2012 IEEE International Conference on Systems, Man, and Cybernetics
October 14-17, 2012, COEX, Seoul, Korea

# Recognizing Temporary Changes on Highways for Reliable Autonomous Driving

Young-Woo Seo and David Wettergreen
The Robotics Institute
Carnegie Mellon University
5000 Forbes Ave, Pittsburgh, PA 15213
{young-woo.seo,dsw}@ri.cmu.edu

Wende Zhang
Electrical and Controls Integration Lab
R&D General Motors
30500 Mound Rd, Warren, MI 48090
wende.zhang@gm.com

***Abstract*—In order to be deployed in real-world driving environments, autonomous vehicles must be able to recognize and respond to exceptional road conditions, such as highway workzones, because such unusual events can alter previously known traffic rules and road geometry. In this paper, we present a set of computer vision methods which recognize the bounds of a highway workzone and temporary changes in highway driving environments through recognition of workzone signs. Our approach filters out irrelevant image regions, localizes potential sign image regions using a learned color model, and recognizes signs through classification. Performance of individual unit tests is promising; still, it is unrealistic to expect perfect performance in sign recognition. Performance errors with individual modules in sign recognition will cause our system to misread temporary highway changes. To handle potential recognition errors, our method utilizes the temporal redundancy of sign occurrences and their corresponding classification decisions. Through testing, using video data recorded under various weather conditions, our approach was able to perfectly identify the boundaries of workzones and robustly detect a majority of driving condition changes.**

***Index Terms*—Highway Workzone Recognition, Computer Vision, Intelligent Transportation Systems**

## I. INTRODUCTION

Recent technological advances in development of self-driving cars have increased the potential of their realization in the near-future. However, many technical problems remain untackled in order for self-driving vehicles to be deployed in real driving environments. One such challenge is successful development of the capability to respond to unexpected occurrences on the road. For example, suppose an autonomous vehicle is guided by a map of driving environments; although it is possible to provide the vehicle a map with detailed information about road geometry and traffic rules in advance, it is not possible to describe unexpected occurrences a priori, such as traffic accidents or road work.[1] The vehicle must be able to effectively handle such events as they can lead to temporary changes in road conditions. For example, suppose that the road lane on which a vehicle is driving is laterally shifted ahead as a result of a road work, whereas the road is depicted on the map as following a straight path. What if the vehicle's braking distance is longer than its sensing horizon?

To effectively handle unexpected events on a highway, an autonomous vehicle should first be able to recognize them. To tackle problem of providing a vehicle with such perception capability, this paper presents a collection of computer vision methods that identify the bounds of a workzone, e.g., the beginning/end of a workzone, and recognizes temporary changes to highway driving conditions, e.g., a decrease in speed or blockage of a lane, through recognition of workzone traffic signs in perspective images. Such detailed information about a highway workzone would help a robotic vehicle, and potentially help human drivers or driving-assistance systems, properly respond to unexpected events on a highway, in turn leading to safe and reliable autonomous and manual highway driving.

Workzone signs are highly constrained by governmental regulations in terms of location and appearance [10]. However, such constraints do not make it easy to recognize signs in images, because high variation still exists in each sign's image appearance. Under perspective imaging, projection of a 3-dimensional traffic sign onto a 2-dimensional image plane distorts most of the sign's geometric properties, such as its angles, distance, and ratios of angles [9]. In addition, the line of sight between a sign and a camera perceptually and computationally changes the color of a workzone sign from that of the sign template. This leads to the problem of intra-class appearance variation, in which the appearance of the same workzone sign varies based on the conditions of the image acquisition process.

To cope with such challenges in recognition of workzone signs, we utilize and improve some of the well-established machine learning techniques: firstly, our approach learns variations of a color in workzone sign images to perform a pixel-wise binary color classification; secondly, our approach identifies blobs to localize sign image regions; lastly, our approach represents a cropped image in a homogeneous feature space in order to reduce variation of geometric shapes to gain more accurate sign classification. Realistically, any sign recognition system will make errors in incorrectly classifying signs or will even miss some signs. These errors would lead to our approach misreading temporary changes on a highway. To address these potential errors, we devise two algorithms. The first algorithm makes use of temporal redundancy of sign occurrences and their corresponding classification decisions, in order to reduce false positives. The second propagate confidence values of classification decisions toward future in order to reduce the impact of false negatives.

