Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 115 (2017) 359–366

Procedia Computer Science

www.elsevier.com/locate/procedia

7th International Conference On Advances In Computing & Communications, ICACC 2017, August 22-24, 2017, Cochin, India

Stress Detection in Working People

Sriramprakash.S *, Prasanna Vadana. D, O. V. Ramana Murthy

[a]Department of Electrical and Electronics Engineering Amrita School of Engineering Coimbatore, India

Abstract

Stress detector classifies a stressed individual from a normal one by acquiring his/her physiological signals through appropriate sensors such as Electrocardiogram (ECG), Galvanic Skin Response (GSR) etc,. These signals are pre-processed to extract the desired features which depicts the stress level in working individuals. Support Vector Machine (SVM) and K-Nearest Neighbour (KNN) are investigated to classify these extracted feature set. The result indicates feature vector with best features having a strong influence in stress identification. An attempt is made to determine the best feature set that results in maximum classification accuracy. Proposed techniques are applied on benchmark SWELL-KW dataset and state-of-art results are obtained.

Peer-review under responsibility of the scientific committee of the 7th International Conference on Advances in Computing & Communications.

Keywords: Stress;ECG;GSR;Machine learning;SVM(support vector machines);K-Nearest Neighbour (KNN);

1. Introduction

Stress management systems play a significant role to detect the stress levels which disrupts our socio economic lifestyle. As World Health Organization (WHO) says, Stress is a mental health problem affecting the life of one in four citizens [1]. Human stress leads to mental as well as socio-fiscal problems, lack of clarity in work, poor working relationship, depression and finally commitment of suicide in severe cases. This demands counselling to be provided for the stressed individuals cope up against stress. Stress avoidance is impossible but preventive actions helps to overcome the stress [2]. Currently, only medical and physiological experts can determine whether one is under depressed state (stressed) or not. One of the traditional method to detect stress is based on questionnaire [3]. This method completely depends on the answers given by the individuals, people will be tremulous to say whether they are stressed or normal. Automatic detection of stress minimizes the risk of health issues and improves the welfare of the society. This paves the way for the necessity of a scientific tool, which uses physiological signals thereby automating the detection of stress levels in individuals.

* Sriramprakash.S Tel.: +919486619140.
E-mail address: sriramprakashkpr@gmail.com

1877-0509
Peer-review under responsibility of the scientific committee of the 7th International Conference on Advances in Computing & Communications
10.1016/j.procs.2017.09.090

The organization of this paper is as follows. Section II gives overview on related state of art. Section III describes the framework carried out in this research. A simple stress prediction algorithm is also presented in this section. Section IV discusses the obtained results on the application of SWELL-KW dataset. Finally conclusions are drawn in section V.

2. Related Literature

Stress detection is discussed in various literatures as it is a significant societal contribution that enhances the lifestyle of individuals. Ghaderi et al. [4] analysed stress using Respiration, Heart rate (HR), facial electromyography (EMG), Galvanic skin response (GSR) foot and GSR hand data with a conclusion that, features pertaining to respiration process are substantial in stress detection. Maria Viqueira et al.[5] describes mental stress prediction using a standalone stress sensing hardware by interfacing GSR as the only physiological sensor . David Liu et al.[6] proposed a research to predict stress levels solely from Electrocardiogram (ECG).

Multimodal sensor efficacy to detect stress of working people is experimentally discussed in [7]. This employs the sensor data from sensors such as pressure distribution, HR,Blood Volume Pulse (BVP) and Electrodermal activity (EDA). An eye tracker sensor is also used which systematically analyses the eye movements with the stressors like Stroop word test and information related to pickup tasks.The authors of [8] performed perceived stress detection by a set of non-invasive sensors which collects the physiological signals such as ECG [9], GSR, Electroencephalography (EEG), EMG, and Saturation of peripheral oxygen (SpO2). Continuous stress levels are estimated using the physiological sensor data such as GSR, EMG, HR, Respiration in [10]. The stress detection is carried out effectively using Skin conductance level (SCL), HR, Facial EMG sensors [11] by creating ICT related Stressors.

