

“© 2007 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.”

Towards Improving Driver Situation Awareness at Intersections

S. Kodagoda, A. Alempijevic, S. Sehestedt, and G. Dissanayake

Abstract— Providing safety critical information to the driver is vital in reducing road accidents, especially at intersections. Intersections are complex to deal with due to the presence of large number of vehicle and pedestrian activities, and possible occlusions. Information available from only the sensors on-board a vehicle has limited value in this scenario. In this paper, we propose to utilize sensors on-board the vehicle of interest as well as the sensors that are mounted on nearby vehicles to enhance the driver situation awareness. The resulting major research challenge of sensor registration with moving observers is solved using a mutual information based technique. The response of the sensors to *common causes* are identified and exploited for computing their unknown relative locations. Experimental results, for a mock up traffic intersection in which mobile robots equipped with laser range finders are used, are presented to demonstrate the efficacy of the proposed technique.

I. INTRODUCTION

INTERSECTION safety is a major concern for most road users and governing bodies. In United States of America alone, there were more than 2.8 million intersection-related crashes occurred in year 2000 [1]. This represents 44% of all reported crashes and includes 8,500 deaths. These crashes result from poor decision making and carelessness as a consequence of the road users receiving conflicting information due to the complexity of the environment and poor visibility.

Fatalities at intersections can be reduced by improved road designs, imposing laws and regulations, sustained education of drivers and road users. Fatalities can be reduced further by improving the situation awareness of the drivers, which can help reduce the potential for side impacts, a leading cause of serious injury or death. Therefore, in this paper, we made an attempt to address the problem of improving driver situation awareness by means of sensing and sensor fusion.

In improving driver situation awareness, it is very important to have a cocoon of sensors around the vehicle. However, occlusions can limit the sensing field of view. This effect is significant at intersections due to the presence

of buildings, trees, surrounding vehicles and other structures. Fortunately, the limited perception achievable by on-board sensors can significantly be improved by utilizing the sensory information of other sources such as infrastructure based and neighboring vehicle based sensors. With the developments in inter vehicular communications [2] and roadway intelligent infrastructure [3-4], it will be feasible to transmit data in real time between infrastructure and neighboring vehicles (Fig.1). Therefore, feasibility of incorporating such sensory signals in driver situation awareness systems is closer than ever before. However, this inherently introduces a major challenge: un-calibrated data registration among moving observers, which is the main focus of this paper.

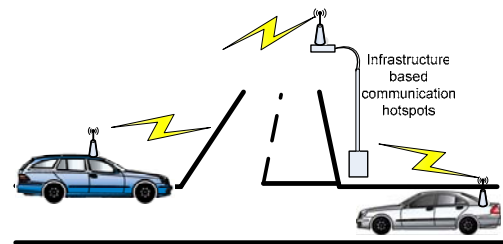


Fig. 1. Communication of information in future roads.

Image data registration is one of the most explored research areas, in remote sensing to biomedical applications [5]. There the problem is to overlay images (two or more) of the same scene taken at different times, from different viewpoints, and/or by different sensors. Registration of biomedical images provides a good insight to un-calibrated multi-modal image data registration. A survey of Mutual Information (MI) based data registration methodologies with CCD, MR or CT images is given in [6]. However, the MI based methodologies reported are not suitable for the vehicular applications due to the requirement of substantial overlaps of the images, unavailability of the knowledge of underlying joint probability density functions in the input space, and high computational complexity. Fisher et al [7] proposed a method to alleviate such problems. They use information theoretic approach for signal level sensor fusion. Appealing experimental results were presented for audio video data registration. The methodology is sound for some specific applications such as pin pointing a mouth of a talking person by analyzing audio and video signals. However the underlying assumption of no relative motion of the speaker in successive image frames is not appropriate when there are dynamic objects in the environment such as

This work is supported by the ARC Centre of Excellence programme, funded by the Australian Research Council (ARC) and the New South Wales State Government. The authors gratefully acknowledge the partial funding of this work by the German Academic Exchange Service (DAAD) with a PhD student scholarship.

