AUTOMATED INVENTORY AND ANALYSIS OF HIGHWAY ASSETS

Final Report

MBTC 2065 Project

Principal Investigator: Kelvin C.P. Wang

Research Staff: Weiguo Gong

Graduate Assistant: Zhiqiong Hou

March 13 2007

Department of Civil Engineering

University of Arkansas

TABLE OF CONTENT

ABSTRACT

The goal of this project was to develop a technology which can automatically detect and localize traffic signs and other roadway assets, in order to provide a comprehensive database for asset inventory purposes in real-time. This report describes the research work on system design, data acquisition, image processing, object tracking, feature extraction, and database recording.

The project work focuses on systems integration of existing technologies to create reliable applications for high performance road data asset inventory. The fundamental methods used for feature extraction, object tracking and retrieval in these applications are well-established. Exploration of the advanced, state-of-the-art or experimental algorithmic development is also made during the project.

The system developed in this project is capable of reliably capturing the Right-Of-Way (ROW) images. The software developed in this project can detect and recognize many of the commonly used road signs with a high accuracy rate in a driving speed up to 60 mph. In certain cases of field tests in urban areas, the accuracy rate for sign detection achieved good result for the 160 signs in the library developed for the project. Included at the end of this report are technical issues relating to errors of sign detection and identification, reasons for not including technologies for automated inventory of assets other than signs, and recommendations for implementation.

Sign detection means that a sign is detected. Sign identification means the sign's content is identified.

INTRODUCTION

Road sign inventory is important in the field of maintenance and management of transportation facilities. A system which can accurately track the total number of signs, the type, the condition and the geographic location of each and every traffic sign is needed for sign inventory and condition survey. However, the current level of the automated processing of the acquired data is not satisfactory. The data analysis in the road sign inventory system is still commonly based on manual interference. Table 1 lists three companies which have products used for road sign inventory. None of the products supports fully automated survey. The commonly used method is to capture the Right-Of-Way (ROW) images by camera or camcorder in the field and then bring the data into the office to do offline processing. The processing is mainly dependent on manual intervention. Normally the road signs or other facilities have to be picked up by the user.

Considering the large amount of image data, manual inventory is tedious, slow, and error-prone. It is desirable to design and implement an automated system which would allow fully autonomous inventory of the road signs with their GPS locations. Therefore, improving the automation and accuracy of the road sign detection and identification is still a research area for road sign inventory.

Table 1 Current Available Technologies in the Market

COMPANY	FULL AUTOMATION	RESOLUTION	GPS/INS
LAMBDA TECH	NO	1600*1200	YES
GEO-3D	NO	1300*1030	YES
ROADWARE	NO	640*480	YES

LITERATURE REVIEW

Automated detection and classification of traffic signs use roadway images acquired from a moving car. Road sign detection is based on computer vision techniques. The methodology of acquiring roadway images is generally referred to as Right-Of-Way (ROW) imaging. The research on the automated road sign detection and identification has been conducted for about 20 years by researchers around the world. Many methods have been proposed. From analyzing single static road scene image to processing real time video sequence or images, both the hardware and the software on automated road sign detection and identification have been improved dramatically. In general, the previous works can be divided into two categories. The first category is using conventional methods based on shape and color analysis. Piccoli et al. (1) described a method for detecting and recognizing road signs in grey level images acquired by a single camera mounted on a moving vehicle. The sign is detected by looking for sets of points of triangular or circular shape through analyzing the edges extracted from a single monochromatic image.

Miura (2) presented a vision system for real-time traffic sign detection using two cameras. One camera with wide angle lens is to detect candidate object, another telephoto camera to capture the candidate in a larger size. Colored candidate region is detected by binarization and area filtering. Boundary lines of rectangular shape are detected from projected histogram peak and circle shape is detected by crossing line peak voting. With an off-the-shelf image processing board, normalized correlation-based pattern matching with database are used to identify sign. Hsu (3) located regions with sign using position, shape and color. The detection phase consists of two processes, training and testing. The training process finds a set of best Matching Pursuit (MP) filter bases corresponding to different road signs to find the best match. However, due to the variety and vagueness of the road signs in many images, fixed threshold and rules may not give high accuracy rate in the environment of changing conditions of light.

Due to the inefficiency of the conventional techniques, since 1990's, using Artificial Intelligence (AI) techniques has become a trend in the road sign detection as the second category of research work. A large amount of papers used Neural Network for this application (4~10). The most popular network is based on back-propagation (BP) method which changes weights by an amount proportional to the error at that unit times the output of the unit feeding into the weight. The main difference among these networks is that they may use different types of input. The input can be the whole pixel image or the features of the image. Other than Neural Network, Genetic Algorithm (11) and Simulated Annealing (12) have also been tested. AI techniques are promising, but they are still under development and not mature enough at this stage for implementation. A well trained neural network might be able to detect a specific type of sign with a limited accuracy rate (10), but it is greatly dependent on the training set and it is hard to refine the performance due to the lack of explainable working rules inside the network. In addition, AI techniques may take more computing cost. For example, the BP network in Luo (4) requires four seconds to recognize an image.

Generally speaking, the progress on fully automated road sign detection and identification is still not satisfactory and far from practical. The traditional technique is weak at handling the vagueness and variety of the sign images. The Neural Network is largely dependent on the training set. In addition, most of the algorithms only applied to one or a few signs. In road sign inventory, detection and identification are needed for a large amount of types of signs in different views and various ambient conditions. The variations in an object appearance due to the variance of the luminance, contrast, scale, distortion and occlusion introduce multitudes of difficulties for fully automated sign inventory. If invariant features of the same object can be found in spite of the aforementioned varieties, the magnitude of the problem can be then reduced. Moreover, the computing time for image processing normally is unacceptably long due to complicated calculation on the pixels. The technique of object tracking can efficiently minimize the searching area in the image thus save the computing cost. So the study on the tracking technique is another area we focused on in this research project.

SYSTEM DESIGN

Since the mid-1990's, research team at the University of Arkansas has engaged in the development of a full digital vehicle for highway data collection, this Digital Highway Data Vehicle (DHDV) has evolved into a new platform shown in Fig. 1. The DHDV is multi-functional and includes pavement surface imaging for condition survey at 1-mm resolution, longitudinal roughness measurement consisting of laser sensors and accelerometers, laser based rutting measurement device with over 1,000 transverse points, ROW images, GPS receiver, DMI, and Gyro sensor. The software system used in the on-board computers of the DHDV employs real-time relational database engine, inter-computer communication techniques, multi-computer and multi-CPU based parallel computing, real-time control of digital sensors, and the generation of multimedia databases.

For right-of-way imaging, there are two digital color frame cameras mounted on top of the DHDV. Each camera is at the resolution of 1300 by 1024. The imaging system can record and archive one frame color image from the camera at user-determined interval up to one frame per 10 feet. Figure 2 shows the structure of the system. Two color cameras and a GPS receiver are mounted on the top of the vehicle. They all connect to the control chassis inside the vehicle. So does the DMI. While the vehicle is in operation, signal sent from the DMI will trigger the data acquisition and the software will process the images and conduct the automated inventory. The system allows for large quantities of spatially referenced data to be stored, managed, analyzed and queried.

