Handling Ambiguous Object Recognition Situations in a Robotic Environment via Dynamic Information Fusion

S. Pourya Hoseini A.
Computer Science and Engineering Department
University of Nevada, Reno
Reno, United States
hoseini@nevada.unr.edu

Mircea Nicolescu
Computer Science and Engineering Department
University of Nevada, Reno
Reno, United States
mircea@cse.unr.edu

Monica Nicolescu
Computer Science and Engineering Department
University of Nevada, Reno
Reno, United States
monica@cse.unr.edu

***Abstract*—Vision is usually a rich source of information for robots aiming to understand activities that take place in their surroundings, where a relevant task can be to detect and recognize objects of interest. In real world conditions a robot may not have a good viewing angle or be sufficiently close to an object to distinguish its features, which can lead to misclassifications. One solution to address this problem is active vision, leading to an improved level of situational awareness in a dynamic environment. In that context, a vision system on the robot actively manipulates the camera to obtain enough discriminating features for the task of object detection and recognition.**

In this paper, an active vision system is proposed that is able to identify a situation with a high possibility of misclassification (for example, partial occlusions) and then to take appropriate action by dynamically incorporating another camera installed on the robot's hand. A decision fusion technique based on a transferable belief model generates the final classification results. Experimental results show considerable improvements in object detection and recognition performance.

Keywords—object detection, ambiguous situation management, robotics, active vision, situational awareness, information fusion, transferable belief model, Dempster-Shafer evidence theory.

I. Introduction

Building vision-based capabilities for robotic systems is an area of interest in the context of a computational framework for intent recognition. This is particularly relevant for situations that involve collaboration among multiple agents or other robotic systems, or detection of situations that can pose a particular threat. For surveillance or military applications, it is highly important to enable understanding the intent of relevant agents in the environment, from their current actions, before any attack strategies are finalized.

A central aspect of this task is the ability to accurately recognize and track objects of interest in a dynamic environment. The success of object recognition, though, is dependent on the distinguishable features of objects that the robot views. If the current viewpoint only allows inadequately distinct features or the object is significantly occluded the object detection system may experience deteriorated performance. A workaround for this problem is to try to view objects from different viewpoints with multiple cameras. However, following this strategy for all objects with every available camera puts unnecessary computational burden on the system, because not all object views are ambiguous, nor every viewpoint provides useful information for classification. The added computation complexity may lead to the inability of the vision system to work in real-time, which is critical in most robotic tasks, and especially when intent recognition is desired. Therefore, active vision, i.e. sensor management by dynamically selecting a few cameras in different viewpoints, might be a more viable approach. A comprehensive survey of the active vision literature can be found in [1]. More details about sensor management are provided in [2][3], and an example of application of sensor management in determining situational awareness is mentioned in [4].

In our work, a dual-camera active vision system working with a PR2 robot was designed for the task of object detection and recognition. There is a Kinect v1 3D sensor (main camera) installed on the robot's head, while an ordinary RGB camera is utilized as the secondary camera on the robot's left hand. The camera pair offers two different views of a scene. The proposed system first detects and classifies objects in the scene viewed by the main camera only. Here, the robot needs to assess if the recognition is reliable given the current viewpoint by computing a confidence measure, then decides whether to incorporate the secondary camera in the recognition process. Only if the vision system determines that the secondary camera would be helpful, another round of object recognition is performed by using the arm-mounted camera. The classification decisions from the two cameras are then fused together via a novel transferable belief model.

The transferable belief model (TBM) [5] is a variation of the Dempster-Shafer evidence theory (DST) [6][7], in which an unnormalized rule of combination is followed [5]. The Dempster-Shafer theory simultaneously takes into account inaccuracy and uncertainty [8]. In the context of Dempster-Shafer fusion we do not deal with singleton probabilities only, instead it is possible to have alternative units with non-empty intersections [9]. The transferable belief model and Dempster-

This work has been supported by Office of Naval Research Award #N00014-16-1-2312.

