

Nasal Cavity Detection in Facial Thermal Images for Non-contact Measurement of Breathing

Dai Hanawa, Toshiki Morimoto, Shota Shimazaki and Kimio Oguchi

Abstract— In this paper, we propose a method of improving the accuracy of detecting nasal cavity location in far infrared images for non-contact measurement of human breathing. We found that although our previous method for far infrared imaging can detect regions that include nasal cavities well, it suffers from high false alarm rates. In order to reduce this rate, we extend our method with a false alarm classification function. Object detection based on a boosted cascade of Haar-like feature classifiers is applied to find the candidates of regions that include the nasal cavities. In false alarm classification, binarization is employed to strictly segment facial area and background. Based on the results of binarization, false alarms on the background can be accurately classified. 5,100 FIR images are collected to train our nasal cavity detector; we evaluate the number of false alarms and detection failures. The results show that the proposed method can reduce false alarm events.

Keywords— Non-contact measurement of breathing, Facial thermal image, Nasal cavity detection, Binarization.

I. INTRODUCTION

In recent years, considerable research has been focused on the techniques for the extraction of vital signals from the human body, including heart rate, blood pressure, body temperature and breathing. By applying these techniques, information and communication services can be created to better support human activities (especially daily life for infants and the elderly). These applications include human activity monitoring, medical treatment, daily care, emotion analysis, and human computer interaction. In order to realize such applications, a technique for monitoring human breathing is one of the important goals, because human breathing is closely related to health, activity, emotion, and so on. However, most of the current major methods for monitoring human breathing take the contact approach. Therefore, development of non-invasive, non-contact type

Manuscript received October 28, 2012; revised December 18, 2012. This work was supported in part by JSPS Grant-In-Aid for Young Scientists (B) (KAKENHI No. 24700123) and the MEXT Grant-in-aid for Building Strategic Research Infrastructure.

D. Hanawa is with the Dept. of Compt. Inf. Sci., Faculty of Sci. and Tech., Seikei University, 3-3-1, Kichijojikitamachi, Musashino, Tokyo, 180-8633, Japan (corresponding author to provide phone: +81-422-37-3732; fax: +42-37-3871; e-mail: hanawa@st.seikei.ac.jp).

T. Morimoto was with the Dept. of Compt. Inf. Sci., Faculty of Sci. and Tech., Seikei University, 3-3-1, Kichijojikitamachi, Musashino, Tokyo, 180-8633, Japan. He is now with Allied Telesis K.K., Japan.

S. Shimazaki and K. Oguchi are with the Grad. School of Sci. and Tech., Seikei University, 3-3-1, Kichijojikitamachi, Musashino, Tokyo, 180-8633, Japan. (e-mail: {dm116221@cc, oguchi@st}.seikei.ac.jp).

method is needed.

In this paper, we study nasal cavity detection in facial thermal images to capture non-contact human breathing. Far Infra-Red (FIR) imaging is receiving attention as an attractive way of realizing this function. We have been studying a machine learning algorithm to detect the nose in thermal images of the user's face [1]. We found that while our previous method can detect regions that include nasal cavities well, the false alarm rate is not insignificant. Based on the results of our previous studies, we develop a false alarm classification technique for integration with our previous method that can reduce false alarms on non-facial regions. We also evaluate the frequency of missed-nose and of false alarms by conducting simulation experiments using 5,100 far infrared (FIR) images.

The rest of this paper is organized as follows. In Section II, we briefly review previous related work. Section III presents details of our proposed method. Section IV describes the results of our experiments. Section V concludes this paper.

II. RELATED WORKS

Several studies on the monitoring of human breathing have been reported [2]-[5]. Most of them do not use the non-contact approach, so that user stress created by the placement of or the detachment of sensor devices by body motion, are unavoidable.

An automatic respiration monitoring system by using rgb imaging was proposed by Nakai et al. [6]. However, this system can monitor only users sleeping on a sensor-equipped bed. The works of K. Abbas et al. [7] and Fei et al. [8], showed that FIR imaging had promise for the monitoring of human breathing. We have also shown that human breathing can be automatically captured by FIR imaging [9]-[11]. However, in these works, including our own, the relative position between the user's face and the FIR camera must be kept stable. Therefore, the detection of breathing becomes difficult when the position of the user's face is unknown.