Figure 1 Digital Highway Data Vehicle (DHDV)

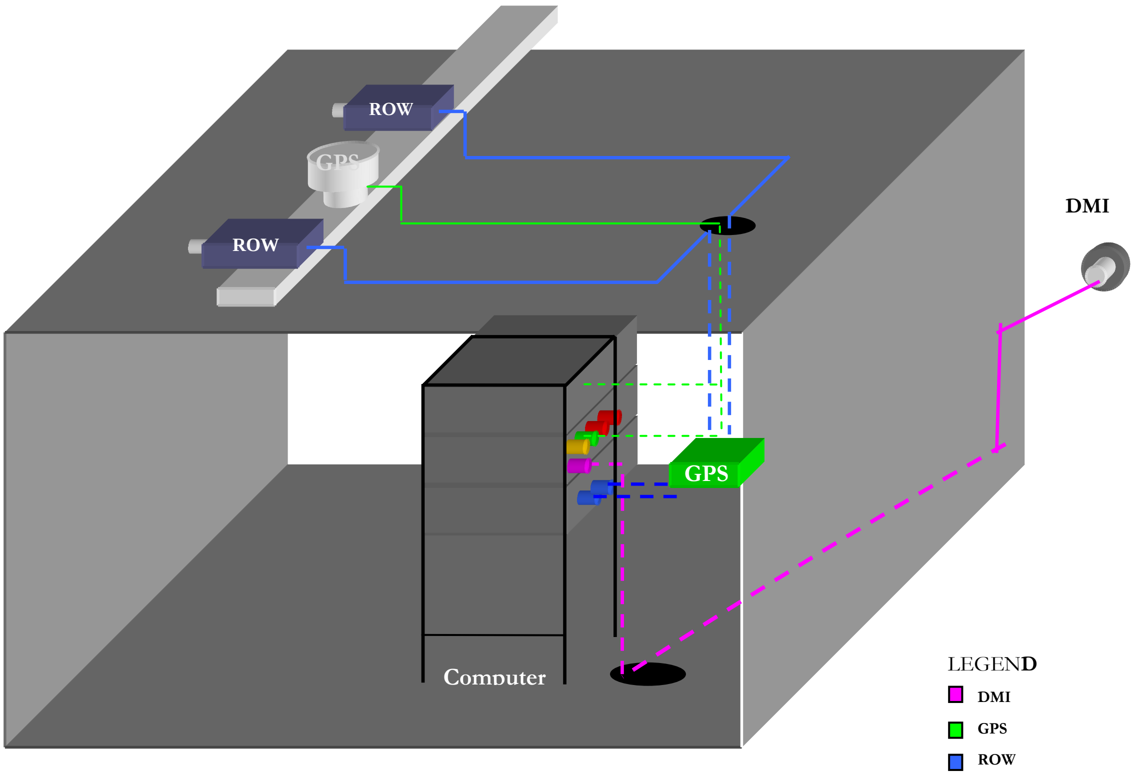


Figure 2 The System Diagram

METHODOLOGIES

The algorithm for the image processing of signs includes the following three steps. The result is given in the form of road sign type and GPS information. Spatially referenced sign objects are recorded in a comprehensive database.

1) Candidate region is searched in the image based on color threshold. The threshold is obtained from experiment.

2) The candidate region from the first step is then tracked among the successive frames using the technique of Kalman filtering. This is used to narrow the searching area and shorten the computing time.

3) The tracked target region is then compared with standard sign by using a SIFT based process to determine sign content.

CANDIDATE REGION DETECTION

Choosing a right color space during the image analysis is critical. The Right-Of-Way (ROW) images captured from Digital Highway Data Vehicle (DHDV) are originally in RGB color space. However, RGB is hardware-orientated and it is non-intuitive. HSL color space is used in the proposed system (Figure 3). HSL stands for hue, saturation and lightness which closely follow human perception of color. Hue specifies the base color, the other two values specify the saturation of that color and how bright the color is. The advantages of using HSL color space for analysis are:

1) It allows a better tolerance to changes in lighting conditions compared to other color models.

2) A specific color can be recognized by matching to a small range of hue value.

3) HSL is more intuitive and is symmetrical to lightness and darkness.

4) It is easy to convert HSL to RGB.

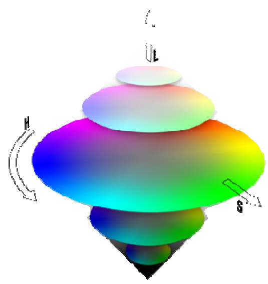
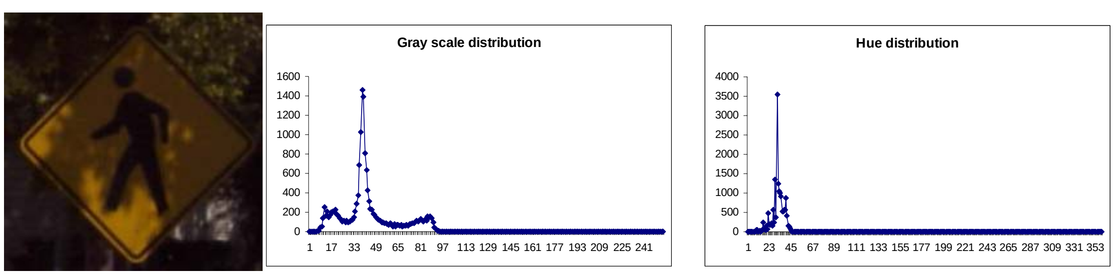


Figure 3 HSL Color Space

A test was conducted to demonstrate the advantage of using HSL color space in the road sign detection problem. The distributions of hue value and the gray scale value for orange background pixels in Figure 4 (a) are plotted in Figure 4 (b) and Figure 4 (c). It shows that although the orange color is partly shaded, the hue value of those background pixels still falls into a narrow range (18~45 out of 360), contrast to 10~92 out of 255 in gray scale value range. It shows that using hue value to segment the signs is better than using gray scale value which is the average of the R, G, B value of each pixel.



(a) (b) (c)

Figure 4 HSL Value of Orange Color Sign

Since there are so many types of road signs, searching each sign color type in one image will take a long time. To speed up the detection of the signs with obvious colors, the raw image is first split into five color bands. It is obtained by labeling the pixels by their hue and saturation values. The hue and saturation threshold values of each of the color bands are obtained by experiment. Right-Of-Way (ROW) images were captured under various weather conditions. The hue and saturation values of the sample points are extracted from the road signs in the testing images, and then analyzed. Each is plotted in a polar coordinate system shown in Figure 5. The radial coordinate stands for the saturation value and the angle coordinate stands for hue value. From the plots, it can be found that each of the sign color types fall into a narrow hue range. Therefore, splitting them into five color bands is applicable. Splitting the pixels into five different color bands narrows the search space for each specific sign type. This approach also provides separate processing steps that can exploit parallel computing techniques to enable real time processing. Moreover, it functions as a pre-classification before the SIFT matching.

