The motion influence on respiration rate estimation from lowresolution thermal sequences during attention focusing tasks

A. Kwaśniewska, J. Rumiński, *Member, IEEE*, J. Wtorek Gdansk University of Technology, Gdansk, Poland

Abstract— Global aging has led to a growing expectancy for creating home-based platforms for indoor monitoring of elderly people. A motivation is to provide a non-intrusive technique, which does not require special activities of a patient but allows for remote monitoring of elderly people while assisting them with their daily activities. The goal of our study was to evaluate motion performed by a person focused on a specific task and check if this motion disrupts estimation of respiration rate. The preliminary results show that it is possible to reliable estimate respiration rate by focusing attention of a patient on a certain activity. The respiratory rate analyzed for silent reading task was estimated with mean error 0.27 breaths per minute, while for reading aloud task with 1.18 bpm. The observed head motion during the reading aloud task was 1.5 higher that for silent reading and about two times smaller for a case in which subjects were not focused on any task.

I. INTRODUCTION

With the 21st century the human population has become an unprecedented aging society. According to the report issued in 2014 by Moody's Investors Service, the number of super-aged societies (populations with more than 20% elderly) is projected to rise from 3 to 13 by 2020 [1]. Global aging and notable demographic shift in the near future has led to a growing expectancy for more elderly people to remain independent as long as possible. Therefore, many societal and health care challenges may arise. The modern health care is supposed to enhance the safety and comfort of elderly patients while maintaining their independence with homebased platforms for indoor monitoring and vital signals extraction [2]. Some solutions for collecting both patientrelated and home environment data in order to recognize the Activities of Daily Living (ADL) and automatically detect changes in the behavioral patterns have already been proposed [3][4]. Other examples utilize portable devices capabilities combined with advanced signal processing algorithms for different assisted-living solutions, e.g. heartbeat rate calculation from a smartphone-captured video [5] or respiration rate estimation with thermal camera [6]. Regardless of the system type, motivation for monitoring elderly people is to provide a non-intrusive technique to assist them with their daily activities, e.g. detecting health conditions during watching TV, reading a book or conducting a videoconference with family. These systems may notify the user when the monitored target object enters dangerous

* This work has been partially supported by NCBiR, FWF, SNSF, ANR and FNR in the framework of the ERA-NET CHIST-ERA II, project eGLASSES – The interactive eyeglasses for mobile, perceptual computing and by Statutory Funds of Electronics, Telecommunications and Informatics Faculty, Gdansk University of Technology.

situations or condition. According to some previous researches, the best results of detecting facial features in order to evaluate vital signals from recorded image sequences are achieved when a patient stays still [7]. One solution for this is to ask a volunteer not to move, however it can be also achieved indirectly by focusing his attention on a specific task, e.g. reading or talking with a caregiver. Taking it into account, the purpose of this study is to detect and evaluate range of motion performed by a volunteer focused on a certain task and check if this motion disrupts estimation of respiration rate in poor quality thermal video sequences.

The rest of the paper is organized as follows: Section II presents existing solutions and methods for motion detection and extraction of a respiration rate. Section III describes methodology used for respiration rate estimation and motion evaluation. In Section IV we demonstrate the experimental results acquired for different measurement scenarios. In Section V we discuss the results, present some ideas of remote health care for elderly people and finally we conclude the paper.

II. STATE OF THE ART

Various approaches for motion search have already been described together with the metrics used to estimate the goodness of the found match, e.g. the mean squared error (MSE) or minimum Sum of Absolute Difference (SAD) [8]. To deal with erratic motions, loss of sharpness, combing artifacts, halos and blockiness metrics are used to estimate the quality of motion compensated video [9]. Other principal characteristics of motion segmentation quality are the spatial deviation from the reference segmentation map or the number of fluctuations of spatial deviations over time [10]. Reliable motion segmentation and determining precise localization of moving objects in each frame can be also achieved by combining the motion information with the mean shift based color segmentation [11]. Scale-Space Extrema Detection algorithm, which allows for searching for stable features across all possible scales was also proposed [12]. In this method, the feature point candidates are selected as local maxima in computed Difference-of-Gaussian (DoG) images. In the next step, vector/gradient magnitude and orientation information is assigned to each of feature and therefore object rotation can be detected. Some other approaches take advantage of raise of High Performance Computing together with low cost computational devices and estimate motion directly by block matching. In this method SAD is computed between current block and reference frame and then the motion information is described by identifying blocks in a current frame and searching for most similar ones in a reference frame [13].

