

Recognition of Human Activity Characteristics based on State Transitions Modeling Technique

Vinayak Elangovan, Amir Shirkhodaie

Center of Excellence for Battlefield Sensor Fusion
Dept. of Mech. & Mfg Engr., Tennessee State University
3500 John Merritt Blvd., Nashville, TN 37209

ABSTRACT

Human Activity Discovery & Recognition (HADR) is a complex, diverse and challenging task but yet an active area of ongoing research in the Department of Defense. By detecting, tracking, and characterizing cohesive Human interactional activity patterns, potential threats can be identified which can significantly improve situation awareness, particularly, in Persistent Surveillance Systems (PSS). Understanding the nature of such dynamic activities, inevitably involves interpretation of a collection of spatiotemporally correlated activities with respect to a known context. In this paper, we present a State Transition model for recognizing the characteristics of human activities with a link to a prior context-based ontology. Modeling the state transitions between successive evidential events determines the activities' temperament. The proposed state transition model poses six categories of state transitions including: Human state transitions of Object handling, Visibility, Entity-entity relation, Human Postures, Human Kinematics and Distance to Target. The proposed state transition model generates semantic annotations describing the human interactional activities via a technique called Casual Event State Inference (CESI). The proposed approach uses a low cost kinect depth camera for indoor and normal optical camera for outdoor monitoring activities. Experimental results are presented here to demonstrate the effectiveness and efficiency of the proposed technique.

Keywords - Human Activity Discovery & Recognition (HADR), State Transitions, Semantic Labeling, Entity-entity Relation, Casual-Events State Inference (CESI)

1. INTRODUCTION

Automatic recognition of human behaviors is an ambitious yet a challenging process in the achievement of a Persistent Surveillance System (PSS). The objective of PSS is to achieve an ability to detect, track, locate, and characterize a potential target, and provide an early warning of potential threats while improving situational assessment in real-time. Significant efforts have been devoted by previous researchers to the understanding, comprehension, and reasoning of the information provided by soft and hard sensors in PSS to identify human interactions with vehicles or objects or with other humans [1-7]. Hard Data is the information computed from hard sensors and usually it provides straight forward solution to analyze and their stability and characteristics are predictable [1]. HD requires low or moderate computing time to analyze the situation. Whereas Soft Data is the information computed from soft sensors i.e. human resources, is strongly based on human intuition and the subjectivity. It represents the ambiguity in human thinking with real life uncertainty. Messages from human's i.e. soft data sometimes provide valuable information or observations which may not be available from hard data like in judging relationships between two individuals, their facial expressions etc. One approach of fusion of hard and soft data is to convert all hard sensory data to linguistic expressions and process them in conjunction with soft sensory information via a linguistic message processing and reasoning engine [4]. By converting hard data to soft messages, the gap between soft and hard data fusion is abridged. Another approach is to correlate soft information with hard data so that hard and soft data integration can be accomplished consistently [12].

Behavioral activity modeling of humans taking place in complex environments has been also investigated by number of researchers [4-12]. However, their proposed approaches are not based on the state transition modeling technique. The task of human activity recognition system is a NP hard decision problem. In literature, human activities are classically considered as a bounded taxonomy task. Supervised training approaches are by and large considered for training of an intelligent system to learn behaviors and patterns of human activities. Popularly, supervised learning methods like Neural Networks, Support Vector Machines, Naïve Bayes decision tree, etc are used effectively for detecting a number of visual concepts [8]. Recognition of human activities can be performed in a manner of different ways. In [5], we proposed a context-based approach for semantic labeling of human activities involved with vehicles. The proposed technique is based on ontological approach that tracks atomic events to construe observations that can be used to generate semantic

annotation of imagery video stream as dynamic human-vehicle interactions take place [5]. Other statistical approaches are based on template-based and state-space methods [8]. In [9-11] human activities are modeled as a Markov process with unobserved (hidden) states. A Hidden Markov Model is a statistical approach based on dynamic Bayesian networks. A rule-based system and state machine framework is proposed in [9] for analysis of videos in a three-level of context hierarchy (i.e. attributes, events, and behaviors) that are used to identify activities and classify human-human interaction as one-to-one, one-to-many, many-to-one and many-to-many using an meeting ontology.