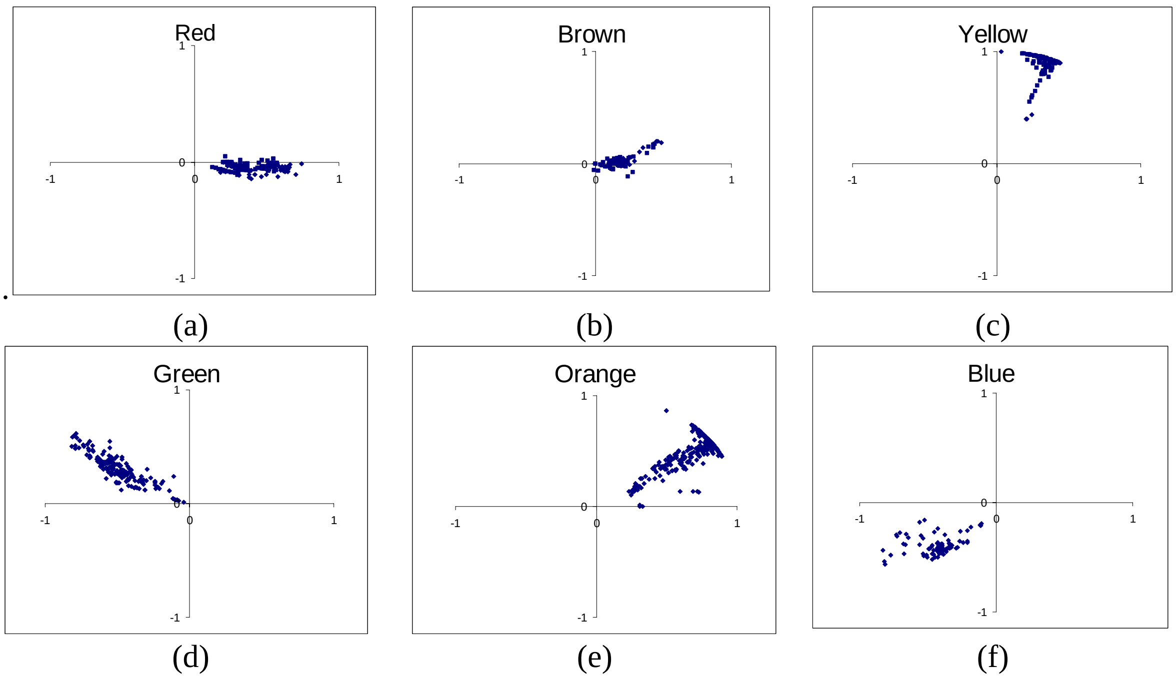


FIGURE 5 Color Distribution in the Sample Images of Road Signs

There are six color bands: red and brown, orange, blue, yellow, and green, but red and brown are merged into one band due to the similar range of the hue and saturation values as shown in Figure 5 (a) and Figure 5 (b). Figure 6 (a) shows a raw image, all the pixels with a hue and a saturation values falling into the red band are labeled and shown in Figure 6 (b). At the same time, white signs and black signs are still detected in the original RGB image because they have a relatively wide hue range. For each raw image, six images are processed, five of them in color bands, one in the original RGB color space.

(a) (b)

Figure 6 Original Image and Image in Red Color Band

Right-Of-Way (ROW) images captured from the DHDV imaging system are in a size of 1280×960. In the detection module, to save the computing cost, images are made smaller by down-sizing. The detected candidate region goes back to the original size in the detection module.

KALMAN FILTER TRACKING

To improve the efficiency of detection process while maintaining a low "false hit" rate, the technique of Kalman filtering is used to predict road sign by camera geometry information. By applying this technique, the detection process can significantly narrow down the search space for each frame and therefore reduce a large number of false hits which may occur in the case of using traditional image processing for each single frame.

First, a road sign candidate region is determined in the current frame, which is accomplished in the previous step of candidate region detection. By computing the geometric relationship of the road sign between adjacent frames, the framework finds the candidate region for the road sign in the next frame. At this time, the image processing step analyzes the candidate region in detail.

It is assumed that the optical axis of the camera is roughly horizontal and the motion of the camera is moving along its optical axis. This assumption is often true in real world settings, particularly, when a camera is mounted on the vehicle and its optical axis is calibrated to be parallel to the horizontal plane of the vehicle. Figure 7 (a) shows the side view of such scenario and Figure 7 (b) illustrates the spatial constraints among road sign planes, image planes and the camera between two successive frames.

The Kalman filter technique includes two phases: time update and measurement update (13). The time update procedure is based on the dynamic equation which is derived from the spatial constraints from the two successive frames. The measurement update is based on the image processing location in the proximity of the predicted candidate region.

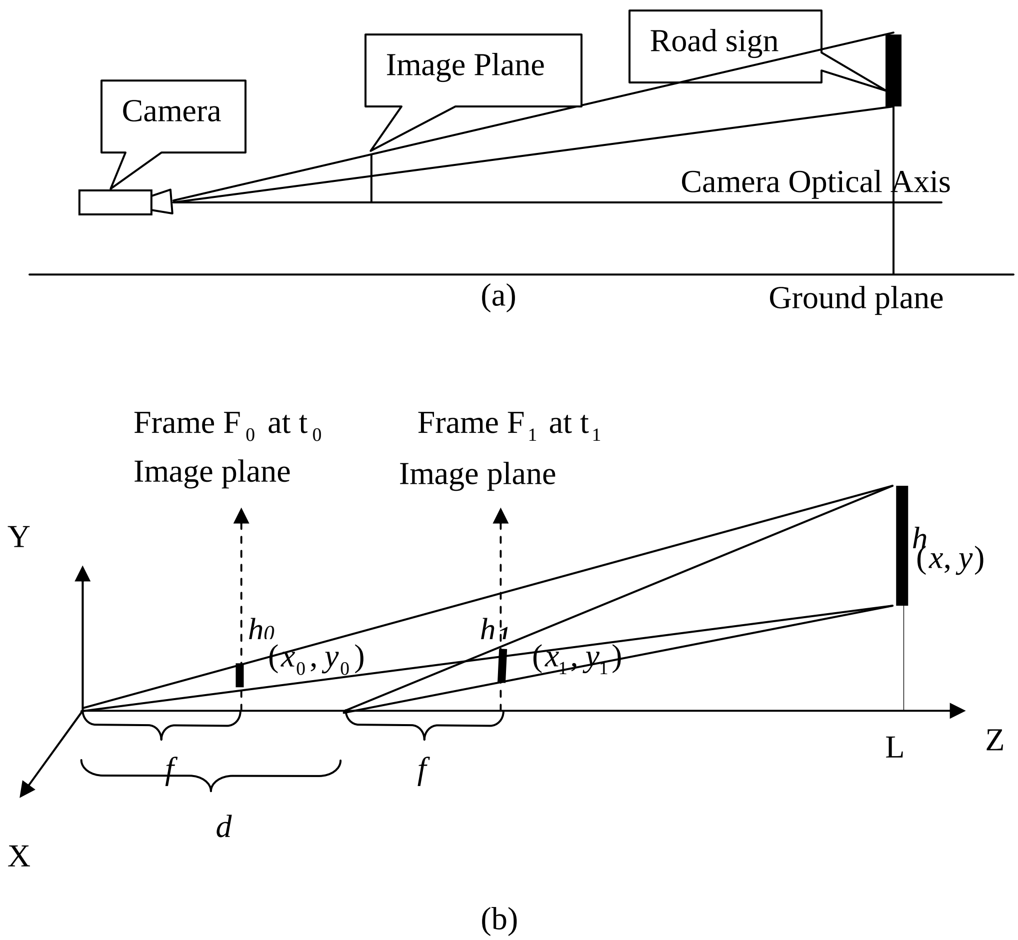


Figure 7 Spatial Constraints in Two Successive Frames

As shown in the Figure 7, the camera coordinate system at time t_0 is taken as the basic coordinate system, Z-axis is the camera optical axis and the X-axis is parallel to the vehicle's horizontal plane. The camera focus length is denoted by f. The camera moves distance d in the traveling direction, which is measured from the on-board DMI.