Motion estimation may be useful for defining the quality of video in some telemedicine solutions based on video processing. One of possible use case is remote estimation of respiration rate from thermal video sequences. Some researches in this area have already been conducted [6][14]. Pavlidis et al. [15] proposed to measure heat transfer related to the moisturized air during expiration in the nostrils or mouth region. In this approach breathing rate was calculated by counting frames labeled as expiratory or nonexpiratory ones. Authors of [16] used a different data analysis method. In their research, average value for each ROI was calculated and normalized. Then, the autocorrelation signal for filtered waveform was calculated and decomposed using Fourier Transform. Finally, power density spectrum (PWS) was calculated. The frequency of the dominated peak in PWS was used as a breathing frequency. In [14] different respiration rate estimators were proposed. They include method based on the dominated peak in the frequency spectrum (eRR sp), estimator based on the number of zero-crossings (eRR zc), estimator based on the number of detected peaks (eRR pk), and estimator based on periodicity of peaks locations for the autocorrelation function (eRR ap). The presented results show that mean absolute error (MAE) was smallest for eRR ap (MAE=0.415bpm). The presented research [14] was conducted assuming that the subjects do not move and do not speak.

Recent technological developments has brought growth in the availability of low-price thermal camera modules that can be used in mobile monitoring systems, e.g. for collecting thermal facial images during patient teleconsultations with a formal or informal caregiver. Remote healthcare systems based on image processing techniques do not require special activities of a patient (like connection of electrodes) and therefore patients feel more convenient while using them. Moreover, such measurements do not have impact on the patient's status allowing the estimation of important vital signs, e.g. the respiratory rate. However, different factors can still have influence on the quality of the estimated parameters. The motion of a patient is very important issue and can be compensated by various automatic, face (or facial features) tracking procedures. However, with a lowresolution thermal camera and keeping the diagnostic system as simple as possible we investigated a situation when a patient is focused on a certain task, like during a potential teleconsultation session. In this way, any involuntary movements can be minimalized, because the patient is more concentrated on a secondary exercise and does not think about being diagnosed. Therefore, we investigated whether motion present during such tasks disrupts estimation of respiration rate using low-resolution thermal video sequences.

III. METHODOLOGY

In our experiment, the respiration rate was calculated from sequences of thermal images recorded with the FLIR Lepton thermal camera module characterized by high dynamic range (14bits), small size (<1cm²) and relatively small spatial resolution (80x60). Subjects were asked to read a text that appeared on the monitor in front of them. Two scenarios were tested during measurements: reading aloud and silent reading. Sequences of frames were captured during

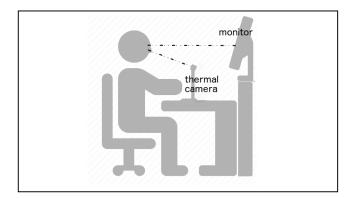


Figure 1. Test bench overview

60s time period (sampling frequency f_s =13Hz) for each scenario from 11 volunteers (age: 31.1±10.6; 9 men, 2 women). At the same time, pressure measurements from the Respiration Monitor Belt (Vernier RMB) were acquired for a reference. To synchronize both signals subjects were asked to hold the breath at the beginning of the measurement. The thermal camera module was placed on a tripod at a distance of 40cm from a subject. The monitor was located at a distance of 70cm from a subject (Fig. 1). To evaluate respiration rate, nostrils (silent reading) or mouth areas (reading aloud) were manually selected. Signals were extracted using the skewness and average parameters as aggregation operators inside the ROI. Produced signals (time series of infrared radiation changes) were filtered using Kaverage filter. Baseline removal was performed using 4thorder high pass Butterworth filter with cutoff frequency equals 0.125Hz. Filtered signals were used to calculate four respiration rate estimators: eRR sp, eRR ap, eRR zc and eRR pk. For signals acquired with Respiration Monitor Belt the same estimators were calculated. Moreover, signals from the belt were visually analysed to manually calculate the respiration rate using the standard definition (number of events in time). The number of respiration events during a time period (N RE) and the total time of all respiration events were calculated (T_RE) . Then, the respiration rate was computed as (N RE*60)/T RE and used as a reference value. The Hijorth parameter Complexity (C) and the Spectral Purity Index (SPI) were calculated to indicate whether produced signals are periodic. If the signal is more similar to sine, the C parameter has values close to 0, while the SPI parameter has values close to 1. Additionally, in order to evaluate range of motions performed by a person focused on a certain task, average Sum of Absolute Differences (SAD) per pixel for reading aloud and silent reading sequences was calculated as:

$$SAD = \sum_{n=0}^{F-1} \frac{\sum_{i=0}^{W} \sum_{j=0}^{h} |P_n(i,j) - P_{n+1}(i,j)|}{w * h * F}$$
 (1)

Where F is a number of frames in a video sequence, n is a number of a current frame, w is a width of a frame, h is a height of a frame and Pn is a pixel grayscale value at a given position. Average SAD per pixel provides an overall estimation about the spatial-temporal activity of the video and can be used for computing moving vectors [17]. As a reference, we calculated the same metric for a sequence, in which a person was not focused on a given task and performed small movements (turn about 30 degrees in up, left and right direction from the straight-ahead position).



CONFIDENTIAL. Limited circulation. For review only.

IV. RESULTS

Evaluated respiration rate estimators (eRR sp, eRR ap, eRR zc and eRR pk), respiration rate calculated manually and two signal measures: Complexity (C) and Spectral Purity Index (SPI) for reading aloud and silent reading scenarios are collected in the Tables 1-2. During reading aloud, for the average operator the mean value of C=0.15±0.014 and the mean value of SPI=0.46±0.056, for the skewness operator C=0.14±0.011 $SPI = 0.51 \pm 0.046$, and for $C=0.31\pm0.154$ and $SPI=0.25\pm0.1$. During silent reading, for the average operator the mean value of $C=0.11\pm0.105$ and the mean value of SPI=0.59±0.103, for the skewness operator $C=0.1\pm0.012$ and $SPI=0.62\pm0.068$, for RMB $C=0.19\pm0.071$ and SPI=0.35±0.154. Table 3 presents mean, standard deviation and mean square error (MSE) of absolute differences between the respiration rates evaluated using the given estimator (for both thermal and RMB signals) and the manually calculated respiration rate. The values of average SAD per pixel during focusing attention and without focusing attention, together with calculated means for these values are presented in the Table 4.

TABLE I. READING ALOUD SEQUENCES - RESPIRATION RATES, COMPLEXITY AND SPECTRAL PURITY INDEX

	6.11		1					
	Subject	eRR sp	eRR ap	eRR zc	eRR pk	Manual	С	SPI
1	Average	13.64	13.55	13.64	15.33		0.15	0.47
	Skewness	13.64	13.92	13.64	15.15		0.15	0.47
	RMB	13.58	13.67	12.78	13.04	14.42	0.25	0.26
	Average	11.69	16.94	13.64	19.61		0.14	0.50
2	Skewness	13.64	13.32	13.64	15.72		0.12	0.54
	RMB	15.97	15.69	14.37	15.71	15.27	0.14	0.45
	Average	13.64	17.32	13.64	13.79		0.15	0.53
3	Skewness	11.69	16.02	11.69	11.39		0.14	0.49
	RMB	11.98	22.70	12.38	8.97	10.52	0.25	0.24
	Average	9.09	12.57	13.64	12.96		0.16	0.37
4	Skewness	10.39	12.67	16.24	13.81		0.14	0.55
	RMB	12.78	20.86	16.77	8.69	14.66	0.51	0.10
	Average	9.74	12.37	14.61	14.09		0.15	0.42
5	Skewness	9.74	12.18	14.61	17.01		0.16	0.42
	RMB	13.58	18.89	16.37	15.39	17.19	0.27	0.29
	Average	9.74	10.26	11.69	16.11		0.18	0.36
6	Skewness	9.74	10.12	14.61	13.36		0.15	0.46
	RMB	9.58	9.28	10.38	10.92	10.35	0.23	0.24
	Average	12.99	12.78	16.24	11.81		0.15	0.48
7	Skewness	16.24	16.70	13.80	17.23		0.14	0.53
	RMB	17.57	17.19	15.57	15.40	14.81	0.27	0.24
	Average	13.64	14.00	15.59	19.28		0.14	0.50
8	Skewness	13.64	13.92	16.56	17.32		0.15	0.55
	RMB	13.38	23.44	18.77	17.66	18.90	0.28	0.29
	Average	15.59	15.18	17.54	15.13		0.15	0.46
9	Skewness	15.59	16.94	15.59	20.24		0.15	0.55
	RMB	14.37	22.51	22.76	15.05	17.07	0.69	0.07
	Average	17.54	16.70	18.51	20.58		0.17	0.50
10	Skewness	15.59	15.48	14.61	17.89		0.15	0.51
-	RMB	14.37	16.24	16.37	16.06	14.87	0.26	0.30
	Average	13.64	13.92	12.67	15.33		0.13	0.50
11	Skewness	13.64	17.85	12.67	12.10		0.13	0.56
	RMB	13.58	14.71	15.17	16.31	15.80	0.26	0.25