S. Kodagoda, A. Alempijevic, S. Sehestedt, and G. Dissanayake are with the ARC Centre of Excellence for Autonomous Systems (CAS), University of Technology, Sydney, AUSTRALIA. (e-mail: S.Kodagoda, A.Alempijevic, S.Sehestedt, G.Dissanayake @ cas.edu.au).

the one addressed in this paper.

In our previous work [8], we have extended Fisher's idea to incorporate dynamic objects in the environment to establish the sensor registration. Dynamic objects provide very little instantaneous information at the signal level. However, tracked dynamic targets provide substantial amount of information in the feature level. Therefore, we have utilized the tracks of targets or their attributes in sensor registration. The target attributes were carefully chosen giving due regards to sensor modality and rate of change of the attribute. The MI based approach was utilized to register the common feature attributes in two sensory signals providing sensor registration. The algorithm was further extended to incorporate the maneuvering targets with possible occlusions as in [9], however this work was restricted to stationary observers. This paper further develops Fisher's algorithm to register and calibrate the sensors with moving observers even when target occlusions are present. We do not assume any prior knowledge of relative localization between the observers.

Target tracking problem at intersections is relatively complex due to the motion of the observer, maneuvering dynamics of the target and occlusions. Multiple model approach, in particular the Interacting Multiple Model (IMM) [10], provides one of the most effective frameworks for tracking maneuvering targets, particularly in combinations with the Sequential Probability Ratio Test (SPRT) [11] which can be used to eliminate false tracks that were initiated due to anomalies in the background and spurious data.

The paper is organized as follows. Section II discusses Fisher's MI based approach for sensor registration. Section III describes how Fisher's algorithm can be adapted to the applications with dynamic observer and dynamic targets. IMM based tracker is presented in Section IV. Experimental results are presented in Section V. Section VI concludes the paper providing future direction of the research.

II. FISHER'S MUTUAL INFORMATION BASED SENSOR REGISTRATION ALGORITHM

In this section we describe an information theoretic approach for sensor registration, which is based on [7]. One important aspect of this method is it does not assume any prior knowledge of the relationships among either sensory signals or sensors.

Let $X^a \in \mathbb{R}^{N_a}$ and $X^b \in \mathbb{R}^{N_b}$ are two high dimensional random variables. Let these be passed through functions, $Y^a = g(X^a, H_a)$ and $Y^b = g(X^b, H_b)$ where, $Y^a \in \mathbb{R}^{M_a}$, $Y^b \in \mathbb{R}^{M_b}$ with $M_a \ll N_a$ and $M_b \ll N_b$. Although the mapping can be done through any differentiable function, here, H_a and H_b are treated as coefficients of linear projections. The mutual information is given by,

$$I(Y^a; Y^b) = h(Y^a) + h(Y^b) - h(Y^a, Y^b) \quad (1)$$

where, $h(\cdot)$ is the differential entropy, which can be defined for a random variable Y with density p_Y as,

$$h(Y) = -\int_{\Omega} p_Y(y) \log(p_Y(y)) dy \quad (2)$$

Our goal is to maximize the mutual information between two random variables Y^a and Y^b . This can be achieved by maximizing the entropies of Y^a , Y^b and minimizing the joint entropy, $h(Y^a, Y^b)$. Entropy is a maximum in a uniform distribution, which is therefore chosen as the desired distribution in the transformed space. Input samples are linearly transformed to the output space and the density of the output samples is estimated using Parzen density estimator [12]. Equation (2) is linearized with a second order Taylor series expanded around the uniform distribution and the difference between entropies of desired and estimated densities are calculated. It can be shown that the updating term becomes,

$$\begin{aligned} \Delta y_i^{(k)} &= b_r(y_i^{(k-1)}) - \frac{1}{N} \sum_{i \neq j} \kappa_a(y_i^{(k-1)} - y_j^{(k-1)}, \sum) \\ b_r(y_i)_j &= \frac{1}{d} \left(\kappa \left(y_i + \frac{1}{d}, \sum \right) - \kappa \left(y_i - \frac{1}{d}, \sum \right) \right) \\ \kappa_a(y, \sum) &= \kappa(y, \sum) * \kappa'(y, \sum) \\ &= - \left(2^{M+1} \pi^{M/2} \sigma^{M+2} \right)^{-1} \exp \left(-\frac{y^T y}{4\sigma^2} \right) y \end{aligned} \quad (3)$$

