International Journal of Bioelectromagnetism *www.ijbem.org*
Vol. 12, No. 4, pp. 158 - 164, 2010

Automatic Signal Appraisal for Unobtrusive ECG Measurements

Johannes Schumm, Sebastian Axmann, Bert Arnrich, Gerhard Tröster
Electronics Laboratory, ETH Zurich, Zurich, Switzerland

Correspondence: Johannes Schumm, Electronics Laboratory, ETH Zurich, Gloriastrasse 35, 8092 Zurich, Switzerland.
E-mail: schumm@ife.ee.ethz.ch, phone +41 44 632 27 49, fax +41 44 632 12 10

Abstract. This paper focuses on facilitating ECG-measurements in our daily life. When measuring ambulatory ECG, the trade-off between signal quality and sensor comfort is obvious. Unobtrusive measurement systems do not achieve the same level of signal quality as systems using conventional wet ECG electrodes. To enable automatic ECG analyses, a continuous appraisal of the signal quality is essential. The presented work describes a method to transform the problem of appraising the quality of an ECG signal into a two-class classification problem. We therefore define an appropriate quality label and train a logistic regression model to predict the signal quality of an ECG signal. We show the working principle using a contactless ECG system incorporated into an airplane seat. The system is evaluated with twelve subjects performing airplane passenger activities. A correct signal quality appraisal rate of 92% is achieved in a leave-one-person-out cross-validation.

Keywords: Unobtrusive ECG Measurement, Signal Appraisal, ECG Quality Modeling

1. Introduction

Current demographic changes lead to higher healthcare costs in many countries. Continuous ambulatory health monitoring, especially the recording and analysis of the cardiovascular system's response are therefore increasingly important [Ottenbacher et al., 2008]. The sensors need to be attached in an unobtrusive way to promote acceptance and hence usage of ambulatory ECG systems. Thus, common systems worn on the body that require wet electrodes attached to the chest are not feasible in this context. Different research activities deal with the integration of ECG sensors into clothing or with the environment. These approaches are facing a trade-off between sensor placement optimized for comfort and signal quality.

With regard to this trade-off, not only the development of the sensors but also the automatic interpretation of the measured signals needs to be addressed. Automatic analyses often depend on the reliable detection of R-peaks. However, when the signal quality is poor the peak-detection algorithms may identify wrong R-peaks. This renders the automatic analysis of ECG impossible.

This problem is addressed by different researchers trying to implement an automatic artifact compensation for the ECG signal. But, as indicated by Such et al., the compensation of artifacts by e.g. adaptive filters or decorrelation methods can lead to a false sense of safety and has to be carried out carefully [Such et al., 2006]. Moreover, if the ECG signal is not reconstructable, compensation is not possible at all. For these reasons, we propose applying an automatic appraisal of the signal quality instead of compensation. We describe a method that transforms the quality appraising of an ECG signal into a two-class classification task. We identify a Quality Label (QL) that indicates parts with high and low signal quality based on the R-peaks. In order to determine this QL we use features from the ECG signal itself, as well as additional sensors to establish a logistic regression model.

As this work is situated in the European SEAT-project (http://www.seat-project.org), the working principle is demonstrated using a contactless ECG system integrated into an airplane seat. The motivation behind the SEAT project is the continuous increase in air-travel. Approximately 1.5 to 2 billion passengers travel with commercial airlines each year [Smith, 2008]. Taking the demographic trend of an aging population into consideration, the likelihood of an older passenger population increases and thus airplanes will soon be a part of our environment in which people may need assistive support [Drummont et al., 2002][Goodwin, 2000].

2. Related Work

In recent years, several research approaches have dealt with the development of cardiovascular sensory systems, enabling continuous ambulatory health monitoring. The trade-off between signal quality and acceptable sensor comfort for the user is highlighted by Such *et al.* in [Such et al., 2006]. The authors point out that an indication of signal confidence would often be preferable over an automatic compensation algorithm. Automatic compensation can lead to plausible but incorrect features and, therefore, to a false sense of safety. Two different approaches can be distinguished concerning automatic artifact detection. The single parameter approach only uses the characteristics of the sensor signal itself to detect artifacts whereas the multiparameter approach uses additional sensor channels to identify these artifacts. An example of the single parameter approach is an algorithm that segments the signal into "artifact-free" intervals before performing a signal analysis.