TABLE II. SILENT READING SEQUENCES - RESPIRATION RATES, COMPLEXITY AND SPECTRAL PURITY INDEX

	Subject							
	Subject	eRR_sp	eRR_ap	eRR_zc	eRR_pk	Manual	C	SPI
	Average	15.59	14.80	13.64	16.24		0.12	0.54
1	Skewness	15.59	14.99	13.64	16.00		0.10	0.64
	RMB	15.97	15.69	14.37	15.71	15.62	0.14	0.45

	Cubicat							
1	Subject	eRR_sp	eRR_ap	piration eRR_zc		Manual	C	SPI
	Average	13.64	14.26	14.61	15.81		0.11	0.58
2	Skewness	15.59	14.71	14.61	16.09		0.11	0.63
	RMB	14.37	14.34	13.18	14.88	14.53	0.21	0.25
	Average	13.64	13.44	12.67	14.83		0.10	0.63
3	Skewness	13.64	13.55	14.61	15.18		0.12	0.60
	RMB	13.58	13.32	11.98	13.53	12.96	0.20	0.26
	Average	13.64	12.78	12.67	14.00		0.11	0.51
4	Skewness	13.64	12.88	12.67	14.75		0.10	0.65
	RMB	15.17	14.43	13.18	14.82	14.47	0.28	0.16
	Average	17.54	18.13	16.56	17.27		0.12	0.61
5	Skewness	13.64	14.26	12.67	13.55		0.10	0.59
	RMB	17.57	17.58	16.77	18.11	17.75	0.14	0.52
	Average	11.69	11.99	10.72	13.52		0.10	0.54
6	Skewness	11.69	11.63	10.72	12.10		0.09	0.56
	RMB	11.98	12.08	11.58	12.36	12.22	0.10	0.51
	Average	13.64	14.09	13.64	12.45		0.15	0.42
7	Skewness	13.64	13.67	11.69	15.41		0.10	0.59
	RMB	14.37	14.00	13.18	14.23	13.80	0.25	0.19
	Average	19.49	19.48	19.48	21.89		0.08	0.79
8	Skewness	19.49	20.78	20.46	21.71		0.11	0.69
	RMB	19.17	18.78	21.16	21.81	21.40	0.17	0.53
	Average	17.54	17.71	17.54	18.50		0.11	0.69
9	Skewness	17.54	17.71	16.56	19.91		0.10	0.71
	RMB	20.76	20.78	18.37	17.69	18.87	0.33	0.20
	Average	11.69	11.06	10.72	12.82		0.11	0.52
10	Skewness	11.69	10.98	9.74	12.82		0.13	0.49
	RMB	10.38	9.99	10.78	10.94	10.85	0.16	0.28
	Average	17.54	17.45	17.54	20.27		0.10	0.68
11	Skewness	17.54	17.32	15.59	17.97		0.09	0.72
1 1	RMB	16.77	16.70	16.37	16.94	16.55	0.13	0.53

