al.* [32] evaluated the Kinect v2 and stated that accurate operation starts at a distance of 0.7 m; available room architectures and previous findings (supplementary material, Kinect One noise distribution) made this distance a plausible choice. The clinical setup also allowed the simultaneous application of time-synchronized PSG.



**FIGURE 2.** Overview of the 3D video processing pipeline. (A) Raw 3D data recorded with the Kinect v2 camera at 30 frames per second and a resolution of  $512 \times 424$  pixels. Resampling with a  $4 \times 4$  kernel and pixel verification resulted in preprocessed depth images (B). We used a convolution filter with a window size  $w = 30$  frames approximating the first derivative to indicate temporal depth change per pixel in false color representation stored in the motion map (C). Next, a one-dimensional signal indicating the movement strength (D, black line) at each frame was calculated. This signal was then classified into segments of movements and nonmovements using proprietary thresholding (D, red lines).

We replicated the clinical setup in our laboratories, allowing for experimental evaluation of the presented 3D-movement detection system. The patient was replaced by a custom-built mechanical lifting table to simulate the body movements of a patient lying in bed, Fig 1B. The Kinect camera was used to obtain 3D depth and 2D infrared videos. Pöhlmann *et al.* [32] also found that the camera's recording characteristics depend on the surface color being recorded. This led to a color-dependent depth precision (= difference of each individual 3D point to a fitted plane, max. 0.05 mm difference between colors), whereas color did not have any influence on the accuracy (= position measured by the Kinect camera compared to physical ground truth position,  $4.0 \pm 3.8$  mm). We therefore covered the lifting table with white fabric similar to the sheets used in sleep laboratories. The movements of the actuator were controlled via a computer connection by setting the amplitude and speed. The speed was limited to a range from 0.8 to 8.5 mm/s, and the amplitude was limited to 35 mm (limited by the actuated table design). The experiments were performed in complete darkness. Temporal synchronization between the lifting table movements and measurements using the camera-based method was achieved manually using a triplet of predefined, clearly detectable marker movements with a 35 mm amplitude and a speed setting of 6 mm/s.

### B. 3D VIDEO PROCESSING PIPELINE OVERVIEW

In the sleep laboratories, the 3D time-of-flight (TOF) camera was mounted on the ceiling above the patient lying in bed (Fig. 1A). For the experiments that used the lifting table, the camera was mounted on the ceiling above the lifting table (Fig. 1B). Each pixel of the recorded 3D depth images encodes the distance between the camera and the nearest surface of an object e.g. bed, patient lying in bed, lifting table. The time-of-flight principle measures the distance between the camera and a surface by determining the time it takes for infrared light, emitted by the camera, to return to the camera after being reflected by the surface of an object. For this measurement amplitude-modulated incoherent near-infrared light at a wavelength of 860 nanometers was emitted. The radiation intensity is far below current safety standards (BS EN 14255-1 2005) for optical radiation.

A pipeline was implemented in Python 3.4 to process these raw 3D depth data to retrieve annotated movements. First, raw depth data (Fig. 2A) were preprocessed using resampling and pixel value verification. Fig. 2B shows color-coded depth images of a person lying in bed and covered by a white blanket of about 3 cm thickness standardly used in the sleep laboratories. A more intense red indicated a smaller distance from the sensor. Next, the motion map was generated (Fig. 2C), indicating the temporal depth change per pixel in false color representation. Here, blue indicated no change, while red and yellow indicated motion. Highlighted areas of the motion map — see the white circle in Fig. 2B & 2C — show pixels with higher temporal depth changes, indicating local movements. With this information, selected areas (e.g., leg area) were analyzed to derive a 2D signal representing the movement strength for each frame (movement strength, Fig. 2D blue line). This signal was then classified into movements and nonmovements using two thresholds (Fig. 2D red line). During this first cycle, parameters were set to detect clear and strong movements.

We found that smooth, reflective surfaces, such as bed bars or the floor, accounted for strong noise in the recorded depth images. To counteract this effect and enhance the signal-to-noise ratio, we implemented a two-cycle approach. After generating the motion map, calculating the movement strength and obtaining annotated movements in the first cycle, we obtained a value indicating noise by calculating the pixel-wise mean of the standard deviation between two annotated movements based on equation (1):

$$f_{i,j}^v = \text{mean}_{t \in s} (m_{i,j}(t)) \quad (1)$$

where  $f$  is the noise value for the pixel at the  $(i,j)$ -position obtained from all frames  $t$  over the segment  $v$  ranging from the end of an annotated movement to the start of the subsequent one;  $m$  indicates the motion value stored in the motion map of the first cycle.

Then, the motion map was normalized by pixelwise division with the obtained noise values  $f$ . The resulting motion map shows normalized noise, setting the basis for obtaining the movement strength with an enhanced signal-to-noise ratio in a second cycle. The normalized motion map enabled thresholds to be set for selected areas (e.g., leg area).

The threshold levels could then be set to lower levels than in the first cycle, as the different noise levels prevalent in the motion map did not cause false positive movement detection, allowing the presented algorithm to classify the movement strength signal in periods of movements and nonmovements with a higher sensitivity.

### C. PREPROCESSING

The raw  $512 \times 424$  pixel-sized images with a frame rate of 30 Hz are resampled using the mean of  $4 \times 4$  kernels. If the Kinect could not compute depth values from the TOF measurements, the API (application programming interface) outputs a zero-value. Resampled pixels were set to zero by our algorithms if the kernel included a zero-valued pixel.

### D. GENERATING THE MOTION MAP

Independent of the direction of movement, if a pixel's depth value changed to a certain degree, this was thought to be a part of a movement. We put this idea into practice using a convolution filter to obtain a signal representing the pixel's motion (motion signal). A predefined window was convolved with the depth value to detect changes in the pixel's depth value. Each pixel was rated individually per frame. Equation (2) shows the mathematical formulation applied to obtain a single pixel's motion signal  $m_{i,j}$ .

$$m_{i,j}(t) = \left| \sum_{k=1}^s \frac{d_{i,j}(t-k)}{s} - \sum_{l=1}^s \frac{d_{i,j}(t+l)}{s} \right| \quad (2)$$

where  $(i, j)$  denotes the pixel's position in the image,  $t$  denotes the frame in which to calculate the motion value and  $d$  represents the pixel's depth value. The calculation of the absolute value is indicated by two vertical bars. We chose a window size of 31 frames ( $s = 15$  frames) to incorporate the preceding and subsequent 0.5 s of the depth signal  $d$ . This window size was thought to be a reasonable choice considering the shortest movement this work was developed to detect were periodic leg movements that have a minimum duration of 0.5 s ( $=15$  frames) [33], [34]. The implications of a 31-frame window are discussed in detail in the discussion section. Fig. 3 illustrates the process of calculating the motion signal. The obtained value approximates the first derivative of the depth signal for each pixel. The calculated motion signal is stored in the motion map (Fig. 2C).

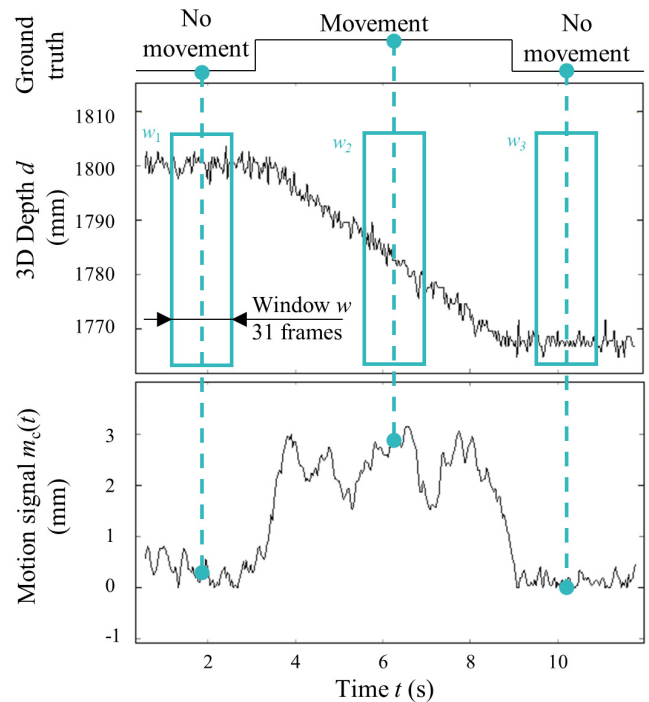
### E. OBTAINING THE MOVEMENT STRENGTH

We defined four criteria,  $C1$  to  $C4$ , for a pixel to be included in the calculation of the movement strength.