Studies for tracking the human nose in thermal images based on particle filters [12] or image processing [13] have been reported. These offer some fixed robustness against head motion. However, in these methods, both naked eyes of the user must be constantly captured by the FIR camera. Moreover, these methods assume that the entire face is captured by the FIR image. How to recognize whether a face exists in an FIR image or not, was not discussed. B. Kaur et al. have proposed a method for tracking region including nasal cavities on thermal image [14]. However, in this method, manual operation to determine the initial position of nose is

needed.

In our study, we try to find whether nasal cavities exist on FIR image automatically, and to extract the region including nasal cavities from FIR images under the simple condition that nasal cavities are captured by FIR camera.

III. PROPOSED METHOD

We propose a method to detect the nose in FIR images by applying a current facial image processing scheme and false alarm classification. A block diagram of proposed method is shown in Fig.1.

A rapid object detection method based on a boosted cascade of rectangular feature (called Haar-like feature) classifiers proposed by Viola and Jones [15] and improved by Lienhart and Maydt [16] is applied in our work. Face-detectors using their method have become popular because of their speed and robustness. In this research, we replace the gray-scale training images with FIR images for nose detection and replace gray-scale input image with an FIR image for detection [1]. GentleBoost is used here to learn classifiers in our method. We introduce the above method to detect the region including nasal cavities. By applying this method, we get n , $x_d(t)_k$, $y_d(t)_k$, $w_d(t)_k$, and $h_d(t)_k$ ($k=1,2,\dots,n$) from a trained nose-detector and input FIR image $f(x, y, t)$ captured at time t . Here, n , $x_d(t)_k$, $y_d(t)_k$, $w_d(t)_k$, and $h_d(t)_k$ denote the number of detected region, horizontal position of the left-upper corner of the detected rectangle region, its vertical position, width of the rectangle and its height, respectively. However, we confirmed that its detection results included regions other than the correct nasal region [1]. The frequency of erroneous nose detection was not negligible.

In order to reduce such errors, we introduce false alarm classification. This classification consists of a binarization process and a classification process with a threshold. We introduce these processes to isolate the human face in the FIR image. In most cases, the human face is hotter than the other areas. Therefore, we can isolate the face by using the suitable threshold described as follows:

$$g(x, y, t) = \begin{cases} 1 & (f(x, y, t) \geq th_t) \\ 0 & (f(x, y, t) < th_t) \end{cases} \quad (1)$$

Where, $g(x, y, t)$ and th_t are the binarized FIR image at time t and the threshold for segmentation, respectively. From $g(x, y, t)$, we calculate the sum of pixels in each detected region by (2).

$$L(p, q, r, s, g(x, y, t)) = \sum_{i=p}^{p+r-1} \sum_{j=q}^{q+s-1} g(i, j, t) \quad (2)$$

Here, (p, q) is the left-upper corner position of the rectangle on $g(x, y, t)$. r and s are its width and its height, respectively. Next, the containing ratio of the facial area in the above rectangle is calculated by (3).

$$R(p, q, r, s, f(x, y, t)) = L(p, q, r, s, g(x, y, t)) / (r \cdot s) \quad (3)$$

Finally, the detected region is classified as false alarm if $R(x_d(t)_k, y_d(t)_k, w_d(t)_k, h_d(t)_k, f(x, y, t))$ is lower than threshold th_r . Otherwise it is classified as nasal region. All detected regions $(x_d(t)_k, y_d(t)_k, w_d(t)_k, h_d(t)_k)$ in each input FIR image $f(x, y, t)$ are classified based on the above algorithm.

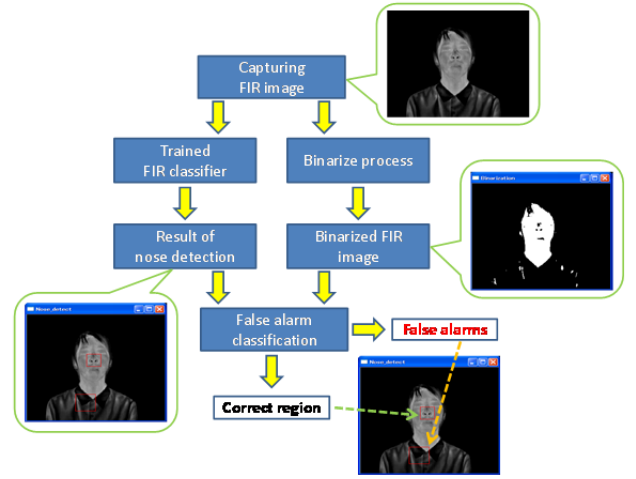


Fig.1. Block diagram of proposed method.