TABLE III. AGGREGATED MEASURES OF DIFFERENCES BETWEEN RR ESTIMATORS VERSUS MANUALLY CALCULATED RR VALUES

	Average			Skewness				RMB				
	Reading aloud											
	sp	ap	zc	pk	sp	ap	zс	pk	sp	ap	zc	pk
Mean	3.14	2.62	2.04	2.78	2.45	2.26	1.84	1.82	2.09	3.38	1.46	1.57
Std.dev	2.16	2.01	1.14	1.92	2.23	2.01	1.15	1.27	1.52	3.54	1.56	1.56
MSE	4.26	3.66	1.18	3.34	4.52	3.66	1.21	1.48	2.10	11.37	2.20	2.21
					Silen	t read	ing					
	sp	ap	zc	pk	sp	ap	zc	pk	sp	ap	zc	pk
Mean	0.76	0.76	1.03	1.27	1.13	0.90	1.77	1.37	0.69	0.61	0.74	0.40
Std.dev	0.55	0.61	0.75	1.00	1.12	0.97	1.27	1.19	0.70	0.86	0.46	0.30
MSE	0.27	0.33	0.51	0.91	1.13	0.85	1.47	1.29	0.45	0.68	0.19	0.08

TABLE IV. AVERAGE SAD/PIXEL FOR DIFFERENT IMAGE SEQUENCES

Cubicat	SAD/pixel									
Subject	Silent reading	Reading aloud	Not focused on a task							
1	0.85	0.58	1.11							
2	0.26	0.31	0.36							
3	0.22	0.31	0.39							
4	0.38	0.60	0.59							
5	0.28	0.48	0.72							
6	0.40	0.73	1.06							
7	0.25	0.49	1.91							
8	0.23	0.45	1.10							
9	0.24	0.40	1.03							
10	0.47	0.65	1.15							
11	0.47	0.71	2.00							
Mean	0.37	0.52	1.04							
Std.dev	0.186	0.147	0.538							

V. DISCUSSION

The aim of this paper was to analyse influence of motion on respiration rate estimation and to determine whether focusing attention on a specific task allows for reducing subject's movement. In our research, four RR estimators



CONFIDENTIAL. Limited circulation. For review only.

were investigated both for thermal-based and belt-based signals. The results show that the smallest mean differences between RR estimated from thermal sequences and reference values during reading aloud were achieved for eRR zc and eRR pk estimators, while during silent reading for eRR sp and eRR ap estimators. However, results obtained for other estimators are also satisfactory. According to preliminary studies, it can be assumed that reliable estimation of respiration rate from thermal image sequences during focusing attention tasks is possible. In this study, the analyzed ROI was specified manually. However, for the fully automatic procedure the region should be automatically extracted and tracked, for example using methods described in [7][18]. According to [14], the averaging operation performed on many pixels may smooth the changes generated by respiration events. In the contrary, waveforms extracted for the skewness operator are not very sensitive to the size of the ROI as long as it contains a nostril region.

In the performed analysis, it was verified that if a person is not concentrated on a specific task, the subjective movement content is higher. Moreover, the results show that the motion is also higher during speaking loudly than during silent reading (mean value of average SAD per pixel during silent reading was 0.37±0.186, during reading aloud 0.52±0.147 and while not being focused on a task 1.04±0.538). Probably it had influence on respiration rate estimation, because, as can be seen in Table 3, calculated mean square errors of absolute differences between the respiration rates evaluated using the given estimator and the manually calculated respiration rates were higher for reading aloud sequences. Furthermore, during silent reading the Complexity parameter had smaller values and the SPI parameter had bigger values than during reading aloud both for thermal and belt based signals. Taking it into account, it can be assumed that signals acquired during the smallest movements and while being focused on a certain task are more similar to sine signal. On the other hand, results for both scenarios (silent and aloud reading) are really promising and acceptable for the needs of remote monitoring of elderly people or for screening purposes (e.g. at airports).

VI. CONCLUSION

In conclusion, the conducted research showed that reliable evaluation of respiration rate is possible using thermal image sequences acquired during performance of a specific task. The results obtained for thermal-based signals are similar and highly correlate to results obtained for reference measurements (RMB). Moreover, captured image sequences can be potentially used for extracting other vital signals (calculation of periodicity of respiration rate, skin temperature changes, emotion-based changes, etc.). Taking it into account, it may be possible to create home based platform for supporting remote medical diagnostic of elderly people, who live alone and want to remain independent.