Shafer theory are typically employed for multi-sensor data fusion [10] and in general information fusion. In [8] and [11] DST is utilized for image segmentation and pattern classification, respectively. TBM is used in [12] for fusing classifications of airborne objects. Fusion of navigation information and vision-based speed limit sign recognition information by TBM in [13] is applied to intelligent speed limit identification. The works of [14] and [15] reported TBM as a reliable data fusion technique for human action recognition and human facial recognition, respectively. In addition, a TBM-based method is presented in [16], which fuses location beliefs of vehicles to verify their actual locations.

In this work, the classifiers corresponding to the two cameras generate the uncertainty information needed by TBM. They are combined along with beliefs about each object category to form the final classification of each object. The contributions of the proposed active vision system are: (1) improved detection and recognition of objects in a robotic environment with capability to switch between two cameras, (2) dynamic ambiguity management for the classification of the detected objects, and (3) a transferable belief model-based decision fusion technique.

In the rest of this paper, the proposed robotic ambiguity management vision system is described in Section II. In Section III, the experimental results are presented and analyzed, while Section IV concludes the discussion.

II. Active Ambiguity Management System For Object Recogniton

The components of the proposed system are shown in the flowchart of Fig. 1. The left vertical bar in Fig. 1 delineates the major phases in data processing. In the following, we first describe the main steps of the method, then we explain the proposed information fusion technique in more detail.

A. *Main Steps of the Proposed Object Detection System with Situational Awareness*

As Fig. 1 shows, in the beginning, every video frame coming from the robot's head camera (main camera) and the hand camera (secondary camera) are captured. At this point, object detection and recognition are only performed on the main camera's frames. For the images taken from the main camera, after a stage of converting color spaces and denoising, a foreground map is obtained by using a Mixture of Gaussians background model and a foreground-background segmentation method [17]. The binary foreground map is then cleared from small noise using morphological operations. In order to detect objects in the scene, a connected component blob detection step is employed.

To classify each detected object, two types of features are computed: a histogram of oriented gradients (HOG) [18] and a color histogram. A Support Vector Machine (SVM) classifier is used to provide the object category. In our method, the SVM classifier outputs a *mass* vector instead of a probability vector, as explained in the next section. Each element in the vector is the belief of the classifier about the similarity of the current object to the respective class. Based on the mass values, a

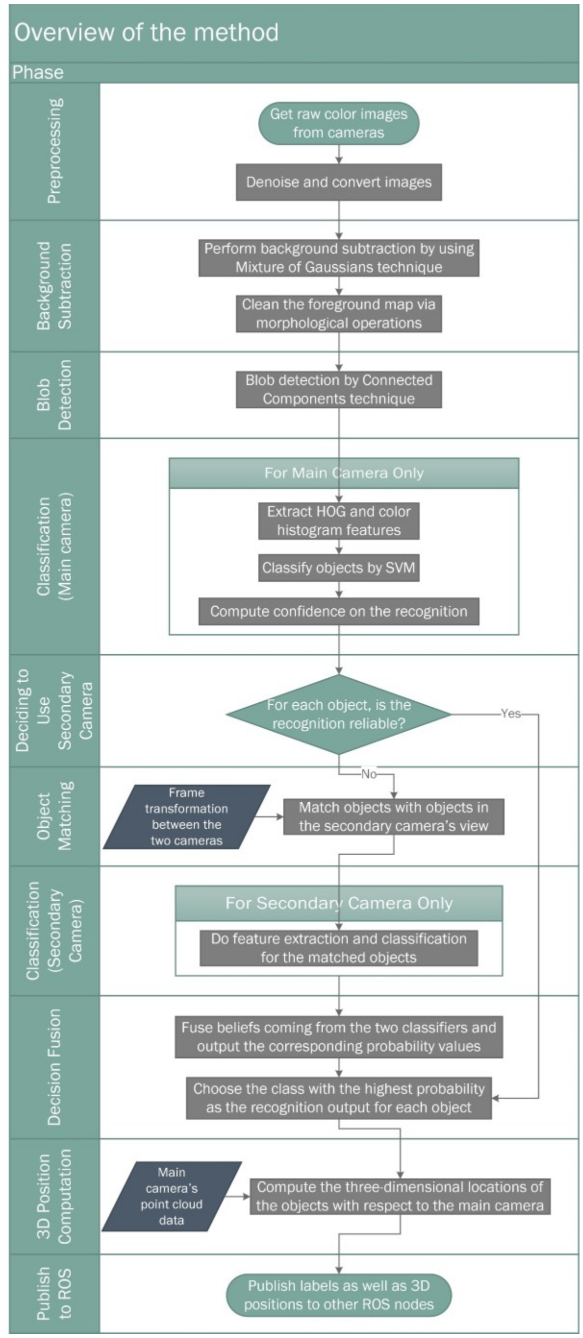

Fig. 1. Flowchart of the proposed active vision system.

