35th Annual International Conference of the IEEE EMBS
Osaka, Japan, 3 - 7 July, 2013

Camera-based System for Contactless Monitoring of Respiration

Marek Bartula, Timo Tigges, and Jens Muehlsteff

Abstract— **Reliable, remote measurement of respiration rate is still an unmet need in clinical and home settings. Although the predictive power of respiratory rate for a patient's health status is well-known, this vital sign is often measured inaccurately or not at all. In this paper we propose a camera-based monitoring system to reliably measure respiration rate without any body contact. A computationally efficient algorithm to extract raw breathing signals from the video stream has been developed and implemented. Additionally, a camera offers an easy access to motion information in the analyzed scenes, which significantly improves subsequent breath-to-breath classification. The performance of the sensor system was evaluated using data acquired with healthy volunteers, as well as with a mechanical phantom, under laboratory conditions covering a large range of challenging measurement situations.**

I. INTRODUCTION

Monitoring of respiration is important in many applications, since deviations in breathing rate or shallowness of breath are an important sign of a person's health. Respiratory disorders are early indicators of physiological deterioration. Despite these facts, it is one of the most seldom measured vital sign in the general ward [1]. Breathing is also a highly relevant sleep parameter, providing insight into the state of relaxation, sleep depth, apnea and snoring events. Accurate monitoring of respiratory rhythm can improve the effectiveness of paced breathing exercises.

Nowadays, established clinical methods for measuring respiration require a sensing device to be attached to the body [2]. Three main approaches of contact measurement focus on (a) nasal/oral air-flow (tidal volume), (b) thoracic impedance changes and (c) chest movement or change of its volume (respiratory effort). The first one requires a mask connected to an air volume measuring device, like a spirometer, or a sensor to be placed in the stream of inhaled/exhaled air. Method (b) usually requires the attachment of electrodes to the skin in the chest area, while for (c) a strain gauge or inductive belt is applied around the chest/abdomen. These methods could compromise the comfort of the monitored person since sensors are inconvenient to apply or wear, and in most cases are connected with cables. While this might be of little relevance in scenarios such as intensive care, body-attached devices are preferably avoided in applications like sleep monitoring or ambient assisted living.

This unmet need for unobtrusive monitoring of respiratory effort has triggered research in contactless solutions. Several methodologies have been investigated in the past, which are usually based on Doppler radar [3], acoustic or imaging sensors. Unlike microphones and radars, cameras provide a 2-dimensional signal with additional context information on the measurement process, relevant for interpretation of the extracted vital signs. This is of particular importance in unsupervised monitoring scenarios. Cameras allow for an automatic selection of region of interest and detection of non-respiratory events such as head or limb movements.

M. Bartula, J. Muehlsteff are with Philips Research Europe, Eindhoven, The Netherlands (e-mail: {marek.bartula, jens.muehlsteff}@philips.com).

T. Tigges is with RWTH Aachen University, Germany (email: timo.tigges@rwth-aachen.de).

II. CAMERA-BASED METHODS

Many attempts have been already made to use imaging sensors for remote measurement of respiration. Thermal infra-red cameras can detect temperature differences and changes in front of the nose/mouth region [4], [5]. However, this method is constrained to situations where the subject's face is visible. Cost of the hardware can be a limiting factor as well. Likewise, time-of-flight (ToF) cameras [6], while allowing to use depth sensing, are an expensive option. Depth reconstruction can be also achieved using multiple cameras (stereo principle) equipped with conventional sensors. However, it involves additional hardware and complex algorithms with limited depth resolution. In contrast, single visible and/or near infra-red (NIR) light camera is much more cost effective.

Respiration signals are usually extracted from video by detection of changes in the image caused by chest/abdomen movements, or direct estimation of that motion. A simple approach consists of subtracting consecutive frames and using the sum of pixels in the difference image as the respiratory signal [7], which is a measure of change in these consecutive images, rather than the actual motion. Its performance strongly depends on the pattern of clothing. It also suffers from noise or other variations in the image, i.e. global light levels, which can significantly distort the signal. The main problem lies in the fact that it does not distinguish inhaling from exhaling, and thus requires heuristically designed post-processing to reconstruct the respiration waveform. Use of more sophisticated algorithms to track respiratory effort, such as movement detection from optical flow [8], improves noise robustness, but still does not reconstruct the respiratory signal and requires high processing power.

Another camera-based solution depends on the projection of light patterns to enhance visibility of subtle respiratory motion [9], but the illumination source needs to be at a different angle with respect to the monitored person than the camera. This eliminates the possibility of having a single device solution which might be a limiting factor for many applications.