$C1$ : The pixel must be part of the predefined region of interest.

$C2$ : The pixel's motion signal must exceed a threshold with the aim of excluding small depth value fluctuations caused by noise.

$C3$ : The pixel's depth value must be between 1.0 m and 2.5 m to exclude pixels not belonging to the target area (e.g., augmented noise from bed bars and the floor).



**FIGURE 3.** The motion signal of a single pixel (bottom) was calculated (equation 2) via convolution of a predefined window (cyan, with size of 31 frames) with the pixel's depth signal (center) at each frame.

$C4$ : The corresponding infrared signal of the pixel must reach at least an intensity of 700 per frame to exclude pixels that show excessive noise due to reflective surfaces.

We then computed the movement strength as a one-dimensional signal  $S$  according to

$$S(t) = \sum_{i,j} m_{i,j}(t) : m_{i,j} \text{ fullfills criteria } l = C1, \dots, C4 \quad (3)$$

where  $m_{i,j}$  is the motion signal of a pixel at frame  $t$  and position  $(i, j)$  fulfilling all four criteria  $l = C1$  to  $C4$  for a pixel to be included in the calculation.

We applied the invalid depth mask to mute all pixels influenced by a Kinect API output for undetermined depth values. Thus, all pixels having a zero value, the Kinect API output for undetermined pixels, within the range of  $\pm 15$  frames were muted. Muting this value was adequate due to the applied convolution filter, which uses a window size of 31 frames. Each zero value thus influences the preceding and subsequent 15 frames.

### F. MOVEMENT CLASSIFICATION

We classified signal periods into movements and nonmovements based on the movement strength using two static thresholds (Fig. 2D). The first threshold  $th_{\min}$  was set immediately above the baseline noise. This was the movement strength amplitude height during the resting state. The start and stop of periods with elevated signal peaks were thus



defined by crossing the  $th_{\min}$  threshold. The definition of the start and stop times were necessary since the movement strength was a positive signal hardly any time reaching a zero value. This method was based on the latest rules defined for the scoring of sleep and related events [33], [34]. The rules defined the resting baseline as right above baseline noise. The start and stop times of leg movements were defined by the EMG amplitude crossing a threshold. For this study, we used the same  $th_{\min}$  value across recordings. This was possible since baseline noise remained the same across recordings due to equal recording site configurations.

The second threshold  $th_{\max}$  checked for the maximum height of previously identified peaks. Peaks not exceeding the  $th_{\max}$  value were defined as noise-induced augmentation, whereas those exceeding  $th_{\max}$  were classified as real movement periods and were annotated as such. The best threshold  $th_{\max}$  was empirically determined by maximizing the F1 score.

We observed that noise-induced augmentations could also exceed the  $th_{\max}$  threshold, leading to false annotations. Thus, we developed a filter allowing for the removal of such noise-induced annotations. These augmentations showed characteristic patterns, with the motion map indicating a movement equally distributed over the whole image. A metric constructed of the maximum, mean and standard deviation values of all motion signals included in an annotated movement determined whether the annotation was marked as noise-induced movement.

### III. EVALUATION

#### A. EXPERIMENTAL DATA ACQUISITION

For the experimental setup the camera was mounted above the lifting table, Fig 1B. In 4 experiments a total of 100 movements were carried out for each pair of speed and amplitude settings. Table 1 summarizes these experiments. Each experiment was conducted twice. Once, the lifting table movements were performed in the center (experiments x.1) and once the lifting table movements were performed in outer regions of the camera's field of view (experiments x.2). Experiment 1 tested for detectable speeds. Therefore, the lifting table executed movements with amplitudes of 30 mm and speeds in the range of 1.5 mm/s to 5.5 mm/s in steps of 0.5 mm/s. Experiment 2 tested for detectable amplitudes. Therefore, the lifting table executed movements with amplitudes in the range of 1.0 mm to 5.0 mm in steps of 1.0 mm and speeds in the range of 3.0 mm/s to 4.5 mm/s in steps of 0.5 mm/s. Experiment 3 evaluated the shortest possible durations detectable by the TOF camera-based approach. Therefore, the lifting table executed movements with amplitudes in the range of 1.0 mm to 4.0 mm in steps of 1.00 mm and speeds of 8.5 mm/s. Experiment 4 recorded the lifting table while not performing any movement for several hours. This experiment aimed at testing whether noise is capable of inducing 3D annotations falsifying the number of 3D-detected movements.

TABLE 1. Experiment overview.

Experiment	Movement region	Speed da/dt (mm/s)		Amplitude $a$ (mm)	
		range	steps	range	steps
1.1	Center	1.5 - 5.5	0.5	30	-
1.2	Foot	1.5 - 5.5	0.5	30	-
2.1	Center	3.0 - 4.5	0.5	1.0 - 5.0	1.0
2.2	Foot	3.0 - 4.5	0.5	1.0 - 5.0	1.0
3.1	Center	8.5	-	1.0 - 4.0	1.0
3.2	Foot	8.5	-	1.0 - 4.0	1.0

Camera noise was not equally distributed across the camera's field of view—increasing from the center to the outer pixels (see supplementary material, Kinect One noise distribution)—which is why we first tested movements located in the image center (experiments x.1) and then tested movements located in the outer regions where, for example, leg movements would occur (experiments x.2). The camera was therefore shifted by 800 mm so that movements of the lifting table occurred in the outer image regions.

#### B. CLINICAL DATA ACQUISITION

We additionally evaluated the proposed 3D movement detection system on data recorded during a clinical study. Leg movement annotations obtained by the 3D system were compared with those obtained by PSG.

A study approved by the Ethics Committees of the Medical University of Vienna (EK-No. 1091/2014) and the state of Upper Austria (EK-No. 254) was conducted collecting polysomnographic and 3D data of patients presenting with various sleep complaints to the Department of Neurology II of the Kepler Medical University of Linz. Written informed consent was obtained from all patients participating in the study. Participants were recorded during one night of at least 8h in bed. Time-synchronized PSG and 3D depth and 2D videos were recorded in accordance to the setup described in Fig. 1A. 3D depth videos and 2D infrared videos were recorded time-synchronized using one Kinect One V2 (Microsoft Corp, Redmond, Washington, USA).

Clinicians performed the patients' PSG testing according to the American Academy of Sleep Medicine (AASM) standards [33]. Leg movements were detected using surface electrodes placed over the tibialis anterior muscles. The Somnoscreen Plus PSG with Domino Software (Somnomedics, Randersacker, Germany) was used to record electrooculography; electroencephalography (F3, F4, C3, C4, O1, O2, M1 and M2 electrodes); cardiorespiratory recordings (single channel electrocardiography); recordings of nasal air flow (thermocouple); nasal pressure cannula data; thoracic and abdominal respiratory movements (piezo); transcutaneous oxygen saturation; electromyography, including at least the mental, submental and both tibialis anterior muscles; Leg movement recording used surface electrodes placed

longitudinally and symmetrically around the middle of the tibialis anterior muscle, 2–3 cm apart. Bipolar surface EMG was recorded with a low-pass filter at 100 Hz, a high-pass filter at 10 Hz, and a sampling rate of 500 Hz. Amplification was set at 10  $\mu\text{V}$  per mm. The impedance of surface EMG electrodes had to be lower than 10 k $\Omega$ .

