© May 2019 | IJIRT | Volume 5 Issue 12 | ISSN: 2349-6002

Thermal Image Usage for Facial Tissue Tracking From Spatial Arrangement

D. Malathi[1], A. Mathangopi[2], Sandra Karunya[3], Dr. G. Rajinigirinath[4]

[1]*P.G, Department of CSE, Sri Muthukumaran Institute of Technology, Chikkarayapuram, Chennai, Tamil Nadu-600069, India*

[2,3]*Asst.Prof. Department of CSE, Sri Muthukumaran Institute of Technology, Chikkarayapuram, Chennai, Tamil Nadu-600069, India*

[4]*Prof. Department of CSE, Sri Muthukumaran Institute of Technology, Chikkarayapuram, Chennai, Tamil Nadu-600069, India*

***Abstract*- Accurate tracking of facial tissue in thermal infrared imaging is challenging because it is affected not only by positional but also physiological (functional) changes. This paper presents a particle filter tracker driven by a probabilistic template function with both spatial and temporal smoothing components, which is capable of adapting to abrupt positional and physiological changes. The method was tested on tracking facial regions of subjects under varying physiological and environmental conditions in 25 thermal clips. It demonstrated robustness and accuracy, outperforming other strategies. This new method promises improved performance in a number of biomedical applications that involve physiological measurements on the face, such as unobtrusive sleep and stress studies.**

Index Terms- Facial tracking, matte, sleep studies, stress studies, thermal imaging.

I. INTRODUCTION

IN the last few years, facial tracking in the thermal infrared spectrum received increasing attention. Initially, applications in surveillance and face recognition were the driving force, where thermal imaging has the distinct advantage of being insensitive to lighting conditions [1] [2]. Later, physiological variables, such as vital signs, proved measurable in this modality [3]–[6], which gave rise to applications in human–computer interaction (HCI) [7], medicine [8], and psychology [9]. The degree of success of such measurements depends on a tracking method that can reliably follow the tissue of interest over time. For example, in sleep studies, if the tracker momentarily loses the nasal region of interest (ROI), the generated breathing signal is far from accurate (see Fig. 1), which affects the ensuing analysis. Thus, the specification of a facial tracker in thermal infrared needs to be quite stringent.

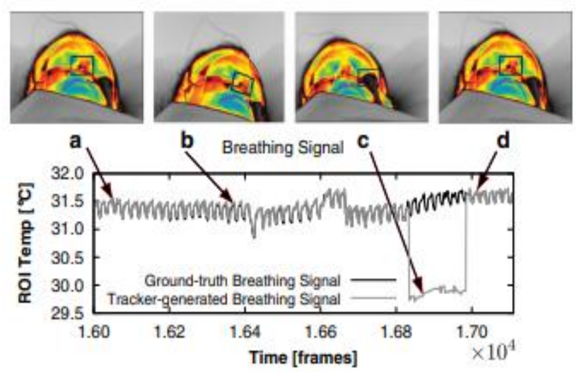

Fig.1. Tracker generated compared with ground-truth breathing signal. (a) Initial frame with the rectangular ROI centered on the nostrils. (b) and (d) When the tracker works well, the generated breathing signal is good. (c) When the tracker loses the ROI, the generated breathing signal is inaccurate.

The proposed method uses a particle-filter tracker, which is driven by a template-based objective function. The choice has to do with the peculiarities of thermal imaging and the needs of the targeted applications. Tracking based on shape models [10] is not very appealing in facial thermal imaging, because the modality images function rather than structure. To give an example, consider the case of tracking nasal tissue in thermal infrared imagery for the purpose of computing breathing function. Under normal conditions, the nose is colder than the surrounding tissue due to convection from nasal air flow. This translates to a characteristic thermal shape similar to the one appearing in visual (structural) images. At some point, an irritant reaches the subject's nasal cavity, there is an allergic reaction that blocks air flow in the nostrils and breathing continues mainly through the mouth. Because air flow is severely curtailed, the temperature over the nasal tissue rises and the nose blends with the surrounding tissue in the imagery. The nose's functionality has dramatically changed and so its characteristic shapes (see Fig. 2). In such cases where stochastic physiological changes affect thermal emission, a shape model tracker may encounter significant difficulties. Note that spatial resolution in thermal imagery is typically lower than that in visual imagery (640 × 512 pixels in our case). Edges in thermal imagery are fuzzy due to diffusion, complicating matters further. None of these factors is conducive to shape-based tracking. Finally, many of the targeted applications are in medicine and HCI. This necessitates a computationally "light" tracker for real-time performance. For example, if the nasal tracking and signal extraction method were not real time, then it would be impossible for a medical technician to optimize the continuous positive airway pressure in a sleep intervention [11].

