Tracking Human Breath in Infrared Imaging

Zhen Zhu, Jin Fei, and Ioannis Pavlidis

Abstract—In this paper, we propose a novel tracker to capture the human breathing signal through an infrared imaging method. Human facial physiology information is used to select salient thermal features on the human face as good features to track. The major component of the tracker is Mean Shift Localization (MSL)-based particle filtering. A special measurement model is designed for particle filtering so that the tracker can handle significant head movement and object occlusion. The breathing signal is achieved based on tracking results. The experiments show that the tracker is robust and stable and the recovered breathing signal is clear enough for breathing functionality computation.

I. INTRODUCTION

BREATHING is an important human physiological phenomenon and a vital signal to a number of diseases. Recently, a new infrared (IR) imaging method has been proposed to detect human breathing and compute the breathing rate by statistical methods [1] and FFT [2]. Unfortunately, both methods only apply on static subjects and constrain the practical use of this novel approach. In this paper, we propose a tracker to tackle this problem.

Our approach to track the breathing signal is inspired by many other methods which are applied to tracking faces and facial features. A lot of work has been done in this area [3][4][5][6][7]. The task is challenging since that the human face is a deformable image object which varies with different poses and the tracker is sensitive to lighting illumination. Most feature-tracking methods are designed to track eye, lip, or mouth movement and are used for application such as human expression analysis and speech recognition. Our problem is two-fold: (1) infrared images are essentially different from those in visible domain; (2) directly tracking breathing flow is difficult. As a result, we need to identify facial features in IR images and track these features so that the breathing air flow region can be indirectly inferred from the tracker's results.

In infrared images, all pixels convey information about heat emissivity. In a thermal image of the human face, we have found that the temperature is relatively high around the eye regions, especially periorbits, small areas between the eye and bridge of the nose. There are other regions on the face exhibiting relatively high temperatures, e.g., the forehead, the corner of mouth. However, it is our observation that hot pixels are consistently found in periorbits throughout all thermal facial images. Meanwhile, we have found that some regions of the face have relatively cold temperatures, for example, the nose and ear lobes. These temperature features are the best candidates for tracking breathing since they are local extreme values. In this paper, we choose the periobital regions and the nose as our Tracking Region of Interest (TROI). The TROI guides us to determine the Measurement Region of Interest (MROI), from which we retrieve the breathing signal.

An algorithm is needed to detect the local extreme values corresponding to these facial features. Mean Shift Localization (MSL) [10] [11] is a suitable way of accomplishing this task. MSL itself, however, is a low level method which can not detect global information and is prone to tracking failure. To fully take advantage of its strength, we use an MSL-based particle filter as our tracking framework. Similar methods have been reported recently [8] [9]. However, single feature point on face regions is difficult to be tracked by this method. The novelty of our approach is that we consider geometric, displacement, and temperature constraints to specially design a measurement model for particle filtering. Experiments show that the TROI can be followed both under significant head rotation and when part of the TROI is occluded.

On determining the TROI, the breathing signal will be adaptively recovered from a region called the Measurement Region of Interest (MROI), which consists of the region below the nose and is expected to have the strongest breath signal. Experiments prove that the breathing signal is fairly clear, and the improved MSL based tracker leads to a more practical modality for breathing measurement.

The paper is organized as follows: in Section II, we determine IR facial temperature variations which supply good features to track. Then we describe principles of MSL based particle filtering in Section III. In Section IV, we will see how to infer the MROI using the TROI. Finally, experiments will be demonstrated in Section V. We conclude the paper in Section VI.

II. FACIAL PHYSIOLOGICAL FEATURES

The human face contains abundant blood vessels. One of the hottest areas of face is periorbital region, the two areas between the bridge of nose and the inner corner of the eyes. We can see this physiological phenomenon clearly in Fig. 1. At the same time, the temperature of skin extending from the bridge of nose to the apex of nose is relatively low

Z. Zhu is with the Computer Science Department, University of Houston, Houston, TX 77204(e-mail: zhu@ cs.uh.edu).

J. Fei is with the Computer Science Department, University of Houston, Houston, TX 77204(e-mail: jinfei@ cs.uh.edu)

I. Pavlidis is with the Computer Science Department, University of Houston, Houston, TX 77204 (phone: 713-743-0101; fax: 713-743-3335; e-mail: pavlidis@ cs.uh.edu).