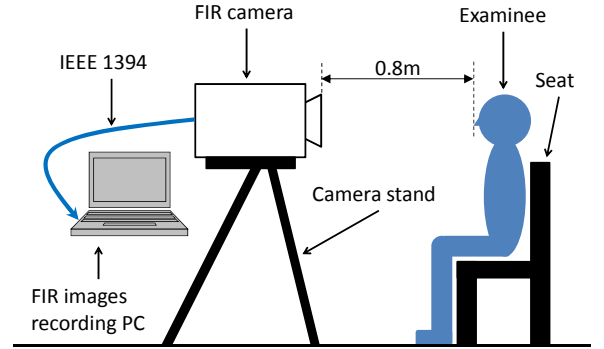


Fig.2. Relationship between FIR camera and subject.

IV. EXPERIMENTS

Experiments were conducted to evaluate the frequency of nasal cavity detection failure and false alarm. Moreover, the relationship between false alarm and threshold th_r was experimentally examined.

A. Recording FIR Image

Facial thermal images were captured by an FIR camera with the following specifications.

- Type: NEC/Avio, TH7102MX
- Capture wavelength: 8-14 μm
- Thermal resolution: 0.06 $^{\circ}\text{C}$
- Size of sensor (active area): 320x240 pixels
- Contrast: 256 levels (8bit), Gray scale
- Frame rate: 30 fps
- Temperature range: 32.0-40.0 $^{\circ}\text{C}$

Room temperature was around 26 $^{\circ}\text{C}$ and remained constant throughout all trials. The subject sat in front of the camera in a room (Fig. 2). Each subject was asked to breathe normally via the nose while being recorded. FIR images from nine subjects (nine males with ages from 22 to 23 years) were acquired. FIR images were recorded for one minute for each subject. 300 continuous frames (ten seconds) wherein the subject breathed stably were selected. Finally, nasal region in each frame was manually extracted for training the FIR classifier for each subject. Minimum and maximum size of

nasal region extracted from the FIR images of each subject were 23x13 pixels and 31x23, respectively. FIR images totaling 2,700 patterns were acquired as positive samples. Moreover, FIR images totaling 2,400 patterns, which included no human face, were collected as negative samples. Examples of positive data and negative data are shown in Fig. 3 and 4, respectively.

B. Training and Evaluation

The nine-fold cross-validation procedure was employed. Here, a set of 2,400 positive samples (8 subjects) and the remaining 300 samples (1 subject) were used as training and test material, respectively. Negative samples for training were fixed, so that 2,400 samples were used. Experiments were performed for one hundred and one (0, 1, ..., 100%) ratios of the facial area. These processes were repeated nine times with each of nine subject's samples used once as the validation data. Threshold th_t for binarization was set at 100. This value (corresponds to about 35 °C) was selected from the results of preliminary experiments.

Experimental metrics were *correct-detection* ratio (denoted as R_c), *miss-detection* ratio (denoted as R_m), and *false alarm* ratio (denoted as R_f). Correct-detection means that one of the detection results indicated the correct nasal region. Miss-detection means that the detection results did not include a correct nasal region. On the other hand, false alarm means that one of the detection results indicated a region other than the correct nasal region. These metrics were calculated as follows.

$$R_c = N_c/N_{all} \quad (4)$$

$$R_m = N_m/N_{all} \quad (5)$$

$$R_f = N_f/N_{all} \quad (6)$$

Here, N_c , N_m , N_f , and N_{all} denote the number of the frames with correct-detection, those with miss-detection, those with false alarm and total number of frames, respectively.

C. Results

At first, results for positive samples were analyzed. Fig. 5 and 6 show the results for training samples and test samples, respectively. Here, average in R_c , that in R_m and that in R_f were calculated for all data sets. In the case that $th_r=0$, the results show the performance of the previous method.