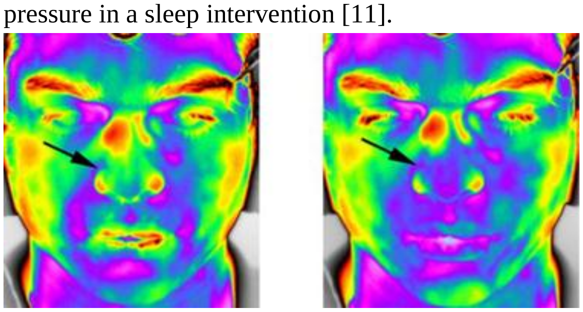

Fig. 2. (a) Thermal image of a subject with normal breathing function. The nasal area is at a contrast with the surrounding tissue. (b) Thermal image of the same subject a few minutes later, when his breathing function is impaired by an irritant. Notice the blending of the nasal area with the neighboring tissue.

Although trackers based on shape models are not appropriate for the problem at hand, tracking methods based on statistical filtering, such as Kalman or particle filters, are quite appealing. In particular, we chose to proceed with particle filtering because in the context of sleep and stress studies of interest, the subjects exhibit infrequent and abrupt turns of the head, which are highly nonlinear. Indeed, particle filtering is not only a general mechanism free of explicit modeling, but it can also handle nonlinear motion in the predict-update loop. We opted to implement the update operation in the predict-update loop through a probabilistic template algorithm. We will demonstrate that this combination of particle filtering with a probabilistic template produces a fast, flexible, and accurate tracker, fulfilling the specifications of the application domain.

II. METHODOLOGY

We use particle filtering to track the ROI's position in the current frame, based on template matching. We denote the motion state of the tracker at time t by Xt and its observations by Zt. The state of the tracker consists of three variables Xt = (xt,yt,θt), where xt and yt are the spatial coordinates of the ROI's centroid and θt is the ROI's rotation angle. The observation Zt refers to the pixels in the current frame. The particle filter tracker uses N = 100 particles (candidate ROIs) in a single iteration per frame. We resample particles for each frame according to

$$X_t = X_{t-1} + V_t \qquad (1)$$

where Vt is a 2-D independent and identically distributed Gaussian noise process. The mean of Vt is zero and the variance is one of the parameters that can be adjusted to ensure an efficient filter performance.

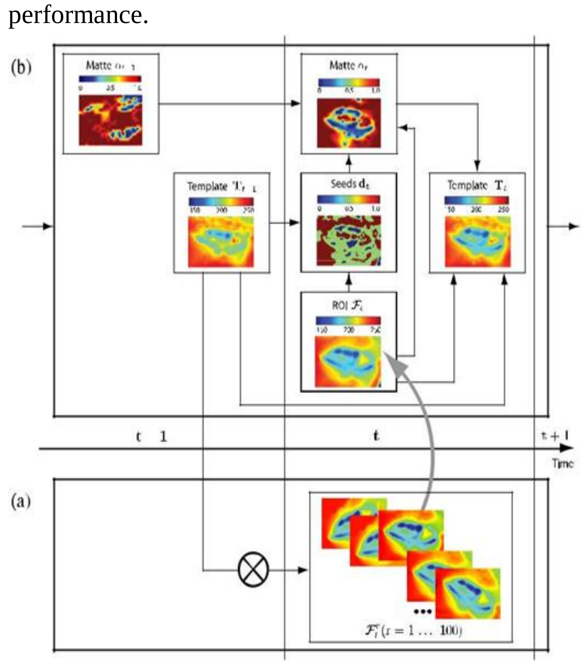

Fig. 3. Illustration of algorithmic flow. (a) First, the likely current ROI is selected based on a MAP

