2010 IEEE Intelligent Vehicles Symposium
University of California, San Diego, CA, USA
June 21-24, 2010

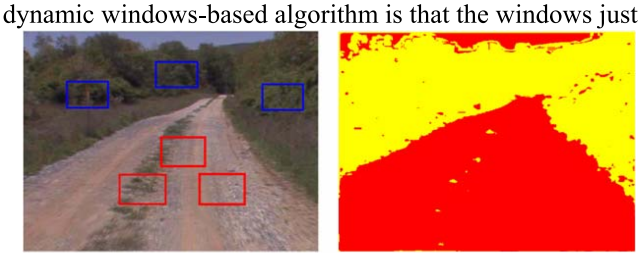
Road Detection Using Support Vector Machine based on Online Learning and Evaluation

Shengyan Zhou, Jianwei Gong, Guangming Xiong, Huiyan Chen and Karl Iagnemma

Abstract— **Road detection is an important problem with application to driver assistance systems and autonomous, self-guided vehicles. The focus of this paper is on the problem of feature extraction and classification for front-view road detection. Specifically, we propose using Support Vector Machines (SVM) for road detection and effective approach for self-supervised online learning. The proposed road detection algorithm is capable of automatically updating the training data for online training which reduces the possibility of misclassifying road and non-road classes and improves the adaptability of the road detection algorithm. The algorithm presented here can also be seen as a novel framework for self-supervised online learning in the application of classification-based road detection algorithm on intelligent vehicle.**

I. Introduction

Road detection is an important requirement for the successful development and employment of intelligent vehicles. In the past decades, the research on vision-based road detection has been an active topic and various methods have been proposed to solve this problem [1]-[5]. In principle, vision-based road detection algorithms can be categorized into three main classes: feature-based technique [1][2], model-based technique [3] and region-based technique [4]-[6]. Generally speaking, feature based technique is more accurate than any others. But it require the detected road having well-painted markings. Otherwise it can be easily interrupted by noise. Model-based technique is more robust than feature-based technique. However, most model-based approaches have some critical and strict geometrical assumptions. Most effective approaches on region-based technique also can be seen as a machine learning problem. Those approaches which allow computers to change behavior based on training set are capable of being robust to noise. In the application of intelligent vehicle, the environment always changes. In that case, a major difficult of those machine learning-based approaches is how to online train the machine to be capable of learning the new environment. J. Wang [4] trained the SVM classifier in initialization and used a voting method to determine class of the block. But it required human supervised learning in every frame. Foedisch [6] selected the training data in each frame by dynamic windows which adaptive adjusted their positions based on the result of last frame. But those algorithms are not adaptive in some situations as shown in Fig.1. In Fig.1, the sky is labeled as road because the pixels in sky region haven't been taken to determine the hyperplane of classifier. The drawback of dynamic windows-based algorithm is that the windows just



(a) (b)

Fig.1 Road detection based on Dynamic Windows Method
(a): Selecting the pixels in the windows as the training data to train the classifier. (b): Classification result by the trained classifier

adjust their positions according to the region but not the property of training data. If the windows fail to cover some training set which determines the hyperplane in real feature space, the result would be disturbed. Fig.2 shows two examples of training set and test set in feature space.

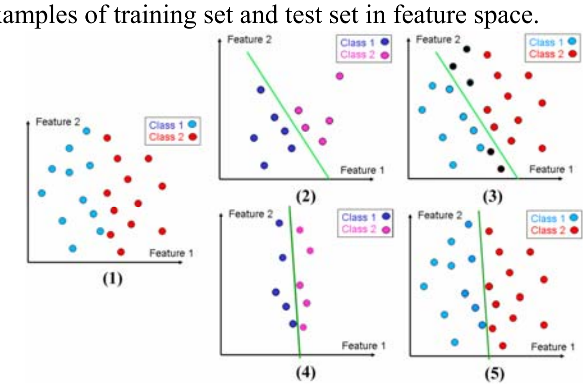


Fig.2 Feature Space and Hyperplane
Pink points and dark blue points in (2) and (4) are separately positive and negative training set. Red points and light blue points in (1), (3) and (5) are separately positive and negative test set. Black points are misclassified data.

