Tracking Human Face Features in Thermal Images for Respiration Monitoring

Farah Q. AL-Khalidi[1]
PhD Research Student,
Faculty of ACES,
Sheffield Hallam University
Sheffield, U.K.
Farah.Q.Al-Khalidi@student.shu.ac.uk

Reza Saatchi[2]
Reader in Computing and Digital Signal Processing
Faculty of ACES,
Sheffield Hallam University
Sheffield, U.K.
r.saatchi@shu.ac.uk

Derek Burke [3]
Consultant in Paediatric Accident and Emergency Medicine, Sheffield Children's NHS Foundation Trust, Sheffield, U.K.
derek.burke@sch.nhs.uk

Heather Elphick[4]
Consultant in Paediatric Respiratory Medicine,
Sheffield Children's NHS Foundation Trust,
Sheffield, U.K.
heather.elphick@sch.nhs.uk

Abstract—**A method has been developed to track a region related to respiration process in thermal images. The respiration region of interest (ROI) consisted of the skin area around the tip of the nose. The method was then used as part of a non-contact respiration rate monitoring that determined the skin temperature changes caused by respiration. The ROI was located by the first determining the relevant salient features of the human face physiology. These features were the warmest and coldest facial points. The tracking method was tested on thermal video images containing no head movements, small random and regular head movements. The method proved valuable for tracking the ROI in all these head movement types. It was also possible to use this tracking method to monitor respiration rate involving a number of head movement types. Currently, more investigations are underway to improve the tracking method so that it can track the ROI in cases larger head movements.**

Keywords facial tracking method, noncontact respiration monitoring; thermal imaging.

I. INTRODUCTION

Monitoring respiration rate is important in medical diagnosis as it has a major role in diagnoses of serious illness [1]. Instruments currently exist to accurately monitor a variety of physiological parameters such as body temperature, heart rate, and blood pressure. However, existing respiration monitoring schemes have some limitations. Several studies were carried out to enhance respiration rate monitoring technology. Storck reported a modality to monitor respiration rate by a nasal temperature probe [2]. Nepal suggested a respiration rate monitoring approach based on an abdominal strain gauge transducer [3]. Respiration rate has also been extracted from electrocardiogram [4]. Werthammer reported an acoustic method of detecting apnea in infants [5]. In all these respiration monitoring approaches, the sensing device needed to be attached to the patient, causing discomfort and possibly altering the respiration rate.

The simplest method of respiration rate monitoring is to manually observe the subject (e.g. chest movements or sensing the air flow from the nose). This is subjective and thus it is not very accurate. Recently, a new infrared (IR) imaging method has been proposed to monitor human breathing and to compute the respiratory rate by statistical [6][7] and fast Fourier transform [8] methods. These methods however required that the subject avoids body movements.

Extensive work has been reported in the tracking area [9] [6][7][8], however most of these were in the visible domain and others have dealt with static subjects.

In our study, a novel method to automatically track a facial region of interest (ROI) related to respiration has been developed. This ROI is the skin area centered on the tip of nose. The method initially detects the physical features of the face in the recorded thermal images and then uses them to locate the ROI. In the following sections, initially an overview of image processing techniques used is provided. Then the methodology used is explained and the results obtained are discussed.

II. IMAGE ENHANCEMENT TECHNIQUES

Image enhancement techniques are used to emphasise and sharpen image features for display and analysis. The principal objective of the enhancement techniques is to process the image so that the resulting image is more informative than the original one. Enhancement is either carried out as a preprocessing step or as post-processing.

The Smoothing spatial filters (LPF) are used for blurring and noise reduction. They select low frequency components and thus reduce or eliminate high frequencies. An increase in the filter or (kernel) size leads to an increase in blurring of the image. Noise reduction can be accomplished by linear or non-linear filters. Smoothing filters are also LPF type. Examples of LPFs are; average filters, order statistical filters and Gaussian filters. Examples of statistical filters include Maximum, Minimum, Median and Mode filters. The filters commonly used for enhancing biomedical images are Median filters; Average filters and Gaussian filters [10][11][12].