Real-time processing of video streams is computationally expensive and continuity of video transmission cannot be guaranteed. When imagery data arrive discretely, this requires a careful examination of events and cross-correlation of pertinent events observable from one frame to another. Certain pertinent events may have a short time response. In order to maintain an incessant perception link from one frame to another, there is a need for a state-transition model that maps discretely generated information per image frame to a time-continuum perception space of events without loss of generality. In this paper, we present an approach for modeling state-transition information. In particular, we describe a robust method for generating ontology-based hypotheses based on observation of spatiotemporal events as they occur in real-time. This paper is organized as follows: Atomic Events Modeling, State Transitions of Human Activities, Discussion & Experimental Results, Conclusions followed by Acknowledgements and References.

2. THEORETICAL IMPLICATION OF STATE-TRANSITION

In control theory, the *state-transition matrix* is considered as a matrix whose product with the state vector x at an initial time t_0 gives x at a later time t . The state-transition matrix can be used to obtain the general solution of linear dynamical systems. It is also known as the matrix exponential. Consider the general linear state space model:

$$\begin{aligned}\dot{\mathbf{x}}(t) &= \mathbf{A}(t)\mathbf{x}(t) + \mathbf{B}(t)\mathbf{u}(t) \\ \mathbf{y}(t) &= \mathbf{C}(t)\mathbf{x}(t) + \mathbf{D}(t)\mathbf{u}(t)\end{aligned}$$

The general solution is given by:

$$\mathbf{x}(t) = \Phi(t, t_0)\mathbf{x}(t_0) + \int_{t_0}^t \Phi(t, \tau)\mathbf{B}(\tau)\mathbf{u}(\tau)d\tau$$

The state-transition matrix $\Phi(t, \tau)$, given by:

$$\Phi(t, \tau) \equiv \mathbf{U}(t)\mathbf{U}^{-1}(\tau)$$

where $\mathbf{U}(t)$ is the fundamental solution matrix that satisfies [16]:

$$\dot{\mathbf{U}}(t) = \mathbf{A}(t)\mathbf{U}(t)$$

is a $n \times n$ matrix that is a linear mapping onto itself, i.e., with $\mathbf{u}(t) = 0$, given the state $\mathbf{x}(\tau)$ at any time τ , the state at any other time t is given by the mapping:

$$\mathbf{x}(t) = \Phi(t, \tau)\mathbf{x}(\tau)$$

If the system is time-invariant, we can define ϕ as:

$$\phi(t, t_0) = e^{\mathbf{A}(t-t_0)}$$

In the time-variant case, there are many different functions that may satisfy these requirements, and the solution is dependent on the structure of the system. Following a similar approach, we show that the state-transition matrices associated with surveillance system can be independently formulated based on perceptual assessment of expert analysts and then, used for modulation of observations to a quantitative measure that represents a level of situational awareness associated with the evolving situation.

In this paper, we present a district form of state transitions matrices representing Casual Events State Inference (CESI). CESI presents a control matrix encapsulating perception of severity associated with state transition of two observable events. Philosophically speaking, such perceptual representation generalizes proposition of implication of two correlated events. As shown later, CESI can be used to infer individual activities. It can also be used as a means for generating effective semantic annotations of events. There are many state transition categories. Each category represents a specific context being considered. Here, we present six of these pertinent state transition categories including: Human state

transitions of Object handling, Visibility, Entity-entity relation, Human Postures, Human Kinematics and Distance to Targets are considered in this paper. A hidden markov model had been developed with help of predefined ontologies for modeling the activities incorporating the severity associated between sequential events through state transition modeling. Developed HHM is not addressed in this paper as a limit to scope of this paper. Behavior models are constructed based on the contextual information's and spatiotemporal relations. A typical human activity scenario generates discrete sequence of intermittent events. For example, consider two sample ontologies listed for Human Vehicle Interaction (HVI) activities below:

Ontology #1: [SAV → OT → POT → CTD → WDS → ODD → GIDS → IE → WFS → TOS] → *A Normal HVI*

Ontology #2: [SAV → OT → POT → WDS → ODD → GIDS → IE → TOF] → *Not A Normal HVI*

The above ontologies depict trailing events of two observable situations. Results from fusion of hard sensors form the atomic components of the ontologies. These observable atomic events, once correlated to each other, can explain each situation sufficiently. For instance, in the first Ontology states: a subject Slowly Approached toward a Vehicle (SAV), Opened the Trunk (OT), Placed an Object inside the Trunk (POT), Closed the Trunk Door (CTD), Walked toward the Driver Side (WDS), Opened the Driver Door (ODD), Got Inside Driver Seat (GIDS), Ignited the Engine (IE), Waited a Few Seconds (WFS) and Took Off Slowly (TOS) a few moment later.

Following a similar interpretation, you will notice in the second Ontology, the driver forgets to close the trunk and takes off the vehicle rapidly, which reveals urgency in his actions. By associating severity in the state transitions of consequent events, suspiciousness involved in HVI can be efficiently identified. In our previous work, we presented a technique for semantic annotation of context based human-vehicle interactions using Zoning of Vehicle (ZoV) for interaction detection and tracking of human(s) around the vehicle and an approach for semantic labeling of acoustic signatures describing human-vehicle interactions [5]. In this paper, through Casual Event State Inference technique, we further elaborate on fusion of soft and hard data in order to raise the degree of confidence and in describing the risk of certainty involved in suspicious human interaction activities. We also note that the developed CESI system can accommodate both raw images and processed data in form of semantic messages which maps relevant ontologies for prediction of suspicious activities. Various experiments have been carried out in determining human activities in normal and dark environments using normal camera and Kinect depth camera. Identification of state transitions between two atomic events in human activity helps in recognizing the severity associated in such actions. State transitions for object handling, human / target visibility, entity-entity relation, human postures, human kinematics and distance to target have been analyzed in this section as shown in Figure 1.

Humans are intellectual and there is a difference in each person's perception towards acquiring decisions for situational awareness. In this work, we employed human perceptions towards the severity associated between two events, given two image frames for decisive action. A survey had been taken from a group people. Two images were given; each individual assessment was taken into consideration and an average assessment has been tabulated. Except for human kinematics transitions, two sequential semantic messages were given to the participants for decisive making since kinematics analysis requires greater extent of image frames. Human activities in general can be an interaction with other human / humans or interaction with objects. In our work, we categorize human interactions into Human-Vehicle Interactions (HVI), Human-Human Interactions (HHI) and Human-Object Interactions (HOI). Observable atomic events from these interactions are identified using Casual-Events State Inference (CESI) engine which considers the spatial-events relationship, for example: Zoning of Vehicle (ZoV) technique as discussed in our previous paper [4].

3. STATE TRANSITIONS OF HUMAN ACTIVITIES

Detection and tracking of human interactions with humans and objects had been an important issue for Homeland Security in identifying suspicious or normal behavior. In our work, pipeline of detected atomic events from each image are matched to an observation through a Hypothesis tracking. State transition modeling is done for each sequential observation pair in the pipeline and an average of transition rate is computed to indicate the level of suspiciousness involved in the activity.