In this paper, we will propose a novel road detection algorithm which is capable of not only doing online evaluation the quality of last frame classification result which is used for determining the online training process in next frame, but also being self-supervised online learning by automatically selecting the training set which has more contribution in determining the hyperplane. The algorithm presented in this paper can also be taken as a novel framework for online learning in the application of vision-based road detection on intelligent vehicle.

Manuscript received January 15, 2010. .This work was supported by the National Natural Science Foundation of China under grant No. 90920304.

The authors Shengyan Zhou, Jianwei Gong, Guangming Xiong and Huiyan Chen are with the Intelligent Vehicle Research Center, Beijing Institute of Technology. (e-mail: zhoushengyan2006@gmail.com).

978-1-4244-7868-2/10/$26.00 ©2010 IEEE

II. ALGORITHM OUTLINE

The proposed algorithm is composed of five components. In the first feature extraction component, a feature vector is extracted from each pixel of input image. Second, the component of dynamic training database (DTD) is filled with training set labeled by a human supervisor in initialization and updated by the new training set online. Third, the component of Classifier Parameters Computing is used to estimate the parameters in SVM classifier. The fourth SVM classifier component is in charge of training and classification which takes the training data and classifier parameters to train the SVM classifier and use the trained SVM classifier to classify image into road/non-road classes. The last component contains two stages: Morphological Operation and Online Learning Operation. The former implements connected region growing and hole filling on the classification result to determine the road region. The latter compares morphological result and classification result to evaluate the quality of current classifier, then select new training set from that comparison and update the DTD.

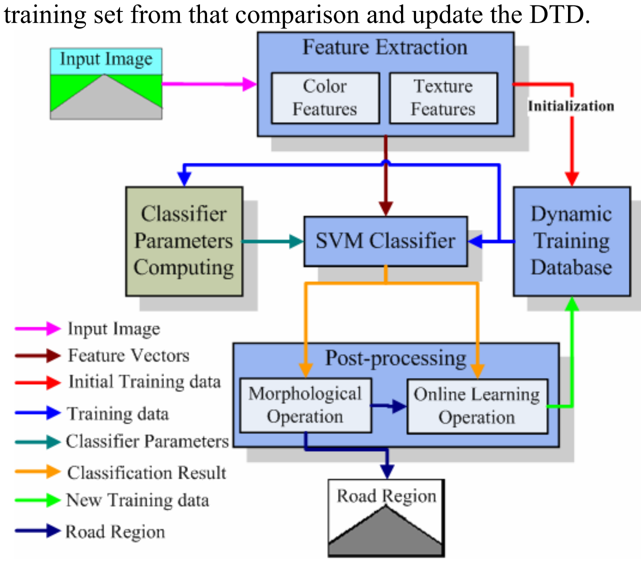


Fig.2 Algorithm flowchart

We will introduce feature extraction, initial DTD, classifier parameters computing, SVM classifier and morphological operation in Section III. Those can be seen as the basic parts of our algorithm. Then, in Section IV, we will discuss Online Learning Operation and Online DTD which are the advanced parts of our algorithm. The advanced parts make our algorithm more robust and adaptively to environment changing. In Section V, we will focus on the experimental results. A conclusion and future work are presented for further research In last Section.

III. BASIC PARTS OF ROAD DETECTION ALGORITHM

The basic parts of road detection algorithm are essential parts in the process of SVM classification task. First, the feature vector in each pixel of image is extracted. Then, two small windows are labeled by human supervisor to sample the training set to initialize DTD. Then, the training data is used for Classifier Parameters Computing and classifier training. At last, the trained classifier is implemented to classify the road/non-road classes and the morphological operation is taken to smooth the road region.