compared to the surrounding skin. This is due to a lack of blood supply to these areas and higher exposure to the ambient environment, allowing a convection effect to cool them down. These features shown in Fig. 1 are consistently found for all subjects in spite of the fact that the exact thermal pattern varies from person to person. These two regions give us good tracking features for two reasons. First, local extreme temperatures provide salient features which are easy to detect. Second, these regions are very close to the nostrils which are the best location to capture the breathing signal.
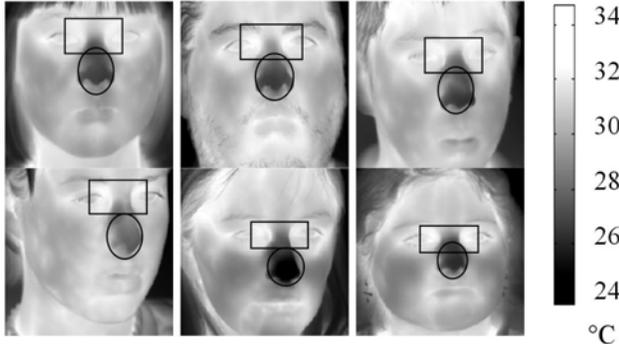

Fig. 1. Thermal faces with salient features. Rectangles cover periorbital regions with high temperature; Ellipses cover nose regions with low temperature.

In terms of our observation, we choose three rectangular areas as our Tracking Region of Interest (TROI). As in Fig. 2, these three rectangles ought to cover the majority of the extreme-valued pixels in the periorbits and nasal apex. Currently, we manually select these three windows in the first frame of the video clip and apply MSL to each of these windows.
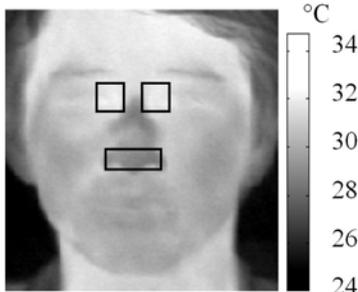

Fig. 2. Selection of three rectangular windows of the TROI

III. MEAN SHIFT LOCALIZATION-BASED PARTICLE FILTERING

In this section, we will focus on the technical side of the breathing tracker. The tracking framework is particle filtering, a Monte Carlo method for Bayesian tracking. We utilize Mean Shift Localization (MSL) to search the local extremes and particle filtering to select the best tracking estimation.

MSL [10] is an algorithm to find the local mode of a distribution and has been successful in tracking visible domain objects [11]. In our problem, the TROI consists of three rectangular windows in Fig. 2. Each of these windows applies MSL independently to converge to modes of the temperature distribution. Since we are interested in extreme temperature values, MSL assigns pixels weights proportional (in the periorbital region) or inversely proportional (in the nasal region) to their temperatures.

A. Bayesian Tracking by Particle Filtering

From a Bayesian point of view, tracking an object through a series of images is a process of probability density propagation [12]. For the breath tracking problem, we need a system evolution model to generate hypotheses and a measurement model to evaluate hypotheses by assigning them different weights in terms of their correctness.

The first step in particle filtering is prediction. The prediction step uses a system evolution model to produce a certain number of hypotheses. The vector X_k for state k is represented as $\left(x_k^{(1)}, x_k^{(2)}, ..., x_k^{(6)}\right)^T$, in which each element represents the coordinates of the three windows of the TROI. The system model is given by a linear equation: $X_k = 3X_{k-1} - 3X_{k-2} + X_{k-3} + v_k$, where X_k is predicted by 3 previous states X_{k-1}, X_{k-2}, X_{k-3} and v_k is a noise component which has a Gaussian form. Then we use these hypotheses as inputs for MSL to find local extremes. After these hypotheses converge to their local extremes, we use a measurement model to assign weights to the hypotheses.