estimate (particle filtering). (b) Then, the template is updated based on the formation of a Matte from stable/unstable seeds, the previous template, and the MAP ROI estimate. The updated template is used in the next time step.
To this end, a mask matrix αt is constructed with the following properties.
1) It has the same size as the template. This one-to-one correspondence between the template and the mask is used to identify which template pixels need to be updated.
2) Each entry in the mask matrix is a number in the range [0, 1] indicating the probability that the corresponding pixel is stable. The larger the probability is, the less the pixel needs to be updated. The most stable and the most unstable pixels in the Matte provide the seed values, which initialize the current Matte computation.
3) The mask takes into account spatial information (i.e., changes occur in regions and not in isolated pixels), providing a smooth outcome on the image plane.
4) The mask also takes into account temporal information (i.e., changes occur in finite time windows), providing a smooth outcome along the time line. The user manually inputs the initial template in the first frame, by selecting a rectangular ROI with the mouse. Next, for each incoming frame, the method automatically performs the template updating according to the following steps (see Fig. 3).
Step 1: Extraction of stable and unstable seeds. Step 2: Computation of the spatiotemporal Matte (STM.
Step 3: Update of the template. The thus formed template correlates with candidate ROIs in the next time step to form the weights in the particle filtering process [see (3)].

III. EXPERIMENTAL DESIGN, RESULTS, AND ANALYSIS

For the purpose of testing the STM template update method in the context of particle filter tracking, we used 25 thermal clips from 24 subjects. The clips were generated as part of sleep studies (subjects identified as Sxx and Lxx) [8], a stress study related to mock-crime interrogation (subjects identified as DxxxIS) [9], and a stress study related to inanimate laparoscopic surgical training (subjects identified as DxxxSS), as per the approval of the appropriate institutional review boards. The set included clips that had at minimum ∼6500 and at maximum ∼49500 frames. At the recoding speed of 30 frames/s, these clips ranged from ∼3.5 to ∼27.5 min in duration. From each thermal clip, challenging segments that featured significant positional and/or physiological change were selected for facial tracking; in some cases, these segments were as long as ∼8000 frames (∼5 min). The targeted facial areas included the nostrils, where vital physiological function is resident, or the periorbital, supraorbital, or maxillary regions where sympathetic activation is manifested. The STM particle filter tracker was compared with the online appearance model (OAM) tracker reported in [13] and the zero-one particle filter tracker reported in [16]. The trackers optimized three state variables, which served as ROI descriptors. These were (x,y) for translation and φ for rotation on the image plane. The templates in all three trackers were formed out of normalized thermal values. All three tracking methods achieved real-time (>25 frames/s) performance on a PentiumIV 4-core computer, with 4 GB memory. A short commentary about each method and the rationale for its inclusion in the benchmarking set is given in the following.
1) STM method: This is the method proposed in this paper that combines the agility of the particle filter framework with the sophistication of a spatiotemporal smoothing template.
2) OAM method: This is an advanced template method applied in visual facial tracking [13]. Thus, it can serve as a representative of noteworthy approaches from the relevant visual imaging literature. The method's statistical template is formulated as a mixture of three components [13], namely a stable component (S), a wandering component (W), and an outlier (L) component. The stable component captures the portion that is stable over time. It follows a normal distribution, the mean and variance of which are updated at every time step. The wandering component represents sudden appearance change. The outlier component is for short time occlusions. The method is probabilistic in nature but has no assumptions about spatial and temporal dependence.
3) Zero-One method: This is a method that combines the agility of the particle filter framework with the simplicity of a deterministic template [16]. It was

recently used for facial tracking in thermal infrared. Therefore, it demarcates progress in the particular domain and can demonstrate the clear benefit of using more sophisticated templating to drive the particle filter loop.

The particle filter mechanisms of the Zero-One and the STM methods featured identical parameterizations. For every subject, all three trackers were tasked to track a selected facial tissue (ROI) from the exact same initial frame.

A. Qualitative Results

In thermal facial imaging, there are two major factors that affect the tracker's performance: head motion and physiological changes. The first factor alters the ROI location, while the second factor affects the pixel values within the ROI. For this reason, our dataset features subjects that exhibited large/small changes in the position or/and the physiology of the ROI. Accordingly, we split the dataset into three groups reflective of three distinct operational scenarios.

1) *Scenario 1*: Subjects that exhibited large changes in position and small changes in physiology.

2) *Scenario 2*: Subjects that exhibited small changes in position and large changes in physiology.

3) *Scenario 3*: Subjects that exhibited large changes both in position and physiology.