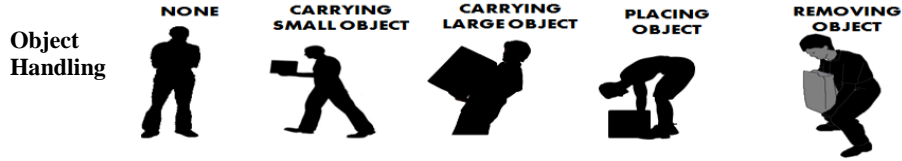


Figure 1a: Human Object Handling State Transition Observation



Figure 1b: Observations in Entity-Entity Relationship



Figure 1c: Observations in Visibility State Transitions



Figure 1d: Human Posture State Transitions Observations



Figure 1e: Human Kinematics State Transitions Observations



Figure 1f: Distance to Target State Transitions Observations

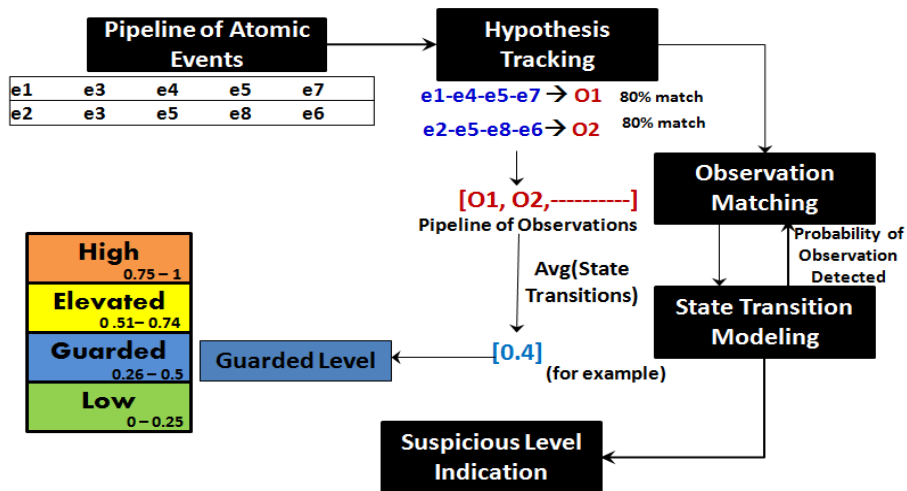


Figure 2: Process of State Transition Modeling

Transitions of states (ST) from six different categories are considered. For example in Human Object Handling state transition: two sequential frames were to be analyzed, observation (O) at frame f(i) evidences subject-1 with no object and observation at frame f(i+1) evidences subject-1 placing an object. A survey from a random group of 25 people had been carried out in depicting the human perceptions in distinguishing normal and suspicious behavior, given two frames. People who involved in the survey were chosen from different engineering streams [2 Professor, 1 post doctor, 3 Ph.D. students, 6 Master’s students, 13 Under Graduate students]. Severity associated between two frames can be computed using

$$.Severity(f(i) \rightarrow f(i + 1)) = ST(O(f(i)) \rightarrow O(f(i + 1)))$$

These sequential observations may involve a suspicious motive from human perspective. Analyzing the following state transitions helps in categorizing the activity into Low, Guarded, Elevated or High based on the probability of risk involved in a suspicious behavior as shown in Figure-2.

The probability of an observation possibly occurring is computed by:

$$P(O(j)_{f(i+1)})=P(O(j)_{f(i)})+(1-P(O(j)_{f(i)}))/2$$

where O(j) is the observation of event j and P(O(j)_{f(i)}) is the current probability of observation j occurring in frame f(i).

In a smooth state transition;

$$P(A) = P(A) + \frac{1-P(B)}{2}$$

where A is the previous sequential observation and B is the current evidential observation. For example in case of an Entity-Entity Relationship, we have

Non Associative (NA) @ frame i → Approaching @ frame i+1 → United @ frame i+2

P(NA) @ f(i) = 0.5 (initial probability); P(Approach) @ f(i+1) = 0.5 and P(NA) @ f(i) = 0.5 + (1-0.5/2) = 0.75 since P(Approach) being an evidential probability for P(NA).

