

BREATHING ANALYSIS USING MS KINECT

*M. Schätz^b, J. Kuchyňka^d, F. Centonze^{b,c}, O. Ťupa^b, O. Vyšata^{b,d},
A. Procházka^{a,b}*

^a Czech Institute of Informatics, Robotics and Cybernetics, Czech Technical University, Prague, CZ

^b University of Chemistry and Technology in Prague, Dept of Computing and Control Engineering, CZ

^c Politecnico di Milano, Department of Mechanical Engineering, Italy

^d Charles University, Department of Neurology, University Hospital in Hradec Kralove, CZ

Abstract

Respiration monitoring in Sleep laboratory using contact method is part of standard Polysomnography (PSG). During examination many biosensors are used and cables from different sensors and their wired connections to computer can affect patients quality of sleep and results of sleep analysis. The method presented in this paper can form the solution of this problem allowing the use of contactless methods instead of respiratory belts. The aim of this contribution is in presentation of contactless measuring of breathing using a relatively inexpensive depth sensor from MS Kinect device.

1 Introduction

Range Imaging is based upon specific computational technologies that forms matrices (images) with their elements (pixels) carrying information about the distance from the sensor to the corresponding image component (also known as the depth map). The beginning of these methods can be dated back to beginning of 20th century when radar was discovered [6]. The very first range imaging sensor method true to its name and way of use was used by the Mars rover in 1977 [9]. Nowadays is mostly used method called time of flight, which measures how much time it takes infrared light to travel from sensor to object and back. The technology using the time of flight is also used in the MS Kinect v2 depth sensor, or in the LIDAR (Light Detection And Ranging) systems.

The breathing rate during wakefulness and sleep is different, during all stages of sleep is significantly more rapid and shallow comparing to that during the awake stage. The most significant reduction occurs during the REM sleep [4]. There exist no significant differences between the sexes either in ventilation or in end-tidal concentrations during any stage of sleep when are values corrected for body surface area. Also there are no significant differences in the levels of ventilation in non-REM sleep, but there can be significant fall of ventilation in the REM sleep. This reduction in ventilation exists because of a breathing pattern which is considerably shallower in all stages of sleep than during wakefulness [4]. Breathing rate is one of the most important vital signs [5]. And the respiratory rate and the respiratory dynamics provide useful information for diagnostics or therapeutic use [8].

There are many methods for measuring breathing with contactless methods. One of good examples is method which uses radio frequency and behaves essentially like a radar system [7, 17, 8]. A very common version of this method uses Doppler radar. It is based on the frequency shift of a signal caused by the relative velocities of a transmitter and receiver. When is transmitted signal reflected from the chest of patient, the signal will change in frequency according to the movement of the chest. Another very interesting method uses Eddy current to uses change in torso conductance, caused by varying quantities of air, blood and fluids during respiration, to measure volume changes of air in the lungs [14].

2 Methods

The MS Kinect v2 depth sensor was used to record Depth maps that follows chest movement, to analyze breathing and to detect its disorders. The official range of this sensor is 4.5m [1], but it is able to measure depth up to 8m with the sampling rate up to 30 FPS (frames per second) [1]. The resulting matrix has 512x424 elements [1, 16], and precision of depth measurement was increased in some depth range when compared with previous version of MS Kinect v1. The depth sensor of the MS Kinect v2 is not affected by outside light and errors from it are not present in depth maps. However single rough errors randomly distributed around edges or in areas out of sensor range still exist and can substantially affect the perception captured scene. The method for reducing of these errors is similar to methods used for filtering the salt-and-pepper noise in standard images. In this case the 2D median filter was used for the depth map.

2.1 Data Acquisition

We considered two methods for processing of depth maps (as on Figure 1), both dealing with the mean of depth over the selected area of M rows and N columns. The first method computes the mean value of selected area in the depth map

$$y(t) = \frac{1}{MN} \sum_i \sum_j DM_t(i, j) \quad (1)$$

where the DM_t is the depth map, acquired in time t , and i, j are indexes of area of interest. The second method evaluates the mean of differences of two consecutive depth maps

$$y(t) = \frac{1}{MN} \sum_i \sum_j (DM_t(i, j) - DM_{t-1}(i, j)) \quad (2)$$

where DM_t is depth map, acquired in time t , and i, j are indexes of area of interest. Resulting signal has zero mean value and the reduced amplitude in comparison with the original signal. It is possible to get better amplitude by computing over selection of torso or over selected pixels with biggest change.