With the information of spatial constraints between two successive frames, the candidate region size of the road sign and its location in the following frame are predicted with a Kalman filter based on the current frame. Fang et al. (14) provided a modeling example using the Kalman filter technique. However, the validity of forming the transition matrix is not suitable for this application. Particularly, stability problems are generated by taking the inverse of an ill-conditional matrix. The modified technique that has been specifically developed in this research has overcome this problem as discussed in the following derivations.

Road sign size prediction:

Assume h is the actual height of the road sign, h_0, h_1 are the heights of the road sign in two successive frames shown in Fig. 7b). Based on triangular principle, the relationship can be found as

$$h_1 = \frac{1}{1 - \frac{dh_0}{hf}} \cdot h_0 \quad (1)$$

$$w_1 = \frac{1}{1 - \frac{dw_0}{wf}} \cdot w_0 \quad (2)$$

Road sign location prediction:

Assume $(x_0, y_0), (x_1, y_1)$ are the centroids of the road signs in two successive frames. Since

$$\frac{y_1}{y_0} = \frac{L}{L-d} = \frac{h_1}{h_0} \quad (3)$$

The relationship can be formed as:

$$y_1 = \frac{1}{1 - \frac{dh_0}{hf}} \cdot y_0 \quad (4)$$

Similarly

$$x_1 = \frac{1}{1 - \frac{dw_0}{wf}} x_0 \quad (5)$$

Here, x, y, h, d, and f are constants. Based on the above relationships, the dynamic equation of the model is constructed as follows.

The five equations for Kalman filter model (15) are:

$$\hat{x}_k^- = A\hat{x}_{k-1} + Bu_{k-1} \quad (6)$$

$$P_k^- = AP_{k-1}A^T + Q \quad (7)$$

$$K_k = P_k^- H^T (HP_k^- H^T + R)^{-1} \quad \text{.......(8)}$$

$$\hat{x}_k = \hat{x}_k^- + K_k (z_k - H\hat{x}_k^-) \quad \text{.......(9)}$$

$$P_k = (I - K_k H) P_k^- \quad \text{.......(10)}$$

Where:

$\hat{x}_k^-$: a priori state estimate at step k with given knowledge from step $k-1$

$\hat{x}_k$:a posteriori state estimate at step k given measurement z_k

P_k^- :a priori estimate error covariance

P_k : a posteriori estimate error covariance

The properties of the road sign in the successive frames at each time instance (frame) can be characterized by its position and size. Let (x_t, y_t, h_t, w_t) represent the road sign pixels position (its centroid) and size at time t. The state vector at time t can therefore be represented as $x = (x_t, y_t, h_t, w_t)^T$. The system can therefore be modeled as

$$x_t = Ax_{t-1} + \omega_t \quad \text{.......(11)}$$

Where:

ω_t represents system perturbation which comes from the drive direction variance, DMI measurement error, etc. It is further assumed that the image processing estimates

$$z_t = Hx_t + v_t \quad \text{.......(12)}$$

Where:

v_t represents measurement uncertainty. Specifically, the position of the current frame t is estimated based on an image processing in the neighborhood of the predicted position (based on the system model). In this manner, the state vector x_t, along with its covariance matrix P_t, can be updated using the system model (for prediction) and measurement model (for updating). However, the test conducted in the research shows that the noise of the location state $[x_t, y_t]$ is much higher than the size state $[h_t, w_t]$. On the other hand, the smaller dimension matrix will have less possibility to be ill conditioned. Due to these two reasons, the state vector $x = (x_t, y_t, h_t, w_t)^T$ is split into two state vectors:

the location state $X = \begin{bmatrix} x_t \\ y_t \end{bmatrix}$ ....(13)

the size state $Y = \begin{bmatrix} h_t \\ w_t \end{bmatrix}$ ....(14)

The calculation of the Kalman filter includes an inverse of the combination of transition matrix A in the middle. The split of the state vector improves the stability of the model because it replaces the 4×4 matrix inverse calculation with a simpler 2×2 matrix. Based on our test, this modification is superior to the formation of the transition matrix in Fang et al. (14) in terms of the model stability and convergence. There are two reasons for the modification. First, their state vector in (14) includes x, y, 1/r, which are not in the same order. The test in the research shows the matrix is always ill conditioned with this method. Second, the location state (x,y) and size state (h,w) should be separated into two models. It is found that prediction model for (x,y) has more errors due to motion of the vehicle. However, the prediction model for size (h,w) is relatively accurate because its possible error is primarily from DMI.

Transition matrix A

$$A = \begin{bmatrix} \dfrac{1}{1 - \dfrac{d \cdot h_{t-1}}{h \cdot f}} & 0 \\ 0 & \dfrac{1}{1 - \dfrac{d \cdot w_{t-1}}{w \cdot f}} \end{bmatrix} \text{..} (15)$$

Measurement matrix H:

$$H = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \text{..} (16)$$

The spatial constraints of the road sign size and the location between two successive frames constitute a dynamic equation which functions as a time update step in the Kalman filter. The goal is to combine tracking with other rules from the color and shape criteria produce a working set of algorithms for sign inventory. Figure 8 shows the flow chart of the Kalman filter based tracking in detail.