A. Feature Extraction

The visual features used in our algorithm are color features and texture features. Color data is directly available from the camera as RGB intensities. However, the illumination intensity affects all three values in a raw RGB representation which can lead to poor classification results. To reduce that effect, a HSV color representation is used. Texture is a measure of the local spatial variation in the intensity of an image. In this paper, the first five Haralick statistical features [7] are exploited. Those three color features and five texture features are combined to form an eight-element feature vector as following:

$$F_{i,j} = \left[f_{t_1(i,j)}, \dots, f_{t_5(i,j)}, f_{c_1(i,j)}, \dots, f_{c_3(i,j)} \right]$$
$$i = 1, \dots, H \quad j = 1, \dots, W \tag{1}$$

where $f_{t_n(i,j)}$ is the nth Haralick statistical feature at the point (i, j), $f_{c_n(i,j)}$ is the nth color feature at the point (i, j) in HSV color space, the H and W are the height and width of image. However, the nature of the visual features is not crucial to our algorithm. The readers can also try other features which might be better than ours.

B. Initial Dynamic training database

There are two stages to build the dynamic training database: which are initial stage and online learning stage. The latter will be discussed in next Section. In initial stage, the training data is labeled by a human supervisor. Two windows are placed on the image by supervisor to select the training data as shown in Fig. 3.


Fig.3 Dynamic Training database initialization Red windows are for positive training set selection. Blue windows are for negative training set selection.

The feature vectors of pixels in these windows are outputted into the DTD. To reduce the computation, the size of DTD we used in our algorithm is limited as 1000. If any window contains more than the maximum size, a random function is used to select samples.

C. Classifier Parameters Computing

Obviously, the relation between road/non-road classes and their feature space is nonlinear. In that case, we couldn't find a linear hyperplane to well separate those two classes in original feature space. A suggestion to use a Gaussian radial basis function (RBF) is recommended by Chang, C and Lin, C [8]. In this paper, a RBF kernel is used as the SVM kernel function. In that case, there are two classifier parameters while using this kernel: C and γ. It is not known beforehand which C and γ are the best for the road detection in a new environment. The goal of Classifier Parameters Computing is to identify good (C, γ) so that the classifier can accurately predict unknown data (i.e., testing data). Therefore, in the component of Classifier Parameters Computing, we use cross-validation over the training data from DTD and 'grid-search' on those two classifier parameters (See [8] for the detail).

The environment of intelligent vehicle is changing continuously while the vehicle is moving. However, because the process of classifier parameters computing is time and computation exhausting, it is not suitable to compute the parameters in every frame. Therefore, given the assumption that the environment does not change drastically in a few consecutive frames, we take the classifier parameters computing as a parallel process with all other components. From the experiments, the values of parameters are updated in every 8-12 frames.

D. SVM Classifier

There are two stages in the component of SVM Classifier: road detection classifier training and road detection classifier classification.

1) Road detection classifier training:

Fig.4 gives an outline of the road detection classifier training stage. As we mentioned above, we use the SVM with the RBF kernel as the road detection classifier. Given the classifier parameters by Classifier Parameters Computing (See Section III-C) and the training data by DTD, the SVM classifier determines the linear separating hyperplane with largest margin in the high-dimensional feature space (See [9] for the detail of SVM).

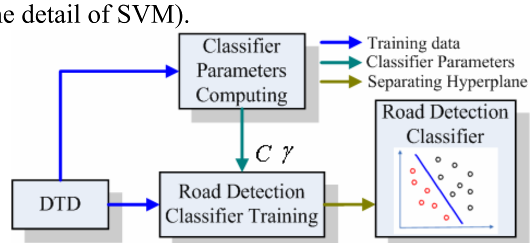


Fig.4 Road Detection Classifier Training Outline

2) Road detection classifier classification:

Fig.5 shows an outline of classification stage. We extract the feature vector on each pixel of the input image. Then, each vector is classified by the trained classifier.

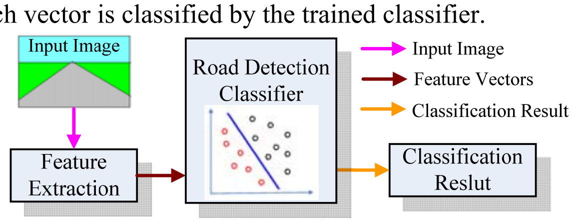


Fig.5 Road detection classifier classification Outline

in the previous stage. Each pixel is classified as either the road class or non-road class. The Fig 6-a shows the sampling windows placed by human supervisor. Then the pixels in the sampling windows are used as training data to train the road detection classifier. The Fig. 6-b gives the classification results.