where, $\kappa_a(\cdot)$ is a kernel: a Gaussian pdf is assumed here. y_i symbolize a sample of either Y_a or Y_b , $M = M_a, M_b$ or $M_a + M_b$ based on the term of (1) that is considered. The j^{th} element of $b_r(y_i)$ is defined as $b_r(y_i)_j$. d is the support of the output space. N is the number of samples. The adaptation procedure includes the update rule (3) followed by a least squares solution for h_a and h_b until a maximum is reached. The stopping criteria used is,

$$\delta = \frac{\max(\Delta_{NN}) - \min(\Delta_{NN})}{\max(\Delta)} \quad (4)$$

where, the term Δ_{NN} is the nearest neighbour distance in the resulting output distribution, $\max(\cdot)$ and $\min(\cdot)$ are the maximum distance and minimum distance between any two points in the output space. In (4) the numerator is a measure of uniformity of the output space and the denominator is a measure of how well the output space is filled.

III. ADAPTATION OF FISHER'S ALGORITHM FOR MOVING SENSOR REGISTRATION WITH DYNAMIC OBJECTS

The Fisher's [7] algorithm estimates the Mutual information between two input signals in the absence of an estimated PDF of either signal. Maximization of Mutual Information is achieved in the locally created subspace. The proposed method has been applied to signal-level fusion of audio-visual information in order to locate a speaker. However, it is assumed that there is no relative motion in speaker position in the image. This assumption is often violated in practical applications, such as the one described in this paper.

In this paper, our goal is to register sensors mounted on moving platforms (moving observers) using commonly seen dynamic targets in the environment. Consider the scenario shown in Fig.2. The vehicle A and B travels towards an intersection, where their side views are mostly blocked by the buildings at the corners. Vehicle A can see most of the areas along the X direction of the intersection, which vehicle B can not see. On the other hand, vehicle B can mostly see the road segment in the Y direction, but not the segment in the X direction. However, if the data can be sent across the vehicles, both the vehicles can benefit by expanded perception. Accuracy of a typical Global positioning system based localizer is not sufficient for the transformation of data from one coordinate system to the other required in this scenario. However, if both of the vehicles observe at least a single moving target (for example a pedestrian crossing the intersection) which can be considered as a *common cause*, the Fisher's algorithm can be exploited for obtaining the sensor registration.

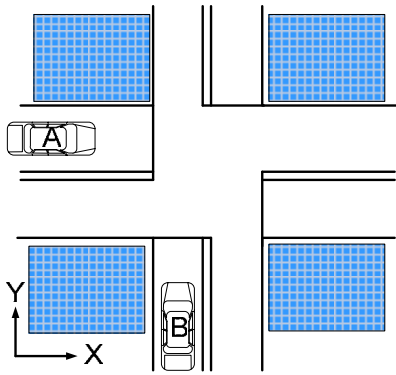


Fig. 2 Road intersection

Moving objects (in this application: pedestrians) wander across the sensory space providing very little instant information in signal level. However, if tracked as they move, they provide substantial amount of information in the feature level. Therefore, we propose adapting Fisher's algorithm to use feature level registration rather than the signal level registration as originally proposed. First the dynamic features (targets) in the scene are extracted and their attributes are estimated. Selection of the feature

attributes depends on the sensor modality. An important aspect in selecting attributes is the requirement that a non-zero rate of change in at least one attribute for all possible movements of the feature. For example, one may choose attributes in image features as bearing to a target, optical flow and area of occupation, whilst another may choose laser feature attributes such as bearing to the target, target speed and the angle of occupation. However, if the observers are moving, the attributes calculated on the tracked objects in the sensory spaces will contain the correlations due to observer dynamics. This causes the same target's attributes calculated from different moving observers to be non informative resulting in the failure of the MI based sensor registration. This problem can be solved, if the targets are tracked in their respective locally defined world frames. Here, we assume the observer localization to a known uncertainty using some external means (eg. SLAM). However, we do not assume the knowledge of relative localizations between the observers.