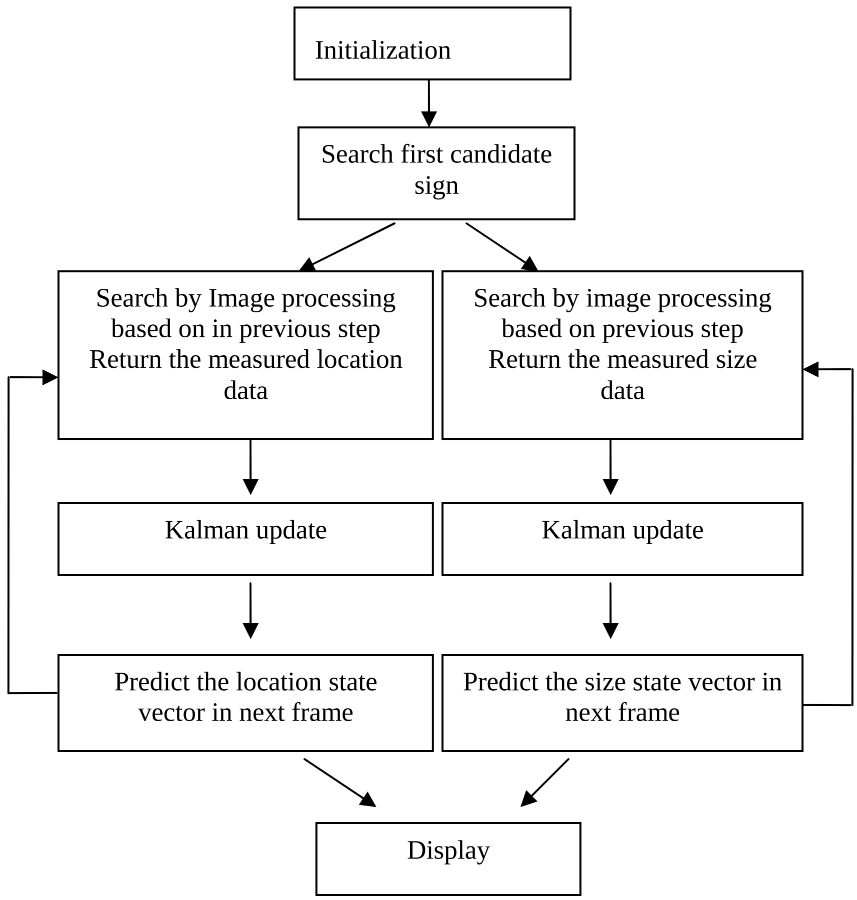


Figure 8 Flow chart of the Kalman filter tracking

SIFT FEATURE EXTRACTION FOR SIGN CONFORMATION AND IDENTIFICATION

Recently, a prominent invariant feature matching technique – Scale Invariant Feature Transform called SIFT (16) and (17) gained reputation on the object detection field. SIFT has been shown to be effective in numerous object detection problems (18) and (19). Mikolajczyk and Schmid (20) have made a performance evaluation for SIFT descriptors versus other invariant feature descriptors. They concluded that SIFT is a very effective imaging technique with respect to the scale, rotation, and illumination changes.

This method recognizes an object by detecting a high level of correspondence between the invariant features from the tested object and from model object. The feature extraction is based on appearance, thus there is no need to extract geometric primitives which are generally hard to reliably detect. Objects can be recognized even if partially occluded (hidden) by other object(s) in the foreground. These features also show good level of invariance to scale, luminance and size irregularities. All these characteristics make SIFT a good candidate tool in solving the road sign detection problem.

SIFT is used mainly for generation of invariant features in an image for object detection, detection and tracking applications. The potential of SIFT lies in the invariance of the generated features to image translation, scaling, rotation, and partially to illumination changes and affine projection.

The following properties of SIFT make it proper to the road sign detection problem:

1) SIFT is invariant to image scale and rotation. It is shown to provide robust matching across a substantial range of affine distortion, change in 3D viewpoint, addition of noise, and change in illumination.

2) The features are highly distinctive. A single feature can be correctly matched with high probability against a large database of features from many images.

3) The detection proceeds by matching individual features to a database of features from known objects near real-time performance.

4) It can be combined with other feature vectors.

The invariant features of SIFT are based on local extrema in scale-space built with Difference-of-Gaussian (DoG) filters. Equation (18) is the Gaussian function. Equation (19) is the scale space of an image.

$$D(x,y,\sigma) = (G(x,y,k\sigma) - G(x,y,\sigma)) * I(x,y) = L(x,y,k\sigma) - L(x,y,\sigma) \quad \text{.......(17)}$$

Whereas:

$$G(x,y,\sigma) = \frac{1}{2\pi\sigma_2} e^{-(x^2+y^2/2\sigma^2)} \quad \text{.......(18)}$$

$$L(x,y,\sigma) = G(x,y,\sigma) * I(x,y) \quad \text{.......(19)}$$

$I(x,y)$ is the input image, σ is the scale space factor in equation (19).

It takes four steps for an image to generate the SIFT features. The first step is the determination of the location and scale of the keypoints. This is done by constructing a Gaussian pyramid, using Gaussian smoothing and sub sampling. Each level in the pyramid is constructed using a Difference-of-Gaussian image, obtained by subtracting the image from its Gaussian smoothed image. Then sub sampling is performed to construct the next level in the pyramid. Feature locations are identified by detecting the maxima and minima relative to surrounding pixels and adjacent scales. With this method, SIFT provides a high level of assurance that the key points are located at regions and scales of high variations, which make these locations stable for characterizing the image.

The second step is to refine the keypoint location. It is done by fitting a 3D quadratic function to the local sample points to determine the interpolated location of the maximum, which provides a substantial improvement for matching stability.

In the third step, the orientation of each keypoint is assigned based on local image properties. The keypoint descriptor can be represented relative to this orientation and therefore achieve invariance to image rotation. So far, the keypoints have been located and each keypoint has parameters of location, scale and orientation. These parameters impose a repeatable local 2D coordinate system in which to describe the local image region, and therefore provide invariance to these parameters.

Finally, a local image descriptor is computed for the local image region that is highly distinctive, yet is as invariant as possible to other variations. Variations of illumination on signs due to shadows are typical examples.

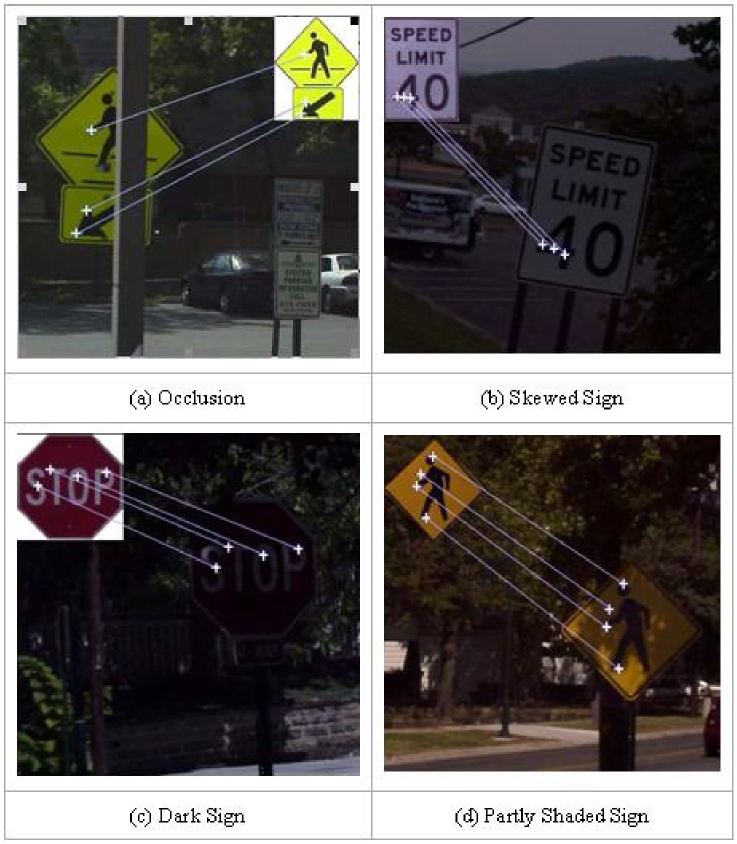
Figure 9 shows some of the matching keypoints in the experiment. The sign templates are put together with the Right-of-Way images to show feature matching. The crosshairs indicate the keypoints detected by SIFT. The line between the keypoints shows the correspondence between the keypoints. Figure 9 (a) shows an occluded pedestrian sign. With some features missing, the left features still contribute to the final detection of the sign. Figure 9 (b) is a skewed sign. Figure 9 (c) shows a stop sign taken in the shaded area with a dark color. Figure 9 (d) shows signs of an unevenly distributed orange color due to shade. Those signs are difficult to recognize using conventional methods. The algorithm based on SIFT correctly extracts the invariant features and match them with the features of standard signs.