A. State Transition of Object Handling:

Five different possible observations made in human handling objects (i.e., none, carrying small object, carrying large object, placing object, and removing object), was considered as shown in Figure 1. The size of object being used is subjective to the user. Kinect images were used in effective detection of objects being removed by performing image processing and color variation – depth mapping. Table 1 presents a sample of state transition matrix associated with object handling. These results are obtained based on statistical perceptual assessment reported by our test group during a simulated surveillance system experiment.

Table 1: Object Handling State Transitions

OBJECT HANDLING STATE TRANSITIONS					
	None	Carry Small Object	Carry Large Object	Placing Object	Removing Object
None	0.02	0.27	0.52	0.73	0.56
Carry small object	0.18	0.25	0.62	0.48	0.61
Carry large object	0.42	0.44	0.35	0.58	0.65
Placing object	0.40	0.60	0.62	0.48	0.58
Removing object	0.63	0.43	0.58	0.49	0.45

B. State Transition of Entity-Entity Relationship:

We consider Entity-Entity relationship between two entities based on five cases, namely, non-associative, approaching, united, cooperating and departing as depicted in Table 2. These observations enable to find the nature of interactions taking place in a group or between two entities. Table 2 shows the state transitions of Entity-entity relationship from each successive observations and the severity associated with these observations.

Table 2: Entity-Entity Relationship State Transitions

ENTITY-ENTITY RELATIONSHIP					
	Non Associative	Approaching	United	Cooperating	Departing
Non-Associative	0.08	0.50	0.62	0.75	0.40
Approaching	0.35	0.28	0.55	0.53	0.65
United	0.57	0.25	0.38	0.52	0.55
Cooperating	0.47	0.35	0.37	0.35	0.59
Departing	0.27	0.53	0.49	0.45	0.47

For a smooth Entity-Entity Relationship state transition, following transitions needed to be followed:

Non Associative → Approaching → United → Cooperating → Departing

C. Visibility State Transitions

Target visibility state modulation is done during target tracking and target behavioral pattern recognition. Suspiciousness here increases during each transition in successive observations. Three observations are deliberated including: normal, obscured and hidden. For example: given two sequential frames, subject-1 is in a visible state at frame $t(x)$ and is hidden at frame $t(x+1)$ when an activity is recorded, then perception of suspiciousness is believed to be 0.83. Weighed observations in visibility state transitions are shown in Table 3. Targets here are assumed to be human or vehicle or a large object.

Table 3: Target Visibility State Transitions

VISIBILITY TRANSITION STATE			
	Normal	Obscured	Hidden
Normal	0.08	0.48	0.83
Obscured	0.37	0.35	0.79
Hidden	0.58	0.63	0.63

D. State Transition of Human Postures:

State transitions of human postures are studied in tracking of target's behavior. Human postures like standing, sitting, bending and laying down are considered from static image observations. Severities associated with sequential observations during the state transitions of human postural behavioral patterns are tabulated in Table 4.

Table 4: Human Posture State Transitions

POSTURE STATE TRANSITION				
	Standing	Sitting	Bending	Laying
Standing	0.07	0.20	0.40	0.73
Sitting	0.23	0.17	0.43	0.65
Bending	0.23	0.38	0.28	0.63
Laying	0.52	0.48	0.53	0.58

E. State Transition of Human Kinematics:

Critical observations in human kinematics behavior are shown in Table 5. Five behaviors are recognized including: motionless, walking, running, crawling and jumping. These observations are made from sequential images in a lively pattern. This transition analysis is done on the received sequential information from efficient semantic labeling. The state transitions and severity colligated from sequential information from generated semantic message are shown in Table 5.