confidence measure is computed for the classification result. The confidence metric is defined as the ratio of the maximum value in the mass vector to the second largest mass value. A low confidence value indicates ambiguity between at least two strong candidate categories for the object, which usually happens when discriminating features cannot be reliably observed, thus producing a flat mass vector.

By comparing the confidence measure with a threshold value, the classification is considered "reliable" or "unreliable". If a reliable confidence measure is obtained the classification of the object is deemed final, otherwise the same procedures up to and including the object detection are repeated for the image taken from the secondary camera. The objects detected in the secondary camera view are then matched with those in the main camera view. The matching process is based on transforming object locations from the main to the secondary camera view, and minimizing the image distances for objects in each matching pair. In the next step, objects in the secondary view that match the unreliable objects in the main view are classified in order to obtain a second mass vector.

To combine the two mass vectors a transferable belief model (TBM) decision fusion is used, where for each object the decision fusion technique outputs a probability vector of all categories, as explained in the next section. Given the final probability vector for each object, a winner class is determined by picking the one with the highest probability value. Note that if the object in the main view had a reliable classification, no matching, second classification or fusion operations would be performed.

Since the proposed vision system is aimed at activity and intent recognition in robotic environments, it is essential for the robot to determine the 3D location of the detected objects. We deployed a 3D localization procedure by incorporating the point cloud data available from the 3D camera (i.e. main camera). The labels of the recognized objects along with their 3D location are then passed to other modules in the robotic system, specifically on a Robot Operating System (ROS) network.

After describing the overall structure of the proposed active vision system, with its inherent dynamic uncertainty management, in the next section we detail the proposed TBM decision fusion method.

B. The Proposed Transferable Belief Model for Fusing Classifications

Let us denote a set of singleton probabilities with Ω. In Dempster-Shafer evidence theory (DST) terminology, Ω is named *frame of discernment*. By singleton we mean that any of the two probabilities in Ω are mutually exclusive, signifying that there is no overlap between them. In our application, Ω characterizes a probability vector coming from a classifier. Instead of merely using Ω, Dempster-Shafer decision fusion alternatively allows for a broader set of probabilities, which is the power set of Ω. The power set of Ω consists of all combinations of singleton probabilities from an empty set to the set universe of Ω. The number of subsets within a power set is 2^{Ω}. In the DST formulation, a value called mass in the range of [0, 1] is attributed to any subset element in the power set. Masses are in some sense a replacement for probabilities in Bayesian fusion, as described in the next paragraphs.

Any subset of Ω with mass value greater than zero is called a *focal element*. Furthermore, the sum of all masses in the power set must be equal to 1, as shown in (1), where Ψ represents an input subset of Ω and *m(.)* is a mass value for it.

$$\sum_{\Psi \subseteq \Omega} m(\Psi) = 1 \tag{1}$$

In our work, a frame of discernment stands for a probability vector coming from one of the classifiers - therefore, we have two frames of discernments to combine. Given a probability vector with n elements (for n object categories), its power set is designed to have (n + 1) focal elements. Out of (n + 1) focal elements, n elements are the same singletons as the original probability vector. The last one is the universe of the probability vector elements (i.e. the complete probability vector itself.) By defining the power set this way, we have masses for each object class that work like probabilities of that object class. Furthermore, there is a mass value for all the trained objects. This extra mass actually represents how a classifier believes an object of interest is similar to all of the objects it is trained for. Accordingly, the mass of the extra focal element corresponds to the probability of "unknown" class for a test object.

For the case of the DST decision fusion in our work, we trained classifiers for (n + 1) categories to get (n + 1) masses instead of n probabilities from a classifier. The first n categories are subject to the normal training routine as before, where each category is for an object class. The last category (class "unknown") is trained by aggregating half of each object class training images as its training set. Half of the training set is used in place of the total set to decrease the training time. Moreover, categories are balanced during training by using weights relative to training set size in the optimization formulae of the SVM classifier.