Once the SIFT invariant features are generated from the candidate region from the detection module, they are compared with the features of the standard road sign which are saved offline in the standard sign database. Final verification of each match is achieved by finding a low-residual least-squares solution for the unknown model parameters. To recognize a sign, the number of key points' correspondence between the road sign candidate and the template in the database is used as the similarity measure.

The number of invariant features or points of interest varies from one sign to another. Therefore, it is not reasonable to use the absolute number of correspondent features. The matching process is performed according to the objective function used by Farag (21).

$$f_i = \frac{n}{\sqrt{n_b n_i}} \qquad (20)$$

Where f_i is the objective value for i^{th} standard road sign, n is the number of matches, n_i is the total number of the invariant features in the i^{th} standard sign and n_b is the total number of the invariant features in the candidate sign.



(a) Occlusion | (b) Skewed Sign | (c) Dark Sign | (d) Partly Shaded Sign

Figure 9 Examples of SIFT keypoints matching

The matching decision is made according to the standard sign with the highest object value. If the highest objective value is less than a certain threshold, the detection is discarded. The described SIFT technique is used for confirmation of sign detection conducted in the previous step, and also for extracting sign content for sign identification.

GPS/GYRO POSITIONING

The final step consists of geo-referencing the detected road sign using GPS/GYRO information. Depending on the GPS receiver, the positioning accuracy varies. The data from GPS receiver is adjusted by local GYRO information. The integration of the GPS receiver and GYRO sensor more accurately defines location of the detected road sign than just using GPS data alone.

SOFTWARE

A Right-Of-Way (ROW) capturing program was developed in this project. It controls the operation of the two cameras mounted on the top of the vehicle which can capture the roadway images over 10 frames a second. And it allows editing for the color settings. The images captured are stored in the local computer hard disk. A comprehensive database which records the spatial information of each image is saved in the same directory.

A road sign detection and identification program was developed based on the described algorithms, which is a substantial achievement in this project since no similar product is available on the market by our knowledge. It can detect many of the common signs in real time. Figure 10 and Figure 11 are screenshots of these programs.

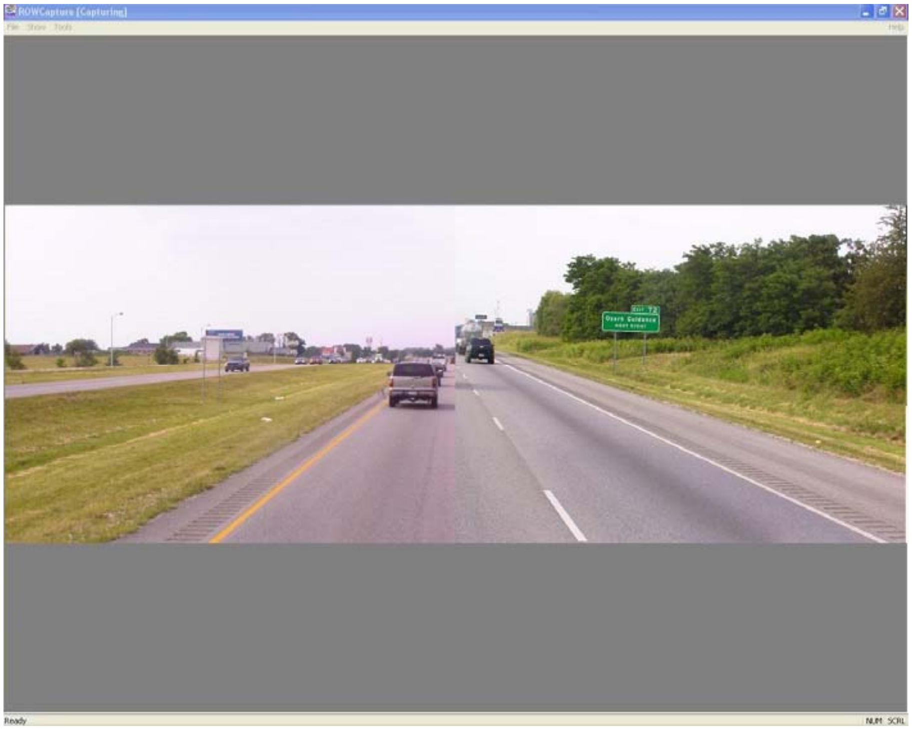


Figure 10: ROW Capture Software

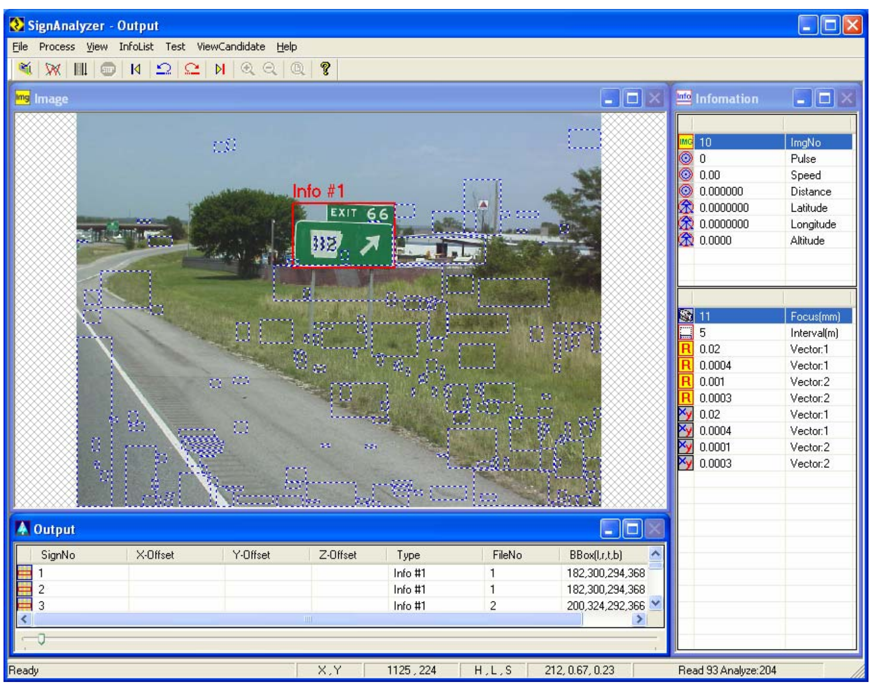


Figure 11 Road Sign Detection Program

The detection program can work with the image acquisition program during real-time capturing or process the data offline. The road signs are detected automatically and corresponding referenced spatial information is saved into a database which can be read and reviewed later. While the program is running, a red bounding box will be drawn around the road sign along with the classification label. Other frames in the program interface show the corresponding information (GPS, tracking parameters and classification result) for each detected sign.

Currently, our program can detect many of the 160 common signs. Some of the more important road signs such as stop sign, yield sign can be uniquely recognized as an individual sign type. Other signs such as diamond shaped signs can only be recognized into categories. Speed information on Speed Limit signs can also be extracted to identify given speed limit. Tables 1 and 2 list the detection and identification capability of the developed software.

