

# Coalitional Tracking in Facial Infrared Imaging and Beyond

Jonathan Dowdall<sup>1</sup>

Ioannis T. Pavlidis<sup>1</sup>

Panagiotis Tsiamyrtzis<sup>2</sup>

<sup>1</sup>Department of Computer Science  
University of Houston  
Houston, TX  
{jbdowdal, ipavliidi}@central.uh.edu

<sup>2</sup>Department of Statistics  
Athens University of Economics  
Athens, Greece  
pt@aueb.gr

## Abstract

*We propose a novel facial tracking method that uses a distributed network of individual trackers whose interactions are modeled using coalitional game theory. Our tracking method maintains a high level of accuracy and can negotiate surface deformations and occlusions.*

*We tested the method on a substantial video set featuring non-trivial face motion from over 40 subjects in both the infrared and visual spectra. The coalitional tracker demonstrated fault tolerant behavior that exceeds by far the performance of single condensation trackers. Our method represents a shift from the typical tracking paradigms and may find broader application in demanding imaging problems across the electromagnetic spectrum.*

## 1 Introduction

The extraction of high-level information from video through the use of computer vision algorithms has become increasingly important over the past decade. A diverse array of applications use this technology including quality control in the manufacturing sector [1][2], surveillance in the security industry [3][4], biomedical measurements for healthcare [5][6][7], and behavioral analysis [8][9][10]. Of key importance to all these computer vision applications is the ability to detect and track objects in their respective input video streams.

The problem of tracking can be cast as guessing how things change over time. Specifically, tracking involves modeling how the parameters of the object modulate in successive input frames by using prior knowledge. Computer vision tracking has been dominated by sequential Monte Carlo methods [11] for the last several years. The most popular Monte Carlo tracking method is the CONDENSATION algorithm, which was introduced by Isard and Blake in 1998 [12][13][14]. Alternative tracking methodologies employ specific models of the target to provide better accuracy [15][16][17].

Unfortunately, this increased accuracy typically comes at the expense of speed and generality.

Work has also been done using multiple trackers that work together to follow *multiple objects* [18][19][20][21]. In contrast, we employ multiple trackers to track a *single object* rather than multiple objects.

Of particular interest is tracking in the infrared spectrum because recent research demonstrated that many vital signs including blood flow [5], pulse [6], and breathing function [7] can be measured in this modality, if there is a reliable tracking method to register the motion of facial tissue. Although, a large body of work has been devoted to facial tracking research in infrared [22][23][24], we found the existing methods insufficient to achieve the high degree of accuracy required in imaging measurements of facial tissue. This was our initial motivation in exploring a novel tracking paradigm.

Furthermore, developing a tracking methodology that is able to accurately track a wide array of targets across imaging modalities is a challenging task and remains elusive. Our effort is a step towards addressing this difficult but worthy cause.

In the remainder of the paper, we first discuss the details of our tracking approach in section 2. Then, we outline our experimental design in section 3, followed by the experimental results in section 4. Finally, we present our conclusions and plans for future work in sections 5 and 6 respectively.

## 2 Tracking Methodology

Our goal is to develop an automated tracking methodology that can *accurately* monitor the motion of the target's *surface* even in the presence of deformation or partial occlusion. Many existing tracking methods monitor *approximately* the target's *outline* (not surface). This is a different and far easier problem.

We arrived at a distributed, fault tolerant tracking method that works both on infrared and visual video. We use a network of trackers that influence each other (see Figure 1). Each individual tracker is unreliable at times, but the combination of many neighboring trackers produces robust performance. The inter-tracker influence is modeled as a coalitional game where each tracker is a



Figure 1: Overview of the tracking method. **Initialization:** consists of the following steps: (a) extraction of the user selected region of interest from the input video, (b) subdivision of the ROI into the tracking network, and (c) individual tracker template creation. **Tracking:** the individual trackers in the tracking network follow their targets. **Target State Estimation:** consists of the following steps: (d) the winning coalition is produced, (e) the deformation mesh is calculated from the winning coalition, and (f) the deformation mesh is used to calculate the target state. The method proceeds from initialization to tracking (arrow 1) to target state estimation (arrow 2) and back to tracking (arrow 3). In the latter transition, the winning coalition is passed back to the tracking stage to distribute the inter-tracker influence.

player, and the goal of the game is to propagate one's influence in subsequent frames of video. Within this framework, the winning coalition of trackers is used to calculate the state of the tracked object.

## 2.1 Tracking Network

We use a network of trackers to achieve fault tolerance. Tracking is maintained even if all but one of the trackers fail in the tracking network. The trackers are each assigned a different portion of the target's surface to track. By default the trackers are configured in a regular grid, although alternative configurations are possible through a feature selection mechanism. Inter-tracker communication allows trackers that are correctly tracking the target to "tip-off" other trackers that have become lost, as to the true location of the target's surface. This inter-tracker influence is realized within a statistical framework and is managed by the coalitional game model described in section 2.2.