Fig. 7 shows the change in the number of non-facial region detection for each threshold. Here, the number of the detected rectangle region as false alarm in each frame was totaled and its average per frame was calculated.

Fig. 8 and 9 show an example of detection by the previous method and that by proposed method using the same input image, respectively. Fig. 10 shows an example of the results yielded by the proposed method including correct-detection and false alarm.

Optimal threshold for the proposed method for each subject was examined based on the experimental results. Table I and II show detailed results for training samples and test samples, respectively. Also the optimal threshold for each subject, and corresponding R_f values, are shown in the two rightmost columns. These were selected under the condition that R_c is equal to that in the previous method so that R_f is smaller.

Next, results for negative samples were analyzed. Fig. 11 shows the results for the training samples. Here, the average R_f for all data sets is calculated. When $th_r=0$, the results show

the performance of the previous method. Fig. 12 shows the change in the number of non-facial regions detected for each threshold. Table III shows the detailed results for training samples. Here, optimal th_r in Table I was used for each data set.



Fig.3. Examples of positive samples.



Fig.4. Examples of negative samples.

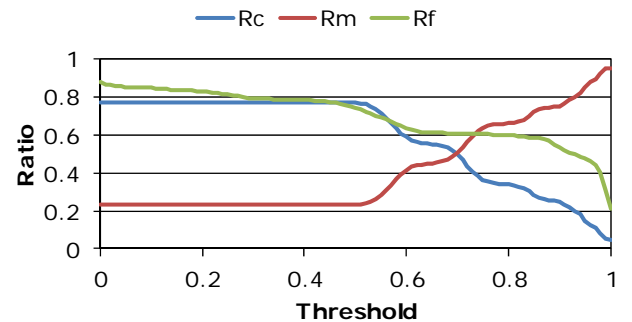


Fig.5 Results for training samples (positive samples).

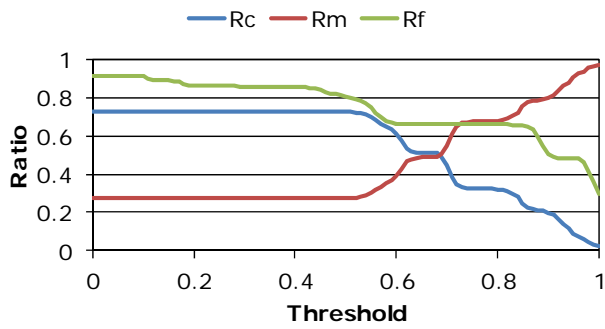


Fig. 6. Results for test samples (positive samples).

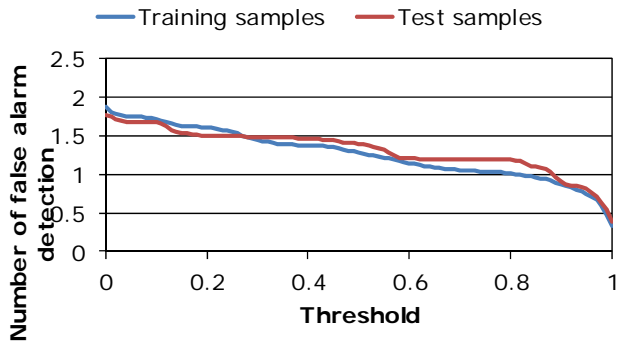


Fig. 7. Average number of false alarms (positive samples).

TABLE I
DETAILED RESULTS FOR TRAINING DATA
(POSITIVE SAMPLES, $N_{ALL}=2400$)

Data set	Previous method			Proposed method	
	R_c [%]	R_m [%]	R_f [%]	R_f [%]	Th_r
Except for A	81.5	18.5	73.3	61.0	0.48
Except for B	68.0	32.0	76.2	62.7	0.50
Except for C	85.7	14.3	89.6	77.3	0.48
Except for D	60.6	39.4	91.6	74.8	0.50
Except for E	84.3	15.7	97.4	87.3	0.48
Except for F	80.9	19.1	86.1	59.4	0.53
Except for G	83.9	16.1	87.8	74.1	0.48
Except for H	65.8	34.2	98.2	94.3	0.51
Except for I	83.0	17.0	90.3	78.6	0.48
Average	77.1	22.9	87.8	77.7	-