The fusion of the two mass vectors from the main and secondary view classifiers is accomplished by using the unnormalized rule of combination [5], as shown in the following:

$$m(\Psi) = \sum_{\alpha \cap \beta = \Psi} m_A(\alpha) * m_B(\beta), \ \ \forall\, \Psi \subseteq \Omega \tag{2}$$

where $m(\Psi)$ is mass of category Ψ, and sets A and B are mass vectors of the main view and the secondary view, respectively. To reiterate, a mass vector here is the set of (n + 1) focal elements of a classifier's power set. Focal elements α and β are each a category from the mass vector of the main view and the secondary view classifiers in order. From (2) and regarding the set of categories in our case, a category α in the mass vector of the main view (A), except the class "unknown", has intersection with two βs. The two intersecting βs are the "unknown" category of the mass vector of the secondary view (B) plus an element in B with the same category as α. The same rationale applies to β.

Through adopting the unnormalized rule of combination, our Dempster-Shafer fusion turns out to be a transferable belief model [5]. To convert the combined mass vector to a probability vector, we utilize the pignistic transformation described in [19]:

$$P(\omega) = \sum_{\omega \in \Psi} \frac{m(\Psi)}{|\Psi|}, \ \ \forall \omega \in \Omega \text{ and } \forall\, \Psi \subseteq \Omega \tag{3}$$

where $P(\omega)$ is the probability of an object class ω, excluding the "unknown" category. Also, Ψ is any non-empty (focal element) of the set of object classes and $|\Psi|$ designates the number of object classes in the subset Ψ. The above equation shows that any mass belief is distributed among its comprising class probabilities [11].

Generally speaking, the two classifiers feeding the DST fusion provide a mass value for an "unknown" category besides actual classes of objects present in the training. This specific mass value can be considered as indicating to the fusion module the uncertainty of the classifier about its recognition results. Since the sum of all values in a mass vector is 1, by an increase in the mass of the unknown category others will receive lower mass values. Considering the combination rule in (2), we see not only that an increase in the mass of the unknown category of a classifier reduces other masses of that classifier, but also that it weighs more toward the masses of the other classifier, with which it has a non-empty intersection. In contrary, when the unknown mass of a classifier is low, this implies a more resolute classifier which contributes more to the final fused decision.

After fusing mass beliefs to form probabilities of each object class, probabilities are normalized to sum to one, then the category with maximum probability value is selected for the object label. Fig. 2 shows a sample object detection situation. There are four discernible red bounding boxes, which specify ambiguous situations in the classifications of the main view (bottom left window). Except the object labeled as Ball, the other three unreliable recognitions are actually incorrect. For the unreliable detections only, the corresponding objects are detected in the secondary view and matched with them (bottom right window). The decision fusion result is depicted in the top left window of Fig. 2. As it can be seen, the errors in the initial detections are corrected after fusing with the secondary view detections.

III. Experimental Results

In this section, the results obtained from nine real world experiments are presented. In all the test benchmarks, a PR2 robot detects objects in table-top settings in front of its main camera and uses its left hand for the secondary view. We performed various tests in different configurations by changing the placement of objects and the relative position and orientation of the two cameras during experiments. Fig. 3 shows the robot in one of the test conditions. In the next subsection, the results are presented and analyzed. Later, we compare the proposed decision fusion method with the Bayesian decision fusion.

A. Results and Analysis

Fig. 4 shows the confusion matrix for the main view classifier. As illustrated in the confusion matrix, eight objects were defined in the training set in the experiments. Anything not defined in the training set is considered "Background". Subsequently, the Background row in the confusion matrix stands for any unintended object classified as one of the known categories. In a similar fashion, any failure in detection of known objects is counted in the Background column. Note that the intersection of Background column and row is intuitively void.