Automated stress detection is made possible by several pattern recognition algorithms. Every sensor data is compared with a stress index which is a threshold value used for detecting the stress level. The authors of [5] collected data from 16 individuals under four stressor conditions which were tested with Bayesian Network, J48 algorithm and Sequential Minimal Optimization (SMO) algorithm for predicting stress. Statistical features of heart rate [12], GSR [3], frequency domain features of heart rate [13] and its variability (HRV) [6], and the power spectral components of ECG [7] were used to govern the stress levels. Various features are extracted from the commonly used physiological signals such as ECG, EMG, GSR, BVP etc., measured using appropriate sensors and selected features are grouped into clusters for further detection of anxiety levels . In [8], it is concluded that smaller clusters result in better balance in stress detection using the selected General Regression Neural Network (GRNN) model. This results in the fact that different combinations of the extracted features from the sensor signals provide better solutions to predict the continuous anxiety level [14]. Frequency domain features like LF power (low frequency power from 0.04 Hz to0.15Hz), HF power (High frequency power from 0.15Hz to 0.4 Hz) , LF/HF (ratio of LF to the HF). and time domain features like Mean , Median, standard deviation of heart signal are considered for continuous real time stress detection in [10].

Classification using decision tree such as PLDA is performed using two stressors namely pickup task and stroop based word test wherein the authors concluded that the stressor based classification proves unsatisfactory [15]. In 2016, Gjoreski et al. created laboratory based stress detection classifiers from ECG signal and HRV features [10]. Features of ECG are analysed using GRNN model to measure the stress level [8]. Heart rate variability (HRV) features and RR (cycle length variability interval length between two successive Rs) interval features are used to classify the stress level [12]. It is noticed that Support Vector Machine (SVM) was used as the classification algorithm predominantly due to its generalization ability and sound mathematical background [16] Various kernels were used to develop models using SVM and it is concluded in [17] that a linear SVM on both ECG frequency features and HRV features performed best, outperforming other model choices [18].

3. System Overview

The proposed model for stress detector as shown in Fig.1 is designed to enhance the generalization ability in stress detection of working individuals. Physiological sensor data plays a vital role in stress detection. GSR detects the different conductance level of the skin when a person is in stress. When a person is stressed, the nervous system responds quickly by releasing sweat. The GSR electrodes are placed under the fingers which act as terminals of resistance.On the contrary, ECG is the one of the dominant identification of long term and short term stress detection. ECG is the measuring of electrical activity of the heart by inferring features of heart rate (variability)[6].

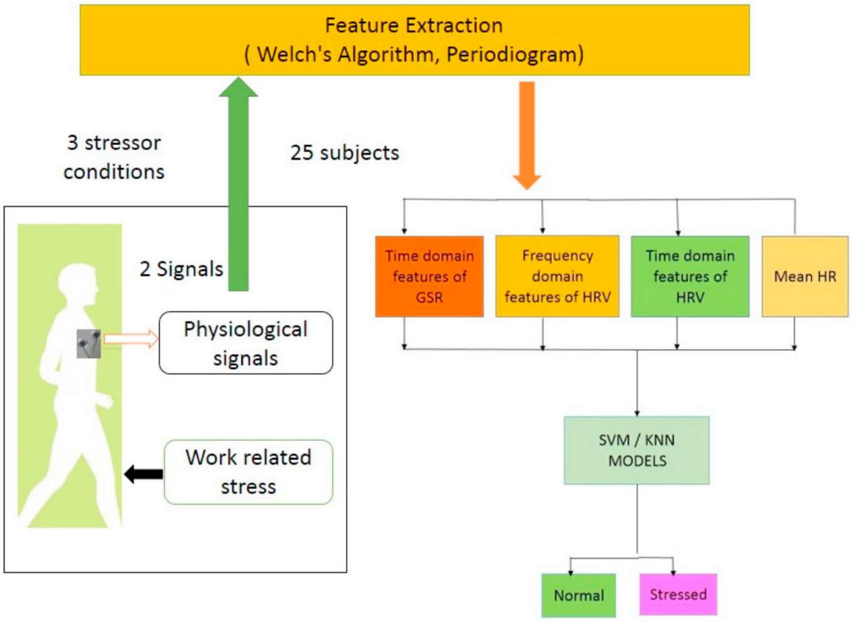

Fig. 1. Overall Framework: Stress Detection-Feature Extraction- Classification

The desired features are extracted from the ECG and GSR sensors which are applied as inputs to non-parametric and parametric classifiers. The non-parametric classifier used here is the K-Nearest Neighbour (KNN) [11]. KNN is applied for classifying the stress levels which purely depends upon Euclidean distance between nearest neighbour of training and testing feature vectors. The parametric classifier used here is the Support Vector Machine (SVM) where classification is performed using two types of kernel functions: (i) linear kernel and (ii) the Radial Basis Function (RBF) which is a non-linear kernel [16]. This algorithm finds superlative hyperplane which reflects to maximize the boundary between two sample groups. The feature vectors adjacent to the ideal hyperplane are called support vectors. SVM results in an optimized unique classifier for the given set of data [16].