The Median filter is a non-linear filter that is used to reduce the noise rather than blurring edges. It is also used to remove impulse noise. The idea of the median filter is to replace each pixel in the image by the median value of the neighboring pixels. The median is based on the size of the kernel used. The pixels values are sorted into their numerical order and then the considered pixel is replaced with the middle pixel value(s). With a kernel size of 3x3, the resulting value is the median (number 5 out of the sorted list of 9).

The Average filter is a linear filter used to reduce the intensity variation between neighboring pixels. Each pixel value in the image is replaced with the average value of its neighborhoods. This has the effect of eliminating pixel values which are unrepresentative of their surroundings as indicated by (1).

$$g(x,y) = \frac{1}{N}\sum_{(n,m)} f(x,y) \quad x, y = 0, 1, n-1 \qquad (1)$$

Where, *g(x, y)* is the filtered image, *f(x, y)* is the original image before it is filtered, and *N* is the total number of pixels in neighborhood.

The Gaussian filter is a linear filter with 2-D convolution operator that is used to blur images and remove noise by Gaussian function. The form of Gaussian function in 2-D is

$$G(x, y) = \frac{I}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \qquad (2)$$

Where σ is the standard deviation and $G(x, y)$ is the filtered image with Gaussian filter. The degree of image blurring is set by σ . The Gaussian filter can be selected from 3x3 up to 7x7 kernel size [10] [12].

III. Automatic Segmentation Techniques

Segmentation can be defined as a process of partitioning an image into component parts, or into separate objects. There are several methods to perform segmentation, including thresholding, edge detection, and region-identification. In this study, we used edge detection segmentation to identify the boundaries of the subjects from the image background [13].

Edge detection is an automatic segmentation technique that can be defined as processes of transforming an input image into an edge map which can be viewed as a line drawing image. The goal of edge detection is to find the boundaries of the object of interest. Edge detection operators are based on the idea that the edge information in an image is found by looking at the relationship a pixel has with its neighbors. In this study we used the Prewitt operator.

Prewitt operator is an edge detection mask that looks for the edges in both the horizontal and vertical directions. It combines this information into a single metric. Its mask is shown by the image neighborhood ,

$$\begin{bmatrix} Z_1 & Z_2 & Z_3 \\ Z_4 & Z_5 & Z_6 \\ Z_7 & Z_8 & Z_9 \end{bmatrix}$$

Row mask $G_x = (Z_7 + Z_8 + Z_9) - (Z_1 + Z_2 + Z_3)$

Column mask $Gy = (Z_3 + Z_6 + Z_9) - (Z_1 + Z_4 + Z_7)$

These masks are each convolved with the image. At each pixel location, there are two numbers: G_x corresponding to the results from the row mask and G_y corresponding to the results from the column mask. These numbers are used to compute the two matrices, the edge magnitude and the edge direction using (3) and (4) [12].

$$\text{Edge Magnitude} = \sqrt{G_x^2 + G_y^2} \qquad (3)$$

$$\text{Edge Direction } \alpha(X,Y) = \tan^{-1}\left[\frac{G_x}{G_y}\right] \qquad (4)$$

IV. Facial Physiological Features

The physical features for human face are represented by the distribution of the blood flow in the vessels. The warmer areas of the face are represented by several regions such as the periobital region (the two areas between the bridge of nose and the inner corner of the eyes). A number of other physical features are represented by the lower facial temperatures. The reason for the cooler areas is lower blood supply and higher exposure to the ambient environment, allowing a convection effect.

There are several regions with the lower temperature such nose, the ear lobes and the checks [13][14].