IV. IMM BASED TRACKER

The dynamics of the objects that are present in urban intersections can be considered as maneuvering targets. One of the best algorithms to track maneuvering targets with stationary/moving observers is the Interacting Multiple Model (IMM) algorithm. Therefore, here we propose to use IMM for tracking maneuvering targets together with SPRT for track initiation, maintenance and deletion. Once the targets are being detected and tracked the attributes of the targets are calculated to be used in MI based sensor registration. The IMM algorithm [10] is illustrated in Fig.3

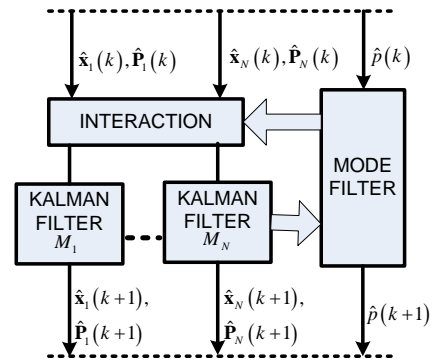


Fig.3 The IMM algorithm [10]

Where, $\hat{\mathbf{x}}_i, \hat{\mathbf{P}}_i$ define the states and their covariance matrices for $i=1, \dots, N$ hypotheses. $\hat{p}(k)$ is the mode probability.

A. Process model

The target dynamics in the world coordinates, W , (Fig. 4) are modelled by a combination of constant speed and turn

rate models.

$$\dot{\mathbf{x}}_w(t) = \mathbf{f}(\mathbf{x}_w(t), \boldsymbol{\omega}, \mathbf{v}(t), t) \quad (5)$$

where, \mathbf{f} is a nonlinear function, $\mathbf{x}_w = [x \ \dot{x} \ y \ \dot{y}]^T$, $\{x, y\}$ are the position coordinates of the target in x - y directions and $\{\dot{x}, \dot{y}\}$ are the respective velocities in the world coordinates.

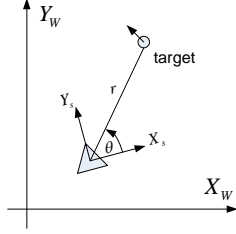


Fig. 4. Coordinate frames: suffix “W” – world coordinates, suffix “S” – sensor coordinates

B. Observation model

Here we assume range and bearing observations. Therefore the observation model in the sensory space is,

$$\mathbf{z}(t) = \mathbf{h}(\mathbf{x}_w(t), \mathbf{x}_s(t), \mathbf{w}(t), t) \quad (6)$$

where, \mathbf{h} is a nonlinear function, $\mathbf{z}(k+1) = [r \ \theta]^T$ are range and bearing to the target in the sensory space. $\mathbf{x}_s(t)$ is the pose of the observer in the world reference frame. $\mathbf{w}(t)$ is the measurement noise (zero mean Gaussian).

False tracks may be initiated due to spurious target detections. The true track confirmation and false track deletion are handled using the ideas from integrated probabilistic data association (IPDA) [13] with SPRT [11]. Using the Markov relationship, the probability of existence of the true target, $P_T(k+1|k)$ before the receipt of data in scan $k+1$ is [11],

$$P_T(k+1|k) = P_{22}P_T(k|k) + P_{12}[1 - P_T(k|k)] \quad (7)$$

where P_{22} is the probability of transition from observable state to observable state, whilst P_{12} is the probability of transition from unobservable state to observable state. Then, the update of the probability of target existence is [11],

$$P_T(k+1|k+1) = \frac{1 - \delta_{k+1}}{1 - \delta_{k+1}P_T(k+1|k)} P_T(k+1|k) \quad (8)$$

where δ_{k+1} and subsequently \bar{V} are defined as,

$$\delta_{k+1} = \begin{cases} P_D P_G & N_{k+1} = 0 \\ P_D P_G \left[1 - \bar{V} \sum_{i=1}^{N_{k+1}} \frac{1}{P_G (2\pi)^{M/2} \sqrt{|S(k+1)|}} e^{-d_i^2/2} \right] & \text{Otherwise} \end{cases}$$