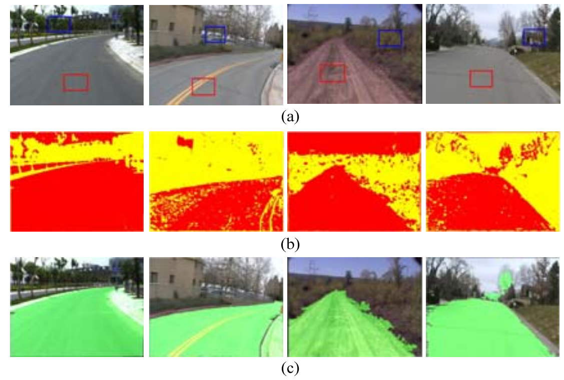


Fig.6 The process of Basic road detection algorithm
(a) The original image and sampling windows. (b) The classification results (Red is road class. Yellow is non-road class). (c) The results of morphological operation (The regions labeled as green are road regions).

E. Morphological operation

Morphological operations [10] are commonly used to understand the structure or form of an image. Morphological operations play a key role in applications such as machine vision and automatic object detection.

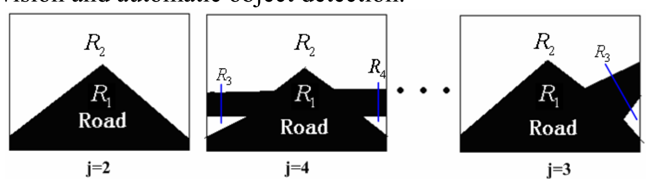


Fig.7 Simply connected road model (R_1 is simply connected road region. $R_n (n > 1)$ is non-road region.)

In this paper, the main morphological operation is flood-fill operation based on the assumption that road region is simply connected (as shown in Fig 7). The morphological operation is implemented to determine the largest connected road region and erode the holes in that connected road region. Then that largest connected road region is labeled as road region and others are labeled as non-road regions. The process of morphological operation is showed in Fig.8. And Fig.6-c shows the results after morphological operation.

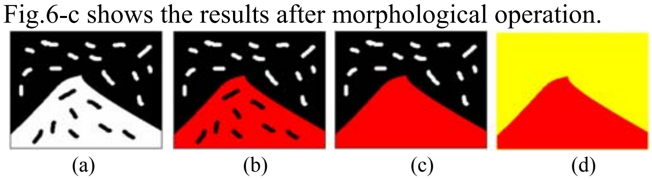


Fig.8 Morphological Operation (a): Classification result (white: represents road; black: represents non-road). (b): Largest connected road region (Red). (c): Erosion operation. (c): Morphological operation result (red is road region and yellow is non-road region)

IV. Advanced Parts of Road Detection Algorithm

The advanced parts of the road detection algorithm can be seen as the crucial parts in the framework of self-supervised online learning of classification-based road detection algorithm. We've already discussed the basic components of our algorithm. From Fig.6-b, we could see that the basic road detection algorithm misclassified many points on the image. Although morphological operation can help us to get rid of most misclassified points (see Fig.6-c). However, in the long run of driving, the misclassified points due to the inadequate learning would bring the potential dangerous situation when the misclassified points connected with the road region. (The curbs were connected with road regions in the first two images of Fig.6-c. The tree was misclassified as road in the fourth image of Fig.6-c.). The key reason of misclassification is because that the training data didn't well represent the test feature vector space due to lack of learning. That is also the main drawback of windows based learning algorithm which the windows just can find the local feature in the whole image. The task of adaptive self-supervised online learning in this paper is to find the interested training samples automatically which haven't be learnt and but are important in road/non-road classes hyperplane determination.

In this section, we will introduce the two crucial components to solve the automatic learning process in our algorithm: Online Learning Operation and DTD in online learning stage.