V. Methodology

The operations to identify and tracking the ROI consisted of the following tasks:

1. Enhancement of the images
2. Segmentation of the subject's face from the image background.
3. Identifying the ROI
4. Tracking the ROI

The thermal images were recorded by using the FLIR A40 type thermal camera. The subject sat on a comfortable chair in front of a wall. The room temperature was constant during the recording at about 25°. The subject was asked to breath normally in a relax manner. The camera was fixed on a

tripod infront of the subject at a distance of about one meter. The camera settings were: the emissivity= 0.92°, reflected temperature=15C° and relative humidity=50%. Thermal sensitivity of the thermal camera at 60 Hz was 0.08 C° at 30 C°. The image recording rate was 50 frames per second.

In order to determine the effectiveness of the tracking algorithm in locating the ROI in successive images, two types of head movements were considered. These were predefined movements and random movements. These are defined as:

- Predefined movements involved head turning to the left and right by about 45°.
- Random movements involved head moving randomly to the left, right, up and down.

The recorded thermal images were enhanced by using the median filter to smooth the image and to reduce the noise. The effectiveness of the median filter to enhance thermal images was investigated. The median filter had a size of 5; Fig.1 shows two thermal images and their corresponding histograms before and after application of the median filter.

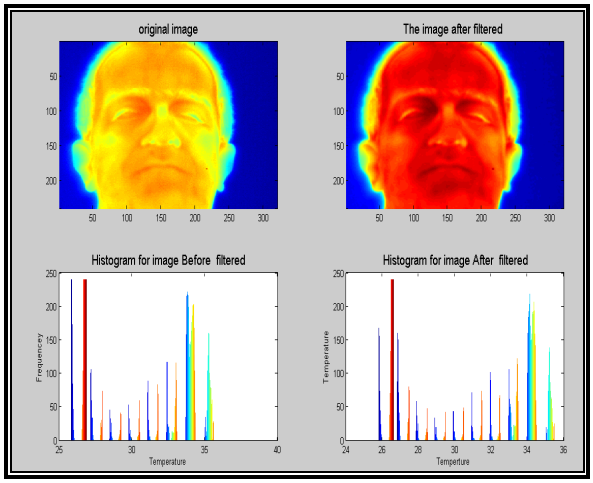

Figure 1. Thermal image enhancement before (left) and after (right) application of the median filter with their corresponding histograms.

The effectiveness of the Prewitt operator to detect the subject's facial boundary was investigated. The two masks for the Prewitt operator were

$$G_x = \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

$$G_y = \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix}$$

These masks were each convolved with the image. The edge magnitude, edge direction were computed by using Eqs.3 and 4. The application of the Prewitt operator to detect subject boundary from the image background is shown in Fig.2.

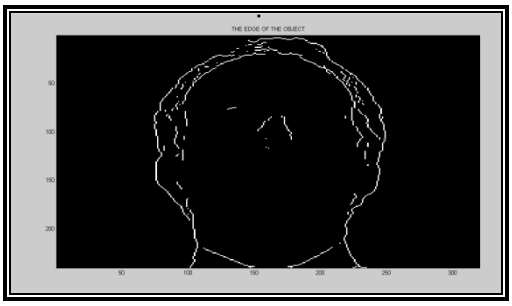

Figure 2. Detection of the subject boundary from the image background

Two methods were used to automatically locate the ROI in successive recorded images. In the first method, once the subject's face was extracted from image background, the center of the resulting image (X_c , Y_c) was identified and the area was represented by a circle; the lowest pixel value within the circled part was obtained. The corresponding location represented the tip of the nose. This ROI were used to estimate the breathing rate. This method worked well in the cases that there were no head movements but it failed when there were head movements as shown in Fig.3.