The idea arose naturally in the effort to address the problem of facial tissue tracking in the infrared. As the subject's head moves (e.g., left and right) part of the facial surface is occluded at times. Trackers that correspond to the occluded part of the face are aided by trackers that correspond to the exposed part.

In our implementation, each tracker in the tracking network is an individual CONDENSATION based tracker. We denote the state of each individual tracker  $i$  at time  $t$  by  $\mathbf{x}_i^{(t)}$  and its associated image observation by  $\mathbf{z}_i^{(t)}$ . The

target tracker's prior,  $p(\mathbf{x}_i^{(t)} | \mathbf{z}_i^{(t-1)})$  will be formed using intra-samples  $\mathbf{s}_{(i,i)}^{(t)}$  from tracker  $i$  and inter-samples  $\mathbf{s}_{(i,j)}^{(t)}$  that correspond to the (inter-tracker) influence of tracker  $i$  from tracker  $j$ . The inter-samples  $\mathbf{s}_{(i,j)}^{(t)}$  are generated based on the initial relationship between the trackers involved in the exchange:

$$\mathbf{s}_{(i,j)}^{(t)} = \mathbf{T}_{\mathbf{x}_i^{(0)}}^{\mathbf{x}_j^{(0)}} \mathbf{x}_j^{(t)}, \quad (1)$$

where  $\mathbf{s}_{(i,j)}^{(t)}$  is the inter-sample generated by tracker  $j$  for tracker  $i$  and  $\mathbf{T}_{\mathbf{x}_i^{(0)}}^{\mathbf{x}_j^{(0)}}$  is the transformation that gives a sample for tracker  $i$  given a state for target  $j$  at time  $t$ . The transformation  $\mathbf{T}_{\mathbf{x}_i^{(0)}}^{\mathbf{x}_j^{(0)}}$  is computed during initialization for every possible tracker pair  $(\mathbf{x}_j^{(0)}, \mathbf{x}_i^{(0)})$ .

## 2.2 Coalitional Tracking

The tracking network is a versatile architecture for tracking object surfaces, but it does not have any intrinsic method to generate the final target state or to manage tracker interaction. The simplest solution would be to allow every tracker to influence all of the other trackers. Unfortunately, this is not an optimal solution, because trackers that have lost their target would be allowed to influence other trackers in the network that have not gone

awry. This also highlights the problem of determining which trackers in the network are correctly tracking their targets and which ones have strayed away. What is needed is an optimization mechanism that can determine the validity of each of the trackers, compute the target's state vector based on the valid trackers, and finally propagate the influence of the valid trackers to keep the network correctly tracking the target surface.

There are many optimization algorithms one can use to manage the network of trackers. We chose to optimize tracker interaction using a game theoretic solution for two main reasons: it naturally fits the problem space and it is relatively simple. Game theory [25][26][27][28] has been successfully used to analyze topics ranging from simple deterministic games, to complex economic models [29][30], and even international negotiations [31][32]. Our adaptation was to view the trackers as players in a cooperative game [33][34], where the objective was to increase their influence by forming coalitions with other trackers. The winning coalition would then be used to compute the state vector of the target and subsequently propagate its influence onto the entire tracking network.

Specifically, the members  $m_j^{(t)}$  of the winning coalition  $C^t$  influence every other tracker  $i$  in the tracking network by adding inter-samples  $\mathbf{s}_{(i,j)}^{(t)}$ . Trackers who are not members of the winning coalition cannot propagate any influence at all.

The intuitive affinity of the problem space to cooperative gaming is apparent in the example of facial tissue tracking. There, the winning coalition is composed mostly of trackers that correspond to the exposed part of the face. These are trackers that feature high quality information and give a 'helping hand' (influence) to the 'clueless' trackers that correspond to the occluded or deformed part of the face.

Assume that our network consists of  $N$  individual trackers. Based on the principles of game theory, the players (trackers) of the game are rational, and therefore will join the coalition that promises them the largest payoff. The size of the coalition will be denoted from now

on by  $k$ , where  $k \in \{1, 2, \dots, N\}$ , and there are  $\binom{N}{k}$

coalitions of size  $k$  giving a total of  $2^N - 1$  possible coalition combinations.