TABLE II
DETAILED RESULTS FOR TEST DATA
(POSITIVE SAMPLES, $N_{ALL}=300$)

Data set	Previous method			Proposed method	
	R_c [%]	R_m [%]	R_f [%]	R_f [%]	Th_r
A	6.7	93.3	93.0	85.0	1
B	100.0	0.0	50.7	7.3	0.58
C	54.7	45.3	100.0	45.3	0.94
D	100.0	0.0	100.0	100.0	0.84
E	100.0	0.0	99.7	99.3	0.90
F	90.7	9.3	94.7	94.7	0.49
G	69.7	30.3	89.3	0.0	0.67
H	30.7	69.3	100.0	100.0	0.80
I	100.0	0.0	96.7	91.3	0.78
Average	72.5	27.5	91.6	69.2	-

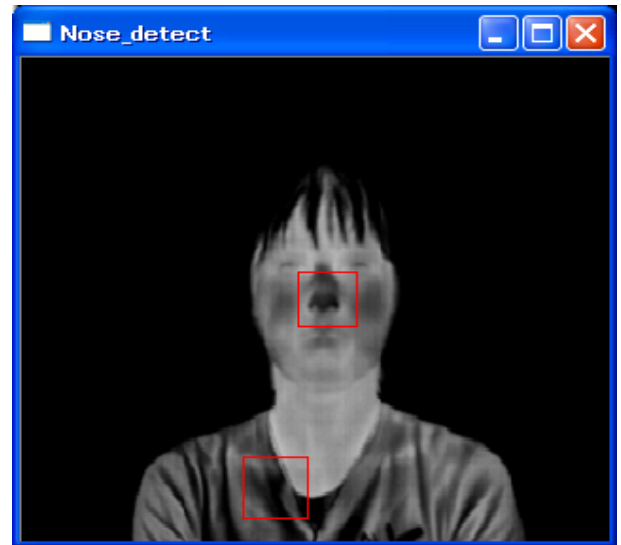


Fig. 8. An example of detection by the previous method. Red rectangles at the center and at the bottom are correct-detection and false alarm, respectively.

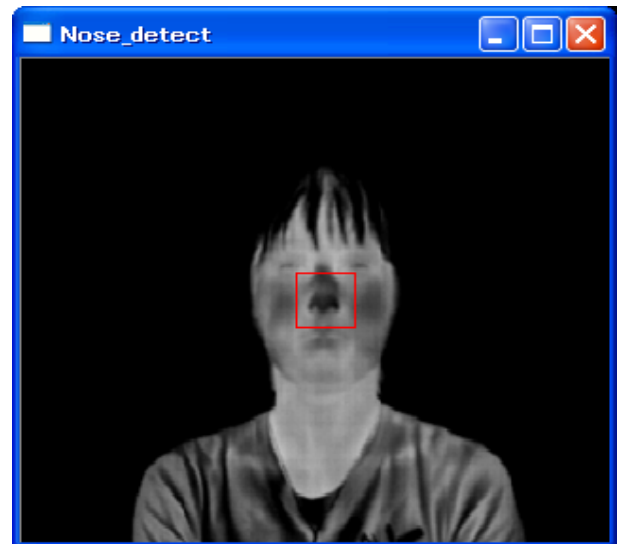


Fig. 9. An example of detection by proposed method. Red rectangle is correct-detection. False alarm in Fig. 8 was eliminated.

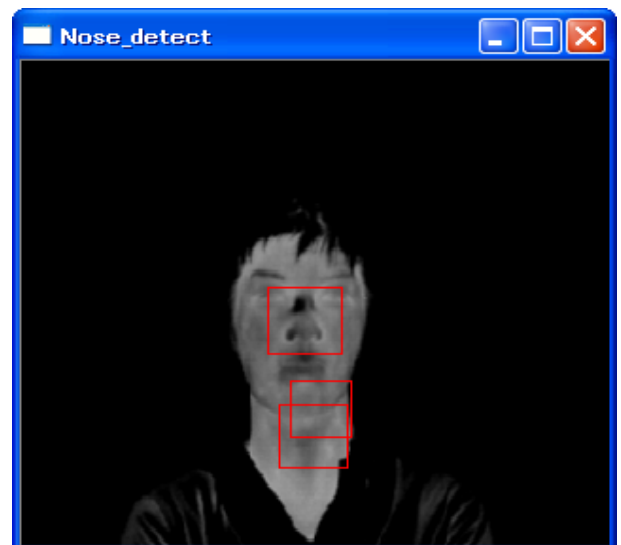


Fig. 10. An example of false alarm by proposed method. Red rectangles other than the one at the center are false alarms.