Ground truth leg movements were derived from recorded EMG signals, which have been manually annotated by the somnologists SS and MB according to the American Academy of Sleep Medicine (AASM) standards [33]. Periods of recordings were excluded if data were incomplete or technical problems such as artifacts or loss of sensors made the signals unusable. The 3D system annotated movements based on the 3D depth videos recorded during this study. Regions of interest (ROIs) were set manually and covered only the patients' legs. Periods of recordings were also excluded if movements of other body parts extensively interfered with the ROI or the legs exited the bed. Full-body movements were marked to be muted for the analysis of PSG and 3D data. Ground truth and 3D-annotated movements were compared by a software written in Python 3.4.

### C. REPORTED METRICS

Reported metrics indicated the performance of the 3D system in detecting movements when compared to the ground truth lifting table movement or EMG activation. Both the ground truth movements and the 3D-detected movements were annotated continuously. Annotations thus might start and stop at any time. We reported the following metrics:

- True Positive (TP): A ground truth annotation temporally overlapping with exactly one 3D annotation, independent of the overlapping time.
- Multiple True Positives (MTPs): A ground truth annotation temporally overlapping with more than one 3D annotation.
- False Positive (FP): A 3D annotation not overlapping with any ground truth movement.
- False Negative (FN): A ground truth movement not overlapping with any movement detected by the 3D system.

We counted a true positive if the ground truth overlapped with a 3D annotation, independent of the overlapping time. This was thought to be a reasonable choice as the 3D system, and the frequently compared EMG annotations relied on different methods of detection (muscle stimuli in PSG vs. depth change in 3D). For example, a movement could result from a body part falling back into its steady position after muscle activation ended. In such a scenario, the detection between methods might only overlap for a short time. EMG would have detected muscle activation, while the 3D system would have detected the actual movement of the leg falling back in its initial positions. However, the EMG and the 3D system targeted the same movement.

We introduced the MTP metric to gain further insight into the detection characteristics of our algorithm. The MTP measure aimed at investigating the number of ground truth

TABLE 2. Detection levels.

Detection level	Description
1	F1 = 100% and no MTPs
2	F1 = 100% and MTP occupation $\geq 95\%$
3	F1 $\geq 95\%$ and no MTPs
4	F1 < 95% or MTP occupation < 95%

annotated movements split into several shorter movements by the 3D system. For MTPs, we reported the occupation as the percentage of the ground truth movement occupied by 3D system detections.

The F1 score (4) shows how accurate the 3D-based detection was when compared to the ground truth.

$$F1 = \frac{2 \cdot (TP + MTP)}{2 \cdot (TP + MTP) + FP + FN} \quad (4)$$

We defined four detection levels based on the F1 score. To account for reliable detection even though ground truth movements were split into several shorter movements by the 3D system, the MTP occupation was incorporated in the detection levels. Table 2 shows the defined detection levels. Detection level 1 was achieved when the F1-score was equal to 100 % and no MTP was scored. Detection level 2 was achieved when the F1-score was equal to 100% but MTPs were scored with a mean MTP occupation  $\geq 95\%$ . Detection level 3 was achieved when the F1-score  $\geq 95\%$  and <100% without the occurrence of MTPs. An F1-score < 95% or a mean MTP occupation < 95% resulted in detection level 4.

In experiment 4, we stated the number of falsely annotated 3D-detected movements induced by noise and the number of noise-induced but correctly filtered annotations (see filter in “II. F Movement Classification”).

Clinical evaluations compared the 3D-system-obtained leg movement detection result to the result obtained by PSG. For each method, we stated the number of detected leg movements. We were interested in the ratio between leg movements detected by the 3D system ( $LM_{3D}$ ) and PSG-detected leg movements ( $LM_{PSG}$ ) calculated by equation (5):

$$R_{3D/PSG} = \frac{LM_{3D}}{LM_{PSG}} \quad (5)$$

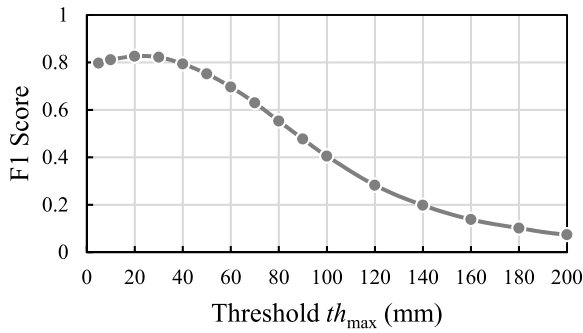
The percentage of annotated PSG leg movements also detected by the 3D system was determined using equation (6):

$$TPR = \frac{TP + MTP}{TP + MTP + FN} \quad (6)$$

Similarly, we provide the percentage of PSG-detected LM not detected by the 3D systems equation (7):

$$FNR = \frac{FN}{TP + MTP + FN} \quad (7)$$

We were also interested in whether a human scorer could visually recognize movements detected only by PSG but that were missed by the 3D system. Two scorers analyzed relevant



**FIGURE 4. Threshold optimization based on the F1 score. The plot shows the F1 score (y-axis) for different settings for  $th_{max}$ . A threshold of 20 mm was reported to achieve the highest F1 score.**

movements and annotated whether the missed movements were:

- not observable in the motion map or in the infrared video,
- only observable in the motion map,
- only observable in the infrared video or
- observable in the motion map and the infrared video.

Pearson’s correlation coefficient determined whether the number of 3D-detected movements correlated with the number of PSG-determined movements. The same coefficient was used to determine the relationship between  $R_{3D/PSG}$  and other PSG parameters (age, sleep efficiency, arousal index, periodic leg movement index, apnea-hypopnea index).

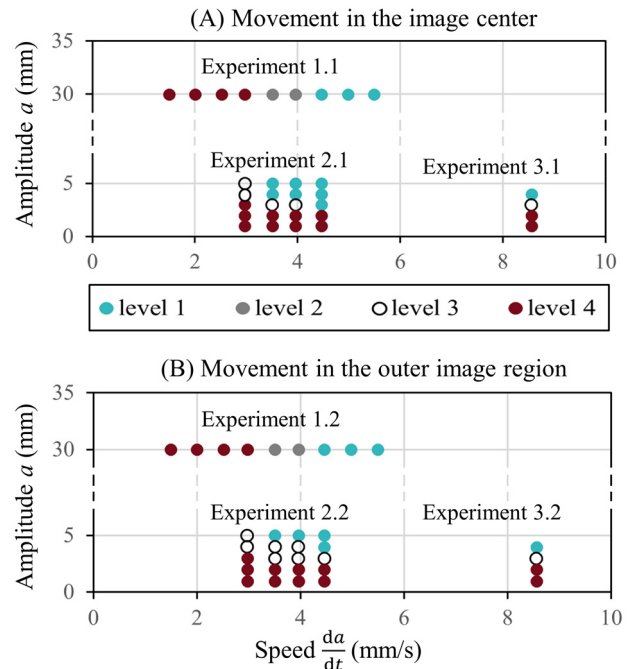
**IV. RESULTS**

**A. EXPERIMENTAL RESULTS**

We determined the best threshold  $th_{max}$  empirically by optimizing the F1 score, as shown in Fig. 4. A threshold of 20 mm was reported to achieve the highest F1 score = 0.83.