Our contributions include successful application of well-established techniques to the problem of highway workzone recognition and development of two new methods for handling sign recognition errors.

[1] In the U.S., traffic authorities advertise workzone plans in advance, but the information is not precisely described in terms of the bounds and changes of road conditions.

978-1-4673-1714-6/12/$31.00 ©2012 IEEE

## II. Related Work

In this section, our approach is further described in comparison with previous work in the area of traffic sign recognition. For any traffic sign recognition method that focuses on vision sensors, an initial requirement is to locate potential sign image regions from an input image. Some systems, including ours use color information to localize signs. In addition, for any sign recognition system, which utilizes color, it is necessary to find an optimal range of target color values because the actual values of the target color vary based on image acquisition processes. These threshold values are often obtained empirically by repeating manual surveys of pixel color values from sample sign images [5], [6]. Because of its simple implementation, such a manual process is attractive, yet tends to be error-prone. By contrast, our approach automatically obtains the limits of optimal color-values through binary pixel-classifier training.

Another dominant approach for traffic sign detection is to use of sign shapes. Some researchers use the geometric property of sign shapes, such as equiangularity, in order to locate the centroids of traffic signs [2]. This approach is intrinsically error-prone because it relies on a geometric property, which is not preserved under perspective imaging, and also because it assumes high contrast in image intensity, which is not easy to acquire from real-world image acquisition. An alternative approach for utilizing the geometric properties of signs is to locate parts (e.g., corners or edges) of a traffic sign and to combine the results of these partial detections. For example, in order to identify potential sign image locations, some researchers have used Haar-like features [13], [14], [15] and their variants, such as a set of rectangular features in particular color channels [1] and non-symmetric dissociated dipoles [3] or a variant of the histograms of oriented gradients (HOG) [12]. This learning approach demonstrated successful performance only when a large number of manually labeled data was available to train the detector on multiple-scales for a long period of time.

Some of the existing methods [12], [14] have demonstrated very impressive recognition results in their experimental setups, e.g., a detection rate of more than 98.8%. However, in general, it is unrealistic to expect perfect performance in sign recognition. Most traffic sign recognition methods may miss a workzone sign or may also incorrectly classify a sign image in a stream of perspective images. Such inevitable errors would cause any sign recognition method to misunderstand the traffic rules and road geometry. To cope with such potential sign recognition errors, we utilize our sign classification output in a twofold manner: first, we propagate classification confidence values toward future to reduce the impact of false negatives; second, we investigate previous classification decisions for the same sign to reduce the number of false positives. To the best of our knowledge, we have not seen such methods, particularly in handling potential sign recognition errors for recognizing temporary highway changes, in this field.

## III. Recognizing Temporary Highway Changes

Our task in this paper is to reliably detect and accurately classify relevant workzone signs through use of a perspective video, in order to acquire detailed information about a highway workzone, such as where a workzone begins/ends and how the work changes the driving conditions. In what follows, we detail our approach for detection and classification of workzone signs and explain our approach for dealing with possible sign recognition errors.

### A. Workzone Sign Detection

Although it is obvious that the color of a workzone sign is orange, it is challenging to correctly identify orange pixels in a given image because of possible variation of the orange color. To effectively deal with such variation, we formulate the learning of the orange color variation as a binary color classification using the Bayesian inference framework.

$$P(\text{sign}|\mathbf{X}) = \eta P(\mathbf{X}|\text{sign})P(\text{sign}) \tag{1}$$

where $\mathbf{X}$ is an image comprised of $|\text{width} \times \text{height}|$ number of $m$-dimensional pixels, $\mathbf{x}_j \in \mathbf{X}$ and $\eta$ is a normalizer for the posterior distribution. In particular, $\mathbf{x}_j$ is a 2-dimensional color vector of which components include hue and saturation values. The posterior probability, $P(\text{sign}|\mathbf{X})$, which assigns a value to the probability that individual pixels are part of workzone signs, is computed by multiplying the likelihood function, $P(\mathbf{X}|\text{sign})$, and the prior probability distribution, $P(\text{sign})$, of traffic sign locations found in image frames. We obtain the prior probability density of workzone sign locations from the ground truth of our training data, which is comprised of several hours of highway workzone video footage and manual annotations. Figure 1 shows the density of workzone traffic