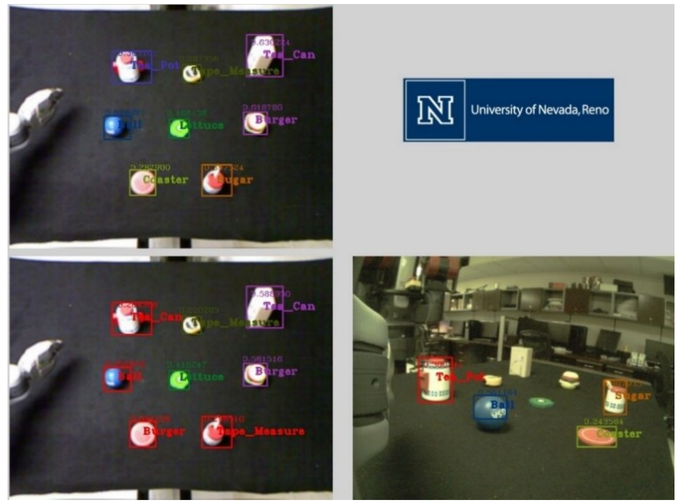

Fig. 2. An illustrative example of the dynamic object detection. Top left: fused view; bottom left: main view; bottom right: secondary view. The red-colored boxes in the main view show unreliable classifications (ambiguous classification situations).

Fig. 3. PR2 robot in one of the test conditions.

The confusion matrix for the final detections is shown in Fig. 5. From the confusion matrices we compute several performance measures: accuracy, precision, recall, and F_1 score. Recall, the ratio of true positives to the ground truth positives, is a measure of a classifier of how well it correctly finds objects of interest. On the other hand, precision, the ratio of true positives to all positives, verifies ability of a classifier in discriminating true objects from false positives. F_1 score is the harmonic mean of recall and precision. In contrary to precision and recall, accuracy provides a balanced insight about the performance of the classifier. Accuracy is defined as the ratio of true detections to all detections.

Table I shows the precision, recall, F_1 score, and accuracy for the main view detections, compared to the detections after fusing information from the secondary view (if dynamically requested). In that table, macro-averaging means the metric is computed by averaging over separate calculations for each object category, in contrast with micro-averaging, which is a way of computing the measure for all the test samples aggregately. Micro-averaging precision and recall are identical to accuracy in the case of a multi-class classifier. For this reason, they are not indicated separately in Table I.

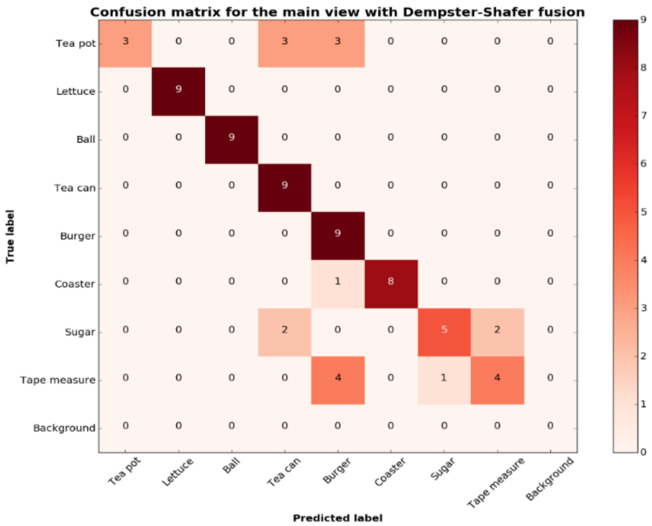

Fig. 4. Confusion matrix for the main view classification only.

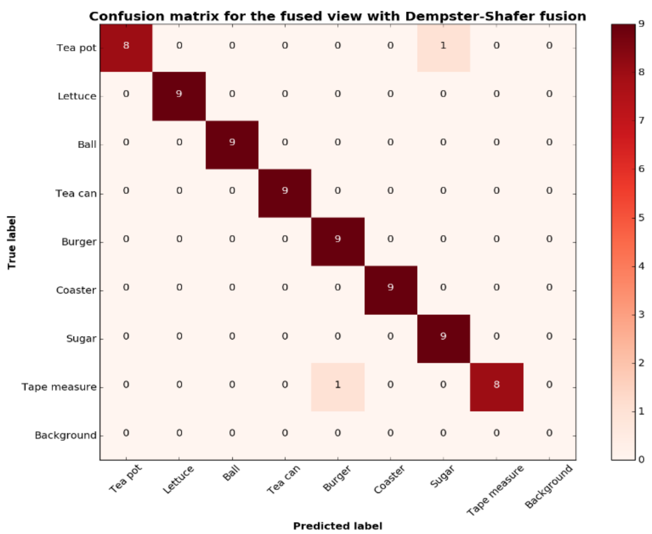

Fig. 5. Confusion matrix for the entire system after ambiguous situation management and the TBM information fusion.