The payoff of each coalition is a function of four scores calculated from the trackers participating in the coalition under consideration. Specifically, these scores at time  $t$  are:

- template match  $\alpha^{(t)}$
- geometric alignment  $\beta^{(t)}$

- inter-frame projection agreement  $\gamma^{(t)}$
- inter-frame membership retention  $\delta^{(t)}$

The payoff scores support the fact that quality tracking is characterized by consistency in the content and geometric configuration of the individual trackers as well as by heeding to continuity constraints. Specifically, the template match score rewards trackers that maintain consistent imaging content. The geometric alignment score favors coalitions whose members have geometric alignment analogous to the original ( $t=0$ ) configuration. The inter-frame projection agreement score is a continuity constraint. It improves robustness by penalizing abrupt (and improbable) changes of the projected state of the target between successive frames. The inter-frame membership retention score is also a continuity constraint. It reflects the tendency of the winning coalition from the previous time step to retain its members.

The template match score  $\alpha_{C_k}^{(t)}$  for a coalition  $C_k$  of size  $k$  at time  $t$  is given by:

$$\alpha_{C_k}^{(t)} = \frac{1}{k} \sum_{i=1}^k \alpha_{m_i}^{(t)}, \quad (2)$$

where  $\alpha_{m_i}^{(t)}$  refers to the template match score (a number in  $[0,1]$ , see section 2.4) of member  $m_i$  in the coalition  $C_k$  at time  $t$ .

For the second and third scores we first need to define the function that measures the geometric alignment between two target projections (see Figure 2), as are computed from samples  $\mathbf{s}_i$  and  $\mathbf{s}_j$ :

$$G(\mathcal{F}(\mathbf{s}_i), \mathcal{F}(\mathbf{s}_j)) = G(\mathbf{S}_i, \mathbf{S}_j) = \omega \times \left[ 1 - \frac{\left( (S_{ix} - S_{jx})^2 + (S_{iy} - S_{jy})^2 \right)^{1/2}}{M_d} \right] + (1 - \omega) \times \left[ 1 - \frac{|S_{i\theta} - S_{j\theta}|}{M_\theta} \right], \quad (3)$$

where  $\mathcal{F}(\mathbf{s})$  is a function that transforms the tracker sample  $\mathbf{s}$  into its corresponding target projection  $\mathbf{S}$ ,  $(S_{ix}, S_{iy})$  are the  $(x, y)$  coordinates of the center of target projection  $\mathbf{S}_i$ , while  $S_{i\theta}$  is the angle of rotation about the center of target projection  $\mathbf{S}_i$ .  $M_d$  is set to the maximum movement allowed by the target in a single

frame while  $M_\theta$  is the (positive) maximum rotation allowed by the target in a single frame. The weight  $\omega$  penalizes appropriately the center and angle discrepancies. Ideally, the target projections in Figure 2(c) should have coincided (perfect alignment), so that the combined projection of the two tracker samples is reminiscent of the original target shape. Note that the upper bound for  $G(.,.)$  is 1 (when the two target projections are identical), but the lower bound is not necessarily 0. This would have been the case, if we chose  $M_d$  and  $M_\theta$  to be the maximum observed values at time  $t$ , but this would have slowed down the computation. Besides, we do not mind giving negative scores to some tracker pairs (i.e., penalizing as opposed to rewarding them), whose geometric alignment is very bad.

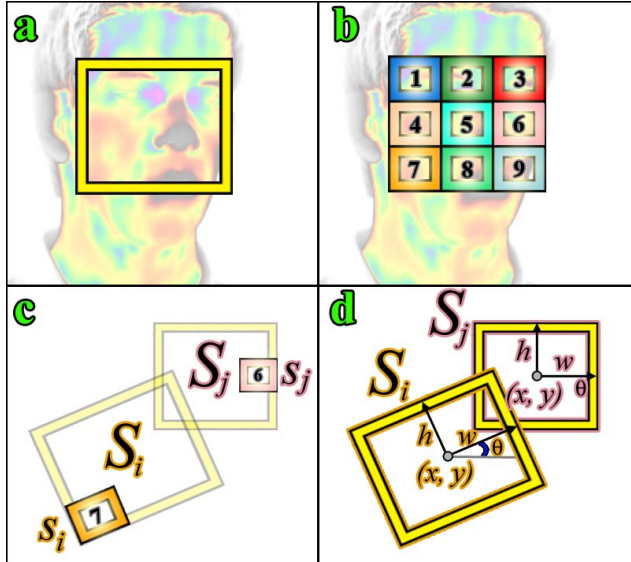


Figure 2: Geometric alignment of tracker's target projections. (a) Target projection at  $t = 0$ . (b) Tracker network overlaid on the initial target projection. (c) Trackers 6 and 7 at a subsequent time  $t$  along with their corresponding target projections. (d) Parameterization of target projections to facilitate measurement of geometric alignment.