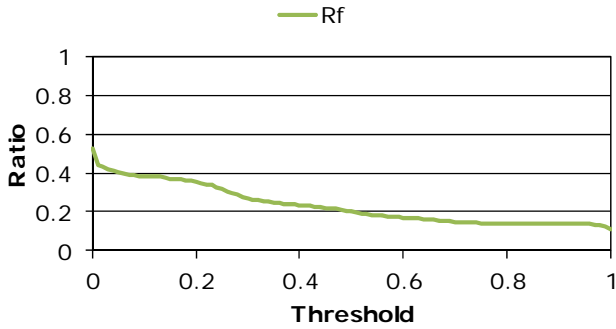


Fig.11. Results for training samples (negative samples).

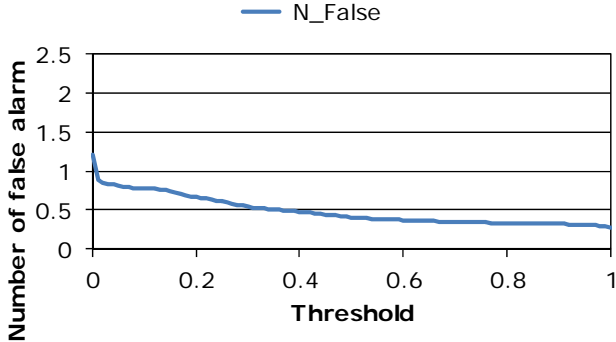


Fig.12. Average number of false alarms (negative samples).

TABLE III
COMPARISON OF FALSE ALARM FOR TRAINING DATA
(NEGATIVE SAMPLES, $N_{ALL}=2400$)

Data set	Previous method [%]	Proposed method [%]
Except for A	47.5	24.5
Except for B	35.7	15.0
Except for C	32.1	6.2
Except for D	53.7	23.8
Except for E	69.2	29.8
Except for F	68.0	2.5
Except for G	52.5	26.9
Except for H	56.5	25.4
Except for I	58.1	30.1
Average	52.6	20.5

D. Discussion

According to Fig. 5 and 6, the average false alarm ratio of the proposed method improved 10% around $th_r=0.5$ for positive data compared to the previous method, while the proposed method basically matches the correct-detection ratio of the previous method. This optimal threshold was mostly fixed for all data sets (Table I). Analysis of the detection result shows false-alarm occurring in non-facial regions were eliminated by the proposed method. These results show that proposed method can reduce the detection of non-nasal regions by using a suitable threshold. In Fig. 8, the previous method yielded one false alarm due to wrinkles on clothing. On the other hand, the proposed method was able to classify it as a false-alarm (Fig. 9). In Fig. 11 and Table III, the proposed method improved the false alarm ratio by 32% around $th_r=0.5$ for negative data compared to the previous method. False alarm events in negative data were classified more accurately. We confirmed that false alarms generated in areas other than the face were eliminated by the proposed method with optimal threshold.

In Fig. 7 and 12, the average number of false alarms detected around $th_r=0.5$ for positive samples and that for

negative samples improved by about 0.5 and 0.8, respectively. It is useful to classify all detection results as either correct-detection or false alarm. However, the proposed method can not overcome false alarms that occur in the facial area such as throat or jaw (Fig. 10). The area including these parts was classified as facial area based on their temperature, because the surface skin was uncovered. We believe that this type of false alarm can be eliminated by observation of the variation in temperature over several frames, since the nasal cavities exhibit cyclic changes in temperature. Moreover, another object detection method could be examined. For example, an object detection method based on support vector machines seems to be useful. Studies on improving the accuracy of nasal region detection are future works.