A. *Online Learning Operation*

1) *Evaluation function:*

Before execute the online learning operation, an evaluation function is implemented to evaluate the performance of current classification and determine whether to activate the Online Learning Operation and train the road detection classifier in next frame. This evaluation functions as shown in the following formulas are based on the assumption we mentioned in previous sections that the road region is simply connected.

$$E_{AFP} = \sum_{r=1}^{H}\sum_{c=1}^{W} V_1(r,c) \Big/ \sum_{r=1}^{H}\sum_{c=1}^{W} R_1^M(r,c) \tag{2}$$

$$E_{AFN} = \sum_{j=2}^{N}\sum_{r=1}^{H}\sum_{c=1}^{W} V_j(r,c) \Big/ \sum_{j=2}^{N}\sum_{r=1}^{H}\sum_{c=1}^{W} R_j^M(r,c) \tag{3}$$

$$E_{AF} = \sum_{j=1}^{N}\sum_{r=1}^{H}\sum_{c=1}^{W} V_j(r,c) \Big/ \sum_{j=1}^{N}\sum_{r=1}^{H}\sum_{c=1}^{W} R_j^M(r,c) \tag{4}$$

$$V_j(r,c) = \begin{cases} 1, if\,(R_j^C(r,c) \neq R_j^M(r,c)) \\ 0, if\,(R_j^C(r,c) = R_j^M(r,c)) \end{cases}$$
$$r = 1,\ldots,H; c = 1,\ldots,W; j = 1,\ldots,N \tag{5}$$

where AFP, AFN and AF refer to Assumption-based False Positive, Assumption-based False Negative and Assumption-based classification False, the N is the number of regions which are defined in Fig.7, the r and c are the row and column, the $R_j^C(r,c)$ is the value of classification result at (r,c), the $R_j^M(r,c)$ is the morphological operation result at (r,c), the j is the number of region. Apparently, $V_j(r,c)$ indicates whether $R_j^C(r,c)$ and $R_j^M(r,c)$ are belonged to the same class.

Given the values of E_{AFP}, E_{AFN} and E_{AF}, we have three thresholds of $T_{E_{AFP}}$, $T_{E_{AFN}}$ and $T_{E_{AF}}$ to tune. If the value of evaluation is larger than its threshold, the retraining process is implemented.

2) *Online Learning Operation:*

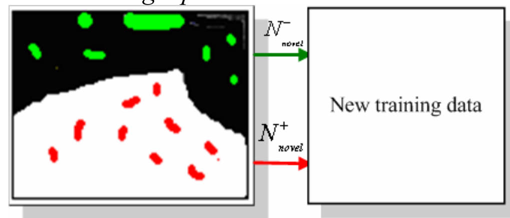


Fig.9 Interested points
Green points are in the non-road region and misclassified as road class. . Red points are in the road region and misclassified as non-road class.

The process of online learning operation in our algorithm is to acquire feature vectors at the interested points. Given that assumption that road region is simply connected, the points classified as road lying in the regions of non-road can be seen as interested points and labeled as negative samples (non-road samples), vice versa. We label those points as new training data (shown in Fig.9). In practice of online learning process, there is one thing should be paid attention which is one doesn't know exactly where is the real boundary of our road region. What we can get is the edge of road region in the morphological operation result. In order to reduce the possibility of mislabeled training data near the road boundary, a threshold M is set as the width of margin near boundary which segments the road and non-road region in morphological operation result as shown in Fig.10. In our experiment, we set M as 40 pixels width.

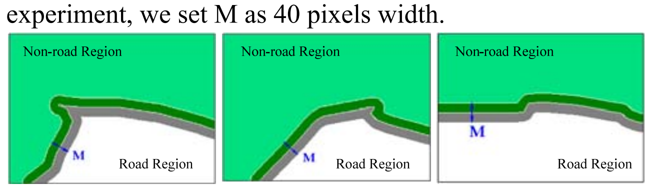


Fig.10 Road region boundary in Morphological Operation Result

B. *DTD in online learning stage:*

In order to reduce the computation of training SVM classifier, the DTD is limited to contain just 1000 training positive and 1000 negative samples labeled by human supervisor in the initial stage. Usually, the number of new training data in each class is more than 1000. So we randomly choose T new training samples in each class and also abandon T old training samples in each class in DTD. T is a threshold to determine the learning speed. Too large value of T will lead to over learning on new misclassified data while too small value of T will make our algorithm low adaptability and robust to environment changing. From the viewpoints of our many experimental results, It is best to set T as 1/10 of the size of DTD (T is 100 in this paper).