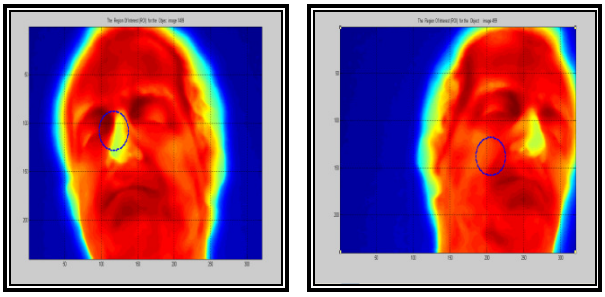

Figure 3. Tracking the ROI using the first tracking method.

The second tracking method overcame the limitation of the first method as explained in the following section. The warmest facial region from the part of the image that had been segmented from its background was determined. This area represented one of the two areas between the bridge of nose and the inner of the eyes. This area was marked. The process was repeated to indentify the second warmest facial region from the part of the image that had been segmented from its background. This area also represented the other of the two areas between the bridge of nose and the inner of the eyes. The lowest average facial temperature from the part of the image that had been segmented from its background below the warmest regions was determined. This corresponded to the tip of the nose.

A circle was placed around the identified tip of the nose. The bottom half of this circle represented the ROI and was used to monitor respiration rate (see Fig4).

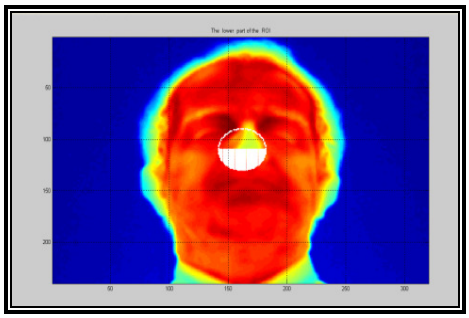

Figure 4. The ROI represented by the lower part of the circle

This ROI best represented the facial temperature changes caused by breathing and was used to determine respiration rate. This method allowed successful tracking of the ROI with small head movements.

The pixel values within the ROI were averaged for each image and the resulting information was plotted against time. In order to remove unwanted noise, this signal was enhance by using bandpass filter with cutoff frequencies 0.05 to 1 Hz.

VI. Results and Discussion

Tracking the ROI with significant head movements was successfully achieved using the proposed second method as illustrated in Fig.5.

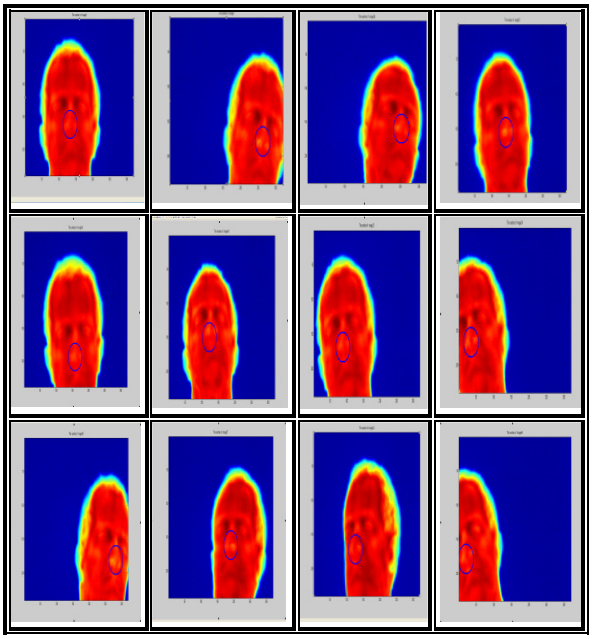

Figure 5.The ROI tracked using the second method.

Visual observations indicated that the temperature was relatively higher around the eyes regions, especially the areas between the bridge of the nose and the inner corner of eyes. Fig.6 shows that the periorbital regions with the highest average temperature.