Having defined the geometric alignment function for a pair of samples (see Equation(3)), we use it to compute the geometric alignment score  $\beta_{C_k}^{(t)}$  of a coalition of size  $k$ :

$$\beta_{C_k}^{(t)} = \frac{f(k)}{\binom{k}{2}} \sum_{i=1}^{k-1} \sum_{j=i+1}^k G(\mathbf{S}_i^{(t)}, \mathbf{S}_j^{(t)}), \quad (4)$$

where  $\mathbf{S}_i^{(t)}$  and  $\mathbf{S}_j^{(t)}$  are target projections corresponding to the samples with the highest template match scores for

coalition members  $m_i$ ,  $m_j$  respectively. Regarding the function  $f(k)$ , we have  $f(1) = 0$  and it is non-decreasing for  $k = 2, 3, \dots, N$ . The  $\beta_{C_k}^{(t)}$  is analogous to the average of the geometric alignment of all possible tracker pairs in the coalition. In general, as the size of the coalition  $k$  increases, the average of the geometric alignment function of the members of the coalition decreases. To compensate for that loss, we introduced the linear function  $f(k)$ , whose role is to reward higher order coalitions as opposed to lower order ones.

We also use the geometric alignment function for a pair of samples (see Equation(3)) to compute the inter-frame projection agreement  $\gamma_{C_k}^{(t)}$  score:

$$\gamma_{C_k}^{(t)} = \frac{1}{k} \sum_{i=1}^k G(\mathbf{S}_i^{(t)}, \mathbf{S}_i^{(t-1)}), \quad (5)$$

where  $\mathbf{S}_i^{(t)}$  is the target projection corresponding to the sample with the highest template match score for coalition member  $m_i^{(t)}$  at time  $t$ ;  $\mathbf{S}_i^{(t-1)}$  is the target projection corresponding to the target state at time  $t-1$  (previous frame).

The inter-frame membership retention score  $\delta_{C_k}^{(t)}$  for a coalition  $C_k$  of size  $k$  at time  $t$  is given by:

$$\delta_{C_k}^{(t)} = \frac{1}{k} \sum_{i=1}^k \Delta(m_i^{(t)}, C^{t-1}) \quad (6)$$

where  $m_i^{(t)}$  is the  $i$ -th member of coalition  $C_k$  at time  $t$ ,  $C^{t-1}$  is the winning coalition from the previous time step, and  $\Delta$  is defined as:

$$\Delta(m, C) = \begin{cases} -1 & \text{if } m \text{ is not a member of } C, \\ +1 & \text{if } m \text{ is a member of } C, \end{cases} \quad (7)$$

where  $m$  is a tracker and  $C$  is a coalition.

Having defined the four scores, we proceed with the definition of the payoff function  $\Pi_{C_k}^{(t)}$  of coalition  $C_k$ :

$$\Pi_{C_k}^{(t)} = \omega_\alpha \times \alpha_{C_k}^{(t)} + \omega_\beta \times \beta_{C_k}^{(t)} + \omega_\gamma \times \gamma_{C_k}^{(t)} + \omega_\delta \times \delta_{C_k}^{(t)}, \quad (8)$$

where  $\omega_\alpha$ ,  $\omega_\beta$ ,  $\omega_\gamma$ , and  $\omega_\delta$  are the weights (values range in  $[0,1]$  and sum to 1) assigned to the four scores. Note that because of the function  $f(k)$  in the geometric alignment score, the payoff score may exceed the value of 1. This may happen in cases where we have higher order coalitions and quite good geometric alignment.

For every size of coalition  $k \in \{1, 2, \dots, N\}$  we have  $\binom{N}{k}$  different coalitions of size  $k$ , out of which we select the one with the highest payoff. Thus, we decide for the winning coalition  $C^t$  at time  $t$  to be:

$$C^t = \arg \max_{C_k} \Pi_{C_k}^{(t)}. \quad (9)$$

### 2.3 Target State Estimation

We compute the final target state  $\mathbf{S}'$  from the winning coalition  $C^t$  in two steps. In the first step, we compute the deformation mesh  $M^t$  from the winning coalition. The deformation mesh  $M^t$  is composed of a set of points  $\mathbf{A} = (\mathbf{a}_1, \dots, \mathbf{a}_m)$ , which are distributed over the selected target region during the initialization step. Each point is linked to anywhere between 1 - 4 trackers depending on its spatial location; 1 on the corners, 2 on the borders, and 4 on the inside. The second step is to compute the final target state  $\mathbf{S}'$  from the deformation mesh  $M^t$  by using the border points  $\mathbf{B} = (\mathbf{b}_1, \dots, \mathbf{b}_4)$  (see Figure 3).

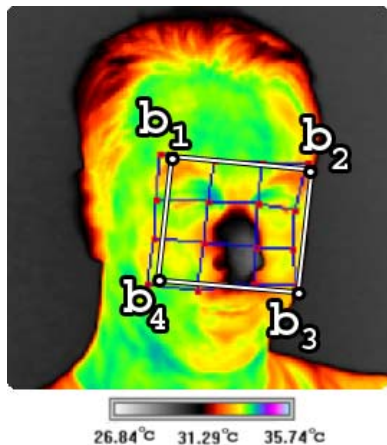


Figure 3: Border points of the target projection. The target projection is shown in white. The deformation mesh is shown in blue and the deformation mesh points are shown in red.