A multiparameter approach for ECG recording with dry electrodes is presented by Ottenbacher *et al.* [Ottenbacher et al., 2008]. They simultaneously measure the ECG and the electrode-skin impedance in order to detect artifacts by changes in the impedance signals. For validation purposes they compare the extracted R-peaks with a Ground-Truth ECG signal measured with wet electrodes. Another example of a multiparameter approach is the use of accelerometers to remove artifacts from the ECG signal of a photoplethysmograph sensor [Gibbs et al., 2007]. The authors use an adaptive filter to compensate for artifacts induced by body motion.

Concerning the integration of cardiovascular monitoring systems into a seat, the contactless measurement of the ECG represents a promising approach in recent publications. Two different research activities in this direction were identified. Steffen *et al.* designed a system for the capacitive measurement of ECG of a sitting person [Steffen et al., 2007]. Another method is shown by Zakrezwski *et al.* [Zakrezwski et al., 2006]. They present the possibility of detecting the heart rate using a radar system. Both measurement principles are unobtrusive but also sensitive to movement artifacts.

3. Methods

Our objective is to predict the signal quality of a contactless ECG recording using a suitable model. In order to achieve this goal the following four tasks need to be addressed: (1) measurement setup, (2) determination of the ground truth quality label, (3) computation of features incorporated into the model and (4) modeling and prediction of the quality label.

3.1 Measurement setup

To unobtrusively measure the ECG, we integrated a contactless capacitive ECG system into the backrest of an airplane seat, shown in Fig. 1 on the left side [Schumm et al., 2010]. The system, referred to as "Contactless-ECG", is a research prototype developed by RWTH Aachen University [Steffen et al., 2007]. It measures the ECG capacitively without direct skin contact. As it is integrated into the seat, it does not disturb the user, but is sensitive to body movements. It only delivers sufficient signal quality if sufficient and constant contact pressure between the upper part of the body and the back of the seat is maintained.

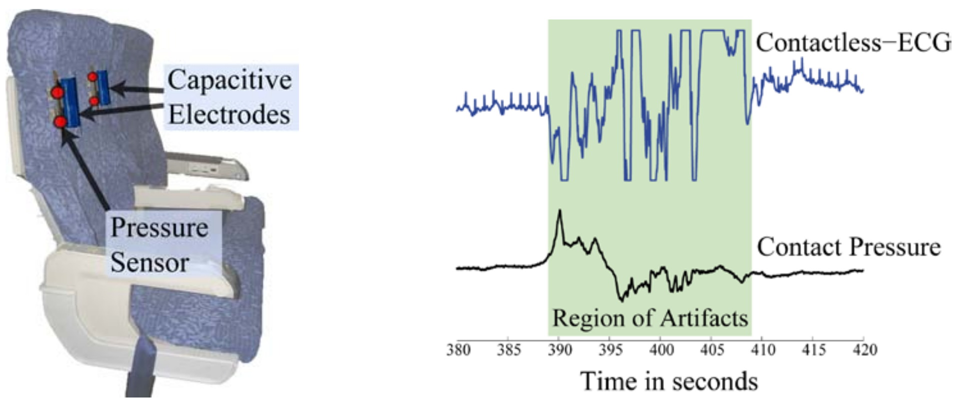

Figure 1. *Left: Airplane seat showing the positions of the contactless capacitive electrodes and pressure sensors. Right: ECG signal and sum of the four pressure sensor signals indicating the overall contact pressure. Changes in the overall contact pressure indicate that the passenger is moving which results in a disturbed ECG signal.*

We therefore decided to incorporate additional pressure sensors into the seat to appraise the ECG signal quality. To find feasible locations for the pressure sensors, the contact pressure at the back of the seat was recorded with a pressure mat from Tekscan (www.tekscan.com) in parallel with the Contactless-ECG signal. Based on a visual inspection of both signal modalities, we decided to place one pressure sensor at the top and another one at the bottom of each electrode as depicted in Fig. 1 on the left side.

An example of a signal of the ECG system and the sum of the four pressure sensor signals is shown in Fig. 1 on the right side. Changes in the overall contact pressure indicate that the passenger is moving which results in a corrupted ECG signal.