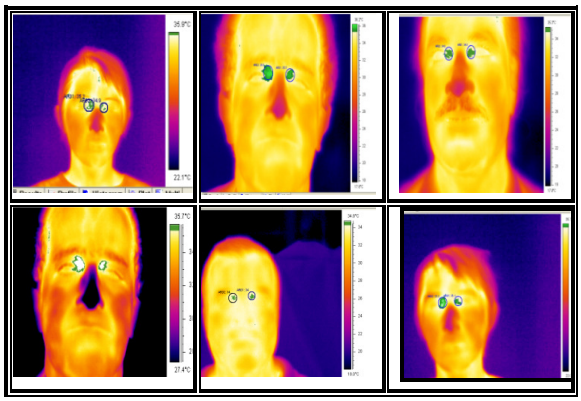

Figure 6. The periorbital regions (marked areas close to the eyes)

It was also observed that the temperature was relatively lowest in the apex of the nose.

Initially the ROI was represented as a rectangle. However, this shape proved to be ineffective in the tracking process. The reasons are indicated in the following section.

For the images that had large diagonal head movements, the rectangular representation of the ROI provided incorrect information as shown in Fig.7.

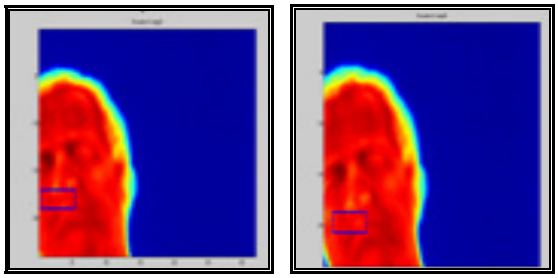

Figure 7. Rectangular representation of the ROI with large diagonal head movement

In some images the nose appeared too closed to the edge of the image as shown in Fig.8. Therefore, the rectangled area fell outside the boundary of the images.

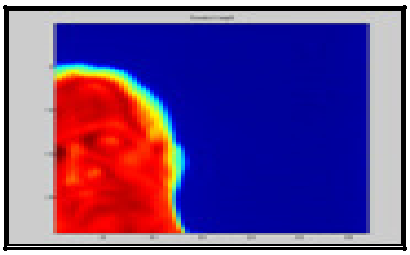

Figure 8. Head movement highlighting the nose very close to the edge of the image.

To deal with the above issues, the ROI was represented by a circle. This dealt with the problems of large diagonal head movements as well as the nose appearing too close to the edge of the image. Fig.9 shows the ROI being represented by a circle. This tracking algorithm was applied to several videos recorded from subjects with different head movements.

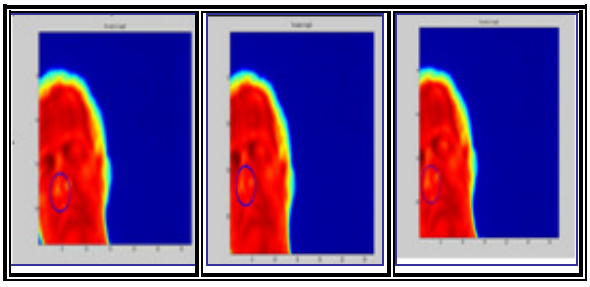

Figure 9. The ROI highlighted by a circle

The tracked ROI for each image in the recorded video was visually inspected. The number of cases where the algorithm did not successfully track the ROI was determined. Based on this observation, an analysis of the accuracy of the tracking method was carried out. The results are summarized in Table1.

TABLE 1 TRACKING ANALYSIS RESULTS FOR DIFFERENT HEAD MOVEMENTS

Recording Number	Number of images	% Failure Rate	Head Movement type
1	6013	0	Random
2	5991	0	Regular
3	1148	0	Random
4	1150	0.5	Random
5	1150	0	Random
6	2580	0.5	Regular
7	550	0	Random
8	50	0	Random
9	50	8.0	Random
10	39	0	Random
11	28	0	Regular
12	35	0	Random

The tacking failure rate was calculated as the total number of failures divided by the total number of images in a recorded video. Tracking failed when a subject opened his/her mouth during the recording or the ROI went too far close to the edge of the image. These issues are currently being investigated to further improve the accuracy of the tracking. The plot of respiration rate information for a subject who performed random head movement is provided in Fig.10. This signal was obtained from a recording consisting of 6000 images recorded over a 2 minutes interval.