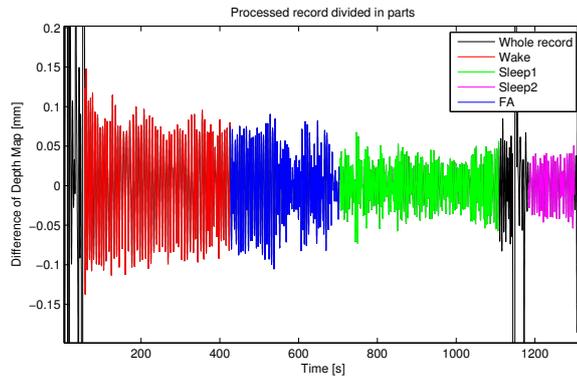
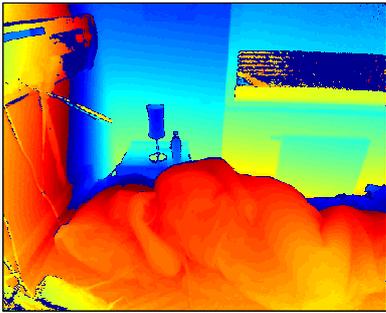


Figure 1: Example of one depth map from MS Kinect v2. Figure 2: Processed signal divided into four parts of *Wake*, *Sleep1*, *Sleep2* and *Falling Asleep (FA)*.

2.2 Filtering

As the MS Kinect depth sensor is not precise enough with its error increasing with the distance of the measured object, the output of any depth map processing will be affected by noise components. Typical respiratory rate for a healthy adult at rest is 12-20 breaths per minute [8, 2, 12, 10], which corresponds to 0.2–0.33 Hz. This fact can be used for the selection of the cut-off frequency of the low-pass FIR filter designed to filter out noise and preserve the slower periodic signal of breathing. The best option is to use a Savitzky-Golay filter of the second order:

$$y(k)_s = \frac{\sum_{i=-n}^n A_i y_{k+i}}{\sum_{i=-n}^n A_i} \quad (3)$$

where A_i are weighting coefficients of filter with width of $(2n + 1)$, which is able to preserve steep changes in the signal while smoothing it. The cut-off frequency of 0.7 Hz (42 breaths per minute) was chosen.

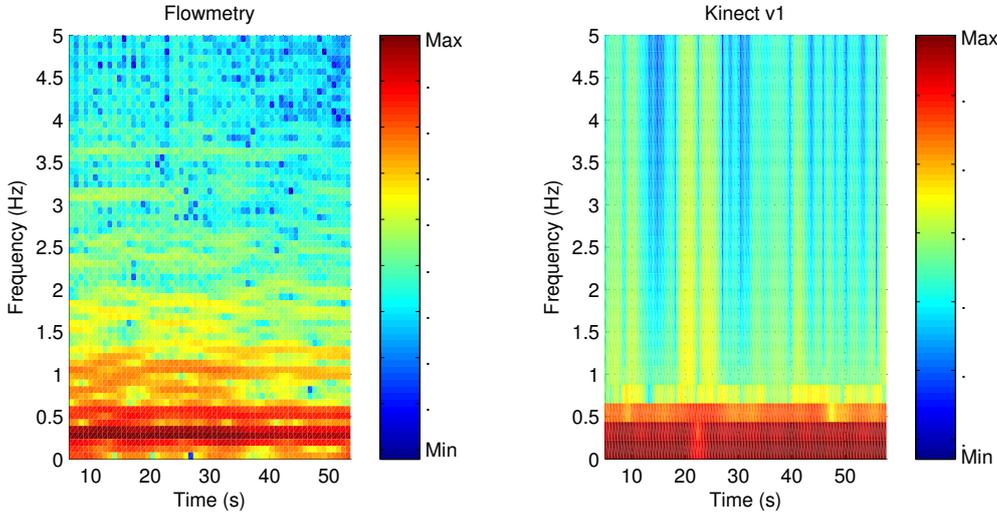


Figure 3: The comparison of spectrograms evaluated from independent data; (a) acquired by the contact method and (b) obtained by the the MS Kinect.

After application of designed filter on recorded data from depth sensor, were results of their time-frequency analysis compared with data acquired by the contact method of the standard polysomnography which is used for recording of breathing during the sleep. To compare both records for confirming that important frequencies of breathing are preserved were computed spectrograms from both signals. Comparison of results obtained by a contact method and by the MS Kinect is presented in Figure 3, where is visible that both spectrograms have most energy between 0.1 and 0.5 Hz, corresponding to breathing rate of 10 to 30 breaths per minute.