978-1-4577-0216-7/13/$26.00 ©2013 IEEE

Alternative approaches focus on tracking a group of feature points; however, it is not guaranteed that a sufficient number of those points can be found. Attaching markers to the monitored subject provides artificial characteristic points that can be easily tracked. On the other hand, an ideal solution should not require any cumbersome preparation before a measurement, as it may limit its applicability.

Motion due to respiration can be very subtle, with maximum excursion of a few millimeters [10]. We have found in our initial experiments that classical motion estimators, such as previously proposed optical flow [11], do not provide sufficient sensitivity to reliably detect those movements. The goal of this work was to develop an algorithm that combines low complexity and high sensitivity, utilizing a single off-the-shelf camera, without on-body markers or projection of light patterns. It should provide directionality information, in terms of rising 'inhale' and falling 'exhale' signal, for a real-time reconstruction of breathing wave without excessive post-processing and noticeable delay.

III. Proposed algorithm

We present a robust breathing monitoring system that allows a cost effective implementation. It consists of a camera and image processing algorithm. We have used a monochrome camera for both visible and/or near infra-red (NIR) light imaging. The luminance channel obtained from the conversion of RGB input may be used as well. The block diagram of our algorithm is presented in Fig. 1.

A. *Extraction of raw respiratory signal*

In the first step, a one-dimensional representation (profile) of the image, or selected region of interest (ROI), is obtained through a projection-like transformation onto a vertical axis. This is based on the observation that natural person-camera geometries show the strongest motion component along that axis. Therefore, it is most of the time sufficient, and computationally least complicated, to build a vertical profile applying the operation along rows of the frame. This can be any function that captures/preserves information about texture, edges or any other detail present in the image. In the current implementation, a combination of mean and standard deviation is used. An example video frame with the corresponding vertical profile is shown in Fig. 2.

The obtained 1D vector is high-pass filtered to enhance edges and to make the system relatively insensitive to changes of global illumination. Final steps in profile pre-processing consist of spatial and temporal low-pass filtering to suppress the noise.

In the next stage, breathing motion is efficiently obtained by correlating the 1D vector of a current image with that of an earlier image. Unlike typical projection-based motion estimators [12], we use cross-correlation rather than phase correlation. This is motivated by the fact that instead of an absolute translatory shift of a rigid object, we are trying to detect deformations and displacements caused by chest expansion during breathing. These may have different appearance depending on multiple factors, e.g. position of the person, camera's distance and angle of view. Therefore, extracted shift indicates a relative change during the breathing cycle and has no absolute meaning.