Table 5: Human Kinematics State Transitions

KINEMATIC STATE TRANSITION					
	Motionless	Walking	Running	Crawling	Jumping
Motionless	0.00	0.26	0.54	0.75	0.68
Walking	0.23	0.14	0.52	0.76	0.73
Running	0.42	0.41	0.43	0.76	0.73
Crawling	0.58	0.53	0.70	0.54	0.75
Jumping	0.53	0.41	0.63	0.79	0.48

F. State Transition of Distance-to-Target:

State transition associated with two images with respect to the distance between the entities to the target is identified using Table 6. By identifying the distance to the target, the entity's kinematics can also be identified through logical reasoning. For example: in frame t(1) subject-1 found far away from the border of target and in frame t(2) subject-1 is found inside an restrictive zone. In this case, it can be reasoned that subject-1 must have moved fast towards the target. For a smooth Distance to Target state transition, a human may attain the state transition of distance-to-target stated below:

Far Away → Near Border → Inside Zone → Nearby Target.

Table 6: Distance-to-Target State Transitions

DISTANCE WITH RESPECT TO ZONE/TARGET				
	Far Away	Near Border	Ins Zone	Nearby Tar
Far Away	0.13	0.37	0.62	0.79
Near Border	0.61	0.30	0.59	0.75
Inside Zone	0.66	0.47	0.57	0.78
Nearby Target	0.70	0.53	0.58	0.75

By normalizing and aggregating the above state transitions with respect to the context of atomic event categories, the model can predict severity of a given situation as human activities take place. Such estimation can be further used as an inference to generating appropriate semantic labels communicated from within a surveillance system to a command control in form of tagged linguistic messages. This approach will improve situation awareness in persistent surveillance systems.

4. EXPERIMENTAL RESULTS & DISCUSSION

Percentage of confidence involved in identifying an event through pre-processing is also fed back to system. For example: event detected from a single source with less confidence, its corresponding event index is removed and fed back to the model for tuning the activity prediction for achieving a better confidence. A probability process of ‘reward’ and ‘punish’ is followed in our work. When a probability of a sequential observation evidentially support a previous observation, then the probability of that previous observations is rewarded with an increase in confidence. Development of State Transition Modeling serves two purposes: 1) To compute the suspiciousness involved in human activity; and 2) Increase the confidence of observation detected via testing the occurrence of smooth transition states. To test the effectiveness of the State transition in severity assesment, a Bag Dropping Scenario was carried in an outdoor environment as shown in Figure 3. Subject-1 (S-1) drives a red semi-truck (V1) to parking lot (PL1), meets Subject-2 (S-2), receives a large object (O1) from S-2, drives to parking lot (PL2), drops O1 at Loc-xx and leaves the scene.

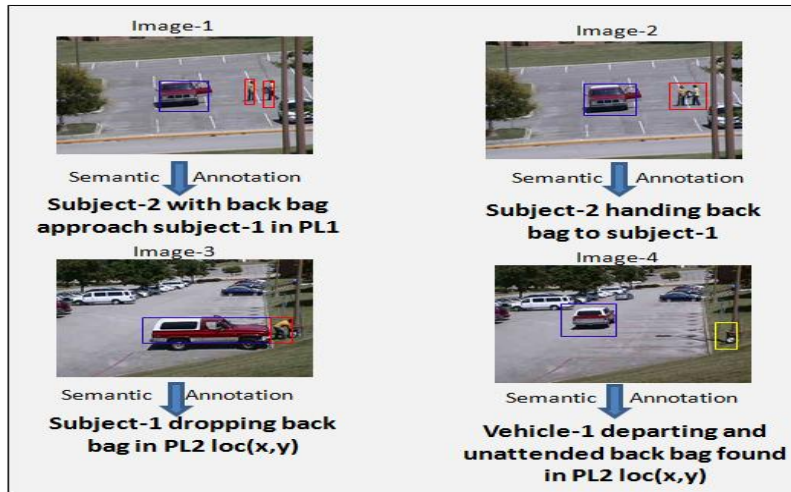


Figure-3: Bag Dropping Senario

Probability of events or observations occurred at each frame is computed and transition rate from one frame to another is computed. Observations from each image are identified. For example, in image-1, an observation of S-2 with back bag (O1) approaching S-1 is detected. These observations are mapped to respective state transitions as discussed above to compute the severity associated with each sequential frame.