TABLE I. PERFORMANCE MEASURES OF THE MAIN VIEW ONLY AND THE ACTIVELY FUSED VIEW.

Performance Measure	Main Camera	Actively Fused View
Macro-Averaging Precision	0.834	0.975
Macro-Averaging Recall	0.777	0.975
Accuracy	0.777	0.972
F_1 Score	0.804	0.975

As shown in Table I, all metrics in the case of dynamically resolving the ambiguous classification situations and fusing information with the proposed transferable belief model are superior to those using the main camera only. The accuracy with the proposed method is 19.5% better, while the recall, precision, and F_1 score experienced 19.8%, 14.1%, and 17.1% improvements, respectively. It is evident that the amount of enhancement is substantial. For example, the imperfection in accuracy went down from 22.3% to 2.8%.

B. *Comparison with Bayesian Fusion*

The focus of this section is to evaluate the proposed transferable belief model decision fusion by comparing it with the standard Bayesian decision fusion – here we report the results obtained from replacing the TBM fusion with the Bayesian fusion technique in the proposed robotic vision system.

Bayesian fusion relies on Bayes' theorem, shown in (4). Here, $\omega \in \{1, \ldots, \Omega\}$ is a class label among the set of known classes with Ω elements. Additionally, $P(.)$ represents probability and A denotes the input probability vector taken from a classifier. Accordingly, an element of the probability vector A for the class ω is in essence equivalent to $P(\omega|A)$.

$$P(\omega|A) = \frac{P(A|\omega) * P(\omega)}{P(A)} \tag{4}$$

$P(\omega)$ is the prior probability of class ω and $P(A)$ is the evidence. In addition, the extension of the Bayes' rule to the two classifiers case is shown in the following.

$$P(\omega|A,B) = \frac{P(A,B|\omega) * P(\omega)}{P(A,B)} \tag{5}$$

In the above equation, B represents the probability vector for the second classifier, while A is the probability vector of the first classifier. Presuming the two classifiers are working independently, we have:

$$\begin{aligned} P(\omega|A,B) &= \frac{P(A|\omega) * P(B|\omega) * P(\omega)}{P(A) * P(B)} \\ &= \frac{\frac{P(\omega|A) * P(A)}{P(\omega)} * \frac{P(\omega|B) * P(B)}{P(\omega)} * P(\omega)}{P(A) * P(B)} \\ &= \frac{P(\omega|A) * P(\omega|B)}{P(\omega)} \end{aligned} \tag{6}$$

We do not place different prior probabilities for the trained classes, so $P(\omega)$ in (6) is a constant equal to one over the number of the trained object categories. After removing the constant value, $P(\omega)$, from (6) we observe that the combined probability of an object category given the probability vectors of the two classifiers is associated to the product of probabilities for that class in the two probability vectors, as shown in (7).

$$P(\omega|A,B) \approx P(\omega|A) * P(\omega|B) \tag{7}$$

Therefore, in the Bayesian fusion we just multiply the corresponding probabilities from the two probability vectors obtained from the main and the secondary view classifiers to compute the probability of a class. We then normalize the fused probabilities to bring their sum back to one.

We performed the same tests with the Bayesian fusion as we had for the TBM fusion. The resulting confusion matrices for the main view only and the complete system are illustrated in Fig. 6 and Fig. 7. Similar to the proposed fusion method, the performance of the Bayesian fusion in terms of accuracy, precision, recall, and F_1 score is evaluated in Table II.