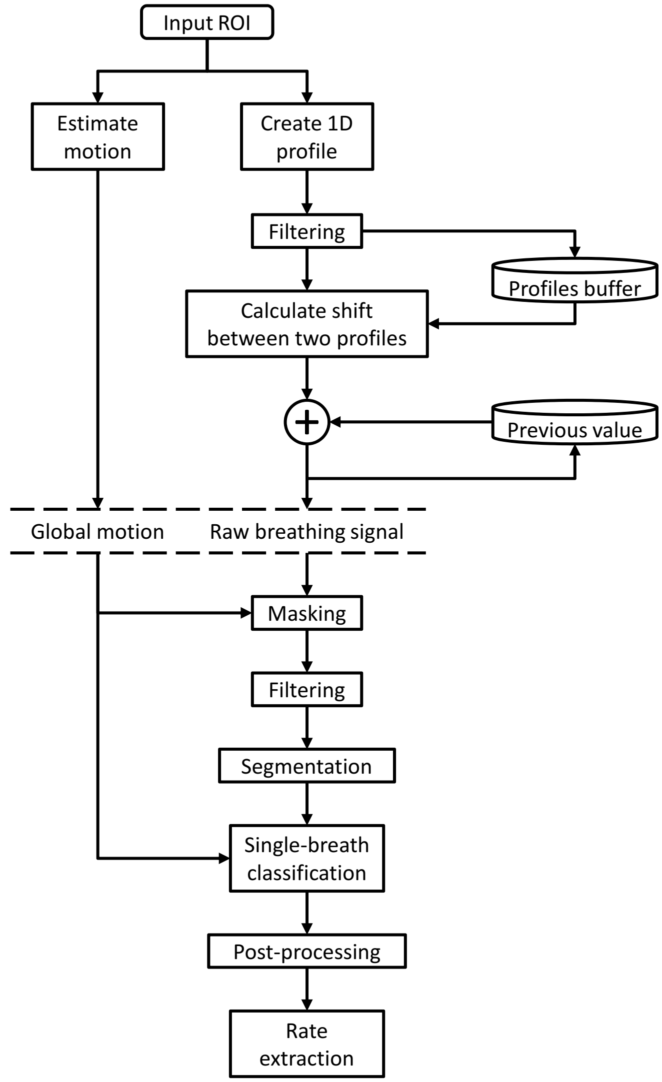

Fig. 1. Block diagram of the proposed algorithm

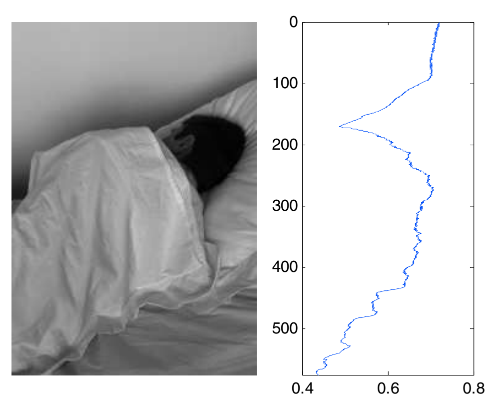

Fig. 2. Example video frame with the corresponding vertical profile

Cross-correlation (r) of the current (p_c) and previous (p_p) profiles is calculated by calculating the inverse Fourier transform $\left(\mathcal{F}^{-1}\right)$ of the product of Fourier transformed current (P_c) and complex conjugate of previous $\left(P_p^*\right)$ profiles

$$r = \mathcal{F}^{-1}\left(P_c P_p^*\right)$$

where $P_c = \mathcal{F}(p_c)$ and $P_p = \mathcal{F}(p_p)$. To suppress boundary effects, a Hann window is applied prior to the Fourier transform.

If we determine the maximum correlation of the shifted 1D vector of the current image with that of the previous image, the obtained translatory offset signal is the derivative of the (chest) position. Since the subtle respiration might appear as a sub-pixel motion, the exact peak location of the correlation function is determined by interpolation, using the samples from its direct neighborhood. At last, the position signal is obtained after numerical integration in the time domain.

B. Non-respiratory (global) motion detection

Given the high sensitivity of our algorithm to subtle breathing motion, it can be easily disturbed by large magnitude motion. This includes a person changing position or waving his/her arms in front of the chest. In such situations it is desired to detect and flag those time segments, to avoid extraction of the rate from a signal that consists of artefacts.

The motion detector is based on a flow algorithm, with meander scan and adaptive thresholding. One binary decision per image block is taken, labeling it as "moving" ('1') or "stationary" ('0'). The final motion signal is a ratio of blocks labeled as "moving" to the total number of blocks in the ROI.

C. Classification of individual breaths

The video processing part of the system has two outputs: the raw breathing signal and the global motion signal. Both extracted signals are used as input for a binary classifier of the raw breathing signal, to take into account only those periods with reliable breathing cycles. The implemented approach is based on a breath-to-breath classification scheme and consist of six essential steps as shown in Fig. 1.

First, the global motion signal is used to identify and exclude intervals containing non-respiratory motion. High and low frequency noise are filtered out by a bandpass filter before further processing, followed by a segmentation procedure which extracts single breath candidates. For each breath candidate, a feature vector is derived and presented to a previously trained decision tree, which assigns one of the two labels: 'good' or 'bad'. Fig. 3 shows an example sequence of the single-breath classification results, where green shading indicates that a breath candidate is labeled as 'good', and red shading indicates label 'bad'.

Only those segments of the signal identified as 'good' are used in rate calculations. This is the major advantage of the rate extraction algorithm in comparison to other frequency estimation methods like e.g. periodograms.

The classification algorithms are trained to maximize the decision accuracy by finding the right label for each breath candidate regardless of its true class. This implies identical

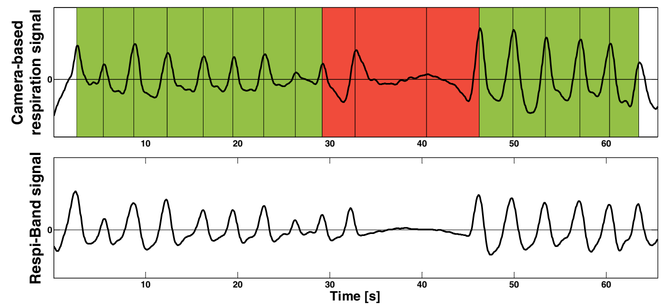

Fig. 3. Example of the classification

misclassification costs of false positive and false negative classification. However, it is assumed that with respect to the subsequent rate calculation, it is worse to include a misclassified 'bad' breath candidate, rather than not including a 'good' one. When information from the order of appearance of breath candidates is utilized in the post-processing stage, the inequality in misclassification costs is taken into account by changing selected 'good' labels into 'bad', but not vice versa. This leads to a significant increase in the classification precision and thus in the precision of calculated rates.