REFERENCES

[1] Moody's Investors Service, "Population Aging Will Dampen Economic Growth over the Next Two Decades", Global Credit Research - 06 Aug 2014, [Accessed: 1/25/2017], Available: https://www.moodys.com/research/Moodys-Aging-will-reduce-economic-growth-worldwide-in-the-next--PR 305951

- [2] M. Bajorek and J. Nowak, "The role of a mobile device in a home monitoring healthcare system," 2011 Federated Conference on Computer Science and Information Systems (FedCSIS), Szczecin, 2011, pp. 371-374.
- [3] M. Kaczmarek, J. Ruminski and A. Bujnowski, "Multimodal platform for continuous monitoring of elderly and disabled," 2011 Federated Conference on Computer Science and Information Systems (FedCSIS), Szczecin, 2011, pp. 393-400.
- [4] M. Popescu, E. Florea, "Linking Clinical Events in Elderly to In-home Monitoring Sensor Data: A Brief Review and a Pilot Study on Predicting Pulse Pressure", Journal of Computing Science and Engineering, Vol. 2, No. 1, March 2008, Pages 180-199
- [5] H. Nejati, V. Pomponiu, T. T. Do, Y. Zhou, S. Iravani and N. M. Cheung, "Smartphone and Mobile Image Processing for Assisted Living: Health-monitoring apps powered by advanced mobile imaging algorithms," in IEEE Signal Processing Magazine, vol. 33, no. 4, pp. 30-48, July 2016.
- [6] J. Rumiński, "Analysis of the parameters of respiration patterns extracted from thermal image sequences", Biocybernetics and Biomedical Engineering Journal, Volume 36, Issue 4, 2016, Pages 731–741
- [7] A. Kwasniewska and J. Ruminski, "Real-time facial feature tracking in poor quality thermal imagery," 2016 9th International Conference on Human System Interactions (HSI), Portsmouth, 2016, pp. 504-510. doi: 10.1109/HSI.2016.7529681
- [8] B. Lei, R. K. Gunnewiek and P. H. N. De With, "Reuse of Motion Processing for Camera Stabilization and Video Coding," 2006 IEEE International Conference on Multimedia and Expo, Toronto, Ont., 2006, pp. 597-600.
- [9] M. Nicolas, J. Roussel and F. Crete, "Metrics to Evaluate The Quality of Motion Compensation Systems in De-interlacing And Upconversion Applications," 2008 Digest of Technical Papers -International Conference on Consumer Electronics, Las Vegas, NV, 2008, pp. 1-2.
- [10] T. Schlogl, C. Beleznai, M. Winter and H. Bischof, "Performance evaluation metrics for motion detection and tracking," *Proceedings of* the 17th International Conference on Pattern Recognition, 2004. ICPR 2004, 2004, pp. 519-522 Vol.4.
- [11] A. Briassouli, V. Mezaris and I. Kompatsiaris, "Joint Motion and Color Statistical Video Processing for Motion Segmentation," 2007 IEEE International Conference on Multimedia and Expo, Beijing, 2007, pp. 2014-2017.
- [12] S. Hong and E. Atkins, "Moving Sensor Video Image Processing Enhanced with Elimination of Ego Motion by Global Registration and SIFT," 2008 20th IEEE International Conference on Tools with Artificial Intelligence, Dayton, OH, 2008, pp. 37-40.
- [13] R. Gaetano and B. Pesquet-Popescu, "OpenCL implementation of motion estimation for cloud video processing," 2011 IEEE 13th International Workshop on Multimedia Signal Processing, Hangzhou, 2011, pp. 1-6.
- [14] J. Ruminski, A. Kwasniewska, "Evaluation of respiration rate using thermal imaging in mobile conditions", chapter in monography, Ng, E. Y. K., EtehadTavakol, M., (ed.), Application of Infrared to Biomedical Sciences, Springer 2017, in press
- [15] R. Murthy and I. Pavlidis, "Non-contact monitoring of breathing function using infrared imaging" Department of Computer Science University of Houston, TX, 77204, USA http://www.cs.uh.edu Technical Report Number UH-CS-05-09 April 09, 2005
- [16] J.Fei, Z. Zhu, I. Pavlidis, "Imaging Breathing Rate in the CO2 Absorption Band", Proceedings of the 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference Shanghai, China, September 1-4, 2005
- [17] J. Joskowicz and J. C. L. Ardao, "A parametric model for perceptual video quality estimation" Telecommun. Syst., vol. 46, p. 14, 2010
- [18] A. Kwasniewska, J. Ruminski, "Face detection in image sequences using a portable thermal camera", Proc. Of the 13th Quantitative Infrared Thermography Conference, Gdansk 2016.