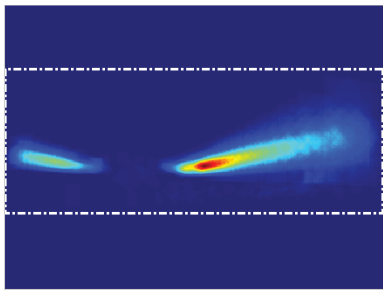


Fig. 1. A heat-image (640×480) showing workzone sign locations where the color closest to red represents the highest density. The (white) rectangular dashed-line defines the region of interest (ROI) for our workzone sign detection. Viewed best in color.

sign location, which is obtained by projection all of the ground truth bounding boxes onto an image, and is used as the prior for workzone sign locations.

We use AdaBoost [7] to learn the likelihood function, $P(\mathbf{X}|\text{sign})$, of a given pixel as a part of a workzone sign. The training data is comprised of a set of workzone images, some of which were downloaded from the web while the rest of which were obtained from our workzone video images. Using this data, we train a set of weak-learners and their weights.

$$P(\mathbf{X}|\text{sign}) = \text{mode}\left(\cup_j g(f(\mathbf{x}_j|sign))\right) \tag{2}$$
$$\text{where,} \quad f(\mathbf{x}_j|\text{sign}) = \sum_{i=1}^{H} \alpha_i h_i(\mathbf{x}_j)$$

where $H$ is the number of weak learners, $h_i$ represents a weak learner implemented by a decision stump and $\alpha_i$ is its weight. We use logistic regression to implement the function $g$ to convert the binary output of AdaBoost into a probabilistic output [8], $g(f(\cdot)) = \frac{\exp(f(\cdot))}{\exp(f(\cdot))+\exp(-f(\cdot))}$.

For a given image, our color classifier evaluates pixels within the ROI, as presented in the Figure 1, and assigns a probability for whether individual pixels are part of an orange workzone sign. Our sign detector runs a connected-component grouping algorithm to identify orange blobs and generates up to $k$ bounding boxes as candidates for a workzone sign. The detector then removes any of bounding boxes with radii[2] smaller or larger than the predefined thresholds and uses non-maximal suppression to select the largest bounding box. The confidence value of the selected bounding box is computed using the mode of the confidence values assigned to all pixels within the bounding box. To detect a regulatory, rectangular workzone sign which includes two colors (orange at the top and white at the bottom), we implement a heuristic for investigating the aspect ratio of a bounding box, in order to extend the height of the bounding box.

To evaluate the performance of our sign detector, we use the performance metrics used for PASCAL object detection challenges [11]. An output bounding box, $o_i$, is considered a potential match to the ground truth bounding box, $g_i$, in a given image frame, $i$, if their area of overlap is greater than a predefined value, $\tau < \frac{Area(o_i \cap g_i)}{Area(o_i \cup g_i)}$. When a potential match is found in a given image, sign detection performance can be further analyzed by measuring the following performance metrics: *precision*=$\frac{Area(o_i \cap g_i)}{Area(o_i)}$ and *recall*=$\frac{Area(o_i \cap g_i)}{Area(g_i)}$.

| | Color-based | | Shaped-based [2] | |
|---|---|---|---|---|
| | *Warning* | *Regulatory* | *Warning* | *Regulatory* |
| Precision | 0.951 | 0.954 | 0.487 | 0.535 |
| Recall | 0.928 | 0.903 | 0.497 | 0.662 |

TABLE I
FOR THIS TEST, WE USED 103 WORKZONE IMAGES, INCLUDING 55 WARNING (OR DIAMOND-SHAPED) SIGNS AND 41 REGULATORY (RECTANGULAR-SHAPED) WORKZONE SIGNS. WE SET $\tau$ 0.5 AS THE VALUE FOR A POTENTIAL MATCH.

Table I presents macro-averages of precision and recall where a macro-average is computed by averaging individual measurements over testing images. We compared the performance of our color-based sign detection approach to Loy and Barnes' method which utilizes geometric shapes of signs to achieve sign detection [2]. Loy and Barnes' method did not perform well for our data because most of our testing sign images have low contrast in image intensity. Figure 2 shows some examples of the sign detection output obtained by our method.