IV. Experimental Results

A. Set-up

During our experiments we used a custom-made data acquisition system that allows for synchronous recording of uncompressed video frames and reference signals (see Fig. 4). Recorded monochrome sequences had a resolution of 752x480 or 768x576 pixels and a frame rate of 20 Hz.

We have conducted two different sets of tests. In the first one, a mechanical breathing phantom was used for the purpose of having a well-defined setup. It was excited with a sinusoidal signal with a frequency ranging from 0.1 to 1 Hz (6-60 bpm). To assess the impact of the camera position and orientation, we recorded the phantom from five positions around the bed, each at three different heights. Additionally, we explored the effects of various lighting conditions, patterns and artificially induced shadows, as well as motion in the scene. In total, approximately 9800 breathing cycles were acquired.

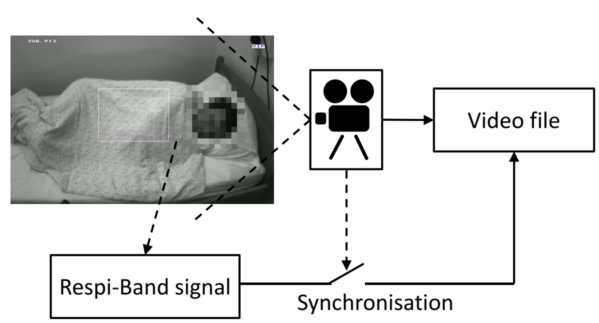

Fig. 4. Overview of the acquisition system

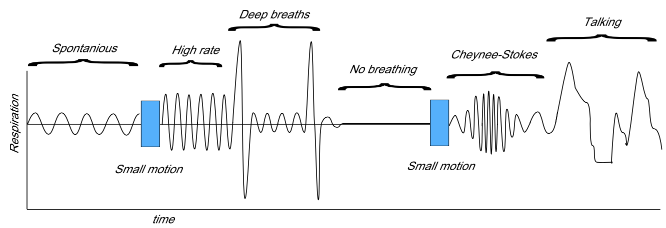

Fig. 5. Respiratory patterns simulated during video recordings

In the second test, five healthy subjects (1 female, 4 male) were asked to follow a sequence of different breathing patterns, orchestrated by a supervisor, while lying down in a bed. During the experiments subjects were covered with a blanket. Thoracic inductance plethysmography served as the reference signal for the subject's breathing effort. The experimental protocol is shown in Fig. 5.

B. Performance

In the experiment with breathing phantom, 7 features extracted from the raw breathing signal were used, achieving accuracy and precision of approx. 90 %. Post-processing strategies have significantly improved the precision, to a level of 95 %. The correlation coefficient between reference and video-based estimation was R = 0.97 (Fig. 6).

For the recordings of human subjects, using on average 8 features gave the accuracy and precision of approx. 85 %. Again, post-processing significantly enhanced performance, with up to 89 % accuracy and 95 % precision. The correlation coefficient was R = 0.98 (Fig. 7).