Fig. 4 shows a panorama of all subjects in the dataset categorized per operational scenario. Two characteristic thermal shots are shown for each subject with the STM tracker reliably tracking a facial tissue of interest. This figure gives a visual insight to the diversity of experimental circumstances that STM can negotiate. Fig. 5 shows comparative tracker performance for a case representative of Scenario 3, where large positional and physiological changes occur. The signals represent the evolution of the translational and rotational errors of STM, OAM, and zero-one regarding the tracking of subject's D009IS nasal ROI. STM performs flawlessly in terms of translational accuracy and exhibits only small rotational errors in a short interval. OAM exhibits moderate translational errors and large rotational errors for extended periods of time. Zero-one maintains translational accuracy but exhibits significant rotational inaccuracy. This figure gives a dynamic sense of tracker performance, associating error numbers with visual impressions.

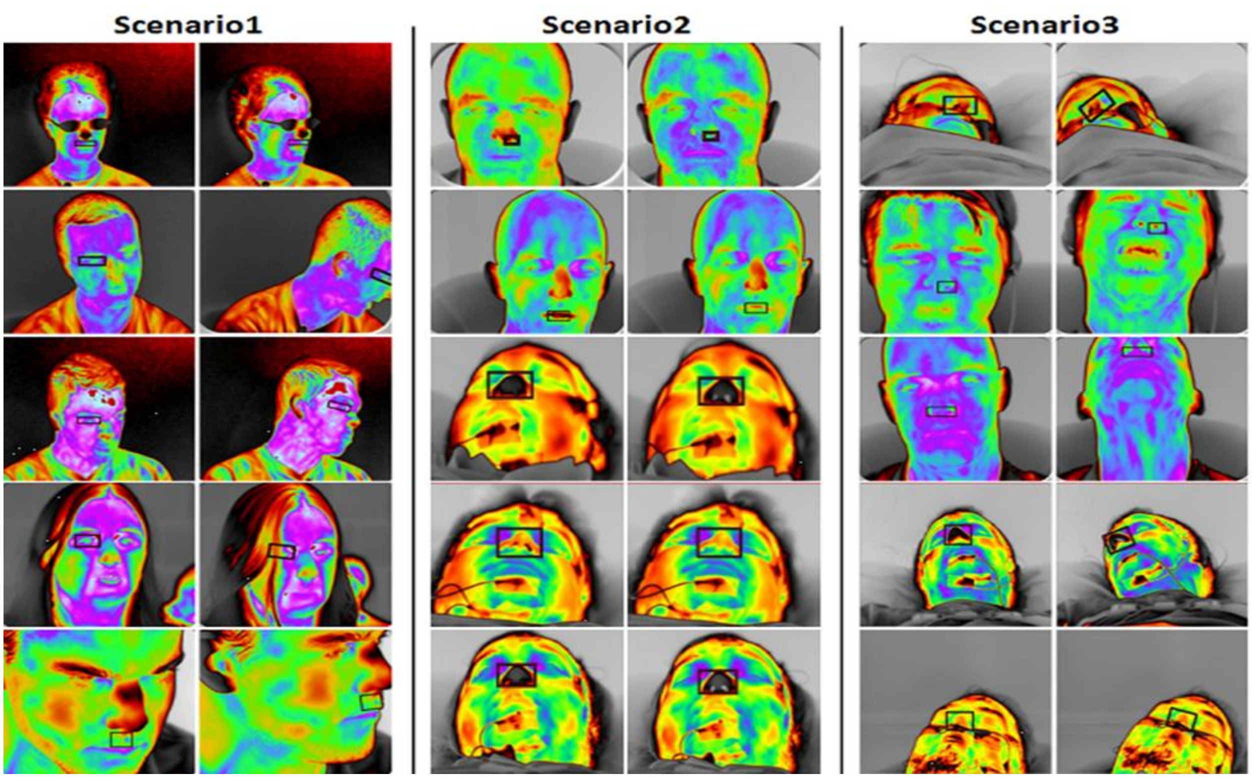

Fig. 4. Panorama of characteristic thermal shots (two per subject) from 24 out of the 25 thermal clips in the dataset, categorized per operational scenario.