TABLE 2 Signs Individually Detected and Identified with the Software

No.	Classification	Number of Types	Label
1	Stop Sign	1	"Stop"
2	Yeild Sign	1	"Yield"
3	Do Not Enter Sign	1	"Do Not Enter"
4	Speed Limit (10,15,20,25,30,35,40, 45,50,55,60,65,70,75,80)	15	eg. "Speed 25"
5	Turn Prohibition (No Left turn, No Right Turn, No U Turn)	3	eg. "No Left Turn"
6	Rectangle (SpeedLimit, Left Turn Arrow, Right Turn Arrow, Straight Arrow, Up Arrow, Down Arrow)	6	eg. "Left Turn Arrow"
7	Direction Sign (North, Left, South,West, JCT, TO, End)	7	eg."North"
8	Center Lane Sign	1	"Center Lane"
9	General Guide Sign (Left, Right, Up, Down, Straight)	5	eg."Left"
10	Exit Sign	1	"Exit"
11	No Parking Sign	3	"No Parking"
12	Pedestrian Sign	3	"Pedestrian signs"

Figure 12 lists types of signs that can be categorized, some of which can be uniquely identified (Table 2), while others can only be categorized (Table 3). The first column shows the category number used in the program. The second column shows the classification of the sign in the detection and identification program. The third column shows how many types in that category can be recognized. The last column shows the information label given and shown for the recognized sign. The category is defined in this way for the convenience of the algorithm developing. This initial result shows the feasibility of the detection and identification system. More detailed categorization can be made in the future.

TABLE 3 Signs Only Detected into Categories

No.	Classification	Number of Types	Label
13	Diamond Shape Sign	--	"Warning Sign"
14	Green Information Signs	--	"Info #1"
15	Blue information Signs	--	"Info #2"
16	Brown Information Signs	--	"Info #3"

R1-1

R1-2

R5-1

R2-1

1.Stop 2.Yeild 3.Do Not Enter 4.Speed Limit

R3-1

R3-2

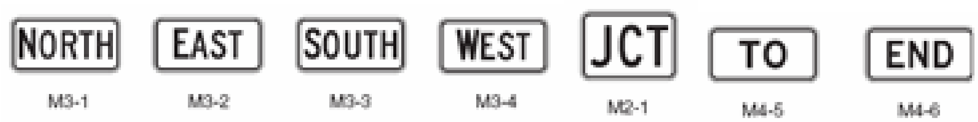
R3-4

5. Turn Prohibition

R3-5

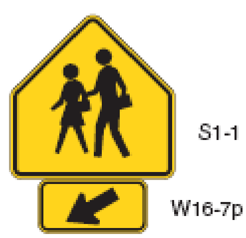
R3-5a

6. Rectangular shape



7. Direction Sign

11. No parking sign



12. Pedestrian sign

13. Diamond Shape Sign

14. Green information sign

15. Blue information sign

16. Brown information sign

FIGURE 12 Road Sign Examples in Table 2 and Table 3

THE DLL USAGE

A Dynamic Link Library (DLL) file is provided as a result of this project. This enables AHTD developer to build their road sign detection and identification software. The DLL supports 3 main functions: Road sign detection and identification, 3D calibration and 3D positioning. Examples of using the DLL is shown as below.

1. Road sign detection and identification function

```
bool RoadSignAnalyzer( int nWidth, //the image width
int nHeight, //the image height
Unsigned char * pImgBuf, //raw image
INPARA*  pInputPara,// input
OUTPRST*  pOutputRest//output
)
```

Whereas INPARA and OUTPRST are structures. They are declared as below

```
Typedef struct
{
   BOOL  bUseKalmanFilter;// This is set as TRUE when the KalmanFilter function is called to improve the accuracy. However, it sacrifices the computing cost.
   BOOL  bNeedSpeedUp;// When this is set as TRUE, the processing speed will be improved while it decrease the accuracy. It is recommend that this is set as TRUE only when the image quality is good.
   STVISONPARA p3DCalRestParas;// This is a parameter used for 3D reconstruction
} INPARA;

Typedef struct
{
   Int nRoadSignNum;// the number of road sign recognized
   ROADSG * pRoadSignList;// Parameter for result
}OUTPRST

Typdef struct
{
   Int    nRSType;// Road sign types
   Char   strCaption;// Road Sign Detection Label
   CRect  rect;// The location in the image
   Int    nImgID;// The image ID
   Float  X,Y,Z;// The 3D coordinates
} ROADSG
```

2. 3D calibration function

```
BOOL  Do3DCalibration( INPUTPTS  pPtList,
                       Int  nPtNum,
                       OUTPUT3DPara pOutputParas)
```

nPtNum: The number of marking points
pPtList : The physical coordinates of the marking points

pOutputParas: The output, referring to the following structure

```
Typdef struct
{
    Int x, y;// The image coordinated in the left image
    Float    X,Y,Z; //Physical 3D coordinates

}INPUTPTS;

Typdef struct
{

Float  F; // Focal length of camera
Float  K;// Radial lens distortion coefficient
Float Cx,Cy;// Co-ordinates of centre of radial lens distortion
Float Sx; Scale factor to account for any uncertainty due to imperfections in hardware timing for scanning and digitization

Float Rx,Ry,Rz;// Rotation angles for the transformation between the world and camera co-ordinates
Float  Tx,Ty,Tz;// Translation components for the transformation between the world and camera co-ordinates

}OUTPUT3DPara
```

3. 3D positioning function

```
BOOL  Calcu3DCoords(OUTPUT3DPara pLeftCamParas,// Calibration parameter of the left camera
OUTPUT3DPara pRightCamParas,// Calibration parameter of the right camera
Int  xL, int yL, //Image coordinates of the marking point  in the left image
Int  xR, int yR,// image coordinates of the marking point  in the right image
Float  &X, Float &Y, Float &Z);// The physical coordinate of the target point.,
```

RECOMMENDATION AND CONCLUSION

The results from this research project demonstrate the feasibility of developing an automated road sign inventory system using image processing techniques. In addition, almost all the target road sign types in the project have been detected by the algorithms correctly, with a few errors during limited field tests. However, considering that the total number of road signs defined in MUTCD is large, the current software may introduce errors resulting in certain signs not detected, detected incorrectly, or that non-sign objects are detected as signs. For example, signs not included in the built-in library will not be identified at all. Additional verification of the developed algorithms and sign library needs to be conducted in terms of network level performance for rural and urban roads, interstate and local roads. Furthermore, the sign library needs to be modified to accommodate MUTCD conventions and classifications.

It should be noted that there are several capabilities that were stated in the proposal of the project, but not developed. The research team met many difficulties in developing the automated sign inventory technology, which has been a primary focus of the project. Due to constraints of time and resources, detection of no-passing lane markings, guard rails and commercial signs cannot be conducted with the developed software.

The following is a list of issues that make it impossible for immediate implementation of the result from this project.

1. Due to time constraint, and unanticipated difficulties in the research phase, this project has been largely an effort in algorithm development; implementation and extensive trials were not conducted. For example, the research team did not anticipate a large portion of time had to be used in testing existing imaging methodologies for the purpose of selecting an optimal, effective and fast approach.

2. Out of the 160 signs that were built into the sign library, the developed software can uniquely identify and extract contents from over 20 common signs for detection and identification. For other signs, only their shape, color and general classification can be identified at this time.

3. Certain signs are not detected, detected incorrectly, or that non-sign objects are detected as signs. One primary reason is that the developed sign library is incomplete. In addition, the developed algorithms should need more testing and refinements.