V. RESULT

Road detection by the proposed algorithm worked well in a variety of test conditions. First, some results are shown in Section A to demonstrate how the adaptive online learning process to acquire training data. Then, a comparison between our algorithm and dynamic windows approach is taken in Section B. At last, the results in sequences which were taken in different conditions are shown in Section C. The system's performance in the following results is compared with manually annotated frames to measure the accuracy.

A. Comparison between offline learning and online learning

The results in four different conditions shown in the Fig.11 give a strong proof of the performance of self-supervised online learning process in our algorithm. In each experiment, two small sampling windows are selected on the image to initialize the DTD. This can be seen as offline learning. Then, our algorithm restudies from the poor classification result, then retrains the classifier and reclassifies the road image.

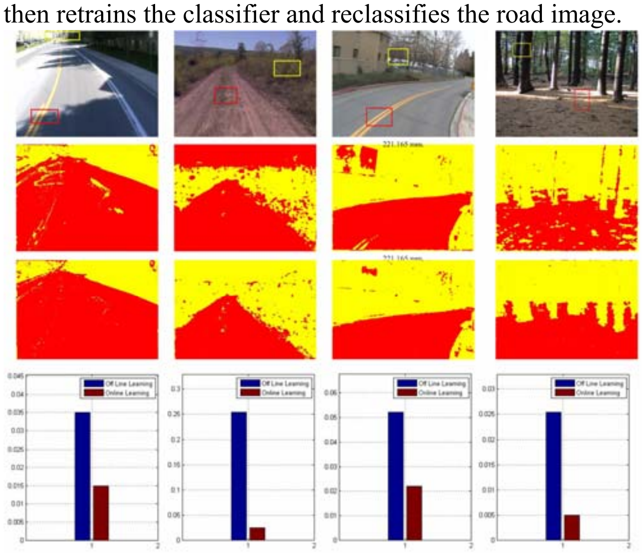


Fig.11 Comparison results between Offline learning and online learning
The first row shows original images. The second row shows the classification results of offline learning. The third row shows the classification results of online learning. The fourth row shows the classification error rates.

The first column shows the results of different learning processes in the shadow road situation. From the offline learning results, some road marks and tree shadows are misclassified as non-road. After online learning process, the number of misclassified points is getting smaller.

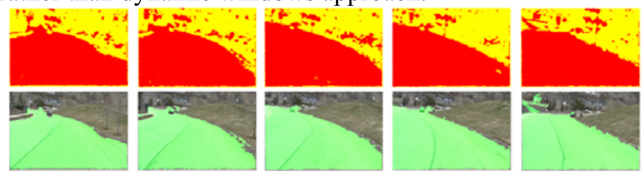
The second column shows the results in the unstructured road situation. Almost all the points in the sky are misclassified as road in the result of offline learning. With the process of learning online, the results become more accuracy.

From the third column we can see that the wall of left building is misclassified as road because lack of learning the samples in offline sampling windows. After our self supervised online learning process, the computer automatically learns the samples it haven't learnt and generates the more accurate result.

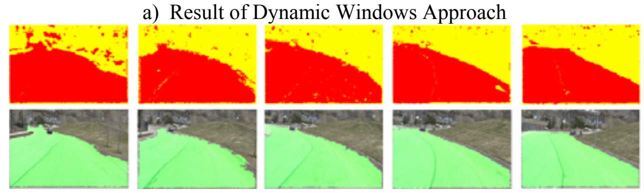
It is more complicated for road detection in the forest. However, from the results in the fourth column we could see that the points are well classified and the classification error becomes smaller after our online learning process.