3.2. Determination of the ground truth quality label

To train and test a model for appraising the signal quality, we first need to determine the true signal quality. Therefore, we define a Quality Label (QL) for an ECG signal, based on the most characteristic feature of the ECG, the R-peaks. For the automatic extraction of the R-peaks, a state-of-the-art algorithm proposed by Hamilton in 2003 is used (http://www.eplimited.com). The algorithm is based on the work published by Hamilton and Tompkins in [Hamilton et al., 1986].

To calculate the QL, we need to record the ECG signal with the Contactless-ECG system together with a reliable ground truth device. For each R-peak detected in the ground truth ECG signal it is determined whether a corresponding R-peak is detected in the Contactless-ECG signal. This comparison between the two R-peaks is based on the ANSI/AAMI Norm for judging R-peak detection algorithms [ANSI-AAMI Norm, 2003]. Fig. 2 depicts for an artificial ECG signal the R-peak comparison for calculating the QL. The red boxes at the bottom indicate the timestamps where peaks were detected in the ground truth signal. Correct hits (True Positives), a missing peak (False Negative) and an erroneously detected nonevent (False Positive) are shown. The interval X in which a peak is still counted as a correct peak is chosen as 150ms according to the ANSI/AAMI Norm. Therefore, when more than one peak is inside the interval X or when a peak is outside the interval this results in an error.

As a consequence, the QL indicates whether the closest R-peak in the ground truth ECG signal was detected correctly or not for each sample in the signal. It represents a class label (0 or 1) of the signal quality (insufficient or sufficient) for each sample.

Drawbacks of this method are errors of the Hamilton peak detection algorithm (e.g. the so called “lucky hits”). Even if the measured signal represents pure noise, the R-peak detection algorithm still identifies peaks. In this case it may happen that detection occurs by chance at the correct position. This sets the QL to one, indicating a sufficient signal quality even though the signal is corrupted at this position. On the other hand the Hamilton peak detection algorithm requires an adaptation time after signal disturbances. During this adaptation time (approx. 5 seconds) the QL is set to one, although R-peaks are clearly visible. Therefore, a post processing step was performed to manually correct all obvious errors.

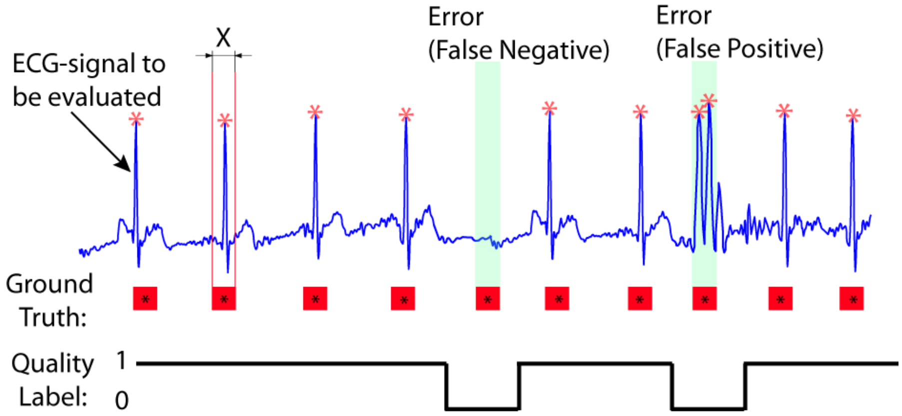

Figure 2. *Example showing the calculation of the Quality Label. The stars at the top indicate the positions where R-Peaks in the Contactless-ECG signal were detected whereas the red boxes at the bottom indicate the positions where R-peaks in the ground truth ECG signal were detected. Errors are indicated with green bars.*

3.3. Computation of features incorporated into the model

To train a model for predicting the QL, features were calculated every 0.8s using the ECG signal and the sum of the four back-pressure signals at equidistant points. All features were computed using a window size of one second and a shift of zero, one, two, three and four seconds into the past. We calculated three basic features for the two signal modalities: the mean value, the variance and the mean of the gradient of the signal in the corresponding window. For the ECG, we also calculated the percentage of samples that are equal to the maximum value of the ADC converter. This feature indicates if the ADC converter has reached saturation. With this procedure we generated a total of 35 features every 0.8 seconds

3.4. Modeling and prediction of the Quality Label

The features from the ECG and pressure signals, together with the corresponding ground truth label, enable us to train a model for predicting the QL. We have selected the logistic regression model that belongs to the top ranked algorithms in terms of mean errors rates [Lim et al., 2000]. To train the model, we applied the forward stepwise conditional logistic regression (SPSS version 17.0). We then applied the default threshold of 0.5 to classify the output of the logistic regression model as 0 and 1 respectively.