4. Automated identification of sign location is not developed in this project. Integration with AHTD work flow is not developed either. For example, MMHIS can be a good platform to use the developed algorithms for sign identification. The completed development does not have a direct interface to work with MMHIS at this time.

5. Image definition and image quality can greatly affect the performance of the developed algorithms. Particularly, when lighting condition is not favorable, such as images are captured facing the sun, or under darker conditions, the developed software may perform worse than images captured under good lighting conditions. In recent weeks, the research team started experimenting with a High-Definition camera with built-in auto iris and auto focus which are not present with the cameras that were used for the research.

There are no serious difficulties preventing implementation of algorithms for highway signs not identified with the developed software. At this stage, only important and common signs are chosen for testing purpose. Other signs can be detected in the similar manner. Nevertheless, implementing all of road signs defined in MUTCD in the current system requires a substantial amount of additional time and efforts. Another potential problem is the high computing cost if a real time processing is required. Accurate identification of sign location depends on the 3D positioning of the sign using stereovision technique as well as GPS data. The location accuracy can be kept around 1 meter using the current system. False sign identification can be decreased by more testing and modification of the algorithms.

In addition, location referencing based on AHTD's log mile system and other used positioning system, such as GPS, DMI, et al. shall be fully integrated into the database. This feature is critical for AHTD to fully utilize the information extracted from MMHIS imagery database for sign inventory. At this time, condition survey of signs is not developed yet. It is a potential research item in a future project.

The following summarizes the necessary additional technical work for the implementation of an automated sign detection and identification system for AHTD:

1. Due to difficulties encountered in the development, it is recommended that future development work on automated roadway asset management be focused on traffic signs only.

2. MUTCD standards shall be included in the built-in database. Efforts shall be made to accommodate as many signs as possible that are in the MUTCD standard, and are also commonly used by AHTD. Categorization of signs with MUTCD conventions must be used.

3. Further improvements of algorithms are needed for improving identification accuracy of signs in the database and new signs that are to be added. Imaging and database compatibility shall be built with MMHIS currently in use at AHTD.

4. Positioning reference shall be integrated into the system database without slowing down the processing. Highway log mile reference methods shall abide by current AHTD conventions.

5. Real-time capability shall be investigated, particularly when High-Definition images are used, which will present new challenges due to its large image size.

6. Extensive tests and trials shall be conducted to accommodate roads in various classifications to establish performance benchmarks.

Despite the fact that there is still no fully automated sign detection and identification technology in the market place, the research team has paved the foundation in this research project to make the technology available. Particularly, when the High-Definition standard based on the HD 1080p format is used as the acquisition platform in the future, it will present new opportunities for automated sign detection and identification. We believe that in the next few years, fully automated asset inventory and condition survey systems will be deployed in the field for various roadway asset management purposes.

ACKNOWLEDGEMENT

The researchers would like to thank Mr. Alan Meadors, head of AHTD Planning and Research, for his insight to provide initial ideas for the project. The project coordinator, Mr. Ron Strickland, research engineer of AHTD, provided valuable input about research directions throughout the duration of the project. Ron has participated in brainstorming sessions and editing technical information for the project. His assistance and guidance is very much appreciated.

REFERENCES

1. Piccioli, G., et al. Robust Road Sign Detection and Recognition from Image Sequences. In *Intelligent Vehicles Symposium*, Paris, 1994, pp. 278-283.

2. Miura, J et al. An Active Vision System for Real-Time Traffic Sign Recognition. In *Conf. on Intelligent Transportation Systems*, Dearborn, MI, 2000, pp. 52-57.

3. Hsu, S.H and Huang, C.L. Road Sign Detection and Recognition Using Matching Pursuit Method. In *Image and Vision Computing*, Vol19, 2001, pp. 119–129.

4. Luo, R. C., Potlapallil, H. and Hislop, D. W. Translation and Scale Invariant Landmark Recognition Using Receptive Field Neural Networks. In *Proc. Int. Conf. Intelligent Robots and Systems (IROS'92)*, 1992, pp. 527–533.

5. Luo, R.C., Potlapalli, H., Hislop, D. Traffic Sign Recognition in Outdoor Environments Using Reconfigurable Neural Networks. In *IEEE International Joint Conference on Neural Networks*, 1993.

6. Estable, S., Schick, J., Stein, F., Jhansen, R., Writter, R. O. H. and Zheng, Y.-J. A Real-time Traffic Sign Recognition System. In *Proc. Intelligent Vehicles Symposium*, Paris, France, 1994, pp. 212–218.

7. Adorni, G., Da´ndrea, V., Destri, G. and Mordoni, M. Shape Searching in Real World Images: a CNN-Based Approach. In *IEEE.Fourth Workshop on Cellular Neural Networks and Their Applications*, 1996.

8. Kim, S.K., Forsyth, D.A. A New Approach for Road Sign Detection and Recognition Algorithm. In *30th International Symposium on Automotive Technology and Automation, Robotics, Motion and Machine Vision in the Automotive Industries*, ISATA, 1997.

9. Escalera, A., Moreno, L., et al. Road Traffic Sign Detection and Classification. In *IEEE Transactions on Industrial Electronics*, Vol. 44, 1997, pp.848-858.

10. Liu, X. H. and Ran, B. Vision-Based Stop Sign Detection and Recognition System for Intelligence Vehicle. In *TRB Annual Meeting*, 2001.

11. Aoyagi, Y. and Asakura, T. A Study on Traffic Sign Recognition in Scene Image Using Genetic Algorithms and Neural Networks. In *Proc. IEEE 22nd Int. Conf. Industrical Electronics, Control, and Instrumentation* (IECON), vol. 3, 1996, pp. 1837–1843.

12. Betke, M., Makis, N. Fast Object Recognition in Noisy Images Using Simulated Annealing. In *IEEE International Conference on Computer Vision*, 1995.

13. Kalman, R. E. A New Approach to Linear Filtering and Prediction Problems. *Transaction of the ASME—Journal of Basic Engineering,* 82(Series D), 1960, pp. 35-45.

14. Fang, C.Y., Chen, S. W. and Fuh, C. S., Road Sign Detection and Tracking. IEEE Trans. On Vehicular Technology, Vol. 52, No. 5, 2003, pp.1329-1341.

15. Kalman, R. E. A New Approach to Linear Filtering and Prediction Problems. Transaction of the ASME—Journal of Basic Engineering, 82(Series D), 1960, pp. 35-45.

16. Lowe, D. G. Object Recognition from Local Scale Invariant Features, ICCV, pp.1150-1157, 1999.

17. Lowe, D. G. Distinctive image features from scale-invariant keypoints. In *International Journal of computer vision*, 2004.

18. Gyuri, D and Schmid, C. Object Class Recognition Using Discriminative Local Features. In *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 2004.

19. Pierre, M. and Pietro, P. Common-Frame Model for Object Recognition. In Lawrence K.Saul, Yair Weiss, and Leon Bottou, editors, *Advances in Neural Information Processing Systems* 17. MIT Press, Cambridge, MA. 2005.

20. Mikolajczyk, K. and Schmid, C. A Performance Evaluation of Local Descriptors. In CVPR, 2003, pp.257-263.

21. Farag, A. A. and Abdel-Hakim, A. E. Detection, Categorization and Recognition of Road Signs for Autonomous Navigation, In *Proceeding of Acivs* , Brussels, Belgium, 2004, Aug, 31-Sept.3.