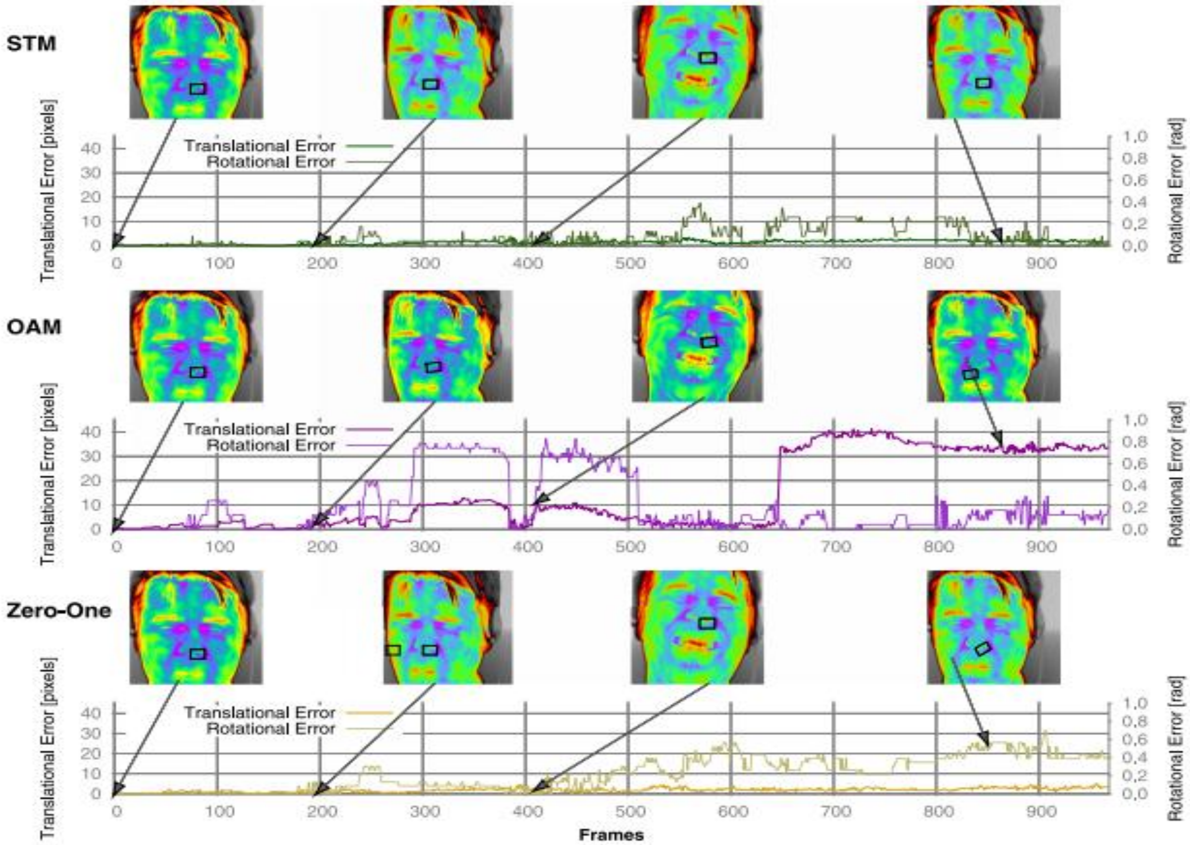

Fig. 5. Evolution of translational and rotational tracking errors for STM, OAM, and zero-one for subject D009IS. Thermal snapshots are indexed to representative points in the error signals to enable association with the nasal tracker's (black box) position in the actual runs.

B. Quantitative Results

To quantify tracker performance, we need to compare the tracker's ROI with the ground-truth ROI throughout the time line. In medical imaging, the ground-truth data are usually obtained by manually segmenting the ROI in each frame. With thousands of frames in the dataset, however, manual ground truthing was not practical. For this reason, we adopted a different strategy. We used each of the three trackers to generate tracking results. We examined the results and where each tracker appeared to have failed, we manually repositioned the tracker and reinitiated tracking from that point onward, to correct the error. At the end, we formed ground-truth trackers as the means of the individual corrected trackers. Tracking performance correlates to the Euclidean distance and angular difference between the ground-truth ROI and the ROI that each of the three competing methods produces. The smaller the Euclidean distance and angular difference are, the better. The 25 clips of the dataset when partitioned according to Scenario 1, Scenario 2, and Scenario 3 provide 8, 9, and 8 clips, respectively. Fig. 6(a) shows a graphical representation of the distribution of translational (Euclidean) errors per tracking method and operational scenario. As the plot indicates, the STM approach outperforms the other two template update strategies in all scenarios. Both the OAM and Zero-One methods appear to have particular difficulties with large combined positional and physiological changes (Scenario 3). Fig. 6(b) shows a graphical representation of the distribution of rotational errors per tracking method and operational scenario. STM still outperforms the other two methods across the spectrum, but its relative error magnitude increased with respect to the translational error. To statistically verify these indications, a series of hypotheses tests (at 0.05 level of significance) were performed to check how the means of the translational and rotational error distributions differ between methods for each of the scenarios in the database. Let us denote by μT S ,μT O , and μT Z the means