From Table II, optimal thresholds for test data varied only slightly among the subjects. Average false alarm ratio was improved by 22% by using the optimal threshold for each subject. On the other hand, optimal thresholds for training data were about 0.5 regardless of the subject. The proposed method seems to provide a fixed performance for known users. How to estimate the optimal threshold for an unknown user should be examined in order to reduce the number of false alarms. This challenge is another future work.

All in all, the above results confirm the validity of our proposed method for reducing the false alarm rate in detecting the nose.

V. CONCLUSION

We proposed to apply false alarm classification into our previous method to detect the nasal region accurately in thermal images of the user's face. The intent is to reduce the false alarm rate of our previous method. First, a method to classify false alarms was proposed; it accurately detects the region that includes the nasal region from thermal facial images. Next, the performance of our proposed method was evaluated by experiments. Our results suggested that the proposed method suffers far fewer false alarm events than our previous method even though it uses simple image processing operations. Moreover, the proposed method minimizes miss-detection if a suitable threshold is used.

However, several technical issues hinder the achievement of breathing detection with few constraints on the user. For example, our method can not eliminate false alarm events that occur in the facial area. A method that can handle arbitrary poses also should be studied. The results described in this paper clarified several technical issues when we apply the object detection method by Viola and Jones to FIR images.

In future work, we will study how to classify the nasal region more accurately from detection results. For example, we plan to study how to choose better training samples. Moreover, we plan to combine the proposed method with observation of the variation in temperature over several frames, or consider the use of another object detection method.

REFERENCES

- [1] D. Hanawa, T. Morimoto, S. Terada, T. Sakai, S. Shimazaki, K. Igarashi and K. Oguchi, "Nose detection on far infrared image for non-contact measurement of breathing," in *Proc. IEEE-EMBS BHI 2012 [USB-Memory]*, Hong Kong and Shenzhen, China, Jan. 2012, pp.878-881.

- [2] Y. Nishida, T. Hori, "Non-invasive and unrestrained monitoring of human respiratory system by sensorized environment," *Proc. IEEE Sensor 2002*, Orlando, FL, June 2002, pp. 62.4(1)-(6).
- [3] Y. Nishida, T. Mori, H. Mizoguchi and T. Sato, "Sleep apnea syndrome diagnosis based on image processing," *J. RSJ*, vol.16, no.2, Mar. 1998, pp. 274-281 (in Japanese).
- [4] Y. Nishida, M. Takeda, T. Mori, H. Mizoguchi and T. Sato, "Unrestrained and non-invasive monitoring of human's respiration and posture in sleep using pressure sensors," *J. RSJ*, vol.16, no.5, July 1998, pp. 705-711 (in Japanese).
- [5] K. Higashikatsuragi, Y. Nakahata, I. Matsunami and A. Kajiwara, "Non-invasive respiration monitoring sensor using UWB-IR", in *Proc. IEEE ICUBW 2008*, Hannover, German, Sept. 2008, pp.101-104.
- [6] N. Nakai, M. Watanabe, Y. Miyake, K. Takeda, K. Yamashita, H. Shinmori and K. Ishihara, "Automatic respiration monitoring system by time-varying image analysis," *IEICE Trans. Inf. & Syst.*, vol.J83-D-II, no.1, Jan. 2000, pp.280-288 (in Japanese).
- [7] K. Abbas, K. Heiman, T. Orlikowsky, S. Leonhardt. "Non-contact respiratory monitoring based on real-time IR-thermography," in *IFMBE Proc. of WC2009*, 25/IV, 2009, pp.1306-1309.
- [8] J. Fei and I. Pavlidis, "Analysis of breathing air flow patterns in thermal imaging," in *Proc. IEEE EMBC 2006*, New York City, NY, Aug. 2006, pp.946-952.
- [9] D. Hanawa, Y. Yaginuma, Y. Enomoto, T. Koide, S. Terada and K. Oguchi, "Automation of non-intrusive nasal breathing detection by using far-infrared imaging," in *Proc. u-Healthcare 2010*, Nov. 2010.
- [10] D. Hanawa, T. Koide and K. Oguchi, "A proposal of nasal breathing detection method by using far infrared imaging in a home health care system," *IEICE Trans. Inf. & Syst.*, vol.J94-D, no.1, Jan. 2011, pp.260-263, (in Japanese).
- [11] T. Koide, S. Yamakawa, D. Hanawa and K. Oguchi, "Breathing detection by far infrared (FIR) imaging in a home health care system," in *Proc. IEEE ISBB 2009 [USB-Memory]*, Melbourne, Australia, Sept. 2009, pp. 206-209.
- [12] Z. Zhu, J. Fei and I. Pavlidis, "Tracking human breath in infrared imaging," in *Proc. IEEE BIBE 2005*, Minnesota, USA, Oct. 2005, pp.227-231.
- [13] F.Q. AL-Khalidi, R. Saatchi, D. Burke and H. Elphick, "Tracking human face features in thermal images for respiration monitoring," in *Proc. IEEE/ACS AICCSA 2010*, Hammamet, Tunisia, May 2010.
- [14] B. Kaur, J. K. Nelson, T. Williams and B. O'Kane, "Adaptive region of interest (ROI) detection and tracking for respiration measurement in thermal video," in *Proc. SPIE*, Vol.8401, May 2012, pp.840117-1-84117-9
- [15] P. Viola and M. Jones, "Robust real-time face detection," *Int. J. Comput. Vision*, vol.57, no.2, May 2004, pp.137-154.
- [16] R. Lienhart and J. Maydt, "An extended set of Haar-like features for rapid object detection," in *Proc. ICIP 2002*, Rochester, New York City, NY, Sept. 2002, pp.900-903.