4. Experiment

As the presented work is embedded in the SEAT-project, we investigated the following typical passenger activities in our experiment: Being entertained, working, reading, sleeping and eating. The chosen activities were based on the statistical distribution of passengers' activities during a long haul flight [SEAT, 2006]. The different activities and the percent of time that they occur on a long haul flight are depicted in Fig. 3. In total, 12 subjects participated in the experiment (8 male, 4 female, mean age: 26.5 years, mean weight: 67kg). Each activity was performed for approximately 10 minutes, resulting in an overall dataset of almost 10 hours of recordings. During the activity "Entertainment", the subjects watched a relaxing movie on a laptop placed on the front-table. For the activity "Working", the subjects had to perform normal office tasks, such as writing and formatting a text. For the activity "Reading", the subjects were reading a journal and during "Sleeping" the subjects were asked to lean back, close their eyes and relax. During the last activity "Eating", the subjects ate a piece of cake and drank some water.

During all activities, the ECG was measured using the Contactless-ECG system and with the Mobi (www.tmsi.com) as ground truth device. The Mobi uses wet electrodes attached to the chest for measuring the ECG signal. In addition, the pressure signals were recorded synchronously.

Figure 3. *Typical passenger activities and their temporal distribution during a long-haul flight, the remaining 2.4% are other activities [SEAT, 2006].*

5. Results

All results presented in this chapter are based on a leave-one-person-out cross-validation performed with twelve subjects. For each run of the cross-validation, the logistic regression generated a slightly different model. On average 17 features were selected, of which 14 were contained in all models. This indicates a high stability of the derived models.

In Table 1, the sensitivity, specificity and the overall accuracy are presented. The sensitivity indicates the proportion of correctly samples with QL=1, whereas the specificity indicates the proportion of correctly samples with QL=0. With our data the logistic regression achieves a higher sensitivity than specificity.

Table 1. *Achieved sensitivity (SE), specifity (SP), and accuracy (ACC) for the different activities.*

	Overall	*Entertain*	*Work*	*Read*	*Sleep*	*Eat*
SE	*93 %*	*99 %*	*89 %*	*95 %*	*98 %*	*86 %*
SP	*84 %*	*76 %*	*88 %*	*84 %*	*80 %*	*90 %*
ACC	*92 %*	*96 %*	*88 %*	*92 %*	*95 %*	*88 %*

Table 2. *The Original QI indicates the pure signal quality; the Model QI indicates the signal quality for the parts, where the classification predicted high signal quality; Signal Length is the time predicted as high signal quality; the Number of Blocks ≥ 2min usable for HRV analysis; all results are upscaled to a 12h flight*

	Overall	*Entertain*	*Work*	*Read*	*Sleep*	*Eat*
Original QI	*73 %*	*89 %*	*41 %*	*68 %*	*84 %*	*34 %*
Model QI	*94 %*	*97 %*	*84 %*	*93 %*	*96 %*	*81 %*
Signal Length	*73 % (8.7h)*	*90 % (158min)*	*44 % (25min)*	*70 % (82min)*	*86 % (226min)*	*36 % (32min)*
Number of Blocks ≥ 2min	*57*	*18*	*1*	*7*	*31*	*0*

For the more “calm” activities (entertainment, reading and sleeping) the overall performance is higher than for the more “active” activities (working and eating). On average, we achieve a correct signal quality appraisal rate of 92% for unseen subjects. To illustrate the effectiveness of the proposed method for use in an airplane application we define the Quality Index (QI) as the ratio of high quality QLs (QL=1) to all QLs over a certain time period.

Table 2 shows the QI achieved for each activity in the first row. Please note that all results in Table 2 are scaled up to a long-haul flight using the time distribution indicated in Fig. 3. The Original QI reflects the overall signal quality achieved using the Contactless-ECG system. With the proposed modeling technique we aim at identifying regions of high signal quality. A perfect model would show a QI of 100% during the predicted high signal quality regions. The second row in Table 2 shows the achieved QI of the regions where our model indicated a high signal quality. For a long-haul flight this results in 94% of sufficient signal quality when considering all activities. The third row shows the percentage of time where a high signal quality was predicted. The number in brackets is the corresponding time for each activity. For a 12 hour flight, this results in 8.7 hours with an overall QI of 94%.