In this work, a large scale multimodal action data titled SWELL-KW [19] dataset generated from knowledge workers, described by SaskiaKoldijket al. [18] is used for the development of intelligent stress detection algorithm. The data are collected in context of working people by creating stressors consisting of computer logging details, expression from face etc. and is evaluated, using Kinect 3D sensor for body postures, heart rate and its variability features from ECG sensor and skin conductance level from body sensors. The data was collected in three intervals of circa 45 mins under three different conditions namely,

- To generate a normal signal
  - Duration : 45 min.
  - Relaxation time: 6 mins (Relaxation is the time taken to start generating the actual data)
- Stressed signals are generated under two different stressors which are mentioned below

- Email interruption
  * Duration : 45 mins
  * Relaxation time: 6 mins
- Time Pressure
  * Duration : 30 mins
  * Relaxation time: 6 mins

This dataset focusses on high task load stress in-terms of mental demand, frustration, and temporal demand in working professionals [18]. The raw and pre-processed signals are available in the SWELL-KW dataset.

4. Feature Extraction

Generally physiological signals used for analysis are often pigeonholed by a Non-stationary time performance, hence features of time and frequency exemplifications are desirable.The feature extraction algorithm provides important information of the original signal into a more condensed lower dimension feature vector. The extracted features explicitly gives the stress index of the physiological signals.The ECG signal is directly assessed by using commonly used peak finder algorithm [20] to obtain the R-R interval. The peaks are counted in a one minute time frame and this represents the heart rate of the individual. The power spectral density of the HRV features from the ECG signal extracted using Welch algorithm dominates the stress detection. The raw ECG is further pre-processed using window average method [19]; pre-processed GSR signal is used to compute the below mentioned features. A total of 17 different features are identified for further classification of stress levels. Table 1 lists the total number features used for classification.

Table 1. Feature Extracted

S.No	Features	Abbreviation	Source& Domain
1	Mean HR	Mean Heart Rate	
2	Median HR	Median Heart Rate	
3	MAD HR	Mean Absolute Deviation	Heart rate Statistical Features
4	STD HR	Standard Deviation of Heart Rate	
5	RMSSD	Root mean square of successive difference in distance	
6	AVNN	Average of NN intervals	
7	SDANN	Standard Deviation of Average of NN intervals	
8	SDNN	Standard Deviation of NN intervals	HRV statistical Features
9	NN50	Number of pairs of successive NN intervals differ by 50ms	
10	PNN50	The ratio of NN50 to the total number of NNs	
11	LF	Low frequency power from 0.04 HZ to 0.15Hz	
12	HF	High frequency power from 0.15 HZ to 0.4Hz	HRV frequency domain features
13	LF/HF	Ratio of LF to HF	
14	Mean GSR	Mean of GSR	
15	Med GSR	Median of GSR	
16	MAD GSR	Mean absolute deviation of GSR	GSR Statistical Features
17	STD GSR	Standard deviation of GSR	

4.0.1. Description of Features

The description of feature gives the detailed overview of selection of Feature to predict stress.

- ***HR and GSR Statistical Features:*** Statistical features are considered from ECG and GSR signal.HR is the rate of existence of cardiac beats per minute. The Heart rate from ECG signal is calculated by taking duration between RR intervals and dividing it by per minute. The mean, median, standard deviation are considered to be first order statistical features of physiological signals. The variation or distribution of a signal is computed

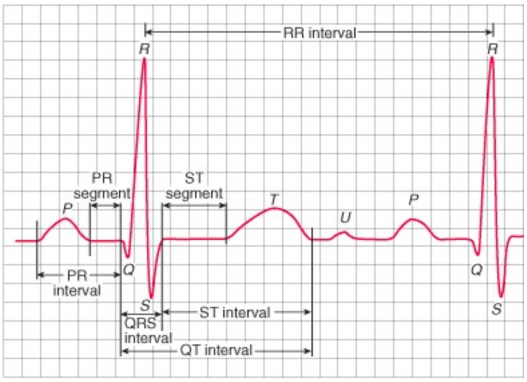

Fig. 2. RR Interval

using standard deviation.