$$\bar{V} = \frac{V_{G_{k+1}}}{(N_{k+1} - P_D P_G P_T(k+1|k))}$$

P_D is the probability of detection, P_G is the gate probability, V_G is the gate volume, N_{k+1} is the number of measurements inside the validation gate, S is the innovation covariance, and d_i^2 is the normalized innovation squared of the i^{th} measurement. The Log Likelihood Ratio, LLR, can now be defined as [11],

$$LLR_{k+1} = \ln\left(\frac{P_T}{1 - P_T}\right) \quad (9)$$

Once the LLR is obtained, track confirmation and termination thresholds are determined using the SPRT [11] as,

$$\begin{aligned} LLR_{k+1} &\geq \ln\left(\frac{1 - \beta_T}{\alpha_T}\right), \text{ declare track confirmation} \\ \ln\left(\frac{\beta_T}{1 - \alpha_T}\right) &< LLR_{k+1} < \ln\left(\frac{1 - \beta_T}{\alpha_T}\right), \text{ continue test} \\ LLR_{k+1} &\leq \ln\left(\frac{\beta_T}{1 - \alpha_T}\right), \text{ delete track} \end{aligned} \quad (10)$$

where, α_T is the probability of false track confirmation, and β_T is the probability of true track termination.

V. RESULTS

Fig. 5 shows a bird’s eye view of the experimental site. It was chosen as a simulated road intersection with surrounding buildings (cubicles). A Pioneer robot is used to simulate one vehicle, whereas a wheelchair robot is used to simulate the other vehicle approaching the intersection. Both robots are equipped with a laser range finder and wheel encoders. Clocks on both platforms were synchronized. A person is walking randomly but slowly in the middle of the intersection.

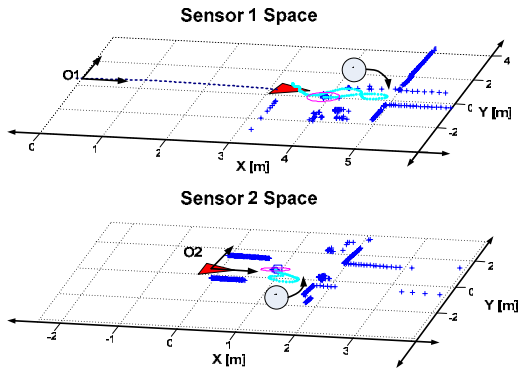


Fig. 5 Experiment site

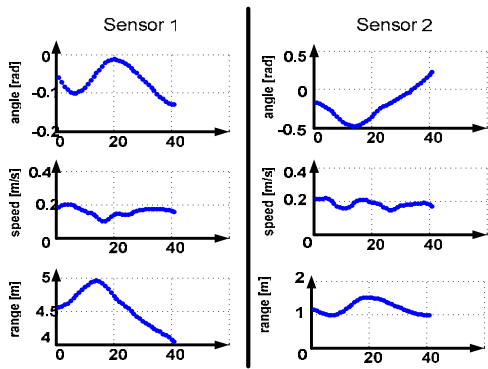
A. Scenario I

In the first scenario, the wheelchair was kept stationary facing the intersection, whilst the Pioneer is towards the intersection. Sick laser range finders mounted on each platform are used to log data. Fig. 6 (a) shows the trajectory of the tracked pedestrian as seen by Pioneer (Sensor 1 space) and the wheelchair (Sensor 2 Space). It is to be noted that only the last laser range scan was overlaid to show the difficulty in carrying out sensor registration, for example using Iterative Closest Point algorithm. Each observer builds a track of the pedestrian in its own coordinate frame without knowing the locations of the other observer.

Fig. 6 (b) shows the extracted attributes of the tracked pedestrian (target). Here we choose the attributes as, angle to the target, speed of the target and the range to the target. Table I shows the calculated MIs with the highest MIs are denoted by the symbol '*'. The algorithm concluded predicting that speed is being the most informative attribute in both sensory spaces. Therefore, the whole tracks of the pedestrian in both sensor spaces were used in a least square sense to calculate the coordinate transformation between tow sensory spaces. Fig. 6 (c) shows the results of transforming sensor 2 data onto sensor 1 coordinate system. The crosses, 'x', represent the laser data seen by sensor 1, whereas, '+' represent the transformed laser data seen by sensor 2 in sensor 1's coordinate frame. As it is obvious from the plot, now the both robots can see more parts of the environment.