Fig. 5 compares the results obtained from the experiments separated for movements executed in the image center (Fig. 4A) and outer image regions (Fig. 4B). The figure shows the detection levels according to Table 2 for the different speed and amplitude settings in different colors. Experiment 1.1 tested for speeds detectable by the TOF camera-based system using amplitudes of 30 mm and speeds ranging from 1.5 to 5.5 mm/s. The experiment demonstrated that for movements with speeds > 4.0 mm/s, the camera-based method achieved detection level 1 (F1 = 100%, MTP = 0). At slower speeds of 4.0 and 3.5 mm/s, 3 MTPs and 19 MTPs were detected, respectively. However, the F1 score reached 100%, and more than 95% of ground truth movement durations were occupied by camera-based detections (MTP occupation  $\geq$  95%), resulting in detection level 2. For slower speeds, MTP occupation fell below 75%, and the majority of movements were either split by the camera-based method or were not detected at all (detection level 4). For speeds  $\leq$  2 mm/s, hardly any movement was detected.

Experiment 1.2 was based on the same protocol as Experiment 1.1 but tested movements in outer regions of



**FIGURE 5. Results of experiments for movements occurring in the image center (A) and outer image regions (B).**

the field of view of the TOF camera. The detection metrics of the camera-based method slightly improved compared to those of Experiment 1.1, especially for movement speeds  $\leq$  2 mm/s. However, this did not cause a shift in the detection levels obtained in Experiment 1.1. The mean number of true positives increased by 3.1%, the mean number of MTPs increased by 3.7%, and the mean number of detected false negative counts decreased by 6.7% (mean TP: +3.1%, mean MTP: +3.7%, mean FN: -6.7%). No false positives were scored throughout Experiment 1.

Experiment 2 tested for the smallest movement amplitude detectable by the image-based method. The movements ranged from amplitudes of 1 to 5 mm and speeds from 3.0 to 4.5 mm/s. No ground truth movement was split into shorter ones by the camera-based method, nor was any false positive movement recorded during any setting of Experiment 2. Experiment 2.1 involved movements in the center of the camera’s field of view and revealed that, according to the speed limits determined in Experiment 1.1, movements with an amplitude  $\geq$  3 mm achieved detection level 1 (F1 = 100%, MTP = 0). In contrast to Experiment 1.1 involving higher amplitudes, the camera-based method detected movements with speeds as low as 3 mm/s according to detection level 1 (using an amplitude of 4 mm). Amplitudes  $\leq$  2 mm only scored F1  $\leq$  95% (detection level 4).

Experiment 2.2 simulated movements in outer regions of the image recorded by the TOF camera and showed lower detection rates than those detected in Experiment 2.1. The mean number of true positive counts decreased by 6.9%, while the mean number of false negative values increased

by 6.8% (mean TP:  $-6.9\%$ , mean FN:  $+6.8\%$ ). With amplitudes of 3 mm, camera-based detection reached detection level 3 for speeds  $\geq 3.5$  mm/s with  $F1 \geq 95\%$ . Amplitudes of 4 mm resulted in detection level 1 ( $F1 = 100\%$ ) only for a speed of 4.5 mm/s and detection level 3 ( $F1 \geq 95\%$ ) for all other speeds. Only movements with amplitudes of 5 mm resulted in detection metrics similar to the results obtained for movements in the image center. Movements with amplitudes  $\leq 2$  mm were detected according to level 4 ( $F1 < 95\%$ ) by the camera-based system.

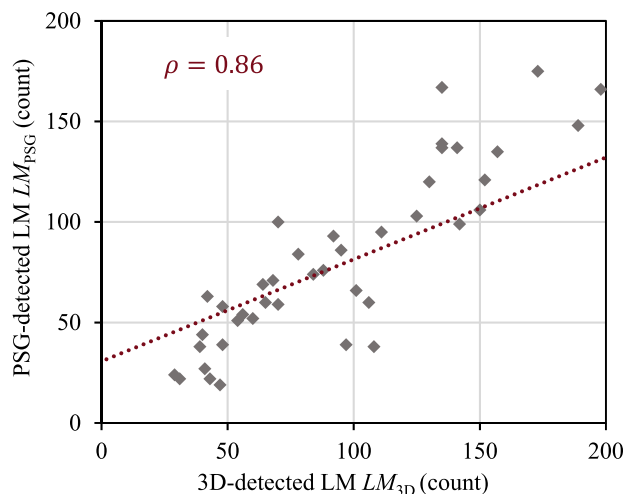
Experiment 3 tested for detectable minimal durations. All settings used the highest possible speed of 8.4 mm/s, and amplitudes ranged from 1 to 5 mm. Movements with amplitudes  $> 3$  mm/s were detected with an F1 score of 100% by the camera-based method (detection level 1). Movements with 3 mm amplitudes were detected according to detection level 3 ( $F1 \geq 95\%$ ) and movements with amplitudes  $\leq 2$  mm were detected according to detection level 4 ( $F1 < 95\%$ ). Using a speed of 8.4 mm/s and an amplitude of 3 mm resulted in 0.35 s, which were the shortest movements detected with a reasonable F1 score in our experiments.

Experiment 4 showed that 49.81 hours of recording the lifting table, without any movement performed, resulted in only one false annotation by the 3D system. On the other hand, our algorithm correctly filtered 17026 noise-induced movements using the filter described in II. F Movement Classification.

## B. CLINICAL RESULTS

PSG and synchronized 3D depth and 2D videos of 44 different patients (18 female, 28 male) were recorded overnight according to the clinical evaluation described in III.B. Patient age ranged from 17 years to 77 years of age ( $53.43 \pm 14.5$  years). Diagnoses included patients with hypersomnia (2), insomnia (5), narcolepsy (3), obstructive sleep apnea (23), restless legs syndrome (5) and no specific sleep disorder (6). In total 79.62 h (mean: 1.81 h, standard deviation: 1.23 h) of sleep were analyzed.

The results revealed that a relationship between the number of 3D-detected leg movements and PSG-detected leg movements was observable (Fig. 6). This relationship was confirmed by Pearson's correlation coefficient stating a correlation of  $\rho = 0.86$  between the 3D system and PSG-detected leg movements. The 3D system found 31.2% more leg movements than annotations based on PSG (total 3D: 4620, total PSG: 3694). A total of 93.1% of PSG-annotated movements were also detected by 3D. PSG-annotated movements where the 3D system missed a movement comprised 6.9% of the total PSG detections. Investigations on these missed movements showed that 4.5% of completely detected PSG movements (165 LM) did not show any visually recognizable deflection in either the infrared video or the motion map. For 1.7% (61 LM) of completely detected PSG movements, the scorer would have scored a movement based on the deflection in the motion map, of which 8 were also clearly visible in the infrared video. Noise-induced but filtered annotations comprised 3146 annotations in the clinical data.



**FIGURE 6.** Leg movement scatter plot comparing the number of PSG-detected leg movements  $LM_{PSG}$  (y-axis) with the number of leg movements detected by the 3D system  $LM_{3D}$  (x-axis) per patient. The red line indicates the linear trend. The methods show significant correlation  $\rho = 0.86$ .

No significant correlations between the percentage of additionally detected 3D-leg-movement  $R_{3D/PSG}$  values and obtained sleep parameters were found (age:  $\rho = 0.06$ , sleep efficiency:  $\rho = 0.24$ , arousal index:  $\rho = 0.07$ , periodic leg movement index:  $\rho = 0.00$ , apnea-hypopnea index:  $\rho = -0.03$ ).

## V. DISCUSSION

We designed and examined a new method for detecting body movements during sleep using a 3D-TOF camera. The proposed method aims to overcome several issues of currently used methods. Namely, automated surveying of the scene in 3D allowed all body movements to be monitored independent of the source muscle. Therefore, the 3D system allowed for the detection of a more complete movement spectrum than EMG with electrodes affixed to certain muscles or actigraphy with sensors affixed to certain limbs. The presented contactless approach also aims to facilitate usage and prevent interference with the patient's natural sleep behavior while recording. This is in line with other clinical applications of 3D-TOF cameras, which have been described in our previous works. In these previous works we demonstrated how to detect leg movements using TOF cameras [2] and how to use the same method to detect rhythmic movements in children [35]. Furthermore, previous works showed that a 3D-TOF camera can also be used for the assessment of respiratory efforts [27], [36].