2.3 Spectral Analysis and Feature Extraction

The record presented in Figure 2 presents processed signal and its partitioning into three sections. The two sections are during sleep, just divided in two parts because recorded person moved during sleep. Spectral analysis of these parts (Figure 4) reveals that there is difference in frequency of breathing while being awake and while sleeping, which is then used as one of features for classification.

Taking into account all changes of processed signal, the most important differences of selected parts of signal include; (a) the change of frequency, (b) the change of the amplitude.

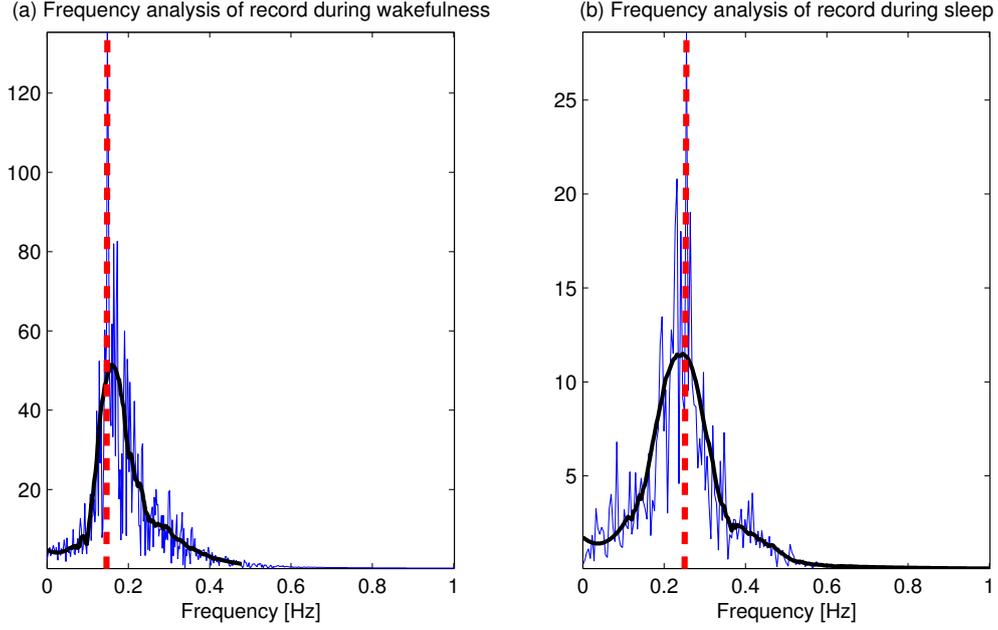


Figure 4: Frequency analysis of selected parts of the record showing frequency components (a) during the wake stage and (b) during the sleep.

During wakefulness is the dominant frequency around 0.15 Hz and the height of amplitude is 0.13, while during the sleep is the dominant frequency around 0.25 Hz with amplitude height of 0.5. The frequencies cannot be directly transferred into breaths per second because of MS Kinects unstable sampling frequency changing from 8 Hz to 30 Hz, with its mean of 18 Hz.

2.4 Bayesian Classification

Bayesian classification is based on the Bayes' theorem, which is stated by relation:

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)} \quad (4)$$

where $P(A)$ and $P(B)$ are independent probabilities, $P(A|B)$ is conditional probability (probability A given B is true) and $P(B|A)$ is conditional probability (probability of B given A is true).

Expression of Equation (4) for $P(B_j|A)$, when B_i is a random event (where $i = 1 \dots k$ and $\sum_{i=1}^k P(B_i) = 1$, because all events B_i are mutually exclusive) is:

$$P(B_j|A) = \frac{P(A|B_j)P(B_j)}{\sum_{i=1}^k P(A|B_i)P(B_i)} \quad (5)$$

The goal of classification is to find an estimate of class \hat{c}_k of an unknown instance \mathbf{x} , expressed as:

$$\hat{c}_k = \max_{c_1, c_2, \dots, c_M} (P(c_k, \mathbf{x}))$$

then probability estimation of \mathbf{x} being in class $\{c_k\}_{k=1}^M$ when put in Equation (5) is:

$$P(c_k|\mathbf{x}) = \frac{P(\mathbf{x}|c_k)P(c_k)}{\sum_{k=1}^M P(\mathbf{x}|c_k)P(c_k)} \quad (6)$$