B. Comparison with dynamic windows approach in consecutive frames

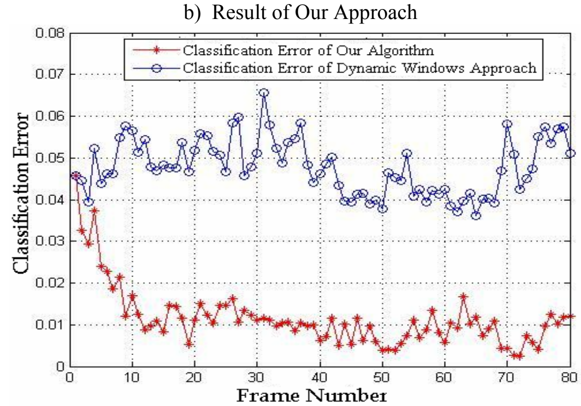
We compare our algorithm with dynamic windows approach [6] in consecutive video frames. Some frames of classification results and morphological operation results are shown in Fig.12-a, Fig.12-b. The classification result in each frame is compared with hand labeled ground true to acquire the classification error (See Fig.12-c). From the results, we arrive to conclusion that our algorithm is more accuracy rather than dynamic windows approach.



a) Result of Dynamic Windows Approach



b) Result of Our Approach



c) Classification Error

Fig.12 Comparison of Two Road Detection Approaches

C. Results of our algorithm in consecutive frames

Our road detection algorithm performs better on detection rate and more robust to environment changing.

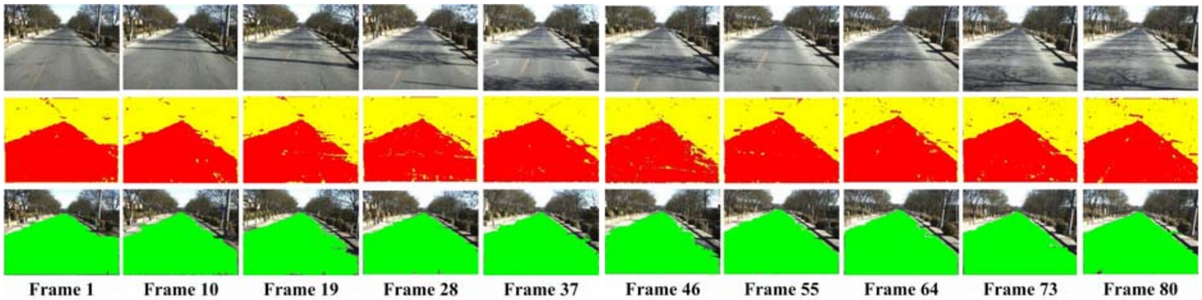
Fig.13 shows the results of our algorithm in consecutive frames in different situations. From 7th and 61st frames in Fig.13-c and 37th and 69th in Fig.13-d we can see, the classification error rates become larger sharply because of environment changing, our proposed algorithm make reduce the error by online learning in the next frames after online learning.

VI. CONCLUSION AND FUTURE WORK

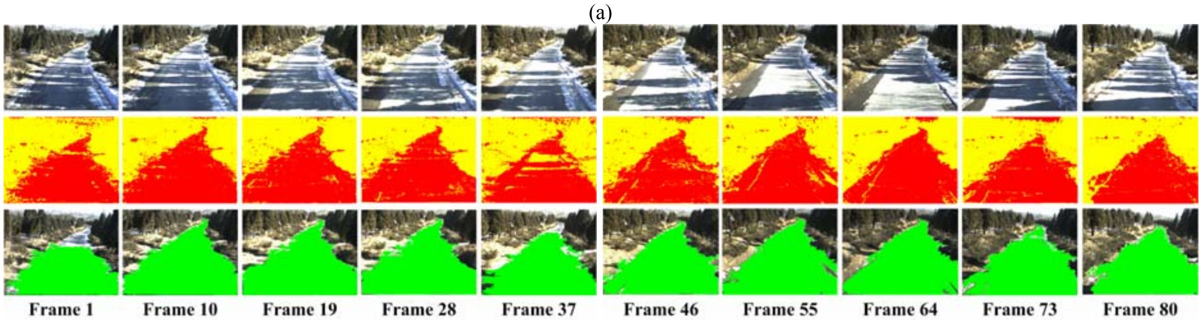
In this paper, we introduced our algorithm which is more adaptive and robust in online continuous learning. In the future, we will combine a LIDAR sensor with the online learning process to instruct the training data acquisition. In order to speed up the learning convergence process, we also will find a method to abandon the old training data which are less important in classifier hyperplane determination in DTD in stead of randomly discarding.