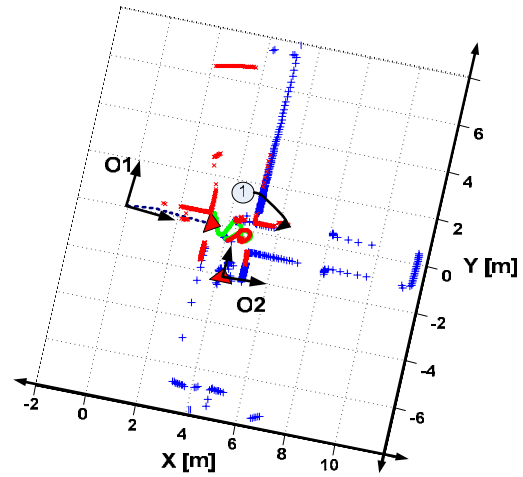
(a) Tracked targets in each sensory space



(b) Target attributes

TABLE I: MI VALUES. a – ANGLE, s – SPEED, r – RANGE, T – TARGET

Sensor/ signal	a. T1	s. T1	r. T1
1	4.5	7.85*	0.34
2	1.15	8.85*	1.17



(c) Sensor 2 data transformed onto reference frame of sensor 1, showing the correct sensor registration. Symbol 'x' denotes the last raw laser range scan obtained by sensor 1, and symbol '+' denotes the last raw laser range scan obtained by sensor 2 after transformation

Fig. 6 Sensor registration with a moving observer and a stationary observer with a single dynamic target.

B. Scenario II

In the scenario II, both the Pioneer and the wheelchair were in motion while one pedestrian ("1" in Fig. 7 (a)) is randomly walking in the intersection. The other pedestrian "2", who is imitated to be crossing the road, however can only be seen by the Pioneer (sensor 1). Again, it can be seen from the plotted last laser scans that it is not feasible to register the two sensors using conventional ICP based technique due to the limited data overlap. The estimated target attributes are shown in Fig. 7 (b). MI learning curve depicted in Fig. 7 (c) shows the that the MI based algorithm found a match in its 38th iteration. Although the sensor 1 could see an additional target, Table II shows that range of target 1 was the most informative in both sensory spaces providing correct registration (Fig. 7 (d)).

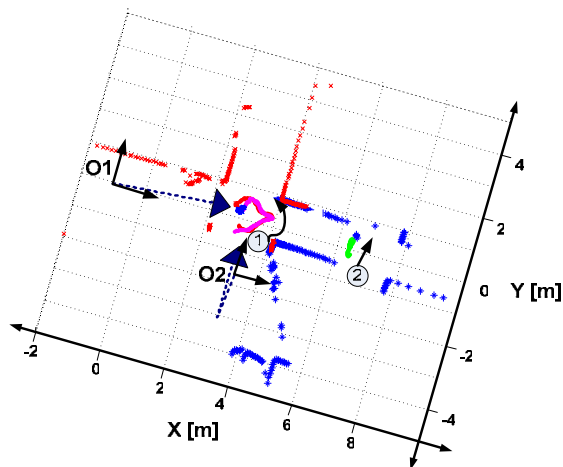
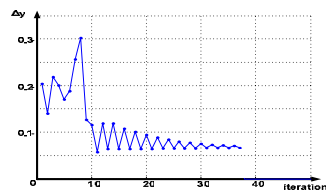
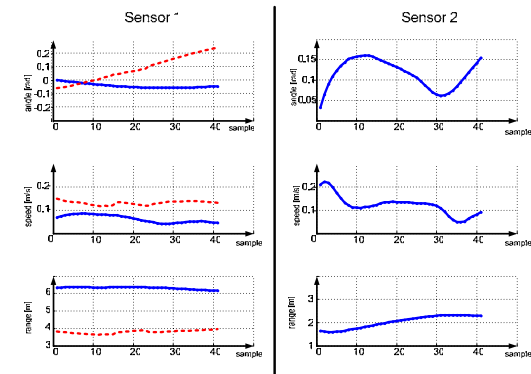
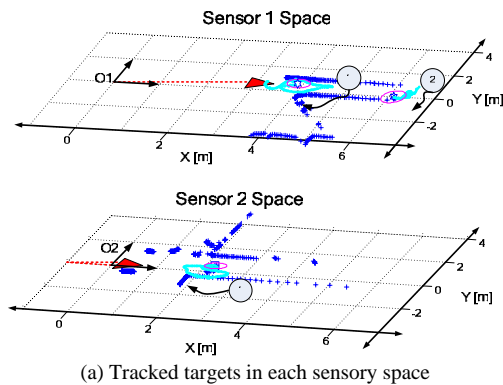