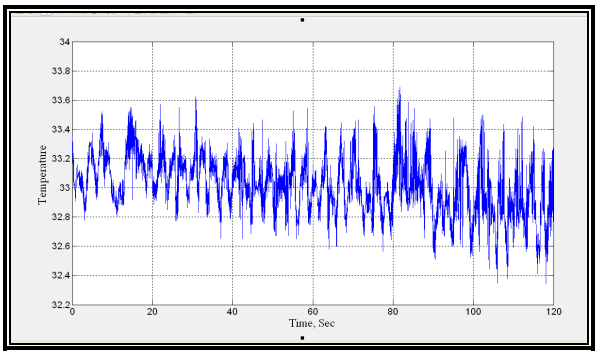

Figure 10. Respiration signal obtained during large head movements

Fig.11 shows the respiration signal shown in Fig.10 following bandpass filtering operation. The filtering has reduced the noise.

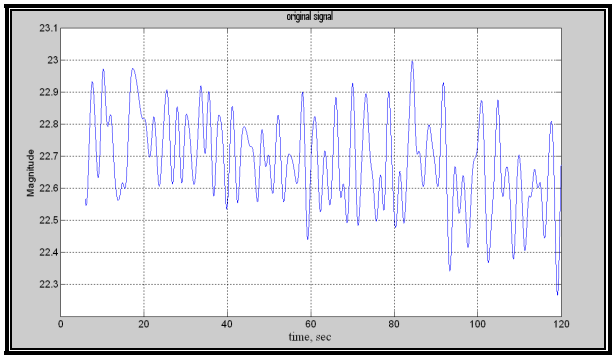

Figure 11. The signal in Figure.10 following band pass filtering

As shown in Fig.11, the increased in temperature represented by breathing out and the decreased in the temperature represented the breathing in are clearly visible. The breathing rate can be calculated by determining the number of breathing cycles per second.

Fig.12 show the respiration signal for a case that the subject did not make head movements during recording. Fig. 13 shows the same information following bandpass filtering. It

can be observed from Fig.13 that the signal is more consistent as compared with Fig.11.

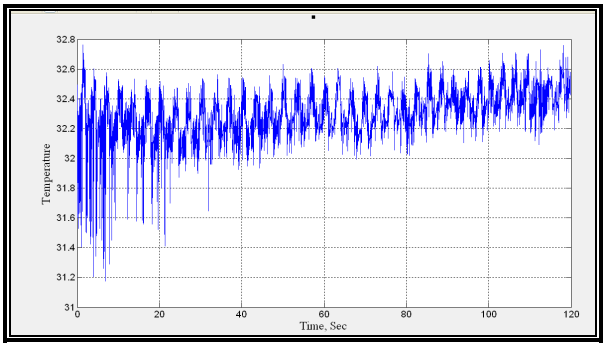

Figure 12. Respiratory signal recorded during no head movements.

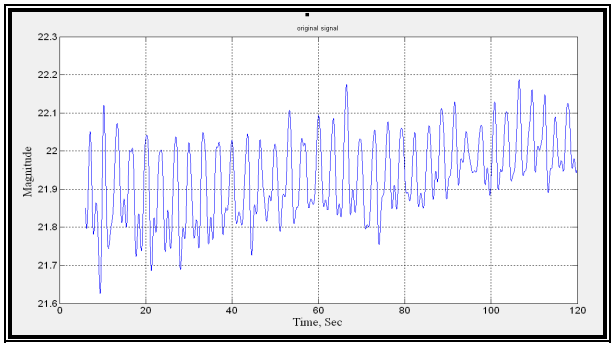

Figure 13. The information in Fig.12 following band pass filtering.