### B. Workzone Sign Classification

An image sub-region localized as a potential workzone sign is given as input for our sign classification module. Our task in this paper is to recognize the bounds of a workzone and temporary changes to highways by classifying workzone signs. In this regard we chose 9 workzone signs as reliable indicators of workzone bounds and driving condition changes and assigned all remaining workzone signs to another class. Table II shows the number of sign image examples used for each target class.

[2]The radius of a polygon is measured by computing the Euclidean distance between a point on a side (or edge) of a polygon and the centroid of the polygon.


Fig. 2. Some sign detection output images are shown. The thick (yellow) rectangles outlining the signs represent the overlap between the ground truth rectangle (green) and the detection output (red). In the top row, from the left, the images represent examples of W20-1, R22-1, G20-2, W21-19, and R2-2-2 workzone signs. The bottom row includes examples of W1-4, W1-4L, W1-4R, W4-2R, and W4-2L. We include these identifiers of workzone signs for completeness and also later represent the sign target classes. Viewed best in color.

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|
| 86 | 55 | 35 | 75 | 36 | 26 | 43 | 19 | 17 | 92 |

TABLE II
THE NUMBER OF SIGN IMAGES FOR EACH TARGET CLASS. FROM THE LEFT, TARGET CLASSES ARE NUMBERED 1 TO 10, CORRESPONDING TO W20-1, R22-1, G20-2, R2-2-2, W1-4, W1-4L, W1-4R, W4-2L, W20-5L, AND ALL OTHER WORKZONE SIGNS.

It is challenging to correctly classify sign images because of the variation between each sign's appearance in an image. To reduce such variation, we use a log-polar transform, which is a method used for transforming an image from a Cartesian coordinate into an image in a log-polar coordinate [4]. This transform is effective in reducing the variation of sign shape and text because it densely samples image intensity values near the center of a sign image where the difference between signs images is relatively small, and then sparsely collects values from sign image boundaries where the geometric distortions are large. For a given image sub-region localized by our sign detector, our sign classifier first normalizes the image to reduce intensity variation [15], then converts the cropped image into a log-polar image based on two parameters, distance between sampling bins and the centroid, represented by $\rho$, and the rotation angles of sampling bins in counterclockwise, represented by $\theta$, finally produces a column vector, $|\rho \times \theta| \times 1$ (e.g., a combination of the parameters, $\rho$=32 and $\theta$=32, produces 32×32 log-polar image, resulting in a column-vector, 1024×1).

Even with such an effective feature representation method, any conventional supervised classifier might still fail to generalize the target function in a high-dimensionality space (e.g., 1024) because of the small number examples (e.g., less than a hundred) for each of the target sign classes. To handle the curse of this dimensionality problem, we further reduce the original dimension of the log-polar image using principal component analysis (PCA). We then build an eigen-space from the labeled training data and project a testing sign image in the log-polar coordinate space onto this eigen-space. The eigen-space is comprised of $k$ eigen bases, all of which represent more than 95% of the total variance in the log-polar image data matrix. Empirically we found that 10 eigen bases

achieved the best performance. Table III shows the results of our workzone sign classification. The hyper-parameters of these classification methods were chosen through cross-validation.[3] Due to the random selection of our training data, we averaged our results over 5 separate runs for each method. To measure the effectiveness of our sign image representation, we compared the results with another representation method, which scales raw-intensity sign images into an image of the same size (e.g.,100×100), converts it into a multi-dimensional vector (e.g., 10000×1) and then reduces the dimensions using the precomputed eigen bases. The last two rows of Table III show the performance of three classifiers that use a raw-intensity image representation. The table demonstrates that log-polar sign image representation helps classifiers achieve better classification results. All evaluation metrics indicate that SVM outperforms the other two methods.

| | *SVM* | *LDA* | *kNN* |
|---|---|---|---|
| Precision | **0.965/0.012** | 0.856/0.035 | 0.285/0.027 |
| Recall | **0.957/0.016** | 0.854/0.040 | 0.387/0.007 |
| Precision | 0.896/0.035 | 0.756/0.030 | 0.252/0.015 |
| Recall | 0.841/0.028 | 0.742/0.030 | 0.377/0.015 |

TABLE III
PERFORMANCE OF THREE DIFFERENT SIGN CLASSIFICATION METHODS MEASURED BY STANDARD METRICS. EACH CELL IN THE TABLE SHOWS THE MEAN AND STANDARD DEVIATION.