B. Measurement Model of Particle Filtering

MSL may leads to hypotheses with different local extremes. Hence, we need to evaluate their correctness through a good measurement model. Unfortunately, a measurement model is not straightforward for our problem. The reasons are: (1) convergent hypotheses are not necessarily correct answers. Other parts of facial area may have hot or cold spots which are very similar to the windows of the TROI; (2) temperature-based evaluation is not reliable. The temperature of the TROI varies with the position of the head so that hypotheses are easily deviated to other local extremes; (3) IR-opaque CO_2 airflow may occasionally block the TROI. This directly leads to tracking failure even without significant head movement. Therefore, we can not evaluate hypotheses using only temperature information.

To tackle the above problems, we design a measurement model which combines geometric constraints, temporal displacement and temperature all together to evaluate the importance of each hypothesis. The major ideas are:

(1) The three windows of the TROI are virtually connected each other to form a "periorbits-nose" triangle. Lengths of sides of the triangle are constrained within a certain range. We define the geometric weight $W_{Geom} = 1$ if within this range, 0 otherwise.

(2) Deformation of this triangle is assumed to be limited from frame to frame. Deformation here includes displacement and distortion. The deformation weight is defined as $W_{Deform} = e^{-s_1 \left\| \Phi_k^{[ijk]} - \Phi_{k-1} \right\|}$, where s_1 is a positive scale

factor and $\Phi_k^{[ijk]}$ is a triangle for the current state hypothesis. The three vertices are given by i^{th} hypothesis for the left periorbit, j^{th} hypothesis for the right periorbit, and k^{th} hypothesis for the nose.

(3) The temperature within the windows of the TROI does not change dramatically from frame to frame. The temperature weight W_{Temp} is: $W_{Temp} = e^{-s_2 \|T_k^{[i]} - T_0\|^2}$, where s_2 is a positive scale factor, $T_k^{[i]}$ is the current mean temperature for the ith hypothesis in one window of the TROI, and T_0 is the original temperature for the same window of the TROI as $T_k^{[i]}$.

Having all these three weights, a measurement model is given by $p(Z_k|X_k) \propto c + W_{Geom} \cdot W_{Deform} \cdot W_{Temp}$, where c is a constant which guarantees that a good hypothesis will survive a short period of measurement failure.

C. Procedures of the MSL-based Particle Filtering

The procedure of MSL-based particle filtering is as followings:

(1) Initialization. Generate M hypotheses x_0 according to the feature selection for the first frame. Repeat following steps (2) ~ (6) for incoming frames.
(2) Prediction. Use system model to generate hypotheses x_{k+1}'.
(3) Convergence. All hypotheses use MSL to find their local extremes, $x_{k+1}*$.
(4) Weighting. Assign a weight to each converged hypothesis according to the measurement model.
(5) Updating. Resample these hypotheses in terms of their weights.
(6) Go back to step (1) and repeat the same procedure.

If we do not have any triangle that is qualified for constraints, we assume at least one window of TROI is lost or blocked. Then, this window will be regenerated according to the geometric positions of other two windows. Currently, we do not consider the scenario that more than one window of the TROI is lost or blocked during tracking. Fortunately, the probability of losing control of two windows is very low.

IV. LOCATING MROI

On locating the TROI, we need to infer the Measurement Region of Interest (MROI). According to weak perspective camera model [13], from the object frame to the camera frame, the object experiences a scaled transformation followed by an orthographic projection to the image frame from the transferred object in the camera frame. Hence, the MROI is projected orthographically and varies in terms of different poses. Through our experiments, the size of the MROI is not very sensitive to the measuring accuracy of the breath signal given that it reasonably covers the breath flow region. Currently, we only compute the mean temperature within the MROI and use it as the breathing intensity for each frame. The series of mean temperature variations will be utilized to compute breath frequency.

V. EXPERIMENTS AND PERFORMANCE ANALYSIS

We capture the video clips by our highly automated IR image processing platform ATHEMOS [14]. The image is taken by adding a 4.3μm narrow band filter on the camera lens to visualize CO_2 airflow. The size of the image frame is 320*254 pixels, and the video is sampled at 50-55 frames/second. The number of hypotheses in particle filtering is M=15 for each window of the TROI. We use four video clip files, which are named D005-030, D005-031, D005-042, and D005-043.