- ***HRV Statistical Features:*** The HRV is defined as the temporal variation between sequences of consecutive heartbeat intervals. The RR interval is described as the period between two adjacent R waves. R-R interval is given in Fig.3. HR and RR intervals are pondered to be reciprocals of each other, measurement unit of HR is Beats per minute (BPM), whereas RR intervals is milli seconds (ms).

  The HRV indices defined below are the measures of Variability in RR peaks only. A forthright and valuable metric of Heart rate (variability), The Root mean square of successive difference in distance (RMSSD), by computing the square root of the mean of the squared differences between successive NN intervals over 24 hour time intervals. The other remarkable statistical features are SDNN and its variants, PNN50 and NN50. SDNN, is stated as normal deviation of NN (Normal RR) intervals. In calculating SDNN, any normal RR interval that instigates or ends with a Ventricular Premature Complexes (VPCs) and Atrial Premature Complexes (APCs) which are additional misperceive in evaluating autonomic regulation of HR is merely obliterated from the sequence. HRV is considered to be a non-stationary process. The variants of SDNN are extracted by dividing temporal variation into 5 minute divisions, AVNN index is the mean of all 5-minute divisions and SDANN index is the standard deviation of mean of 5-minute divisions of all Normal to Normal RR (NN) intervals.

  The short term variations are epitomised by the below declared features which depends entirely on comparisons between successive heart beats per minute. The NN50 exemplifies number of pairs of successive NN intervals differ by 50ms. PNN50 reckons the proportion of differences between number of pairs of successive NN intervals differ and are greater than 50ms over a 24 hour time.

- ***Frequency Domain Features of HRV:*** The magnitude of autonomous heart beat in temporal variations helps in computing time domain features, whereas frequency domain features offers spectral arrangement of physiological signal. Frequency domain features generally be determined by Power spectral density (PSD) of NN intervals. Non-parametric PSD analysis is deliberated using Welchs method. Spectral density of power delivers how the power is distributed as a frequency function.

  The spectral analysis is carried using following procedure

  – The ECG signal is split into data segments, with overlapping segments of length (L/2)
  – The hamming window is applied to the overlapped segments
  – The periodogram is calculated by Fast Fourier transform, and it is averaged which results in array of frequency and power

The spectral components are further congregated into two bands: Low frequency (LF) and High frequency (HF). The aggregate of power spectral in the Low (0.040.15Hz frequency range) and high (0.150.40Hz range) bands reflects the sympathetic variations and vagal variations of heart activity. The ratio of (LF/HF) is used as significant observable index in detecting stress.

5. Results of Classification

Classifying the pattern of features into the apposite class is the chief role of simulated classifiers. Classification attempts to analyse the best feature, to improve the computational swiftness of classification. As a first step, every feature is analysed for stress using the classifiers. The features are analysed separately to find the best features which are responsible for stress analysis. The extracted features are well enough to predict the stress.

5.1. Individual feature Analysis:

The extracted features mentioned in Table 1 are considered as single inputs and classifier models are generated to analyse the best dominant features to predict stress. Simulations are carried out with different SVM kernels (RBF, linear, polynomial and sigmoid) which showed significant results for RBF kernel with 10-fold cross validation. Individual feature performance is shown in the Table 2 where the best features with the following criteria are highlighted.

- *Precision analysis*
  - Individual feature performance
  - Sorting the best feature according to classification accuracy greater than a chosen threshold accuracy.
- Assortment analysis
  - Splitting all features into various groups based on classification accuracy
  - Feature combination by analysing best dominant features

Table 2. Individual feature performance and best feature selection

S.No	Features	KNN	SVM
1	***Mean HR***	66.59%	79.40%
2	Median HR	37.21%	70.67%
3	***MAD HR***	73.61%	85.85%
4	STD HR	43.39%	70.69%
5	***RMSSD***	68.92%	76.78%
6	***AVNN***	63.38%	80.83%
7	***SDANN***	73.79%	75.48%
8	SDNN	49.93%	66.87%
9	***NN50***	69.68%	86.36%
10	***pNN50***	75.67%	77.48%
11	***LF***	72.56%	83.57%
12	***HF***	76.96%	81.47%
13	LF/HF	41.42%	57.73%
14	***Mean GSR***	78.62%	79.66%
15	***Med GSR***	68.23%	80.10%
16	MAD GSR	53.28%	68.37%
17	***STD GSR***	72.57%	82.40%