It is possible to find fundamental characteristics of each attribute x_j , including mean μ_{c_k, x_j} and variance σ_{c_k, x_j}^2 for each class c_k , assuming Gaussian distribution of their values and independence of their features. Gaussian distribution of each attribute x_j and class c_k is then defined by relation:

$$P(x_j|c_k) = \frac{1}{\sqrt{2\pi\sigma_{c_k, x_j}^2}} \exp\left(-\frac{(x_j - \mu_{c_k, x_j})^2}{2\sigma_{c_k, x_j}^2}\right)$$

It is possible to find out how likely to see observation \mathbf{x} for class c_k by relation:

$$P(\mathbf{x}|c_k) = \prod_{j=1}^R P(x_j|c_k)$$

where R is number of features.

3 Results

Figure 2 presents processed depth data as signal and divided into parts from which are created features for three classes (two for sleep). These classes are *Wake*, *Sleep1* and *Sleep2*.

Table 1: DISTRIBUTION OF FEATURES (1)

Class	Count	Percent
Wake	186	43.06%
Sleep1	208	48.15%
Sleep2	38	8.80%

The distribution of selected features which consists of signal amplitude and dominant frequency are presented in Table 1. The classifier can be then defined to distinguish between the stage of wakefulness and sleep.

Classes for the second classification include *Wake*, *Falling Asleep (FA)* and *Sleep1*. The selected parts of processed recordings prepared for feature extraction are presented in Figure 2, with distribution of features in Table 2. The classifier defined for these features also shows possibility to classify some states in between of wakefulness and sleep.

Table 2: DISTRIBUTION OF FEATURES (2)

Class	Count	Percent
Wake	186	35.43%
Sleep1	208	39.62%
FA	131	24.95%

As is shown in Figure 3, the depth map can be processed with introduced methods using Equations (1) and (2) so breathing rate is brought out out of noise, errors of depth sensor, and movement of recorded person. The whole process made of three steps, denoising of depth maps (2D median filter), processing of depth maps (Equation (1) or Equation (2)), and the low-pass filtering of resulting signal (FIR filter Equation (3)).

4 Discussion

The analysis of record selected parts of signal revealed changes in frequency of breathing rate (Figure 4) and amplitude of signal. The breathing rate during sleep is more rapid and more shallow than during state of wakefulness (as it should be [4]). This change of breathing in processed signal, and its use as one of feature for classification shows that used method is sensitive enough to be useful.

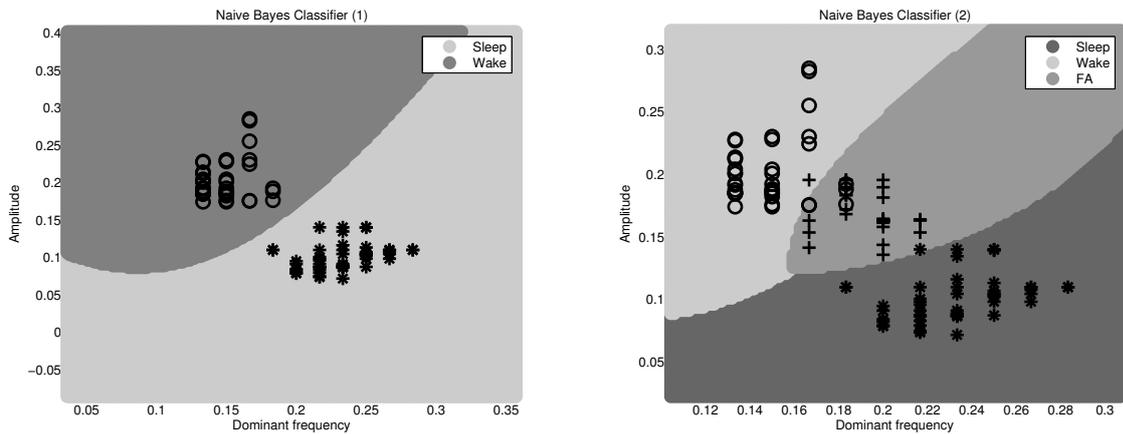


Figure 5: Classification (1) with classes for Figure 6: Classification (2) with classes for Wake and Sleep1. Wake, Falling Asleep (FA) and Sleep1.