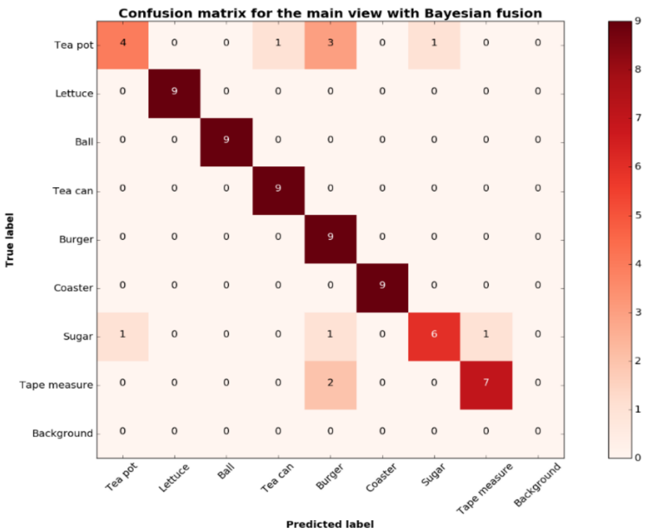

Fig. 6. Confusion matrix for the main view only with Bayesian fusion.

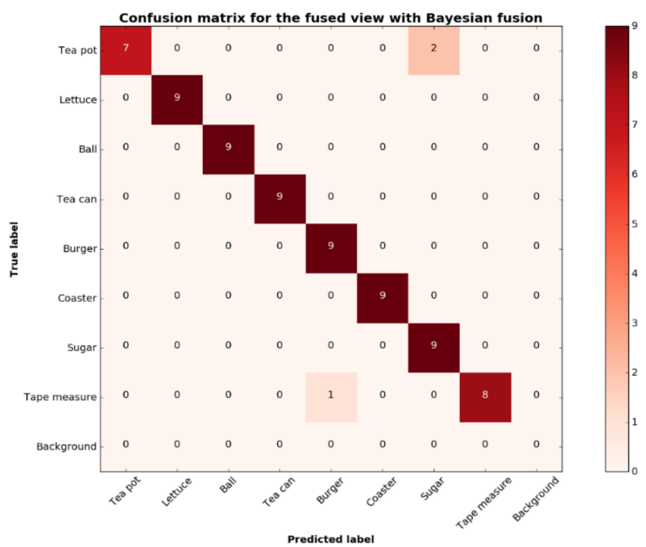

Fig. 7. Confusion matrix for the entire system after ambiguous situation management and Bayesian information fusion.

TABLE II. PERFORMANCE MEASURES OF THE MAIN VIEW ONLY AND THE ACTIVELY FUSED VIEW WITH BAYESIAN FUSION.

Performance Measure	Main Camera	Actively Fused View
Macro-Averaging Precision	0.879	0.964
Macro-Averaging Recall	0.861	0.958
Accuracy	0.861	0.958
F_1 Score	0.869	0.960

As shown in Table II, the active vision system works well with the Bayesian fusion too. There is 9.7% enhancement in accuracy. Moreover, precision, recall, and F_1 score increased 8.5%, 9.7%, and 9.1% compared to a single camera setup. However, by comparing the performance metrics after decision fusion, we realize that the proposed TBM fusion method outperforms Bayesian fusion by 1.1%, 1.7%, 1.4%, and 1.5% for precision, recall, accuracy, and F_1 score, respectively. However, an interesting point is that with the proposed novel transferable belief model decision fusion, the performance of the classifier for the single camera scenario is inferior to the system with Bayesian fusion, but it is effectively better with the active fusion added. It should be noted that single classifiers with TBM and Bayesian fusions are not identical, as the TBM classifier has an added "unknown" object category in its training and classification phases.

IV. Conclusion

In this paper an active vision, dual-camera object recognition system for robotic environments, capable of dynamic management of recognition ambiguity and a novel belief-based information fusion policy were presented. The whole vision system was implemented and tested on a PR2 robotic platform.

The reliability checking mechanism filters out ambiguous classifications made for the objects viewed by the main camera. With the assistance of an object matching procedure, the secondary camera is requested to perform a second classification routine. Finally, the two classifications are combined through an adaptation of a transferable belief model, which in turn is founded on the Dempster-Shafer evidence theory.

Experimental results reveal that the proposed method improves the object detection performance by a large margin. An increase of 19.5% in accuracy compared to the case of a traditional single camera model led to a 97.2% overall accuracy in our benchmarks. In addition, the performance of the transferable belief model fusion surpasses the standard Bayesian fusion, offering 1.4% higher accuracy.