Kimio Oguchi received the B.E. and M.E. degrees in applied physics, and Dr. Eng. degree in electrical engineering in 1978, 1980 and 1995, respectively, from Nagoya University, Nagoya, Japan.

He joined the Yokosuka Electrical Communication Laboratory, NTT, Japan in 1980. From April, 1980 to March, 2004, he was engaged in the research and development of optical systems, in particular, fiber-optic local area networks, advanced fiber-optic digital systems, ATM-link systems, photonic transport networks, applications of advanced photonic devices, and large contents distribution systems. He was involved in the standardization activities in both ITU-T SG's 13 and 15. He now is a Professor of the Information Networking Laboratory, Graduate School and Science and Technology and Faculty of Science and Technology, Seikei University, Tokyo, Japan. His research interests include information networking issues, applications utilizing e.g. sensor networks and optical networks, and vital functions, and contents distribution networks.

Dr. Oguchi is a senior member of the IEICE, a member of IEEE Communications, IEEE Photonics, IEEE Computer, IEEE EMB Societies, OSA, ICST, Society for Running, Japanese Society for Emergency Medicine, Clinical Gait Analysis Forum of Japan, Digital Cinema Consortium of Japan, and CineGrid. He is recipient of the OECC'97 Best Paper Award, the NOC'97 Best Paper Award, the OECC'98 Best Paper Award, the Best Presentation Award at the 2001 and 2005 Annual Conference of Society for Running, 2008 ICT Accomplishment Award of the ITU Association of Japan and IEICE MVE Prize 2009.



Dai Hanawa received the B.E., M.E., and Dr. Eng. degrees from Ibaraki University in 1999, 2001, and 2005, respectively.

From 2005 to 2007, he was a Research Associate of the Department of Computer and Information Sciences at the Tokyo University of Agriculture and Technology. Since 2007, he has been a Research Associate, and from 2008 he has been an Assistant Professor of the Department of Computer and Information Science at Seikei University. His research interests are in the area of human communications, especially the networked virtual environment, sensor network, vital information sensing and their applications.

Dr. Hanawa is a member of IEEE Computer, IEEE Communications, ACM, IEICE, the Virtual Reality Society of Japan and Human Interface Society. He received the first IEICE HCG Prize and IEICE MVE Prize 2009.

Toshiki Morimoto received the B.E. degree from Seikei University in 2012.

From 2012, he has been with Allied Telesis K.K., Japan. His research interests include vital information sensing and its application.

Mr. Morimoto is a member of IEICE.



Shota Shimazaki received the B.E. degree from Seikei University in 2011. From 2011, he has been a Master Student of Graduate School of Science and Technology, Seikei University. His research interests include fiber-optic broadband networks.

Mr. Shimazaki is a student member of IEICE.