### C. Handling Workzone Sign Recognition Errors

Although the previous two sections demonstrated promising results for our sign detector and sign classifier, it is realistic to conjecture that our approach makes mistakes in recognizing some workzone signs. When either a miss or incorrect classification occurs, our methods might fail to acquire detailed information about a highway workzone.

To handle such potential sign recognition errors, we devise two algorithms that utilize the sequence of previous sign classifications. These methods rely on the accuracy of our sign recognition method, which is able to accurately recognize the majority of the target class signs.

Our sign classifier produces a sign classification decision and its confidence value in cases where the sign detector produces a bounding box as a potential sign image. These confidence values represent the level of confidence in our approach's representation in terms of determining whether the cropped images are instances of target workzone signs. We can thus use the magnitude of the confidence value to infer whether our vehicle is driving in a workzone. However, a problem with using these values directly is that a sparsity of confidence values exists, as we cannot obtain such evidence from workzone regions; where no workzone signs are posted or from true workzone signs that are misclassified as other objects. The underlying aim of our algorithm is to propagate confidence values over time in order to hold non-zero values while driving in a workzone, even when the system does not have direct observation of a workzone sign or miss any workzone signs. While spreading these values, the effect of propagation should decay over time, in order to prevent an over-estimation of the true state.

To implement our idea for driving region inference, we use Gaussian smoothing of the confidence values over a specific time domain. Classification confidence at the $i$th time step, $\delta_i$, is propagated to adjacent time periods as far as the value of $\sigma$.

$$[\delta_i * w_j]_{j=-\sigma,\ldots,-1,1,..,\sigma}, w_j = \exp\left(-\frac{i-j}{2\sigma^2}\right) \quad (3)$$

Assuming that the driving speed is 50 mph and the frame rate is 15 per second, an image frame in a video represents a distance of 1.4 meters of driving. In this case, if we set $\sigma$ to 150 (or 150 image frames), the confidence value will be propagated over 210 meters, both toward future and past time steps. Note that, although it is unnecessary to propagate confidence values toward the past, we propagated them in both directions for convenient implementation, without paying any extra computational cost. The choice of $\sigma$ is critical for production of a smooth inference curve. If $\sigma$ is not optimal, either discontinuity or inflation will appear in the resulting curve. We define the value of $\sigma$ based on the rough estimation of distance between workzone signs. The function for estimating a likelihood of driving in a workzone is then computed by adding the current classification confidence value, $\delta_i$, and the propagated confidence values accumulated at time step, $i$, obtained from the neighboring time frames, $score_i = \delta_i + \hat{\delta}_i$, where $\hat{\delta}_i$ represents the confidence values propagated to the time step, $i$.

In a workzone video, a workzone sign appears multiple times before it disappears from the camera's field of view. Our approach utilizes such temporal redundancy of sign occurrences to improve classification accuracy, particularly reducing the rate of false positive. Specifically, when the system makes a classification decision, it refers to previous classification outputs.

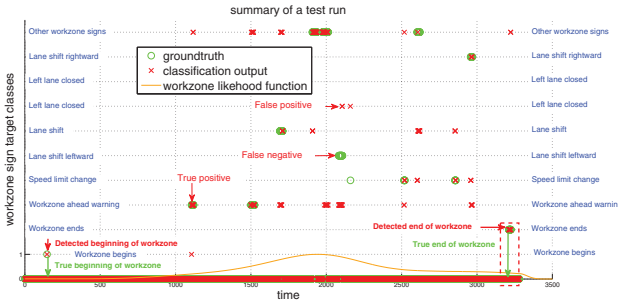
$$y_t(\mathbf{o}_t) = \arg\max_c \left\{ h_t(\mathbf{o}_t, c) + \sum_{l=1}^{T} \gamma^l h_{t-l}(\mathbf{o}_{t-l}, c) \right\} \quad (4)$$