This established property of record was used to train Bayesian classifiers, using as features dominant frequency of signal and maximal change in amplitude. The first classifier was trained on two classes (*Sleep* and *Wake*), and has 100% accuracy (Figure 5) with cross-validation of 0, calculated by leave one out method. The second classifier was trained on three classes (*Sleep*, *Falling Asleep* and *Wake*), with purpose to test possibilities of Bayesian classification for this kind of problem, and has 97% accuracy (Figure 6) with cross-validation of 0.0248, calculated by leave one out method.

References

- [1] Kinect for Windows features. <http://www.microsoft.com/en-us/kinectforwindows/meetkinect/features.aspx>, 2015.
- [2] K. E. Barrett, S. M. Barman, S. Boitano, and H. Brooks. Ganong's Review of Medical Physiology, 24th Edition.
- [3] G. Benchetrit. Breathing pattern in humans: Diversity and individuality. *Respiration Physiology*, 122:123–9, 2000.
- [4] N. J. Douglas, D. P. White, Ch. K. Pickett, J. V. Weil, and W. Clifford. Respiration during sleep. pages 840–844, 1982.
- [5] J. F. Fieselmann, M. S. Hendryx, C. M. Helms, and D. S. Wakefield. Respiratory rate predicts cardiopulmonary arrest for internal medicine inpatients. *Journal of general internal medicine*, 8(7):354–360, 1993.
- [6] C. Hulsmeyer. Hertzian wave projecting and receiving apparatus adapted to indicate or give warning of the presence of a metallic body, such as a ship or a train. *in the line of projection of such waves. [patent]. U.K.*, 13:170, 1904.
- [7] O. Kaltiokallio, H. Yigitler, R. Jantti, and N. Patwari. Non-invasive respiration rate monitoring using a single COTS TX-RX pair. *IPSN 2014 - Proceedings of the 13th International Symposium on Information Processing in Sensor Networks (Part of CPS Week)*, pages 59–69, 2014.
- [8] Y. S. Lee, P. N. Pathirana, Senior Member, and Ch. L. Steinfort. Respiration Rate and Breathing Patterns from Doppler Radar Measurements. (December):8–10, 2014.

- [9] R. A. Lewis and A. R. Johnston. A scanning laser rangefinder for a robotic vehicle. In *Proceedings 5th International Joint Conference on Artificial Intelligence*, pages 762–768, 1977.
- [10] W. Lindh, M. Pooler, C. Tamparo, and B. Dahl. *Delmar’s Comprehensive Medical Assisting: Administrative and Clinical Competencies*. Cengage Learning, 2009.
- [11] S. Ostadabbas, Ch. Bulach, D. N. Ku, L. J. Anderson, and M. Ghovanloo. A passive quantitative measurement of airway resistance using depth data. In *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 5743–5747. IEEE, August 2014.
- [12] Ch. Patel. Breathing. In *The Complete Guide to Stress Management*, pages 141–161. Springer US, 1991.
- [13] A. Prochazka, M. Schätz, O. Tupa, M. Yadollahi, O. Vysata, and M. Walls. The MS kinect image and depth sensors use for gait features detection. In *2014 IEEE International Conference on Image Processing (ICIP)*, pages 2271–2274. IEEE, October 2014.
- [14] A. Richer and A. Adler. Eddy current based flexible sensor for contactless measurement of breathing. In *Instrumentation and Measurement Technology Conference*, pages 257–260, May 2005.
- [15] M. Schätz. Statistical Recognition and Spatial Monitoring of Breathing Using Data Acquired by the MS KINECT. Master’s thesis, ICT Prague, 2015.
- [16] R. Smeenk. Kinect V1 and Kinect V2 fields of view compared. <http://smeenk.com/kinect-field-of-view-comparison/>, 2014.
- [17] T. Taheri and A. S. Anna. Non-Invasive Breathing Rate Detection Using a Very Low Power Ultra-wide-band Radar. pages 78–83, 2014.
- [18] S. V. Vaseghi. *Advanced Digital Signal Processing and Noise Reduction*. John Wiley & Sons Ltd, West Sussex, 2006.
- [19] Ch. W. Wang, A. Hunter, N. Gravill, and S. Matusiewicz. Unconstrained video monitoring of breathing behavior and application to diagnosis of sleep apnea. *IEEE Transactions on Biomedical Engineering*, 61(2):396–404, 2014.

Martin Schätz
Martin.Schatz@vscht.cz

Aleš. Procházka
Ales.Prochazka@vscht.cz