Fig. 6 Sensor registration with a two moving observers and two dynamic targets.

TABLE I: MI VALUES. a – ANGLE, s – SPEED, r – RANGE, T – TARGET

Sensor/ signal	a. T1	a. T2	s. T1	s. T2	r. T1	r. T2
1	18.1	4.4	4	1	20.5*	2.2
2	0.55		4		5.8*	

VI. CONCLUSION

In this paper, we have addressed the problem of moving observer sensor registration in the context of improving the situation awareness at intersections. The problem is challenging, as the observers are dynamic, they have limited perception due to occlusions, and the fact that they are unaware of relative locations of the other observers. We proposed to solve the problem using a mutual information based techniques, which provides us with encouraging results. We are now planning to test the algorithm in real intersections with vehicles and pedestrians. We are also investigating the use of other sensor modalities, such as cameras in the implementation

ACKNOWLEDGMENT

The authors would like to thank Tarek Taha, Zhengzhi Zhang and Dalong Wang for their help in setting up the experiment.

REFERENCES

- [1] <http://safety.fhwa.dot.gov/intersections/interbriefing/01prob.htm>
- [2] Tank T. and Linnartz J.M.G., "Vehicle to vehicle communications for AVCS platooning", IEEE Trans. Veh. Technol., vol. 46, pp. 528-536, May 1997.
- [3] Kenward, M., "Intelligent Moves", The IEE Review, Vol.52, No.2, Feb. 2006, pp.22-23.
- [4] Bishop, R., "Arizona I-19 Wi-Fi Corridor: Assessment of Opportunities for Probe Data Operations", Report TRQS-02, October 2005, <http://www.azdot.gov/TPD/ATRC/publications/QuickStudies/TRQS-02.pdf>
- [5] Zitova, B and Flusser, J., "Image registration methods: a survey", Image and Vision Computing 21 (2003) 977–1000.
- [6] Plum, J.P.W., Antoine, J. B., and Viergever A., "Mutual information based registration of medical images: a survey", IEEE Trans. Medical Imaging, Vol.22, No.8, Aug. 2003, pp. 986-1004.
- [7] Fisher III J.W., Wainwright, M.J., Sudderth, E.B. and Willsky, A.S., "Statistical and Information-theoretic methods for self organization and fusion of multimodal networked sensors", The International Journal of High Performance Computing Applications, Vol.16, No.3, 2002, pp.337-353.
- [8] Alempijevic A, Kodagoda S, Underwood J, Kumar S, Dissanayake G (2006), Mutual Information based Sensor Registration and Calibration, Proceedings of the IEEE/RSJ International Conference on Intelligent robots and Systems (IROS 2006), Beijing: 25-30
- [9] Kodagoda K R S, Alempijevic A, Underwood J, Kumar S, Dissanayake G (2006), Sensor Registration and Calibration using Moving Targets, Proceedings of ninth International Conference on Control, Automation, robotics and Vision (ICARCV2006): 830-835
- [10] Blom H A P, Bar-Shalom Y (1988), The interacting multiple model algorithm for systems with Markovian switching coefficients, The IEEE Transactions on Automatic Control, Vol.33, No. 8: 780-783.
- [11] Blackman S, Popoli R (1999), Design and Analysis of Modern Tracking Systems, Boston, Artech House
- [12] Parzen E (1962), On estimation of probability density function and mode, Annals of Mathematical Statistics, 33: 1065 - 1076
- [13] D. Musicki, R. Evans and S. Stankovic, "Integrated probabilistic data association", IEEE Trans. Automat. Contr., Vol. 39, No. 6, pp.1237-1241, June 1994.