where $h_t(\mathbf{o}, c)$ and $h_{t-l}(\mathbf{o}, c)$ represent the classification outputs for an image sub-region, $\mathbf{o}$, at time steps, $t$ and $t-l$, for the class, $c$, and $\gamma$ is a discounting factor that determines contributions of previous decisions to the current classification decision. Note the sign detector applies non-maximum suppression in order to ensure that image sub-regions, $\mathbf{o}_t$ and $\mathbf{o}_{t-l}$, represent the same object in different scales. By investigating previous classification decisions on the same sign in different scales, our approach offers the opportunity to alter its current classification decision, which has $1 - 0.965$ chance of producing a false positive based on Table III. When a classification decision is made, the system propagates its classification confidence over adjacent time frames based on $\sigma$.

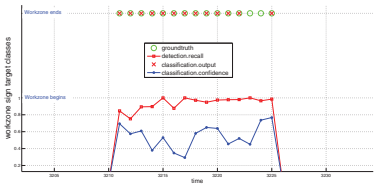
## IV. EXPERIMENTS

This section details experiments conducted to investigate the robustness and the reliability of our workzone recognition method, using images acquired under various illumination conditions.

[3]For SVM, we used the LIBSVM http://www.csie.ntu.edu.tw/~cjlin/libsvm/. We found that a SVM with an RBF kernel ($\sigma = 0.125$) worked best and used an one-against-one scheme for multi-class classification. Linear Discriminant Analysis (LDA) used the weight vector that best performed against the validation set for testing. Our $k$NN implementation worked best when it used the top 10 closest neighbors in terms of Euclidean distance in the eigenspace.




(a) A graph showing the summary of the test results of video data, $D$. For this test we set $\sigma$ as 800 (about 1,120 meters).




(b) This subfigure magnifies the dashed rectangle in Figure 3(a) and shows two additional information.

Fig. 3. Results of a highway workzone recognition test. This figure is best viewed in color.

We collected several hours of video footage of various highway driving experiences and prepared 5 videos out of these as testing data, where each of the five videos showed a vehicle's perspective when driving on a normal highway, passing a workzone, and driving on another normal highway. Each of these videos was decompressed into a set of images. The top row of Table IV gives detailed information about the test video data. Three numbers in each cell correspond to the sum of images, the number of images workzone signs appearing on, and the number of images containing other traffic signs. These video data were acquired under various weather conditions. The first two videos (A&B) were recorded in winter with snow accumulation in the background, while the following two videos, (C&D) were obtained in spring, under fairly gentle illumination conditions (i.e., sunny and clear skies), and the last was recorded on a rainy day in spring.

For each of the test video data, a stream of images was given to our system, which was required to localize signs, if any were present, and classify them, if necessary. For the sign detector and classifier, we used the best-performing learners as described in the previous sections. We empirically found that the temporal smoothing worked best when $\gamma$ was 0.9 and $T$ was 5. Based on the observation that the average driving speed during video acquisition was 50 mph, we set $\sigma$ to a value in the range of 350 (490 meters) to 800 (1,120 meters) for Gaussian smoothing, based on the highway scale and the maximum inter-distance between signs, as described in [10]. The performance of individual modules in workzone sign recognition is somewhat similar to the ones reported earlier and omitted here due to space constraints.

Figure 3 details one of the experimental results, i.e., video data $D$, where the $x$-axis represents the number of image frames organized by time and the $y$-axis represents the target class labels. An instance of sign recognition was counted as correct whenever a (green) circle, representing a ground truth, overlapped with a (red) "x," representing the output of the sign classifier. Figure 3(b) magnifies the dashed rectangle in Figure 3(a) where the "end-of-workzone" signs appeared 12 times before they disappeared from the camera's field of view. Two additional pieces of information about the recall of sign detection are depicted (in red), which are not available to our system during the testing phase, and the confidence values of sign classification (in blue). While it is true that the dimension of a sign in an image gets bigger as it appears closer to bounds of camera's viewpoint, because of motion blur and unavoidable recognition errors, the larger sign dimensions do not always lead to performance improvement. In our case, the values of detection recall and classification confidence increased as the sign became bigger, however the fluctuation of these numbers were observed to be a result of recognition error. Although two out of the last five classification decisions were incorrect, the discounted sum of the confidence values concluded that the system recognized the "end-of-workzone" sign and turned the flag off, indicating that our vehicle was leaving a highway workzone. The (orange) curve represents the estimated function value of the likelihood of driving in a workzone. As shown, the values of this function are greater than zero within a workzone. Although the estimated curve slightly overestimated the actual workzone bounds, this function can be used to inform our vehicle of the likelihood of driving in a workzone, even when our approach misses signs that indicate the beginning or end of a workzone.

