International Journal of Innovative Computing, Information and Control
Volume **9**, Number **7**, July **2013**

ICIC International ©2013 ISSN 1349-4198
pp. 2919–2939

MULTI-STAGE WITH NEURO-FUZZY APPROACH FOR EFFICIENT ON-ROAD SPEED SIGN DETECTION AND RECOGNITION

Hsin-Han Chiang[1], Yen-Lin Chen[2,*] and Tsu-Tian Lee[3]

[1]Department of Electrical Engineering
Fu Jen Catholic University
No. 510, Zhongzheng Rd., Xinzhuang Dist., New Taipei City 24205, Taiwan
hsinhan@ee.fju.edu.tw

[2]Department of Computer Science and Information Engineering
National Taipei University of Technology
No. 1, Sec. 3, Zhongxiao E. Rd., Taipei 10608, Taiwan
*Corresponding author: ylchen@csie.ntut.edu.tw

[3]Department of Electrical Engineering
Chung Yuan Christian University
No. 200, Chung Pei Rd., Chung Li City 32023, Taiwan
ttlee@ntut.edu.tw

Received April 2012; revised August 2012

Abstract. *This paper presents an automatic speed sign recognition system to provide speed limit awareness functions for driver assistance. To reduce the influence of digital noise caused by ambient lighting conditions and sign pollutions, an efficient segmentation method based on pan-red color information was applied to extracting the shape of a speed sign. Based on the edge gradient information of a circular shape, a radial symmetry detection strategy is proposed for fast detection of the speed sign candidates from various urban road scenes. The recognition of the information of a speed sign is achieved through the proposed fuzzy adaptive-kernel-based learning vector quantization (FAKLVQ) approach, which also verifies each candidate to eliminate non-target blobs. Experiments demonstrated the feasibility and effectiveness of the proposed system under a wide variety of outdoor conditions.*

Keywords: Speed sign recognition, Color segmentation, Radial symmetry detection, Fuzzy, Learning vector quantization (LVQ)

1. **Introduction.** Most traffic accidents are caused by driver carelessness under traffic conditions, and one of the promising approaches is the driver-assistance of detecting on-road traffic environment that supports drivers in taking safe driving precautions. Recent studies on vision-based driver assistance systems attempt to identify traffic signs, vehicles, obstacles, lane marking, pedestrians, and other patterns in on-road traffic scenes from captured image sequences using image processing and pattern recognition techniques [1-5]. Traffic signs worldwide provide the driver with crucial information for safety and efficiency. Among them, speed signs are critical for alerting drivers to changes in the speed limit. It was reported that excessive speed increases the risk of crashing or vehicle instability, causing injury and even death [6]. The main reason is failure to notice a speed sign because of driver distraction or adverse conditions that impede visibility of speed signs [7]. Moreover, drivers are easily distracted or blinded by the headlights of oncoming vehicles or digital billboards at night. Vision systems aid humans in the performance of several vision related tasks, which led to crucial research in computer vision studies of driving assistance systems (DAS) [8,9].

A two-stage process is widely used to recognize speed signs, namely, detection and recognition. Two main methods of color- and shape-based segmentation were applied in the detection stage. Several studies exploited the assumption that the wavelength arriving at the camera from a traffic sign object is invariant to the intensity of the incident light; therefore, the transformation into an image can preserve this relation [10-13]. A number of other works often relied on shape-based segmentation in association with color [14-17]. Once the candidate regions have been extracted from the image, the candidate objects of the traffic signs must be classified or recognized to obtain their meaningful information. Numerous recognition studies applied soft-computing techniques, such as neural networks (NNs) and support vector machines (SVMs) [18,19]. Although current techniques exhibit satisfactory performance for speed sign detection and recognition in normal conditions, improvements with regard to computational costs, scale variance, and partial occlusions must be addressed. For example, a crucial issue for driver assistance is the quick identification of the candidate object from the image for classification into the correct category under the vast range of ambient lighting conditions and environmental complexity, which occurs in various road scenes. Moreover, to inform the driver of the maximal allowed speed, the digit number recognition lacks a technically satisfactory solution to manage the shape and appearance variations of the digits, especially in small scale and resolution.