### 2.4 Template Matching

We opted to use a variant of the well-known technique of template matching [35], as the measurement vehicle for the CONDENSATION trackers. Template matching is applicable to both visual and thermal band imagery [23]. The fundamental idea behind template matching is to create the template and then to find the region in subsequent frames of video that most closely resembles the template. In our case, the template is a sub-sampling of

pixels from the target object. The underlying assumption when using templates to represent the object of interest is that the appearance of the object will remain relatively constant throughout the course of the video. This limits the effectiveness of template matching in visual imaging under dynamic lighting conditions, unless more complex methods are employed to compensate, such as updating the template, or using model based light mapping. Because our main focus for this paper was tracking in thermal infrared video, we opted not to implement one of these methods, and to simply constrain our visual band input to static lighting conditions.

## 3 Experimental Design

For the purpose of testing the tracking algorithm on thermal infrared video we selected a dataset that was used in previous publications [23]. It consists of 39 video clips, each containing a human subject undergoing an interview. We chose to track 1000 frames of video from each of the subject clips, for a total of 39000 frames of video. The chosen video segments featured a temporary occlusion of the main subject by another subject who was passing through the field of view. We chose a single CONDENSATION tracker [23] to compare against the coalitional tracker. Both the single CONDENSATION and coalitional trackers were used to track the face of each subject.

To demonstrate that the tracking methodology can also be applied to visual band video, we performed experiments on a series of visual videos, each containing a different subject who was freely moving his head.

## 4 Experimental Results

The results from the thermal infrared experiment (see Table 1, Figure 4, and Figure 5) clearly show that the coalitional tracker provides superior tracking over the normal CONDENSATION tracker.

Reason for Failure	Coalitional Tracker Failures	Single Tracker Failures
Target Rotation	1	18
Partial Occlusion	2	9
No Recovery	1	2
Total	4	29

Table 1: Causation of tracking failures in the thermal data set

The proposed method proved robust in typical (see Figure 6) and difficult (see Figure 7) operational scenarios. The few failures of the coalitional tracker were mainly caused by significant out of plane rotation or substantial occlusion of the target (see Figure 8). The coalitional tracker also performed robustly in several visual band experiments (see Figure 9).

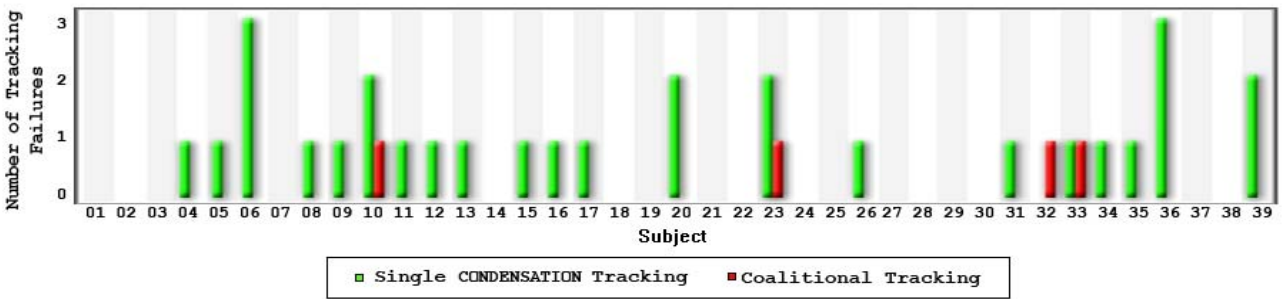


Figure 4: Tracking failure graph for the 39 video clips in the thermal data set. For each clip the number of single CONDENSATION and coalitional tracking failures is shown in green and red respectively. The absence of red bars in some video clip entries indicates perfect performance of the coalitional tracker. All 39 tracking videos can be viewed at [http://www.geocities.com/anonymous\\_author2005/OTCBVS06/results.htm](http://www.geocities.com/anonymous_author2005/OTCBVS06/results.htm)

The algorithm operated at ~15 frames per second in a C# implementation on a dual Xeon processor PC.

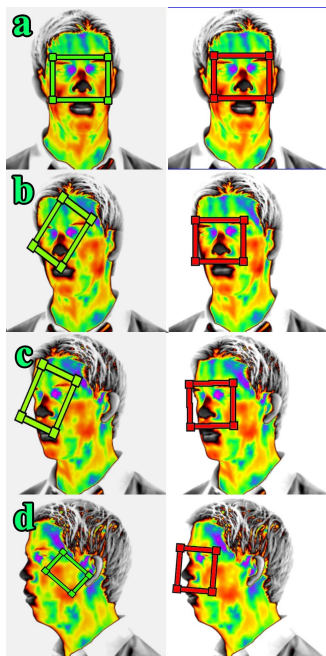


Figure 5: Out-of-plane rotation comparison. The left column shows the single-CONDENSATION tracker (green) and the right column shows the coalitional tracker (red). (a) Initial frame. (b) and (c) Intermediate frames. (d) Final frame in a 1 min thermal clip. The poor performance of the single CONDENSATION tracker is evident.