Nevertheless, consecutive R-peaks are needed to apply sophisticated analyses of the ECG signal. For example, to obtain the heart rate variability (HRV), a minimum block of two minutes is needed [Malik et al., 1996]. Therefore, we calculated the amount of blocks where the model indicated a consecutive high signal quality (QL=1) lasting two minutes or more. The last row in Table 2 indicates the number of those blocks. This results in 57 usable blocks (~4.7h) for HRV analyses.

6. Summary and Outlook

We presented a method that deals with low signal quality in unobtrusive ECG sensing to enable automatic medical decision making. The proposed approach transforms the appraisal of ECG signal quality into a two-class classification task.

To achieve person-independent results we applied a leave-one-person-out cross-validation scheme. The resulting models are able to distinguish between sufficient and insufficient signal quality parts with an average accuracy of 92%. We observe a class-dependent signal quality and modeling performance. The lower signal quality for the activities “Working” and “Eating” are due to the fact that the subjects sometimes lent forward, resulting in less contact pressure between the body and the electrodes. To enable analyses of the ECG signal such as HRV analysis, we estimated the number of blocks with consecutive high signal quality (QL=1). In future work it is promising to combine the presented work with a plausibility analysis used for artifact removal in medical ECG measurements.

Acknowledgements

This project is funded by the EU research project SEAT (www.seat-project.org), contract number: 030958, all views here reflect the author's opinion and not that of the commission.

References

ANSI-AAMI Norm: Association for the Advancement of Medical Instrumentation, "Testing and reporting performance results of cardiac rhythm and ST segment measurement algorithms," ANSI-AAMI EC57:1998/(r) 2003.

Drummond R, Drummond A., "On a wing and a prayer: medical emergencies on board commercial aircraft," Canadian journal of emergency medical care, vol. 4, no. 4, pp 276-289, 2002.

Gibbs P., Asada H., "Reducing motion artifact in wearable bio-sensors using MEMS accelerometers for active noise cancellation," in Proceedings of the American Control Conference, pp. 1581-1586, 2005.

Goodwin T., "In-flight medical emergencies: an overview," British Medical Journal, vol. 321, no. 7272, pp. 1338-1341, 2000.

Hamilton P.S., Tompkins W. J., "Quantitative investigation of QRS detection rules using the MIT/BIH arrhythmia database," IEEE Transactions on Biomedical Engineering, vol. BME-33 , no. 12, pp. 1157-1165, 1986.

Lim T.S., Loh W.Y., Shih Y.S., "A comparison of prediction, accuracy, complexity, and training time of thirty-three old and new classification algorithms," Machine Learning, vol. 40, no. 3, pp. 203-228, 2000.

Malik M., for the Task Force of the ESC and NASPE, "Heart rate variability: standards of measurement, physiological interpretation and clinical use," European Heart Journal, vol. 17, no. 3, pp. 354-381, 1996.

Ottenbacher J., Kirst M., Jatoba L., Grossmann U., Stork W., "An approach to reliable motion artifact detection for mobile long-term ECG monitoring systems using dry electrodes," IV Latin American Congress on Biomedical Engineering, vol. 18, no. 1, pp. 440-443, 2008.

Schumm J., Setz C., Bächlin M., Bächler M., Arnrich B., Tröster G., "Unobtrusive Physiological Monitoring in an Airplane Seat," Pervasive and Ubiquitous Computing, 2010.

SEAT: European Project (Nr. 030958), "Deliverable D5.1: IFE and Comfort in Aircraft Cabin, State of the Art," 2006.

Smith L., "An otolargyngologist's experience with in-flight commercial airline medical emergencies: three case reports and literature review," American Journal of Otolargyngology – Head and neck Medicine and Surgery, 2008.

Steffen M., Alexandrovicz A., Leonhardt S., "Mobile non-contact monitoring of heart and lung activity," IEEE Transactions on Biomedical Circuits and Systems, vol. 1, no. 4 pp. 250-257, 2007.

Such O., Muehlsteff J., "The challenge of motion artifact suppression in wearable monitoring solutions," 3rd IEEE/EMBS International Summer School on Medical Devices and Biosensors, pp. 49-53, 2006.

Zakrzewski M., Kolinummi A., Vanhala J., "Contactless and unobtrusive measurement of heart rate in home environment," 28th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 2060-2063, 2006.