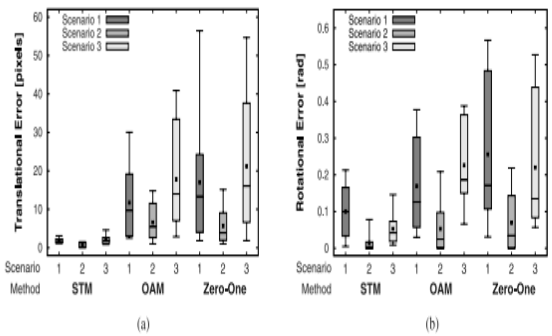

Fig. 6. (a) Boxplots of the translational (Euclidean) error distributions per tracking method and operational scenario for all subjects. (b) Boxplots of the rotational error distributions per tracking method and operational scenario for all subjects. (Scenario 1: large position and small physiology changes, Scenario 2: small position and large physiology changes, Scenario 3: large position and large physiology changes).

of the translational error distributions for the STM, OAM, and zero-one strategies, respectively. Accordingly, let us denote by μRS ,μRO , and μRZ the means of the rotational error distributions for the STM, OAM, and zero-one strategies, respectively. The only instances where the tests fail to reject the null hypotheses are for subjects L02, L06, and L08, where the three competing trackers appear to perform on par. In all other cases, mean performance differs significantly among trackers, with STM clearly outperforming the other two. Note that L02, L06, and L08 are all sleep study cases, where tracking is not as challenging as in stress study cases. Fig. 7 gives a graphical representation of the mean positional and rotational tracking errors per subject for the three competing tracking methods; it visually correlates with the outcome of the statistical tests.

A. Benefit of Temporal Smoothing

To specify the beneficial effect of temporal smoothing, a simulation was run where a thermal nasal region was translated only in the x-direction, while the y-direction and angle of rotation φ were kept constant. The region featured semiperiodic fluctuation in temperature akin to the effect of breathing. This region was tracked first with a particle filter tracker driven by the classical Matte formula with spatial smoothing only. Then, it was tracked with the same particle filter tracker but driven by STM, i.e., the modified Matte formula with both spatial and temporal smoothing. The trajectory results in Fig. 8 demonstrate the fault oscillation introduced in the y and rotational dimensions by the classical Matte method.

B. Sensitivity Analysis

In the Matte template strategy, the values of the nuisance parameters λ1 and λ2 (λ1 < λ2) determine the most stable and unstable pixels in the current frame, which are used as seeds [see (5)]. In this study, the following parameterization was used: λ1 = 5 and λ2 = 20. Note that the temperature values were normalized in the 0–255 range. An experiment was performed to test the sensitivity of the STM method with respect to the λ parameters.

TABLE I
MEAN TRANSLATIONAL ERROR FOR VARIOUS CHOICES OF (λ_1, λ_2)

(λ_1, λ_2)	$D011_{IS}$-nasal	$D225_{IS}$-periorbital	$D016_{IS}$-nasal
(2, 15)	1.614	2.815	1.162
(2, 20)	1.605	2.745	1.177
(2, 25)	1.605	2.718	1.171
(5, 15)	1.606	2.712	1.174
(5, 20)	1.515	2.947	1.088
(5, 25)	1.611	2.643	1.178
(8, 15)	1.613	3.245	1.144
(8, 20)	1.612	2.773	1.170
(8, 25)	1.599	2.707	1.163

Specifically, a representative clip from each of the three operational scenarios (Scenario 1: D011IS -nasal, Scenario 2: D225IS -periorbital, and Scenario 3: D016IS -nasal) was selected. The tracker of the STM template update method was run on these clips multiple times, varying at each iteration the values of λ1 and/or λ2 . Values for the tuple (λ1 , λ2) were drawn from the sets λ1 ∈ {2,5,8} and λ2 ∈ {15,20,25}, providing nine different pairs. Thus, nine runs for every selected subject were produced, each one providing an error distribution of the Euclidean distance and rotational difference between the tracked and ground-truth ROIs. Tables I and II give the mean translational and rotational errors for each of the three clips and every case of the (λ1 , λ2) parameter choices, respectively.