## 5 Conclusion

We have proposed a novel tracking method. Our method uses a distributed network of trackers whose interactions are modeled using coalitional game theory. The output of the method provides high accuracy, even in the presence of multi-dimensional target transformation.

We tested our method in thermal and visual video sets featuring faces and objects. We compared the performance of the proposed coalitional tracker with that of a single CONDENSATION tracker. The coalitional tracker

exhibited superior performance in both regular and challenging tasks. The strength of the method comes from the redundancy that is elegantly encoded in its game theoretic structure. The method appears to be general and flexible enough to be of use in imaging applications across the electromagnetic spectrum.

## 6 Future Work

The particular adaptation of Game Theory to tracking presented in this paper is but one of many possible approaches that might be adopted. For example, the problem of tracking could be alternatively viewed as a non-cooperative game where the trackers compete with each other, and the final solution could then be modeled as a Nash (strategic) equilibrium [26]. Additionally, active research areas in Game Theory, such as Stochastic and Differential Games [36], could potentially be adapted for use in tracking.

In the near future we are planning to update our target parameterization to allow it to model 3D transformations, which will aid in tracking more effectively out of plane rotation.

## 7 Acknowledgments

We would like to thank the National Science Foundation (grant # IIS-0414754) and Dr. Ephraim Glinert, for their support and encouragement during the lifetime of this burgeoning research project. The views expressed by the authors in this paper do not necessarily reflect the views of the funding agency.

## 8 References

- [1] D.A. Gonzalez, F.J. Madruga, M.A. Quintela, and J.M. Lopez-Higuera, *Defect assessment on radial heaters using infrared thermography*, NDT & E International, 38(6):428-432, September 2005.
- [2] M. Burrell, *Computer vision for high-speed, high-volume manufacturing*, in Proceedings of the 1993 International Conference on Systems, Man, and Cybernetics, 3: 349-354, October 17-20, 1993.