A. Tracking with Significant Head Movement and TROI Occlusions

First we see how the tracker deals with significant head rotation (Fig.3. (a-i)). In frame 83, the TROI is clearly visible so that all hypotheses converge to similar points. In frame 102, the hypothesis for the right periorbit begins to slightly be distracted from its unique location because the visibility decreases as the head turns. In frame 137, most of the right periorbit is invisible and all hypotheses for that window converge to the left periorbit and forehead. Thanks to the geometric constraints, the tracker is aware of the loss of control for the right periorbit and therefore generates a new window position for the right periorbit. We see that by this artifactual interference, we still can infer the MROI approximately by locating the TROI. The same situation repeatedly occurs in frames 183, and 242. When the head turns back, we see in frame 245 that the tracker regains control of the right periorbit. In frame 249, the window for the right peri-orbit is lost again because the TROI is occluded by CO_2 airflow. After that, the tracker recovers all of three windows in frame 253. We see that our tracker works very well under significant head movement. The necessary condition however is that two windows of the TROI are under control of the tracker.

Using a similar lost-and-recover mechanism, we see in Fig. 3. (j-q) that how the tracker recovers from TROI occlusions by the subject's hand. By frame 108, we see that the tracker has fully recovered its control over the three windows of the TROI. The hand print left on the face is still visible.

B. Breathing Signal Through MROI

We test the trackers ability to recover the breathing intensity signal from the video file D005-30, in which the subject moves his head slowly followed by some quick movements. Since the CO_2 flow absorbs emissivity, the MROI has a lower temperature when expired flow comes through. We record the mean temperature in the MROI as given by the white quadrilaterals in Fig. 3 and plot its variation against time as in Fig. 4. To visualize the breathing cycles, we manually indicate the expiration cycle by a shadowed bar with the sequence number. For the first five cycles, the average temperature of the MROI almost ideally

COMPUTER SOCIETY

reflects the expiration cycle. After that, the temperature signal fluctuates with the increasing amplitude of movement. Overall, the temperature does drop during expirations, which can be used to represent a breathing cycle. After cycle number 8, quick movement obviously brings a lot of noise to the signal.

C. A Comparison Between Trackers With and Without Constraints

To show the importance of the constraints introduced for our measurement model and provide a quantitative evaluation of our breathing tracker, a comparison is conducted between trackers with and without constraints. The tracker without constraints uses only temperature to evaluate hypotheses. The comparison results are given in Table I. We divide each video file into equally long clips and count the number of failures for each clip. For shorter files D005-030 and D005-031, the length of clip is 200. For D005-042 and D005-043, it is 1000. The rate of failure is calculated as the total number of failures divided by the total number of frames. We define failure when the TROI triangle violates a geometric constraint, which defines the size range and orientation of the triangle. From Table 1, we see that the tracker with constraints considerably improves the tracking performance. Occasional failures for tracker with constraints are due to the subject being out of the viewing scope, moving too fast, and other artifactual reasons.

TABLE I
COMPARISON BETWEEN TRACKERS WITH AND WITHOUT CONSTRAINTS

	D005-030		D005-031		D005-042		D005-043	
Length	2835		2614		17416		17577	
Item	WiC.	WoC.	WiC.	WoC.	WiC.	WoC.	WiC.	WoC.
Rate of Failures (%)	3.04	10.61	1.58	57.62	0.2	87.6	1.84	46.92

Note: WiC.: with constraints; WoC: without constraints

VI. CONCLUDING REMARKS

We have developed a tracker to automatically follow the human breathing signal using an infrared imaging system. The tracker is robust under different circumstances and is able to effectively recover the breathing signal. This has been achieved by selecting human facial physiology features as a TROI and tracking by MSL-based particle filtering. We impose geometric and displacement constraints on the windows of the TROI and design a measurement model in particle filtering, which makes the tracker perform reliably under different head poses and when part of the TROI is occluded. Experiments show that the specially designed measurement model can prevent a number of tracking failures which are inevitable in a tracker without these constraints.

Further work is needed in the near future: (1) determine the optimal number of hypotheses in particle filtering to save as many computational resources as possible; (2) further investigate how the head position and movement influences the breathing signal sampling from the MROI; (3) automatically recover from tracker failure even when more than one window of the TROI is lost; (4) signal processing for breathing analysis when noise is introduced by the tracker.