The mean distance value of the tracker's ROI from the groundtruth position appears to be rather stable for a wide range of parameter choices, indicating that the method is not very sensitive. The sensitivity increases (and the performance deteriorates) when the values for the λ parameters get close. In this case, the method identifies many of the ROI pixels as seeds and starts losing its probabilistic (smoothing) advantage. An example is the case of the pair (8,15) in Table I that features the highest mean errors in all selected subjects.

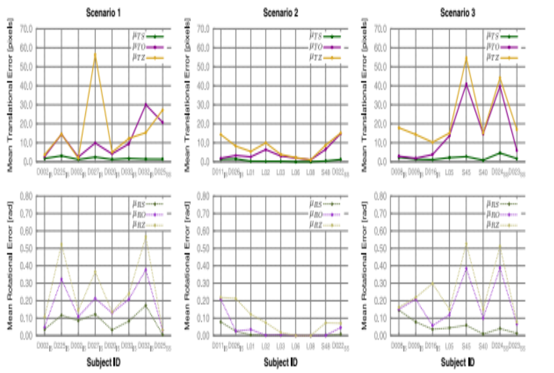

Fig. 7. Mean translational and rotational errors per subject for each competing tracking method grouped by operational scenario. The STM method consistently yields the smallest mean errors.

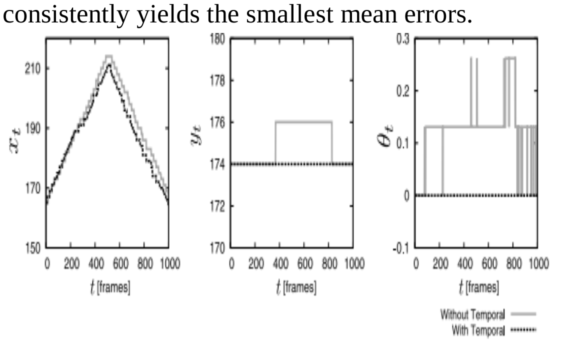

Fig.8.Comparative nasal ROI trajectories (decomposed into the x,y, and φ dimensions) in a controlled simulation experiment. Light solid trajectories were produced by a particle filter tracker operating on a legacy Matte with spatial smoothing only. Dark dotted trajectories were produced by a particle filter tracker operating on a Matte with both spatial and temporal smoothing.

TABLE II
MEAN ROTATIONAL ERROR FOR VARIOUS CHOICES OF (λ_1, λ_2)

(λ_1, λ_2)	$D011_{IS}$-nasal	$D225_{IS}$-periorbital	$D016_{IS}$-nasal
(2, 15)	0.0028	0.0812	0.0370
(2, 20)	0.0028	0.0737	0.0383
(2, 25)	0.0029	0.0780	0.0373
(5, 15)	0.0028	0.0703	0.0380
(5, 20)	0.0028	0.1087	0.0361
(5, 25)	0.0028	0.0722	0.0379
(8, 15)	0.0028	0.1084	0.0376
(8, 20)	0.0028	0.0720	0.0382
(8, 25)	0.0028	0.0722	0.0370

IV. CONCLUSION

This paper presents a new probabilistic template update method that when drives a particle filter tracker is capable of producing sophisticated tracking behavior in thermal facial imaging. Specifically, the method can cope with both large positional and physiological changes, something that other methods from the thermal or visual domain fail to do. The power of the method stems from the spatial and temporal smoothness components of the template that capture well natural thermo physiological characteristics. The new approach was tested on a dataset consisting of 25 thermal clips, thousands of frames each, featuring a variety of conditions that naturally occur in practice. The method promises improved performance in a number of biomedical applications, where unobtrusive physiological measurements on the face are preferred (e.g., sleep and stress studies). Note that in the case the subject's head exhibits frequent and significant out of plane rotation and movement, the resulting physiological signal will not be accurate as this is outside the tracker's operational scenario. Fortunately, for the targeted applications, this is rarely the case. This is obvious for sleep studies. It is also true for stress studies, where the subject stays put, maintaining (by design) directional attention to the stimulus or the interviewer.

REFERENCES

[1] S. Kong, J. Heo, B. Abidi, J. Paik, and M. Abidi, "Recent advances in visual and infrared face recognition—a review," Comput. Vis. Imag. Understand, vol. 97, pp. 103–135, 2005.

[2] C. Eveland, D. Socolinsky, and L. Wolff, "Tracking human faces in infrared video," Imag. Vis. Comput., vol. 21, pp. 579–590, Aug. 2003.