- [3] I. Pavlidis, V. Morellas, P. Tsiamyrtzis, and S. Harp, *Urban surveillance systems: from the laboratory to the commercial world*, Proceedings of the IEEE, 89(10):1478-1497, October 2001.
- [4] R.T. Collins, A.J. Lipton, H. Fujiyoshi, and T. Kanade, *Algorithms for cooperative multi-sensor surveillance*, Proceedings of the IEEE, 89(10):1456-1477, October 2001.
- [5] M. Garbey, A. Merla, and I. Pavlidis, *Estimation of blood flow speed and vessel location from thermal video*, in Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1:356-63, June 27 – July 2, 2004.
- [6] N. Sun, M. Garbey, A. Merla, and I. Pavlidis, *Imaging the cardiovascular pulse*, in Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2: 416-21, June 20-25, 2005.
- [7] J. Fei, Z. Zhu, and I. Pavlidis, *Imaging breathing rate in the CO<sub>2</sub> absorption band*, in Proceedings of the 27th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, September 1-4, 2005.
- [8] J. Levine, I. Pavlidis, and M. Cooper, *The face of fear*, *The Lancet*, 357(9270), June 2, 2001.
- [9] I. Pavlidis, N.L. Eberhardt, and J. Levine, *Human behavior: seeing through the face of deception*, *Nature*, 415(6867):35, January 3, 2002.
- [10] I. Pavlidis and J. Levine, *Thermal image analysis for polygraph testing*, *IEEE Engineering in Medicine and Biology Magazine*, 21(6):56-64, November-December 2002.
- [11] A. Doucet, N. DeFreitas, and N. Gordon, editors, *Sequential Monte Carlo Methods in Practice*. Springer-Verlag, 2001.
- [12] M. Isard and A. Blake, *Condensation - conditional density propagation for visual tracking*, *International Journal of Computer Vision*, 19(1):5-28, 1998.
- [13] M. Isard and A. Blake, *ICONDENSATION: Unifying low-level and high-level tracking in a stochastic framework*, in Proceedings of the 5<sup>th</sup> European Conference on Computer Vision, 1:893-908, June 2-6, 1998.
- [14] J. MacCormick and M. Isard, *Partitioned sampling, articulated objects, and interface-quality hand tracking*, in Proceedings of the 7<sup>th</sup> European Conference on Computer Vision, 1843: 3-19, 2000.
- [15] Y. Shi and W. Karl, *Real-time tracking using level sets*, in Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2:34 – 41, June 20-25, 2005.
- [16] C. Zimmer and J. C. Olivo-Marin, *Analyzing and capturing articulated hand motion in image sequences*, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(11):1838 – 1842, November 2005.
- [17] S. Goldenstein, C. Vogler, J. Stolfi, V. Pavlovic, D. Metaxas, *Outlier rejection in deformable model tracking*, in Proceedings of the 2004 Conference on Computer Vision and Pattern Recognition, June 19 – 26, 2004.
- [18] C. Cheng, R. Ansari, and A. Khokhar, *Multiple object tracking with kernel particle filter*, in Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1:566 – 573, June 20-25, 2005.
- [19] Y. Ting and W. Ying, *Decentralized multiple target tracking using netted collaborative autonomous trackers*, in Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1:939-946, June 20-25, 2005.
- [20] M. Isard and J. MacCormick, *BraMBLe: A Bayesian multiple-blob tracker*, in Proceedings of the 8<sup>th</sup> IEEE International Conference on Computer Vision, 2:34-41, July 7-14, 2001.
- [21] J. MacCormick and A. Blake. *A probabilistic exclusion principle for tracking multiple objects*, *International Journal of Computer Vision*, 39(1):57-71, 2000.
- [22] C. Eveland, D. Socolinsky, and L. Wolff. *Tracking human faces in infrared video*, *Image and Vision Computing*, 21:578-590, July 2003.
- [23] P. Tsiamyrtzis, J. Dowdall, D. Shastri, I. Pavlidis, M.G. Frank, and P. Ekman, *Lie detection – recovery of the periorbital signal through tandem tracking and noise suppression in thermal facial video*, in Proceedings of SPIE Sensors, and Command, Control, Communications, and Intelligence (C3I) Technologies for Homeland Security and Homeland Defense IV, E. M. Carapezza, editors, 5778, March 29-31, 2005.
- [24] S. Krotosky, S. Cheng, and M. Trivedi, *Face detection and head tracking using stereo and thermal infrared cameras for "smart" airbags: a comparative analysis*, in Proceedings of the 7th International IEEE Conference on Intelligent Transportation Systems, 1:17-22, 2004.
- [25] T.Ferguson, Game Theory. [http://www.math.ucla.edu/~tom/Game\\_Theory/Contents.html](http://www.math.ucla.edu/~tom/Game_Theory/Contents.html)
- [26] K. Ritzberger, *Foundations of Non-Cooperative Game Theory*. Oxford University Press, 2002.
- [27] A. Rapoport, *N-Person Game Theory: Concepts and Applications*. University of Michigan, 1978.
- [28] E. Rasmusen, *Games & Information: An Introduction to Game Theory*. Blackwell Publishers, 1989.
- [29] T. G. Fisher, et al, *Managerial Economics: A Game Theoretic Approach*. Routledge, 2002.
- [30] C. Schmidt, Editor, *Game Theory and Economic Analysis: A Quiet Revolution in Economics*. Routledge, 2002.
- [31] P. Ordeshook, *Game Theory and Political Theory: An Introduction*. Cambridge University Press, 1986.
- [32] S. Brams, *Game Theory and Politics*. Free Press, 1975.
- [33] S. Hart, editor, *Cooperation: Game-Theoretic Approaches*, Springer-Verlag, 1997.
- [34] M. Mareš, *Fuzzy Cooperative Games*. Physica-Verlag, 2001.
- [35] S. Baker and I. Matthews. *Equivalence and Efficiency of Image Alignment Algorithms*, in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 1:1090-1097, 2001.
- [36] M.Bardi, T. Raghavan, T. Parthasarathy, Editors, *Stochastic and Differential Games: Theory and Numerical Methods*. Annals of the International Society of Dynamic Games, Birkhäuser, 1998.

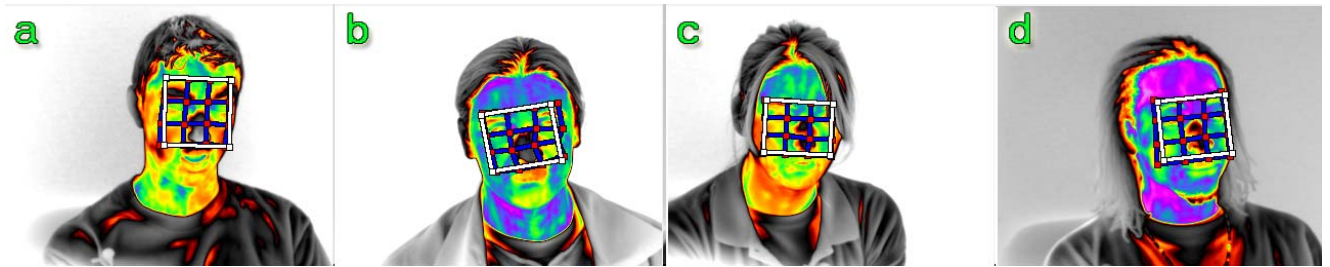


Figure 6: Typical facial tracking examples from the thermal data set. The selected subjects represent different ethnicities and both genders.