REFERENCES

[1] R. Murthy and I. Pavlidis, "Non-contact monitoring of breathing function using infrared imaging," to appear in IEEE Engineering in Medicine and Biology Magazine.

[2] J. Fei, Z. Zhu, and I. Pavlidis, "Imaging breathing rate in the CO_2 absorption band," in Proceedings of the 27th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, (Shanghai, China), September 1-4, 2005.

[3] Z. Zhu and Q. Ji, "Robust Real-time Eye Detection and Tracking Under Variable Lighting Conditions and Various Face Orientations," Computer Vision and Image Understanding, Vol. 38, No. 1, pp.124-154, 2005.

[4] M.L. Cascia, S. Sclaroff, and V. Athitsos, "Fast, Reliable Head Tracking under varying illumination: An approach based on registration of texture-mapped 3D models," IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 22 pp. 322-336, 2000.

[5] G. Edwards, C.J. Taylor, and T.F. Cootes, "Learning to Identify and Track Faces in Image Sequences," in Proceedings of IEEE Int. Conf. Automatic Face and Gesture Recognition, 1998.

[6] N. Olivier, A. Pentland, and F. Berard, "Lafter: lips and face real-time tracker," in Proc. IEEE Conf. Computer Vision and Pattern Recognition, 1997.

[7] M.H. Yang, D.J. Kriegman, and N. Ahuja, "Detection faces in images: a survey," IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 24 pp. 34-58, 2002.

[8] B. Han, D. Comaniciu, Y. Zhu, and L. Davis, "Incremental density approximation and kernel-based Bayesian Filtering for Object Tracking," IEEE Conf. Computer Vision Pattern Recognition, 2004.

[9] C. Chang, R. Ansari, and A. Khokhar, "Multiple Object Tracking with Kernel Particle Filter," IEEE Conf. Computer Vision Pattern Recognition, 2005

[10] K. Fukunaga, and L.D. Hostetler, "The estimation of the gradient of a density function, with applications in pattern recognition," IEEE Trans. Information Theory, Vol. 21, pp. 32–40, 1975.

[11] D. Comaniciu, V. Ramesh, and P. Meer, "Real-time tracking of non-rigid objects using mean shift," in IEEE Conf. Computer Vision and Pattern Recognition, 2000.