Following we present separate discussions of experimental results (laboratory setup) and clinical results (comparison to the gold standard polysomnography).

The experiments in the laboratory were designed to characterize movements detectable by the 3D system. We found that for movement speeds  $> 4.0$  mm/s and amplitudes  $> 3.0$  mm, movements detected by the presented method



exactly matched those executed by the lifting table. Even speeds as low as 3.5 mm/s and amplitudes of 3.0 mm were detected with high sensitivity, showing an F1 score  $\geq 95\%$ . Amplitudes  $< 3$  mm were not reliably detected by the camera-based system, irrespective of the movement speed. This limit in resolution was attributed to the base noise profile of the Kinect One v2 (supplementary material, Kinect One noise distribution) having a standard deviation of  $\sim 1.5$  mm at the image center and increasing up to 4.0 mm at the periphery. Thus, for a limb movement to be reliably detected, it must exceed the amplitude of approximately 3.5 mm.

Another aim of this paper was to show that the presented approach using a TOF camera would be sensitive enough to detect movements relevant to sleep-related movement disorders. Guidelines define the strengths of sleep-related movements in terms of their duration rather than by amplitude or speed. For example, according to the ICSD [37], relevant limb movement durations occurring in periodic leg movements are in the range of 0.5 to 10 s, while myoclonic jerks have the shortest durations ranging from 50 ms to 150 ms. Using the highest possible speed in our experiment (8.5 mm/s, limited by the construction of the lifting table) and the smallest detectable amplitude change of 3 mm, the shortest detectable movement (Experiment 3) had a duration of 350 ms. This duration is suitable for the detection of PLM but is not sufficient to resolve myoclonic jerks of shorter duration. However, this limitation results from the lifting table not allowing for the simulation of faster movements. It is expected that even shorter durations can be reliably detected by the 3D system. In detail, the convolution filter averages the 15 frames to the left and the 15 frames to the right of the central frame to estimate the gradient of the movement strength (window size = 31 frames). Thus, movements longer than 0.5 s are not affected by averaging. The filter perhaps affects movements having a duration shorter than 0.5 s depending on whether the movement ends in its initial position. Movements lasting less than 0.5 s ending in their initial position might not be recognized when having a small amplitude. Due to averaging, this peak might not have sufficient power to elevate the average value into detectable limits. However, movements not ending in their initial position result in a prolonged change in the depth map and thus cause a change in the depth values no matter how short the duration of the movement is. Thus, it is expected that movements independent of their duration are recognized by the presented 3D system if their ending position differs from their initial position. Moreover, the delocalization of a body part or a finger by at least 3 mm is thought to be easily exceeded during sleep. Hence, it could be hypothesized that even myoclonic jerks could be reliably detected using the automated 3D detection system.

The results showed that movements carried out with the same speed of 3 mm/s but different movement amplitudes were split into several shorter movements (MTPs). MTPs exclusively occurred when the lifting table executed higher amplitudes (30 mm, Experiment 1.1, 29 MTP), whereas no MTPs occurred when executing smaller amplitudes

(5 mm, Experiment 2.1, 0 MTP). A possible explanation is that control inaccuracies or mechanical force delays of the motor occurred when performing prolonged movements, leading to nonlinear movements of the lifting table.

Furthermore, the results showed differences between movements within the image center and outer regions. Interestingly, Experiment 1 revealed that detection scores improved for movements made in the outer image regions, primarily for settings where a high number of MTPs occurred. However, reliability levels did not change. These variations could be explained by measurement uncertainties. The outcome of experiment 2 was as expected, and movements in the outer image regions were detected with slightly lower sensitivity due to augmented noise levels.

The evaluation of the 3D system based on data derived from the clinical routine showed that 93.1% of PSG-detected leg movements were also detected by the 3D system. The two methods showed a significant correlation of  $r = 0.86$  (an example video of a scene in the clinical environment where a movement is detected by the 3D system and PSG can be found in the supplementary materials, video 1). Even though this result suggests a better performance than that presented in the recently published work by Veauthier *et al.* [29] (72.8% of PSG-detected leg movements were also detected by the proposed 3D system with moderate correlation), this may not be the case since the cohort composition differed from the one that we used. Additionally, reported results of the same literature were based on the detection of periodic leg movements, while reported numbers in our work were based on all leg movements performed.

Movements not detected by the 3D system but by PSG accounted for 6.9% of the total PSG annotations. Visual investigation showed that 4.5% of total PSG annotations were not visible in the infrared video or the motion map. Most likely, such missed movements were caused by induced artifacts due to poor electrode contact or muscle contractions not being accompanied by visible movements on the surface (supplementary material, video 2). On the other hand, 1.7% of total PSG annotations were visible in the motion map and 0.2% were visible in the infrared video, suggesting that movement deflection was not in the detectable limits of our method. Notably, the motion map provided a useful tool to better detect movements not observable by infrared video alone.

We found that the 3D system detected 31.2% more movements than PSG. Although we did not inspect all annotated movements visually, Experiment 4 confirmed that noise-induced annotations were filtered by our algorithms. The PSG setup only detected movements activating the electrode-attached tibialis anterior muscles, while movements not activating this muscle were simply not detectable by PSG. On the other hand, the presented method using 3D recordings allows for the detection of all movements independent of the activating source muscle (supplementary material, video 3 and 4). Few additional detected movements might have resulted from the blanket moving in the leg region

initiated by movements of other body parts. However, the number of such false positive detections is expected to constitute only a small portion of the data since interference with the leg region of other body parts or full body movements were excluded by manual annotations. The augmented number of leg movements detected by the 3D system compared to the number of PSG movement detections is in accordance with the work of Provini *et al.* [38], who investigated periodic leg movements and stated that only three-fourths of all detected movements included activation of the tibialis anterior muscles. In Veauthier *et al.* [29], a number representing movements only detected by their proposed 3D method was not found, thus not allowing for further comparison.

This work used the static threshold  $th_{\min}$  as the definition of the start and stop of peaks standing out from baseline noise. In this study, using the same threshold across recordings was reasonable due to identical setups and similar baseline noise levels. However, to allow for slight changes between recording sites, an adaptive threshold based on the resting baseline standard deviation would be most desirable. We chose a thresholding approach to decide between movements and nonmovements, as it is most similar to techniques commonly used in the field of sleep-related movement analysis. The standard rules for the scoring of sleep and related events define the onset and offset of leg movements based on the crossing of the EMG amplitude and two thresholds [33], [34]. Learning algorithms as opposed to the threshold approach should be considered in future studies.

The best threshold  $th_{\max}$  was determined by optimizing the F1 score (Fig. 4). Even the best reported F1 score of 0.83 suggested only moderate classification performance. However, this score was based on all data recorded during the laboratory experiments, even those that were not detectable movements were executed.

Both the ground truth movements and the 3D-detected movements were annotated continuously. Thus, annotations might start and stop at any time. Even though discretization of the signals into fixed epochs was previously considered, the following points supported our decision of using continuous signals. First, PSG analyses use discretization only for the definition of sleep stages. The latest rules defining how sleep and related events should be scored rely on continuous signals [33], [34]. The onset and offset of leg movements are defined by the EMG amplitude exceeding a threshold value. Discretization would thus elongate or trim annotations when compared to standard methods. Additionally, discretization would have led to an additional parameter being necessary for evaluation. This parameter had to be used to determine whether an epoch would be scored as positive or negative depending on the amount of overlap with a positive/negative annotation. With the measures provided in this study, we aimed to best represent the agreement between methods.