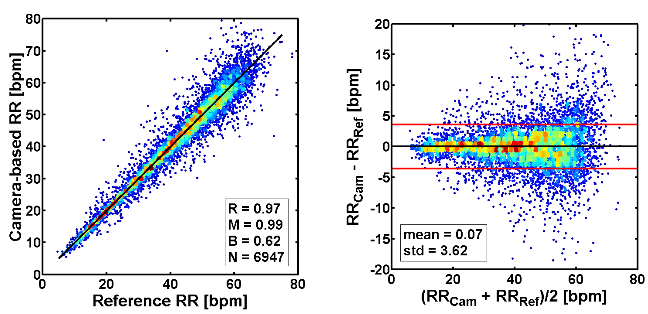

Fig. 6. Respiratory rate estimation based on phantom recordings

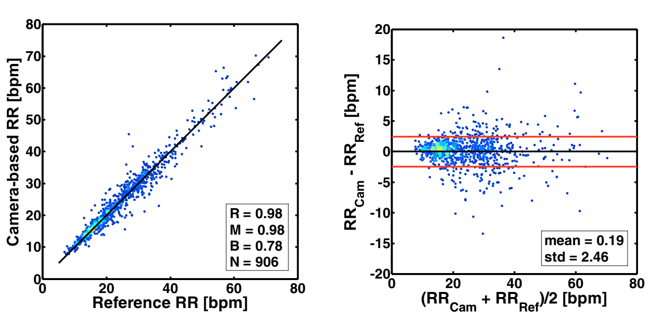

Fig. 7. Respiratory rate estimation based on human subject recordings

V. Conclusions

Camera-based monitoring is an attractive new sensing option which offers comfortable and convenient measurements of the respiration rate. Based on the experiments with healthy volunteers in a laboratory environment and a mechanical phantom mimicking realistic measurement conditions, reliable and robust respiration rate extraction using a smart sensor is feasible. The solution includes an efficient algorithm to extract the raw breathing signal and non-respiratory motion, followed by a classifier to include only the valid breaths in a subsequent rate estimation. A particular advantage of the camera solution is the access to context information, such as global motion, which turned out to be an essential input for the accurate classification.

The low complexity of our algorithm enables efficient real-time implementation even on a relatively simple platform using low-cost cameras.

References

[1] F. McGain, M. A. Cretikos, D. Jones, S. V. Dyk, M. D. Buist, H. Opdam, V. Pellegrino, M. S. Robertson, and R. Bellomo, "Documentation of clinical review and vital signs after major surgery," *Medical Journal of Australia*, vol. 189, no. 7, pp. 380–383, Oct. 2008.

[2] M. Folke, L. Cernerud, M. Ekström, and B. Hök, "Critical review of non-invasive respiratory monitoring in medical care," *Medical and Biological Engineering and Computing*, vol. 41, pp. 377–383, 2003.

[3] J. Muehlsteff, R. Pinter, and G. Morren, "Comparison of respiration rate monitoring with a low-cost doppler-radar sensor and inductive thorax-plethysmography," in *World Congress on Medical Physics and Biomedical Engineering, September 7-12, 2009, Munich, Germany*, ser. IFMBE Proceedings, vol. 25/7. Springer Berlin Heidelberg, 2009.

[4] R. Murthy, I. Pavlidis, and P. Tsiamyrtzis, "Touchless monitoring of breathing function," in *Engineering in Medicine and Biology Society, 2004. IEMBS '04. 26th Annual International Conference of the IEEE*, vol. 1, Sept. 2004.

[5] A. Abbas, K. Heimann, K. Jergus, T. Orlikowsky, and S. Leonhardt, "Neonatal non-contact respiratory monitoring based on real-time infrared thermography," *BioMedical Engineering OnLine*, vol. 10, p. 93, 2011.

[6] J. Penne, C. Schaller, J. Hornegger, and T. Kuwert, "Robust real-time 3D respiratory motion detection using time-of-flight cameras," *International Journal of Computer Assisted Radiology and Surgery*, vol. 3, pp. 427–431, 2008.

[7] K. Tan, R. Saatchi, H. Elphick, and D. Burke, "Real-time vision based respiration monitoring system," in *Communication Systems Networks and Digital Signal Processing (CSNDSP), 2010 7th International Symposium on*, Newcastle upon Tyne, UK, July 2010, pp. 770 –774.

[8] M. Frigola, J. Amat, and J. Pagès, "Vision based respiratory monitoring system," in *Proceedings of the 10th Mediterranean Conference on Control and Automation - MED2002*, Lisbon, Portugal, July 2002.

[9] H. Aoki, K. Koshiji, H. Nakamura, Y. Takemura, and M. Nakajima, "Study on respiration monitoring method using near-infrared multiple slit-lights projection," in *Micro-NanoMechatronics and Human Science, 2005 IEEE International Symposium on*, Nov. 2005.

[10] A. De Groote, M. Wantier, G. Cheron, M. Estenne, and M. Paiva, "Chest wall motion during tidal breathing," *Journal of Applied Physiology*, vol. 83, no. 5, pp. 1531–1537, 1997.

[11] K. Nakajima, A. Osa, S. Kasaoka, K. Nakashima, T. Maekawa, T. Tamura, and H. Miike, "Detection of physiological parameters without any physical constraints in bed using sequential image processing," *Japanese Journal of Applied Physics*, vol. 35, no. 2B, pp. L269–L272, 1996.

[12] A. Sibiryakov and M. Bober, "Real-time multi-frame analysis of dominant translation," in *Pattern Recognition, 2006. ICPR 2006. 18th International Conference on*, vol. 1, Hong Kong, Aug. 2006.