Table IV summarizes the experimental results in terms of recognition accuracy of temporary changes in driving conditions. The first two rows show the accuracy of workzone bound recognition and the remaining five rows show that of driving condition change recognition. The four symbols in each cell correspond to success ($\circ$) or failure ($\times$) of particular events/number of corresponding sign images/precision/recall of workzone sign classification, respectively. Our approach demonstrated excellent performance in identification of workzone bounds. For example, for the test video data $E$, there are 3 images that contain "workzone-begin" signs (i.e., *R22-1* in Figure 2). Although the performance of workzone sign classification on these sign occurrences was not impressive (i.e., 0.083 as precision and 0.333 as recall), our sign detector and classifier successfully recognized one of the three signs with high confidence. The correct classification in fact, happened at the middle of the three sign appearances and the first and the last (or latest, in terms of time elapse) classification decision on the same sign were incorrect. If our system only considered the latest classification decision without looking into previous classification decisions, the system would miss an important workzone sign and eventually, at best, underestimate

| | A | B | C | D | E |
|---|---|---|---|---|---|
| Number of images | 3,305/447/36 | 4,232/603/89 | 874/234/21 | 3,148/451/68 | 3,280/477/62 |
| Workzone Begins | ○/9/0.727/0.889 | ○/8/0.5/0.875 | ○/1/0.5/1.0 | ○/5/0.625/1.000 | ○/3/0.083/0.333 |
| Workzone Ends | ○/12/1.0/0.333 | ○/31/1.0/0.290 | ○/15/1.0/0.867 | ○/12/1.0/0.833 | ○/15/1.0/0.933 |
| Speed limit change | N/A | ○/12/1.0/0.083 | ○/13/0.5/0.461 | ○/9/0.857/0.666 | ×/7/0.0/0.0 |
| Lane shift | N/A | N/A | ○/20/0.166/0.200 | ○/52/0.444/0.077 | ○/53/0.333/0.170 |
| Lane shift leftward | N/A | N/A | ×/25/0.0/0.0 | ○/24/0.876/0.876 | ○/31/0.875/0.903 |
| Lane shift rightward | N/A | ○/30/0.882/0.5 | ○/12/1.0/0.250 | N/A | ×/9/0.0/0.0 |
| Lane closed | N/A | ×/22/0.0/0.0 | N/A | N/A | N/A |

TABLE IV
RESULTS OF PERFORMANCE TESTS ON THE DETECTION OF CHANGES IN DRIVING CONDITION.

the bounds of the workzone. But one of our recognition error handling methods utilized these consecutive classification decisions as explained in equation 4, enabling the system to turn on the flag to indicate whether our vehicle was driving in a workzone. Without these methods, we may see inconsistent sign classification decisions on the same sign in different time frames (or scales) and miss some of the workzone signs which are important for determining the bounds of a workzone. Thus, while there were fluctuations in workzone sign classification performance, the overall trend was similar to this example, resulting in our system's recognition of all of the highway workzone bounds in the test video.

Our methods did, however, make some mistakes in detecting temporary changes in highways conditions. There were 14 occurrences of temporary changes to highway driving environments in our test videos, and four of them were not recognized. This was primarily a result of our sign detector being unable to localize signs in the under- or over-exposed images of the test video data. This resulted in zeroes of the sign classifier's performance because our workzone sign classifier did not receive potential sign images from the detector for classification and incorrectly classified some of the potential sign images.

## V. CONCLUSIONS AND FUTURE WORK

This paper presents a set of computer vision methods that localize, detect, and classify workzone signs in video data in order to obtain detailed information about highway workzones, such as the bounds of a workzone and temporary highway changes caused by road work. Our experimental results are promising in that our approach is capable of identifying workzone bounds and of recognizing the majority of changes in driving conditions. The contributions of this paper include successful application of existing machine learning techniques to our highway workzone recognition task and development of two new methods for handling potential sign recognition errors.