The lack of evaluation of the detection limits of horizontal movements is a limitation of this work. However, we believe that a simulation of horizontal movements would not result

in a profound assessment of the detection limits of the 3D system due to the following considerations. Measuring horizontal motion relies on an indirect measure of the resulting changes in vertical depth as the leg moves in this plane. When objects move horizontally, the resulting depth changes depend on the angle of the measured surface to the image plane. The flatter the surface (small angle), the more the object must move to evoke a detectable depth change in the 3D system. In sleep-related movement analysis, the object most frequently measured is the blanket. The blanket covers the moving body parts. Many factors have an impact on the resulting surface geometries. As an example, when a body part moves, the effect of this movement on the surface of the blanket relies on the textile characteristics, namely, its propensity to stretch and clinch. As a result, with regards to the scope of this paper, the resulting surface geometry of the blanket is too complex to be considered for simulations.

Another limitation of this work is the manual component still required for excluding full-body movements from the analysis. However, the required manual effort diminishes with the use of the 3D system as only detected movements have to be checked for being part of full body movements. Future studies should incorporate the automatic detection of such full body movements to provide more robust results without the necessity of manual annotating full body movements.

## VI. CONCLUSION

We designed a method for the automated detection of movements during sleep and evaluated its performance in terms of the limits of movement speed and amplitude. The experimental investigation showed that within these limits, the system exhibited sufficient sensitivity to provide a reliable detection of sleep-related movements. There were no false positive detections. We evaluated the performance of the method using 3D video data obtained from the clinical routine. We found the method to be sensitive and reliable, and the fact that detection considers movements of all body parts could make the presented 3D system an interesting method for the detection of movements during sleep. By overcoming the current issues and thus improving day-to-day movement analysis, our approach might offer benefits for the clinical routine in sleep diagnosis.

## CONFLICT OF INTEREST STATEMENT

M. Gall reports grants from The Austrian Research Promotion Agency (FFG, nr. 860159) during the conduct of the study.

In addition, M. Gall, H. Garn and B. Kohn have an Austrian patent application AT520863A1 pending for “Method for detecting body movements of a sleeping person”.

## ACKNOWLEDGEMENT

The authors would like to thank Marion Böck from the Medical University of Vienna for supervising manual PSG annotations.