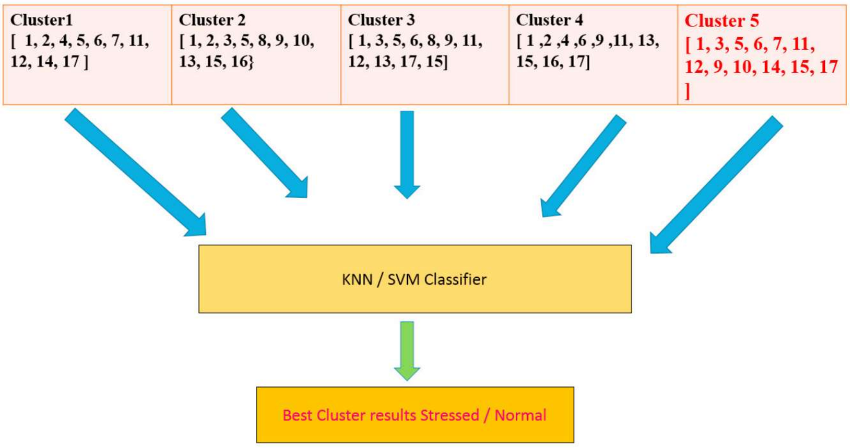

Fig. 3. The cluster combination of features

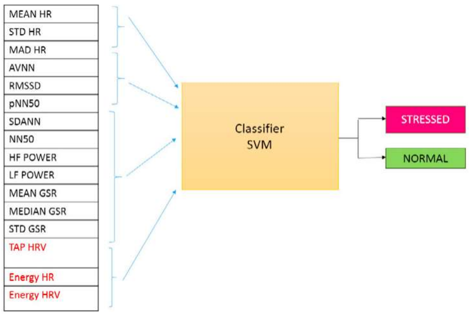

Fig. 4. Consideration of Features For Improved Classification

5.2. Feature combination Analysis

- ***Dominant feture identification:*** Assessing individual feature using classifier contributes to the separation (Normal vs. stressed) by providing a dominant feature. Combination of features as one feature vector is made by the motivation of classification results of individual features. Besides, the combinational feature outperforms the individual feature for stress analysis. The dominant feature are selected by considering highest classification accuracy of individual feature. The clusters are formed based on the observation from assortment analysis of features which can predict stressed or normal of an individual subject. The prediction can change based on the individual subject, hence combination of dominant feature results better classification accuracy.

- ***Combinational Analysis :*** As a second step, best dominant features are clustered as shown in Fig.4 and it is analysed using the classifiers producing the results shown in Table 2.The choice of best cluster is influenced by the classification accuracy. Hence, it is evident that the cluster C5 is chosen to be the best cluster of dominant features for stress detection with a classification accuracy of 66.52% Using KNN and 72.82% Using SVM. The cluster number represents the S.no of Table 1.

5.3. Enhancement using feature combination

The third step aims in improving the classification accuracy of the stress detector by introducing the following new features viz.,

- Total average power(TAP) represents 2-norm or maximum singular value of a ECG HR features.
- Energy of Heart represnts the energy of the HR signal .
- Energy of Heart variability represents the Energy of HRV features.

Now, the enhanced feature set is shown in Fig 5. Upon training with overlapping data, the SVM classifier with RBF kernel, a higher accuracy of 92.75% is obtained for the enhanced feature set. This indicates that GSR, Heart rate and its variability features are immediate response for stress prediction.

6. Conclusions

This paper corroborates how to select dominant features and fuse overlapping technique to extract the features from physiological sensors under conditions of work stress and context recognition to determine the stress in working individuals. It is evident from the classification results that the time and frequency domain features of HR, HRV, and GSR are sufficient to predict the stress. A hardware model for stress detector can be realized with the chosen features and the simulated model in an appropriate platform which results in a complete stress detector device.