Semantic annotations are generated using CESI as shown in Figure 3. State transition from img#1observation to img#2 observation:

- State Transition (approach → cooperate) = 0.53
- State Transition (carry object → handing over object) = 0.35
- State Transition (target visibility) = 0.08
- State Transition (kinematics) = 0.26

Therefore cumulative ST from image-1 to image-2 is equal to 0.3. Observing the State Transition from image-1 to image-2, we conclude the probability of risk involved of this suspicious activity is Guarded, that is, an alert is provided to close monitor the situation. Transition rate of each observations are calculated and being supported by the probability of observations. The state transitions and probability of sequential observations are calculated using the State Transition Tables discussed in section III.

From the observations made from image-1 to image-4, the overall state transitions are found to be 0.54. State transitions in placing an object, leaving it unattended and departing the scene in a quick fashion elevated the Situation Awareness and the Situation Assessment was ranked up to the Elevated Level. The severities of State Transitions are assessed here

only based on the provided images as shown in Figure-3. As more observations are becoming available, the time frame between two image increases, the severity involved in state transitions decreases respectively, that is

$$\text{Time factor } \alpha(f(i) \rightarrow f(i + 1)) \uparrow, \\ \text{then State Transition } (f(i) \rightarrow f(i + 1)) \downarrow$$

In our work, we considered $\alpha=1$. Furthermore, a measure of situation awareness can be appropriately calculated by modulating the spatiotemporal information and according to specific events of interest to the user. For example, the transition of observation (walking \rightarrow running) elevates more in a vigilance environment than in a parking lot. The characteristics of the assessment made are fed to the developed Hidden Markov Model (HMM) in recognizing the type of activity taking place [18].

5. CONCLUSIONS

In this paper, we presented a state transition technique for predicting human interactions with other human, vehicles and objects. A Casual Events State Inference approach is employed with six different categories of state transitions to describe perceptual relevancy of HVI, HOI and HHI activities that are spatiotemporally correlated. The process of increasing the confidence in detecting an observation through smooth state transitions is also presented here. An outdoor experiment involving a suspicious bag dropping is demonstrated. State transition modeling in conjunction with Casual Events Space Inference technique can efficiently and effectively measure the severity associated in a human activity and can facilitate generation of semantic annotations for reporting sensor observed activities.

ACKNOWLEDGMENT

This work has been supported by a Multidisciplinary University Research Initiative (MURI) grant (Number W911NF-09-1-0392) for “Unified Research on Network-based Hard/Soft Information Fusion”, issued by the US Army Research Office (ARO) under the program management of Dr. John Lavery.

REFERENCES

- [1] Ovaska, S.J., VanLandingham, H.F. and Kamiya, A., “Fusion of soft computing and hard computing in industrial applications: an Overview,” IEEE Transaction of System, Man, and Cybernet, Part C: Appl. Rev. v32. 72-79, (2006).
- [2] Marco A. Pravia, Olga Babko-Malaya, Michael K. Schneider, James V. White, and Chee-Yee Chong, “Lessons Learned in the Creation of a Data Set for Hard/Soft Information Fusion,” 12th International Conference on Information Fusion, Seattle, WA, July 6th – 9th, (2009).
- [3] D.L. Hall, J. Llinas, M. McNeese, and T. Mullen, “A Framework for Dynamic Hard/Soft Fusion,” Proc. 11th Int. Conf. on Information Fusion, Cologne, Germany, July (2008).
- [4] Vinayak Elangovan, and Amir H. Shirkhodaie, “A Survey of Imagery Techniques for Semantic Labeling of Human-Vehicle Interactions in Persistent Surveillance Systems,” SPIE Defense, Security and Sensing, Orlando, Florida, April (2011).
- [5] Vinayak Elangovan, and Amir H. Shirkhodaie, “Context-Based Semantic Labeling of Human-Vehicle Interactions in Persistent Surveillance Systems,” SPIE Defense, Security and Sensing, Orlando, Florida, April (2011).
- [6] Joshua Candamo, Matthew Shreve, Dmitry B. Goldgof, Deborah B. Sapper, and Rangachar Kasturi, “Understand Transit Scenes: A Survey on Human Behavior – Recognition Algorithms” in IEEE Transactions on Intelligent Transportation Systems, Vol.11, No.1, March (2010).
- [7] N. Sulman, T. Sanocki, D. Goldgof, and R. Kasturi, “How effective is human video surveillance performance?”, in Proc. Int. Conf. Pattern Recognition, pp. 1–3, (2008).
- [8] Rababaah H., and Shirkhodaie A., “A Survey of Intelligent Visual Sensors for Surveillance Applications,” the IEEE Sensor Application Symposium, February 6-8, (2007).