## REFERENCES

- [1] S. Stranges, W. Tigbe, F. X. Gómez-Olivé, M. Thorogood, and N.-B. Kandala, "Sleep problems: An emerging global epidemic? Findings from the INDEPTH WHO-SAGE study among more than 40,000 older adults from 8 countries across Africa and Asia," *Sleep*, vol. 35, no. 8, pp. 1173–1181, Aug. 2012, doi: [10.5665/sleep.2012](https://doi.org/10.5665/sleep.2012).
- [2] H. Garn, M. Gall, B. Kohn, C. Wiesmeyr, G. Kloesch, M. Wimmer, A. Stefanic-Kejik, M. Boeck, M. Mandl, O. S. Ipsiroglu, and S. Seidel, "0678 contactless 3D detection of leg movements in sleep," *Sleep*, vol. 41, no. 1, p. A251, Apr. 2018, doi: [10.1093/sleep/zsy061.677](https://doi.org/10.1093/sleep/zsy061.677).
- [3] V. Ibáñez, J. Silva, and O. Cauli, "A survey on sleep assessment methods," *PeerJ*, vol. 6, p. e4849, May 2018, doi: [10.7717/peerj.4849](https://doi.org/10.7717/peerj.4849).
- [4] B. Frauscher, A. Iranzo, B. Högl, J. Casanova-Molla, M. Salamero, V. Gschliesser, E. Tolosa, W. Poewe, and J. Santamaria, "Quantification of electromyographic activity during REM sleep in multiple muscles in REM sleep behavior disorder," *Sleep*, vol. 31, no. 5, pp. 724–731, May 2008, doi: [10.1093/sleep/31.5.724](https://doi.org/10.1093/sleep/31.5.724).
- [5] S. Shustak, L. Inzelberg, S. Steinberg, D. Rand, M. D. Pur, I. Hillel, S. Katzav, F. Fahoum, M. De Vos, A. Mirelman, and Y. Hanein, "Home monitoring of sleep with a temporary-tattoo EEG, EOG and EMG electrode array: A feasibility study," *J. Neural Eng.*, vol. 16, no. 2, Apr. 2019, Art. no. 026024, doi: [10.1088/1741-2552/aafa05](https://doi.org/10.1088/1741-2552/aafa05).
- [6] R. Ferri, M. Zucconi, M. Manconi, O. Bruni, S. Miano, G. Plazzi, and L. Ferini-Strambi, "Computer-assisted detection of nocturnal leg motor activity in patients with restless legs syndrome and periodic leg movements during sleep," *Sleep*, vol. 28, no. 8, pp. 998–1004, Aug. 2005.
- [7] H. Moore, E. Leary, S.-Y. Lee, O. Carrillo, R. Stubbs, P. Peppard, T. Young, B. Widrow, and E. Mignot, "Design and validation of a periodic leg movement detector," *PLoS ONE*, vol. 9, no. 12, Dec. 2014, Art. no. e114565, doi: [10.1371/journal.pone.0114565](https://doi.org/10.1371/journal.pone.0114565).
- [8] S. Kye, J. Moon, T. Lee, S. Lee, K. Lee, S.-C. Shin, and Y. S. Lee, "Detecting periodic limb movements in sleep using motion sensor embedded wearable band," in *Proc. IEEE Int. Conf. Syst., Man, Cybern. (SMC)*, Banff, AB, Canada, Oct. 2017, pp. 1087–1092, doi: [10.1109/SMC.2017.8122756](https://doi.org/10.1109/SMC.2017.8122756).
- [9] H. Garn, B. Kohn, C. Wiesmeyr, K. Dittrich, M. Wimmer, M. Mandl, G. Kloesch, M. Boeck, A. Stefanic, and S. Seidel, "3D detection of the central sleep apnoea syndrome," *Current Directions Biomed. Eng.*, vol. 3, no. 2, pp. 829–833, Sep. 2017, doi: [10.1515/cdbme-2017-0174](https://doi.org/10.1515/cdbme-2017-0174).
- [10] G. J. Bellone, S. A. Plano, D. P. Cardinali, D. P. Chada, D. E. Vigo, and D. A. Golombek, "Comparative analysis of actigraphy performance in healthy young subjects," *Sleep Sci.*, vol. 9, no. 4, pp. 272–279, Oct. 2016, doi: [10.1016/j.slsci.2016.05.004](https://doi.org/10.1016/j.slsci.2016.05.004).
- [11] E. Sforza, M. Johannes, and B. Claudio, "The PAM-RL ambulatory device for detection of periodic leg movements: A validation study," *Sleep Med.*, vol. 6, no. 5, pp. 407–413, Sep. 2005, doi: [10.1016/j.sleep.2005.01.004](https://doi.org/10.1016/j.sleep.2005.01.004).
- [12] V. Gschliesser, B. Frauscher, E. Brandauer, R. Kohlen, H. Ulmer, W. Poewe, and B. Högl, "PLM detection by actigraphy compared to polysomnography: A validation and comparison of two actigraphs," *Sleep Med.*, vol. 10, no. 3, pp. 306–311, Mar. 2009, doi: [10.1016/j.sleep.2008.03.015](https://doi.org/10.1016/j.sleep.2008.03.015).
- [13] M. Kobayashi, K. Namba, E. Ito, S. Nishida, M. Nakamura, Y. Ueki, N. Furudate, T. Kagimura, A. Usui, and Y. Inoue, "The validity of the PAM-RL device for evaluating periodic limb movements in sleep and an investigation on night-to-night variability of periodic limb movements during sleep in patients with restless legs syndrome or periodic limb movement disorder using this system," *Sleep Med.*, vol. 15, no. 1, pp. 138–143, Jan. 2014, doi: [10.1016/j.sleep.2013.08.790](https://doi.org/10.1016/j.sleep.2013.08.790).
- [14] A. Stefani, A. Heidebreder, E. Brandauer, M. Guaita, L.-M. Neier, T. Mitterling, J. Santamaria, A. Iranzo, A. Videnovic, C. Trenkwalder, F. Sixel-Döring, G. K. Wenning, A. Chade, W. Poewe, O. S. Gershanik, and B. Högl, "Screening for idiopathic REM sleep behavior disorder: Usefulness of actigraphy," *Sleep*, vol. 41, no. 6, Jun. 2018, Art. no. zsy053, doi: [10.1093/sleep/zsy053](https://doi.org/10.1093/sleep/zsy053).
- [15] A. Ulate-Campos, F. Coughlin, M. Gaínza-Lein, I. S. Fernández, P. L. Pearl, and T. Loddenkemper, "Automated seizure detection systems and their effectiveness for each type of seizure," *Seizure*, vol. 40, pp. 88–101, Aug. 2016, doi: [10.1016/j.seizure.2016.06.008](https://doi.org/10.1016/j.seizure.2016.06.008).
- [16] S. Akbarian, G. Delfi, K. Zhu, A. Yadollahi, and B. Taati, "Automated non-contact detection of head and body positions during sleep," *IEEE Access*, vol. 7, pp. 72826–72834, 2019, doi: [10.1109/ACCESS.2019.2920025](https://doi.org/10.1109/ACCESS.2019.2920025).
- [17] M. Scatena, S. Dittoni, R. Maviglia, R. Frusciante, E. Testani, C. Vollono, A. Losurdo, S. Colicchio, V. Gnoni, C. Labriola, B. Farina, G. D. Marca, and M. A. Pennisi, "An integrated video-analysis software system designed for movement detection and sleep analysis. Validation of a tool for the behavioural study of sleep," *Clin. Neurophysiol.*, vol. 123, no. 2, pp. 318–323, Feb. 2012, doi: [10.1016/j.clinph.2011.07.026](https://doi.org/10.1016/j.clinph.2011.07.026).
- [18] J. U. Bak, N. Giakoumidis, G. Kim, H. Dong, and N. Mavridis, "An intelligent sensing system for sleep motion and stage analysis," in *Proc. Int. Symp. Robot. Intell. Sensors (IRIS)*, vol. 41, 2012, pp. 1128–1134, doi: [10.1016/j.proeng.2012.07.292](https://doi.org/10.1016/j.proeng.2012.07.292).
- [19] A. Heinrich, D. Geng, D. Znamenskiy, J. P. Vink, and G. de Haan, "Robust and sensitive video motion detection for sleep analysis," *IEEE J. Biomed. Health Inform.*, vol. 18, no. 3, pp. 790–798, May 2014, doi: [10.1109/JBHI.2013.2282829](https://doi.org/10.1109/JBHI.2013.2282829).
- [20] R. A. Clark, Y.-H. Pua, K. Fortin, C. Ritchie, K. E. Webster, L. Denehy, and A. L. Bryant, "Validity of the microsoft kinect for assessment of postural control," *Gait Posture*, vol. 36, no. 3, pp. 372–377, Jul. 2012, doi: [10.1016/j.gaitpost.2012.03.033](https://doi.org/10.1016/j.gaitpost.2012.03.033).
- [21] L. F. Yeung, K. C. Cheng, C. H. Fong, W. C. C. Lee, and K.-Y. Tong, "Evaluation of the microsoft kinect as a clinical assessment tool of body sway," *Gait Posture*, vol. 40, no. 4, pp. 532–538, Sep. 2014, doi: [10.1016/j.gaitpost.2014.06.012](https://doi.org/10.1016/j.gaitpost.2014.06.012).
- [22] E. E. Stone and M. Skubic, "Fall detection in homes of older adults using the microsoft kinect," *IEEE J. Biomed. Health Inform.*, vol. 19, no. 1, pp. 290–301, Jan. 2015, doi: [10.1109/JBHI.2014.2312180](https://doi.org/10.1109/JBHI.2014.2312180).
- [23] S. Bauer, A. Seitel, H. Hofmann, T. Blum, J. Wasza, M. Balda, H.-P. Meinzer, N. Navab, and J. Hornegger, and L. Maier-Hein, "Real-time range imaging in health care: A survey," in *Time-of-Flight and Depth Imaging. Sensors, Algorithms, and Applications* (Lecture Notes in Computer Science), vol. 8200, no. 11, M. Grzegorzec, C. Theobalt, R. Koch, and A. Kolb, Eds. Berlin, Germany: Springer, 2013, pp. 228–254.
- [24] T. Grimm, M. Martinez, A. Benz, and R. Stiefelhagen, "Sleep position classification from a depth camera using bed aligned maps," in *Proc. 23rd Int. Conf. Pattern Recognit. (ICPR)*, Dec. 2016, pp. 319–324, doi: [10.1109/ICPR.2016.7899653](https://doi.org/10.1109/ICPR.2016.7899653).
- [25] J. Lee, M. Hong, and S. Ryu, "Sleep monitoring system using kinect sensor," *Int. J. Distrib. Sensor Netw.*, vol. 11, no. 10, Oct. 2015, Art. no. 875371, doi: [10.1155/2015/875371](https://doi.org/10.1155/2015/875371).
- [26] M.-C. Yu, H. Wu, J.-L. Liou, M.-S. Lee, and Y.-P. Hung, "Multiparameter sleep monitoring using a depth camera," in *Biomedical Engineering Systems and Technologies*. Berlin, Germany: Springer, Feb. 2012, pp. 311–325, doi: [10.1007/978-3-642-38256-7\\_21](https://doi.org/10.1007/978-3-642-38256-7_21).
- [27] C. Coronel, C. Wiesmeyr, H. Garn, B. Kohn, M. Wimmer, M. Mandl, M. Glos, T. Penzel, G. Kloesch, A. Stefanic-Kejik, M. Boeck, E. Kaniusas, and S. Seidel, "0460 detecting respiratory events by respiratory effort derived from 3D time-of-flight camera and SpO2," *Sleep*, vol. 42, no. 1, p. A185, Apr. 2019, doi: [10.1093/sleep/zsz067.459](https://doi.org/10.1093/sleep/zsz067.459).
- [28] C. Yang, G. Cheung, K. Chan, and V. Stankovic, "Sleep monitoring via depth video compression & analysis," in *Proc. IEEE Int. Conf. Multimedia Expo Workshops (ICMEW)*, Jul. 2014, pp. 1–6. [Online]. Available: [http://ieeexplore.ieee.org/xpls/abs\\_all.jsp?arnumber=6890645](http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6890645)
- [29] C. Veauthier, J. Ryczewski, S. Mansow-Model, K. Otte, B. Kayser, M. Glos, C. Schöbel, F. Paul, A. U. Brandt, and T. Penzel, "Contactless recording of sleep apnea and periodic leg movements by nocturnal 3-D-video and subsequent visual perceptive computing," *Sci. Rep.*, vol. 9, no. 1, p. 16812, Dec. 2019, doi: [10.1038/s41598-019-53050-3](https://doi.org/10.1038/s41598-019-53050-3).
- [30] C. S. Bamji, P. O'Connor, T. Elkhatib, S. Mehta, B. Thompson, L. A. Prather, D. Snow, O. C. Akkaya, A. Daniel, A. D. Payne, T. Perry, M. Fenton, and V.-H. Chan, "A 0.13  $\mu\text{m}$  CMOS system-on-chip for a 512 $\times$ 424 time-of-flight image sensor with multi-frequency photodemodulation up to 130 MHz and 2 GS/s ADC," *IEEE J. Solid-State Circuits*, vol. 50, no. 1, pp. 303–319, Jan. 2015, doi: [10.1109/JSSC.2014.2364270](https://doi.org/10.1109/JSSC.2014.2364270).
- [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 Sens.*, vol. 7, no. 10, pp. 13070–13097, Oct. 2015, doi: [10.3390/rs71013070](https://doi.org/10.3390/rs71013070).
- [32] S. T. L. Pöhlmann, E. F. Harkness, C. J. Taylor, and S. M. Astley, "Evaluation of kinect 3D sensor for healthcare imaging," *J. Med. Biol. Eng.*, vol. 36, no. 6, pp. 857–870, Dec. 2016, doi: [10.1007/s40846-016-0184-2](https://doi.org/10.1007/s40846-016-0184-2).