[12] S. Arulampalam, S. Maskell, N.J. Gordon, and T. Clapp, "A tutorial on particle filters for on-line non-linear/non-Gaussian Bayesian tracking," IEEE Transactions of Signal Processing, Vol. 50 pp.174-188, 2002
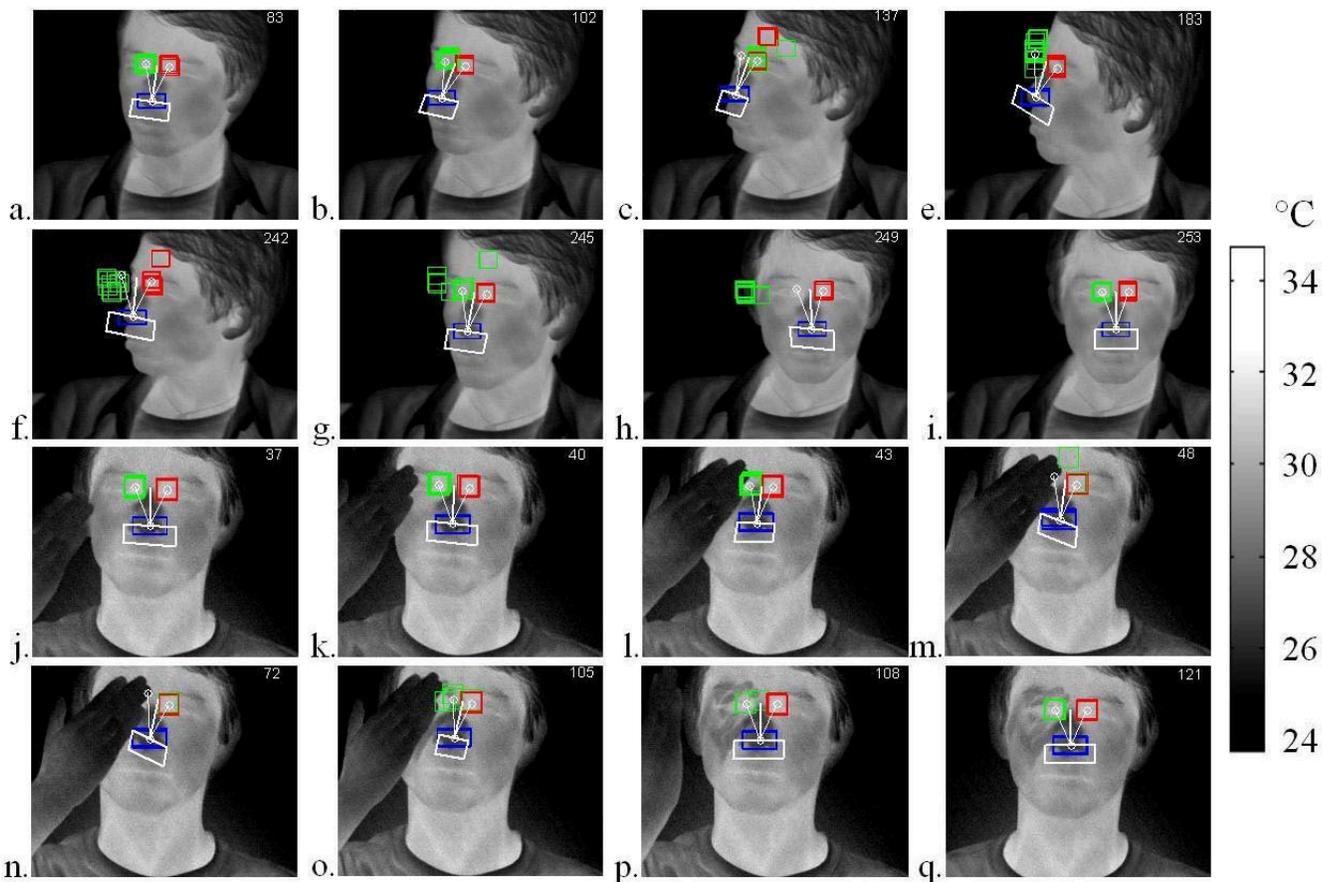
IEEE COMPUTER SOCIETY

Fig. 3. Tracking with significant head moving (a-i) and occlusion (j-q). Green rectangles are for the right periorbit, red ones for the left periorbit, and blue for the nose. The white circles represent the best estimation of the TROI and the white quadrilateral is the MROI. The number on top right is the frame index. All images are from video clip D005-043
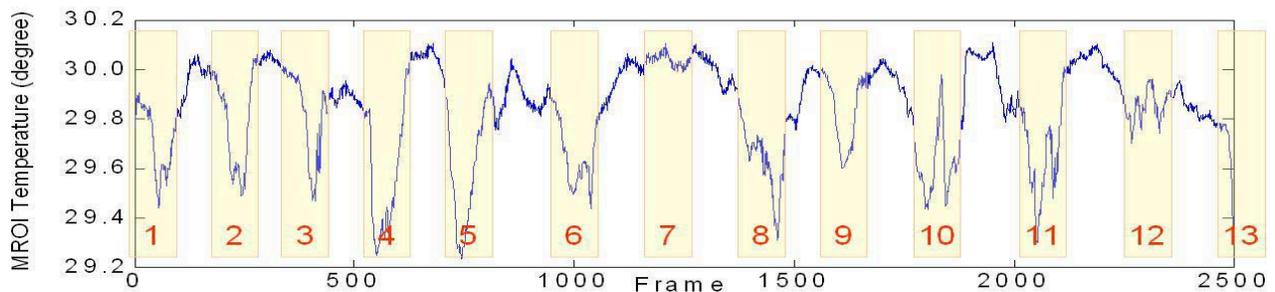

Fig. 4. . Breathing signal against time. The curve is given by mean temperature in MROI. The labeled shadowed bars indicate each inspiration cycle, during which the CO_2 occludes the emissivity from facial regions to lower the temperature.