The contributions of the presented work are highly relevant in the context of robotic systems that need to function in a dynamic environment, while recognizing objects of interest, understanding activities and inferring intentions of other agents present. We plan to extend our work to activity recognition from multiple cameras, with either redundant or missing information in the presence of occlusion, while providing capabilities for optimal camera selection and information fusion.

References

[1] S. Chen, Y. Li, and N. M. Kwok, "Active vision in robotic systems: a survey of recent developments," The International Journal of Robotics Research, vol. 30, no. 11, pp. 1343-1377, August 2011.

[2] A. O. Hero and D. Cochran, "Sensor Management: Past, Present, and Future," IEEE Sensors Journal, vol. 11, no. 12, pp. 3064-3075, December 2011.

[3] Foundations and Applications of Sensor Management, A. O. Hero, D. Castanon, D. Cochran, and K. Kastella, Eds. Springer US, 2008.

[4] D. K. Prasad, C. K. Prasath, D. Rajan, L. Rachmawati, E. Rajabally, and C. Quek, "Maritime situational awareness using adaptive multi-sensor management under hazy conditions," arXiv:1702.00754 [cs.CV], February 2017.

[5] P. Smets, "The combination of evidence in the transferable belief model," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 12, no. 5, pp. 447–458, May 1990.

[6] Classic Works of the Dempster-Shafer Theory of Belief Functions, R. R. Yager and L. Liu, Eds. Springer-Verlag Berlin Heidelberg, 2008.

[7] J. B. Yang and D. L. Xu, "Evidential reasoning rule for evidence combination," Artificial Intelligence, vol. 205, pp. 1-29, December 2013.

[8] B. Scheuermann and B. Rosenhahn, "Feature quarrels: the Dempster-Shafer evidence theory for image segmentation using a variational framework" in Computer Vision – ACCV 2010. ACCV 2010. Lecture Notes in Computer Science, vol. 6493, R. Kimmel, R. Kletter, and A. Sugimoto, Eds. Springer, Berlin, Heidelberg, 2011.

[9] D. Koks and S. Challa, "An introduction to Bayesian and Dempster-Shafer data fusion," DSTO Systems Sciences Laboratory, Edinburgh, Australia, August 2003.

[10] F. Delmotte and P. Smets, "Target identification based on the transferable belief model interpretation of Dempster-Shafer model," IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans, vol. 34, no. 4, pp. 457-471, June 2004.

[11] T. Denoeux, "A neural network classifier based on Dempster-Shafer theory," IEEE Transactions on Systems, Man, and Cybernetics—Part A: Systems and Humans, vol. 30, no. 2, March 2000.

[12] G. Powell, D. Marshall, P. Smets, B. Ristic, and S. Maskell, "Joint Tracking and Classification of Airbourne Objects using Particle Filters and the Continuous Transferable Belief Model," International Conference on Information Fusion, Florence, Italy, July 2006.

[13] J. Daniel and J. P. Lauffenburger, "Fusing navigation and vision information with the Transferable Belief Model: Application to an intelligent speed limit assistant," Information Fusion, vol. 18, pp. 62-77, July 2014.

[14] E. Ramasso, D. Pellerin, and M. Rombaut, "Belief scheduling for recognition of human action sequence," International Conference on Information Fusion, Florence, Italy, July 2006.

[15] Z. Hammal, M. Arguin, and F. Gosselin, "Comparing a novel model based on the transferable belief model with humans during the recognition of partially occluded facial expressions," Journal of Vision, vol. 9, no. 2, February 2009.

[16] D. K. Sheet, O. Kaiwartya, A. H. Abdullah, Y. Cao, A. N. Hassan, and S. Kumar, "Location information verification using transferable belief model for geographic routing in vehicular ad hoc networks," IET Intelligent Transport Systems, vol. 11, no. 2, pp. 53-60, March 2017.

[17] C. Stauffer and W. E. L. Grimson, "Adaptive background mixture models for real-time tracking," IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), Fort Collins, USA, June 1999.

[18] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), San Diego, USA, June 2005.

[19] P. Smets and R. Kennes, "The transferable belief model," Artificial Intelligence, vol. 66, no. 2, pp. 191–243, April 1994.