VII. CONCLUSION

In this study, a method to automatically track human face in thermal images was developed. The algorithm proved effective for dealing with static, regular and random head movements.

During recordings, under different head movement types, it was possible to successfully detect the region of interest (ROI) associated with respiration process. This region was used to determine respiration rate in a non-contact manner. This was achieved by selecting human facial physical features. Moreover, image analysis and image processing operations were carried out to enhance the selected ROI.

The system in the future will be modified to deal with situation that the subject's mouth is open during the recordings. Further developments in image processing and image analysis to deal with images that failed by using this method will be developed. Also, more subjects will be included in the study and the accuracy of the thermal imaging based respiration monitoring against conventional contact based methods (such as the one based on the use of nasal thermister) will be investigated.

REFERENCES

[1] J. Fieselmann , S. M. Hendryx , M. C. Helms., D. S. Wakefield, "Respiratory rate predicts cardiopulmonary arrest for internal medicine patients", The Journal of General Internal Medicine, vol.8, pp.354-360. , (1993).

[2] K. Storck , M. Karlsson, P. Ask and D. Loyed,"Heat transfer evaluation of the nasal thermistor technique" , IEEE Transactions on Biomedical Engineering, vol. 43, no. 12, pp. 1187-1191, (1996).

[3] K. Nepal, E. Biegeleisen , and T. Ning, "Apnea detection and respiration rate estimation through parametric modeling", Proceedings of the 28th IEEE Annual Northeast Bioengineering Conference, Philadelphia, Pennsylvania. pp.277-278., , (2002).

[4] B. G. Moody , G.R. Mark , A. M. Bump ,S.J. Weinstein , D.A.Berman ,E.J. Mietus and L.A. Goldberger , "Clinical validation of ECG-derived respiration (EDR) technique" IEEE Computers in Cardiology, vol.13, pp.507-510., (1986).

[5] J. Werthammer , J. Krasner, J. DiBenedetto and A. R.Stark., "Apnea monitoring by acoustic detection of airflow ", Pediatrics vol.71, pp.53-55. ,(1983).

[6] C. Eveland , D. Socolinsky and L. Wolff ,"Tracking human faces in infrared video", Image and Vision Computing, vol.21 pp.579–590, July (2003).

[7] R. Murthy and I. Pavlidis , "Non-contact monitoring of respiratory function using infrared imaging," in IEEE Engineering in Medicine and Biology Magazine.vol.25,pp.57-57., (2006).

[8] J. Fei , Z. Zhu and I. Pavlidis , "Imaging respiratory rate in the CO2 absorption band," in Proceedings of the 27th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, (Shanghai, China), pp.700-590, September 1-4. (2005).

[9] D. Comaniciu, V. Ramesh and P. Meer ,"Real-time tracking of nonrigid objects using mean shift" in IEEE Conf. Computer Vision and Pattern Recognition,vol.2, pp.142-149 , (2000).

[10] N. S. Bankman, "Handbook of medical image processing and analysis". Johns Hopkins University , (2009).

[11] H. Hashim , R. Jailani and M. Nasir , "A visual record of medical skin disease imaging using matlab tools", IEEE Student Conference on Research and Development proceeding, Biomedical Engineering-Malaysia.pp.40-44., (2002)

[12] R. C. Gonzales , R. E. Woods and S. L. Eddins, ."Digital image processing using MATLAB", Pearson Education, the United States of America. (2004).

[13] N. A. Diakides and J. D. Bronzino , "Medical infrared imaging", CRC, published by Taylor & Francis Group, (2008).

[14] V. Pavlidis , J. Fei, Z. Zhu, "Tracking human breath in infrared imaging" in Proceedings of the 5th Annual International Conference of the IEEE Symposium on Bioinformatics and Bioengineering. pp.227-231., (2005).