- [9] A. Hakeem and M. Shah, "Ontology and taxonomy collaborated framework for meeting classification," Proc. of the IEEE ICPR, pp. 219–222, (2004).
- [10] W. Hu, T. Tan, L. Wang, and S. Maybank, "A survey on visual surveillance of object motion and behaviors," IEEE Trans. Syst., Man, Cybern.C, Appl. Rev., vol. 34, no. 3, pp. 334–352, Aug. (2004).
- [11] Shih-Fu Chang, Wei-Ying Ma, Arnold Smeulders, "Recent Advances and Challenges of Semantic Image/Video Search," in IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Hawaii, USA, April (2007).
- [12] Shunsuke Kamijo, Masahiro Harada and Masao Sakauchi, "An Incident Detection System Based on Semantic Hierarchy," IEEE Intelligent Transportation Systems Conf., Oct. 3rd–6th, 2004, Washington, DC.
- [13] Ismail Haritaoglu, David Harwood and Larry S. Davis, "W4: Real-Time Surveillance of People and Their Activities," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.22, No.8, August (2000).
- [14] Shirkhodaie, A., and Rababaah, H. " Multi-Layered Impact Modulation for Context-based Persistent Surveillance Systems," in the SPIE 2010 Defense, Security, and Sensing Conference, Multi-Sensor, Multi-Source Information Fusion: Architectures, Algorithms, and Application, Orlando, FL., April (2010)
- [15] Rababaah, H., and Shirkhodaie, A., "Feature Energy Assessment Map (FEAM): a Novel Model for Multi-Modality Multi-Agent Information Fusion in Large-Scale Intelligent Surveillance Systems, Networks," SPIE 2009 Defense, Security, and Sensing Conference, Multi-Sensor, Multi-Source Information Fusion: Architectures, Algorithms, and Application, paper 7345-19, Orlando, FL., April 13-17, (2009).
- [16] Rababaah H., and Shirkhodaie A., "Human Posture Classification for Intelligent Visual Surveillance Systems," in the SPIE Defense and Security, Visual Analytics for Homeland Defense and Security, Orlando ., March 17-20, (2008).
- [17] Brogan, W.L. (1991). Modern Control Theory, Prentice Hall.
- [18] Vinayak Elangovan, and Amir H. Shirkhodaie, "Recognition of Human Activity Characteristics based on State Transitions Modeling Technique," SPIE Defense, Security and Sensing, Baltimore, Maryland, April (2012).
- [19] Amir Shirkhodaie, Vinayak Elangovan, Aaron Rababaah, "Acoustic Semantic Labeling and Fusion of Human-Vehicle Interactions", SPIE Defense, Security and Sensing, Orlando, Florida, April (2011).
- [20] T. Ko, "A survey on behavior analysis in video surveillance for homeland security applications ", Applied Imagery Pattern Recognition Workshop, 37th IEEE AIPR pp1-8 Oct 2008.