References

1. Communications, N.. World health report. 2001. URL: http://www.who.int/whr/2001/media_centre/press_release/en/.
2. Bakker, J., Holenderski, L., Kocielnik, R., Pechenizkiy, M., Sidorova, N.. Stess@ work: From measuring stress to its understanding, prediction and handling with personalized coaching. In: *Proceedings of the 2nd ACM SIGHIT International health informatics symposium*. ACM; 2012, p. 673–678.
3. Deng, Y., Wu, Z., Chu, C.H., Zhang, Q., Hsu, D.F.. Sensor feature selection and combination for stress identification using combinatorial fusion. *International Journal of Advanced Robotic Systems* 2013;**10**(8):306.
4. Ghaderi, A., Frounchi, J., Farnam, A.. Machine learning-based signal processing using physiological signals for stress detection. In: *2015 22nd Iranian Conference on Biomedical Engineering (ICBME)*. 2015, p. 93–98.
5. Villarejo, M.V., Zapirain, B.G., Zorrilla, A.M.. A stress sensor based on galvanic skin response (gsr) controlled by zigbee. *Sensors* 2012; **12**(5):6075–6101.
6. Liu, D., Ulrich, M.. Listen to your heart: Stress prediction using consumer heart rate sensors 2015;.
7. Nakashima, Y., Kim, J., Flutura, S., Seiderer, A., André, E.. Stress recognition in daily work. In: *International Symposium on Pervasive Computing Paradigms for Mental Health*. Springer; 2015, p. 23–33.
8. Xu, Q., Nwe, T.L., Guan, C.. Cluster-based analysis for personalized stress evaluation using physiological signals. *IEEE journal of biomedical and health informatics* 2015;**19**(1):275–281.
9. Tanev, G., Saadi, D.B., Hoppe, K., Sorensen, H.B.. Classification of acute stress using linear and non-linear heart rate variability analysis derived from sternal ecg. In: *Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE*. IEEE; 2014, p. 3386–3389.
10. Gjoreski, M., Gjoreski, H., Luštrek, M., Gams, M.. Continuous stress detection using a wrist device: in laboratory and real life. In: *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*. ACM; 2016, p. 1185–1193.
11. Palanisamy, K., Murugappan, M., Yaacob, S.. Multiple physiological signal-based human stress identification using non-linear classifiers. *Elektronika ir elektrotechnika* 2013;**19**(7):80–85.
12. Widanti, N., Sumanto, B., Rosa, P., Miftahudin, M.F.. Stress level detection using heart rate, blood pressure, and gsr and stress therapy by utilizing infrared. In: *Industrial Instrumentation and Control (ICIC), 2015 International Conference on*. IEEE; 2015, p. 275–279.
13. Sioni, R., Chittaro, L.. Stress detection using physiological sensors. *Computer* 2015;**48**(10):26–33.
14. Zenonos, A., Khan, A., Kalogridis, G., Vatsikas, S., Lewis, T., Sooriyabandara, M.. Healthyoffice: Mood recognition at work using smartphones and wearable sensors. In: *Pervasive Computing and Communication Workshops (PerCom Workshops), 2016 IEEE International Conference on*. IEEE; 2016, p. 1–6.
15. Selvaraj, N.. Psychological acute stress measurement using a wireless adhesive biosensor. In: *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. 2015, p. 3137–3140.
16. Vadana, D.P., Kottayil, S.K.. Energy-aware intelligent controller for dynamic energy management on smart microgrid. In: *Power and Energy Systems Conference: Towards Sustainable Energy, 2014*. IEEE; 2014, p. 1–7.
17. Rajagopalan, S.S., Murthy, O.R., Goecke, R., Rozga, A.. Play with memeasuring a child's engagement in a social interaction. In: *Automatic Face and Gesture Recognition (FG), 2015 11th IEEE International Conference and Workshops on*; vol. 1. IEEE; 2015, p. 1–8.
18. Koldijk, S., Neerincx, M.A., Kraaij, W.. Detecting work stress in offices by combining unobtrusive sensors. *IEEE Transactions on Affective Computing* 2016;.
19. Koldijk, S., Sappelli, M., Verberne, S., Neerincx, M.A., Kraaij, W.. The swell knowledge work dataset for stress and user modeling research. In: *Proceedings of the 16th International Conference on Multimodal Interaction*; ICMI '14. New York, NY, USA: ACM. ISBN 978-1-4503-2885-2; 2014, p. 291–298. URL: http://doi.acm.org/10.1145/2663204.2663257. doi:10.1145/2663204.2663257.
20. Yoder, N.. Peakfinder. *Internet: http://www mathworks com-/matlabcentral/fileexchange/25500* 2011;.