ACKNOWLEDGMENT

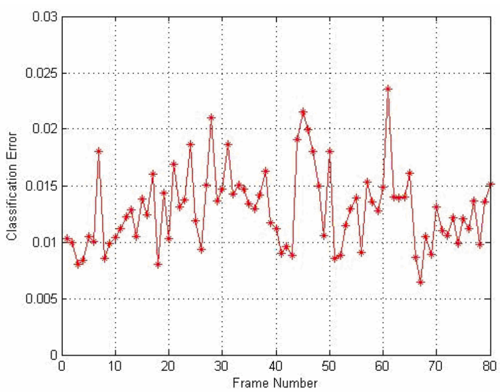
The authors would like to thank Yanhua Jiang from Beijing



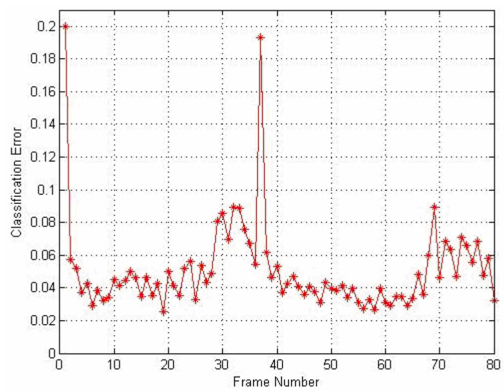
(a)



(b)



(c)



(d)

Fig.13 the experimental results of our algorithm in different situations
(a). Original frames (first row), classification results (second row) and morphological results (third row) in shadow road.
(b). Original frames (first row), classification results (second row) and morphological results (third row) in shadow and snow road.
(c). Classification error rate in shadow road. (d). Classification error rate in shadow and snow road.

Institute of Technology and Takayuki Nishihata from Massachusetts Institute of Technology for their help with the data collection and experiment.

References

[1] Tsai, L.W., Hsieh, J.W., Chuang, C.H., Fan, K.C.: Lane detection using directional random walks. In: IEEE Intelligent Vehicles Symposium. 2008.
[2] Li, Q., Zheng, N., Cheng, H.: A prototype autonomous vehicle and its
[3] Y. Wang, E. Teoh, and D. Shen, Lane detection and tracking using B-snake, Image Vision Computing. vol. 22, no. 4, pp. 269–280, Apr. 2004.
[4] Jian Wang; Zhong Ji; Yu-Ting Su . Unstructured road detection using hybrid features. Machine Learning and Cybernetics, 2009 International Conference on Volume 1, 12-15 July 2009 Page(s):482 – 486.
[5] Sha Yun; Zhang Guo-ying; Yang Yong, A Road Detection Algorithm by Boosting Using Feature Combination. Intelligent Vehicles Symposium, 2007.
[6] Foedisch, M; Takeuchi, A, Adaptive road detection through continuous environment learning. Applied Imagery Pattern Recognition Workshop, 2004.
[7] Haralick, R.M. Statistical and Structural Approaches to Texture. Proceedings of the IEEE, 67, pp. 786-804, 1979.
[8] Chang, C., & Lin, C. (2008, October 29). LIBSVM: a Library for Support Vector Machines. Retrieved November 20, 2008, from http://www.csie.ntu.edu.tw/~cjlin/papers/libsvm.pdf.
[9] Christopher J. C. Burges. A Tutorial on Support Vector Machines for Pattern Recognition. Data Mining and Knowledge Discovery 2:121–167, 1998.
[10] Soille, P., Morphological Image Analysis: Principles and Applications. Springer-Verlag, 1999, pp. 173-174.