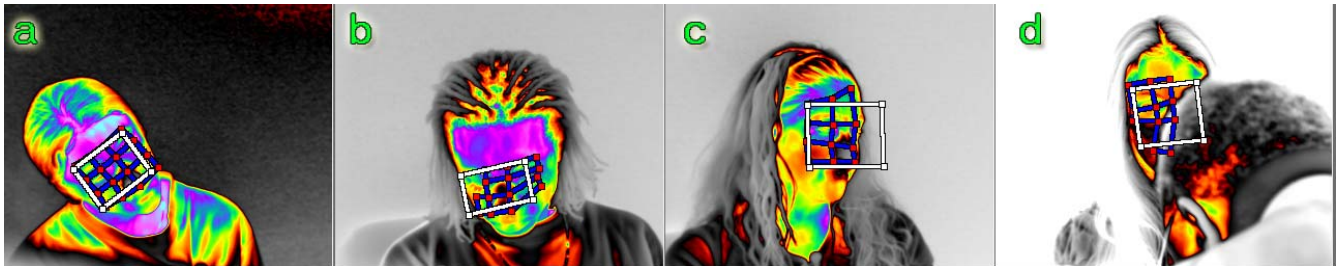


Figure 7: Successful coalitional tracking in the presence of difficult circumstances in the thermal spectrum. (a) The target is rotating in-plane. (b) The target is rotating out-of-plane. (c) The target is rotating out-of-plane. (d) The target is partially occluded.

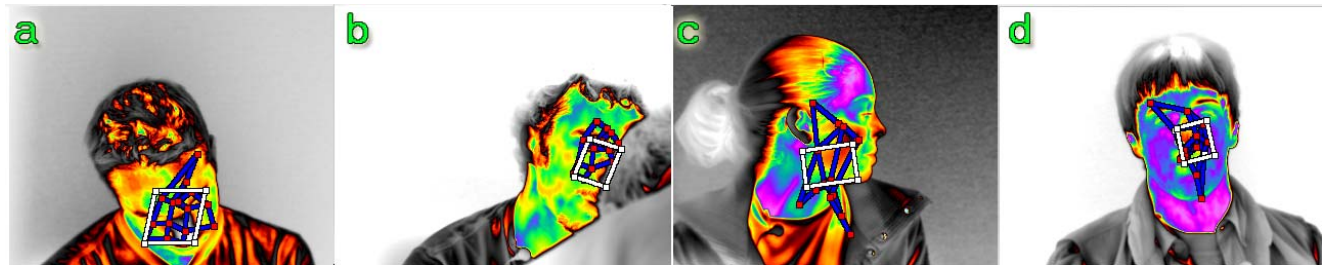


Figure 8: Tracking failures in the thermal spectrum. (a) and (c) The target has rotated out of plane beyond the tracker's ability to compensate. (b) The original target (periorbital area) is largely occluded. (d) The target has undergone extreme physiological changes relative to the initial tracking frame.

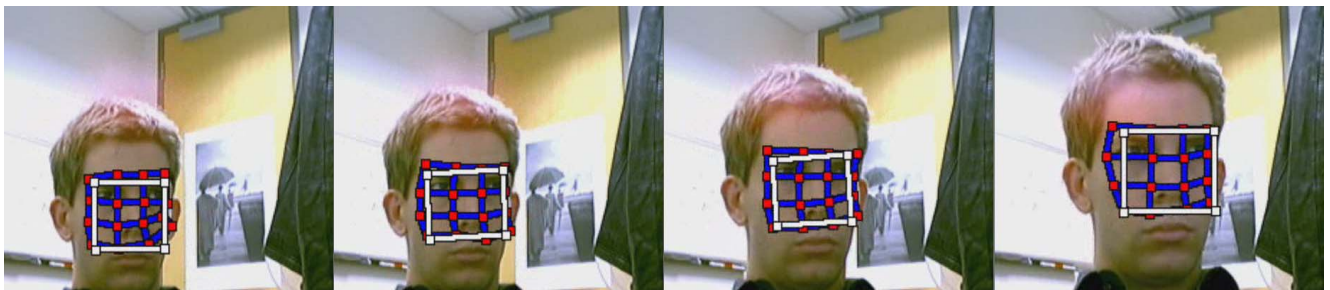


Figure 9: Example of tracking a face experiencing scaling, and translation in the visual spectrum. The frames are shown chronologically from left to right. The deformation mesh is shown in blue and the white rectangle represents the projected target state.