Although we demonstrate the effectiveness of utilizing color in detecting signs in perspective images, there are some cases in which a color-based sign detector may not work in practice (e.g., where a variation of the target color has not been observed during the training phase). For future work, we would like to investigate an approach that combines color information with shape information. Our color-based sign detector failed to detect some workzone signs when their images were under- and over-exposed. In order to handle such images, we would like to investigate a method that estimates the illumination response function of our vision sensor. In addition, our tests for evaluation of the acquisition of highway workzone information may not be exhaustive. In this regard, in future work, we would like to collect more video data in order to include other events and to extensively evaluate our approach's workzone recognition capability.

## VI. ACKNOWLEDGMENTS

This work is funded by the General Motors-Carnegie Mellon University Autonomous Driving Collaborative Research Laboratory (AD-CRL).

## REFERENCES

[1] Claus Bahlmann, Ying Zhu, Visvanathan Ramesh, Martin Pellkofer, and Thorsten Koehler, A system for traffic sign detection, tracking, and recognition using color, shape, motion information, In *Proceedings of IEEE Symposiums on Intelligent Vehicles*, pp. 255-260, 2005.

[2] Nick Barnes, Alexander Zelinksky, and Luck S. Fletcher, Real-time speed sign detection using radial symmetry detector, *IEEE Transactions on Intelligent Transportation Systems*, 9(2): 322-332, 2008.

[3] Xavier Baro, Sergio Escalera, Jordi Vitria, Oriol Pujol, and Petia Radeva, Traffic sign recognition using evolutionary adaboost detection and forest-ECOC classification, *IEEE Transactions on Intelligent Transportation Systems*, 10(1): 113-126, 2009.

[4] Serge Belongie, Jitendra Malik, and Jan Puzicha, Shape matching and object recognition using shape context, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(24):509-522, 2002.

[5] Marcin L. Eichner and Toby P. Breckon, Integrated speed limit detection and recognition from real-time video, In *Proceedings of IEEE Symposiums on Intelligent Vehicles*, pp. 626-631, 2008.

[6] Arturo de la Escalera, Jose Maria Armingol, Jose Manuel Pastor, and Francisco Jose Rodriguez, Visual sign information extraction and identification by deformable models for intelligent vehicles, *IEEE Transactions on Intelligent Transportation Systems*, 5(2): 57-68, 2004.

[7] Yoav Freund and Robert E. Schapire, Experiments with a new boosting algorithm, In *Proceedings of International Conference on Machine Learning* (ICML-96), pp.148-156, 1996.

[8] Jerome Friedman, Trevor Hastie and Robert Tibshirani, Additive logistic regression: A statistical view of boosting, *Annals of Statistics*, 28(2): 337-407, 2000.

[9] Richard Hartley and Andrew Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, 2003.

[10] U.S. Department of Transportation, Federal Highway Administration, *Manual on uniform traffic control devices for streets and highways*, http://mutcd.fhwa.dot.gov/, 2009 Edition.

[11] J. Ponce, T. Berg, M. Everingham, D. Forsyth, M. Hebert, S. Lazebnik, M. Marszalek, C. Schmid, C. Russell, A. Torralba, C. Williams, J. Zhang, and A. Zisserman, Dataset issues on object recognition, In *Towards Category-Level Object Recognition*, pp. 29-48, 2006.

[12] Gary Overett and Lars Petersson, Large scale sign detection using HOG feature variants, In *Proceedings of IEEE Intelligent Vehicles Symposium*, pp. 326-331, 2011.

[13] Jeffrey Schlosser, Michael Montemerlo, and Kenneth Salisbury, Intelligent road sign detection using 3D scene geometry, In *Proceedings of IEEE Intelligent Robots and Systems*, pp. 740-745, 2010.

[14] Radu Timofte, Karel Zimmermann, Luc Van Gool, Multi-view traffic sign detection, recognition, and 3D localisation, In *Proceedings of Workshop on Applications of Computer Vision*, pp. 1-8, 2009.

[15] Paul Viola and Michael J. Jones, Robust real-time face detection, *International Journal of Computer Vision*, 57(2): 137-154, 2004.