[3] N. Sun, M. Garbey, A. Merla, and I. Pavlidis, "Imaging the cardiovascular pulse," in Proc. IEEE Comput. Soc. Conf. Comput. Vis. Patt. Recognit, San Diego, CA, 2005, vol. 2, pp. 416–421.

[4] N. Sun, I. Pavlidis, M. Garbey, and J. Fei, "Harvesting the thermal cardiac pulse signal," in Medical Image Computing and Computer-Assisted Intervention, (Lecture Notes in Computer Science Series). vol. 4191, New York: Springer-Verlag, 2006, pp. 569–576.

[5] R. Murthy and I. Pavlidis, “Noncontact measurement of breathing function,” IEEE Eng. Med. Biol. Mag., vol. 25, no. 3, pp. 57–67, May/Jun. 2006.

[6] S. Chekmenev, A. Farag, and E. Essock, “Thermal imaging of the super-ficial temporal artery: An arterial pulse recovery model,” in Proc. IEEE Conf. Comput. Vis. Patt. Recognit., Minneapolis, MN, Jun. 17–22, 2007, pp.1–6.

[7] I. Pavlidis, J. Dowdall, N. Sun, C. Puri, J. Fei, and M. Garbey, “Interacting with human physiology,” Comput. Vis. Imag. Understand, vol. 108, no. 1– 2, pp. 150–170, Oct./Nov. 2007.

[8] J. Murthy, S. Faiz, J. Fei, I. Pavlidis, A. Abeulhagia, and R. Castriota, “Remote infrared imaging: A novel non-contact method to monitor airflow during polysomnography,” in Proc. Chest Meet. Abstr., Chicago, IL, Oct. 20–25, 2007, vol. 132, no. 4, p. 464.

[9] P. Tsiamyrtzis, J. Dowdall, D. Shastri, I. Pavlidis, M. Frank, and P. Ekman, “Imaging facial physiology for the detection of deceit,” Int. J. Comput. Vis., vol. 71, no. 2, pp. 197–214, Oct. 2006.

[10] B. Faser and J. Luettin, “Automatic facial expression analysis: A survey,” Pattern Recognit., vol. 36, no. 1, pp. 259–275, 2003.

[11] M. Bureau and F. Seri´es,` “Comparison of two in laboratory titration methods to determine effective pressure levels in patients with obstructive sleep apnea,” Thorax, vol. 55, pp. 741–745, 2000.

[12] L. Matthews and T. Ishikawa, “The template update problem,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 26, no. 6, pp. 810–815, Jun. 2004.

[13] A. Jepson, D. Fleet, and T. E.-Maraghi, “Robust online appearance models for visual tracking,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 25, no. 10, pp. 415–422, Oct. 2003.

[14] S. Zhou, R. Chellapa, and B. Moghaddam, “Visual tracking and recog-nition using appearance-adaptive models in particle filters,” IEEE Trans. Imag. Process., vol. 13, no. 11, pp. 1491–1506, Nov. 2004.

[15] J. Dowdall, I. Pavlidis, and P. Tsiamyrtzis, “Coalitional tracking,” Comput. Vis. Imag. Understand., vol. 106, no. 2–3, pp. 205–219, May 2007.

[16] J. Dowdall, “Tracking tissue in thermal infrared video,” Ph.D. dissertation, Dept. Comput. Sci., Univ. Houston, Houston, TX, 2006.

[17] A. Levin, D. Lischinski, and Y. Weiss, “A closed form solution to natural image matting,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 30, no. 2, pp.228–242, Feb. 2008.

[18] Y. Zhou, P. Tsiamyrtzis, and I. Pavlidis, “Tissue tracking in thermo-physiological imagery through spatio-temporal smoothing,” in Medical Image Computing and Computer-Assisted Intervention, (Lecture Notes in Computer Science Series). vol. 5762, New York: Springer-Verlag, 2009, pp.1092–1099.

[19] Y. Li, J. Sun, C. Tang, and H. Shum, “Lazy snapping,” ACM Spec. Interest Group Comput. Graph. Interactive Tech., vol. 23, no. 3, pp. 303–308, Aug. 2004.

[20] C. Rother, V. Kolmogorov, and A. Blake, “GrabCcut—interactive fore-ground extraction using iterated graph cuts,” ACM Spec. Interest Group Comput. Graph. Interactive Tech., vol. 23, no. 3, pp. 309–314, Aug. 2004.

[21] J. Shi and J. Malik, “Normalized cuts and image segmentation,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 22, no. 8, pp. 888–905, Aug. 2000.