[33] R. B. Berry, C. L. Albertario, S. M. Harding, R. M. Lloyd, D. T. Plante, S. F. Quan, M. M. Troester, and B. V. Vaughn, "The AASM manual for the scoring of sleep and associated events: Rules, terminology and technical specifications, version 2.5," Amer. Acad. Sleep. Med., Darien, IL, USA, 2018.

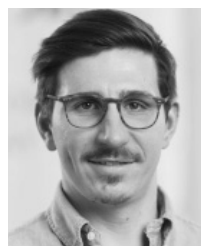
[34] R. Ferri, S. Fulda, R. P. Allen, M. Zucconi, O. Bruni, S. Chokroverty, L. Ferini-Strambi, B. Frauscher, D. Garcia-Borreguero, M. Hirshkowitz, B. Högl, Y. Inoue, A. Jahangir, M. Manconi, C. L. Marcus, D. L. Picchiotti, G. Plazzi, J. W. Winkelman, and R. S. Zak, "World association of sleep medicine (WASM) 2016 standards for recording and scoring leg movements in polysomnograms developed by a joint task force from the international and the European restless legs syndrome study groups (IRLSSG and EURLSSG)," *Sleep Med.*, vol. 26, pp. 86–95, Oct. 2016, doi: [10.1016/j.sleep.2016.10.010](https://doi.org/10.1016/j.sleep.2016.10.010).

[35] M. Gall, B. Kohn, C. Wiesmeyr, R. M. van Sluijs, E. Wilhelm, Q. Rondei, L. Jäger, P. Achermann, H.-P. Landolt, O. G. Jenni, R. Riener, H. Garn, and C. M. Hill, "A novel approach to assess sleep-related rhythmic movement disorder using automatic 3D analysis," *Frontiers Psychiatry*, vol. 10, p. 709, Oct. 2019, doi: [10.3389/fpsy.2019.00709](https://doi.org/10.3389/fpsy.2019.00709).

[36] C. Coronel, C. Wiesmeyr, H. Garn, B. Kohn, M. Wimmer, M. Mandl, M. Glos, T. Penzel, G. Klösch, A. Stefanic-Kejik, M. Böck, E. Kaniusas, and S. Seidel, "Measurement of respiratory effort in sleep by 3D camera and respiratory inductance plethysmography," *Somnologie*, vol. 23, no. 2, pp. 86–92, Jun. 2019, doi: [10.1007/s11818-019-0203-y](https://doi.org/10.1007/s11818-019-0203-y).

[37] M. J. Sateia, *International Classification of Sleep Disorders*, 3rd ed. Darien, IL, USA: American Academy Sleep Medicine, 2014.

[38] F. Provini, R. Vetrugno, S. Meletti, G. Plazzi, L. Solieri, E. Lugaresi, G. Coccagna, and A. P. Montagna, "Motor pattern of periodic limb movements during sleep," *Neurology*, vol. 57, no. 2, pp. 300–304, Jul. 2001, doi: [10.1212/WNL.57.2.300](https://doi.org/10.1212/WNL.57.2.300).



**MARKUS GALL** was born in Vienna, Austria, in 1989. He received the M.Sc. degree in biomedical engineering from the Graz University of Technology, Graz, Austria. He is currently pursuing the Ph.D. degree in biomedical engineering with the Austrian Institute of Technology (AIT) GmbH, Vienna. His current research interests include medical computer vision and bio signal analysis to develop automated monitoring devices for patients having sleep related movement disorders.



**HEINRICH GARN** (Senior Member, IEEE) received the degree in communication engineering from the Vienna University of Technology. In 1985, he started his professional career as a Scientist with the Austrian Research Centre Seibersdorf GmbH. His research was on biological effects of electromagnetic radiation. Since 1992, he has been teaching electromagnetic compatibility with the Vienna University of Technology. He founded the Austria Chapter of the IEEE Electromagnetic

Compatibility Society and served as the Chairman, from 1993 to 1998. He is an Accredited Court Expert and an Expert for laboratory accreditations. Since 1994, he has been the Head of the Industrial Electronics Department. From 2000 to 2008, he was the Head of the Information Technologies Division, Austrian Research Centres GmbH. He founded the Advanced Computer Vision GmbH, a Kplus competence center for digital image processing and served as the Scientific Managing Director, from 2002 to 2008. The competence focused on ultrasound imaging in medical applications (cooperation with GE Medical Systems Kretztechnik, Austria), surveillance and industrial inspection using ultrasound, video, and 3-D sensor systems. Since 2009, he has been the Head of the Sensing and Vision Solutions Business Unit, Austrian Institute of Technology (AIT) GmbH. He has authored 130 reviewed publications in scientific journals and contributions to international conferences. His research interests include 1-D, 2-D, and 3-D sensor and biosignal processing. From 1993 to 2000, he served as an Associate Editor for the IEEE TRANSACTIONS ON ELECTROMAGNETIC COMPATIBILITY.



**BERNHARD KOHN** received the Ph.D. degree in physics from the Universities of Cologne and Göttingen, in early 1999. Following his studies, he worked for four and a half years as a Chief Developer of the Bond Department, Viscom AG, Hanover. He was particularly responsible for project management and the development of new image processing algorithms and optical systems. Since March 2004, he has been with the Austrian Institute of Technology (AIT) GmbH responsible for the development of algorithms and especially as a Technical Co-Ordinator of the sleep lab activities. As part of his scientific work, he became the author and a coauthor of over 50 scientific publications.



**KATHARINA BAJIC** was born in Cologne, Germany, in 1991. She received the bachelor's degree in biomedical engineering from the University of Applied Science, Vienna, in 2016. She is currently pursuing the degree in medical informatics with the Vienna University of Technology. She is also working with the Austrian Institute of Technology (AIT) GmbH, Vienna, Austria. Her current research interests include machine-learning methods and biosignal processing.



**CARMINA CORONEL** was born in Quezon City, Philippines. She received the M.Sc. degree in biomedical engineering from the Vienna University of Technology, where she is currently pursuing the Ph.D. degree. She is also working with the Austrian Institute of Technology (AIT) GmbH. Her current research interests include biosensors and biosignal processing.



**STEFAN SEIDEL** was born in Linz, Austria, in 1980. He received the M.D. degree from the Medical University of Vienna, Austria, in 2004. He is currently working as the Head of the Sleep Laboratory, Department of Neurology, Medical University of Vienna. His research interest includes the phenotyping of sleep disorders, especially narcolepsy and restless legs syndrome as well as the contactless approach to sleep-related phenomena.



**MAGDALENA MANDL** was born in Prague. She received the Ph.D. degree in psychology, philosophy, and human biology from the University of Vienna. She worked with the Sleep Laboratory on Psychiatric Medicine, University of Vienna, from 1980 to 2008. Since 2008, she has been the Head of the Sleep Laboratory, University Clinic Linz, Austria, concerning pulmonary and neurological medicine.



**EUGENIJUS KANIUSAS** was born in Lithuania, in 1972. He graduated from the Faculty of Electrical Engineering and Information Technology, Vienna University of Technology (VUT), in 1997. He received the Dr. Techn. degrees in body sounds and habilitated (venia docendi) and in bioelectrical engineering, in 2001 and 2006, respectively. Since 1997, he has been with the Institute of Fundamentals and Theory of Electrical Engineering, VUT. He is currently the Head of the Research Group Biomedical Sensing/Theranostics, Institute of Electrodynamics, Microwave and Circuit Engineering, VUT. He has authored more than 180 publications and a book series of three single-author volumes by Springer.

...