Several proposed detection methods used a shape-based approach because of its robustness against weather and varying lighting conditions; however, these methods experience difficulties in recognizing speed signs from several objects with similar shapes. Therefore, the recognition strategy must be further coordinated to verify the correct speed signs. Neural Networks (NNs) and pattern matching are two commonly used tools for sign recognition. An automatic road-sign detection and recognition system was proposed in [20] based on a computational model of human visual recognition processing. Satisfactory recognition results were obtained because the tracking system was augmented. The advantage of their applied NNs is its feasibility for establishing a mapping between objects and their class labels. However, it resulted in a limited recognition performance because of the fixed architecture of network, and new patterns cannot be recognized without retraining the entire network. An advantage of pattern matching is that it can be easily extended by adding new patterns to the existing list. In [21], the recognition was conducted by directing the tele-camera to the target to capture the sign in a larger size. This causes a problem of pattern matching because the computational complexity increases considerably in conjunction with the number of patterns. The other problem is the memory usage of data base which also restricts the system validity.

The learning vector quantization technique (LVQ) was first proposed by Kohonen [22] in 1990, who demonstrated its ability to fine-tune the class boundaries of a self organizing mapping (SOM) network based on a competitive learning scheme. The SOM technique has been well applied on the retrieval system of specific image content [23], and also widely implemented to identify handwritten cursive words. However, LVQ experiences considerable problems for its unsupervised clustering. An improved LVQ was proposed by modifying the LVQ objective function to provide the winner and non-winner with differing importance [24]. Although this method results in the minimized loss function of measuring the locally weighted error of the input vector with respect to the winning prototype, the network behaves erratically for particular scaling of the input pattern. To alleviate this problem, the fuzzy LVQ (FLVQ) was proposed to be used as a batch algorithm with the learning rate derived from fuzzy memberships [25]. The so-called prototype under-utilization problem, that is, only the winner neuron will be updated for each input, can be overcome with the hybridization of the fuzzy c-means (FCM) clustering algorithm and the LVQ. However, the FLVQ uses the Euclidean distance measure, which

is only effective when a data set is distributed in a spherical-like manner, but not robust against noises and outliers. The evaluation results in various literature have indicated the evidence that the kernel-based clustering approach performs better than its original algorithm for the same data sets used in the experiment. Incorporating the kernel-based approach with fuzzy c-means algorithm is proposed in [26,27], and the experimental results demonstrate its superior performance to none-kernel-based algorithms. In this paper, we advance the competitive learning and incorporate the kernel-based approach to the batch type of fuzzy competition with the updating of the kernel parameter. We call it a fuzzy adaptive-kernel-based LVQ (FAKLVQ), which efficiently improves the FLVQ, and thus, the robustness of recognizer against noises and outliers can be considerably enhanced.

An automatic speed sign detection and recognition system is presented in this paper. The detection and recognition system consists of three stages:

1) Candidate blobs of circular signs are extracted from the input image by using pan-red thresholding in HIS color space.
2) The obtained blobs from segmentation are identified by the proposed radial symmetry detection strategy and a coarse-to-fine scheme. The candidates of speed sign can be found.
3) The recognition process is based on the proposed FAKLVQ, which determines a recognized digital number associated to the specified digit features.

The overall structure of our algorithm is illustrated in Figure 1. The images were obtained from real traffic scenes using a digital camera, and the detection and recognition process started by searching objects with similar colors to speed signs. To suppress the effects of possible noise caused by ambient lighting conditions and pollutions, an efficient segmentation method based on pan-red color information was applied to extract the shape of speed signs. Based on the edge gradient information of circular-shaped objects of potential speed signs, the proposed radial symmetry detection strategy was subsequently conducted for fast detection of the speed sign candidates from various urban-road scenes. Once the objects of possible candidates were extracted, the features of the geometrical properties of speed signs were analyzed to discard false candidates. Thereafter, the clustered candidate blobs were recognized by the proposed batch FAKLVQ approach, which verifies the inner area of each candidate to eliminate non-target blobs. The subsequent frame was processed when no more suitable objects were recognized as speed digits. The current recognized speed limitation numbers were subsequently displayed to inform the drivers.

2. **Pan-Red Color Segmentation.**

The RGB color space, commonly known as tri-color space, is often applied, and the hypothesis that each color is formed by the red, green, and blue components in different proportions is widely adopted in the color-based image processing applications. Due to the indivisible correlations between the RGB color components, it is hard to extract and analyze the discriminative features of colored object regions directly in RGB color space, ex. the darkness and lightness features of the colors. Thus, transforming color components in RGB space into HSI color space can offer significant efficiency for perceiving and analyzing applicable color features for region segmentation. The HSI color space is composed of three color components, hue (H), saturation (S) and intensity (I). The H component can reflect the chromatic attribution of a color, the concentration of a color can be determined from the S component, and the luminance feature can be obtained from the I component. Therefore, we transformed the colors of the image from RGB color space into HSI color space, and found the object regions with pan-red colors in the image through the HSI color space. The conversion

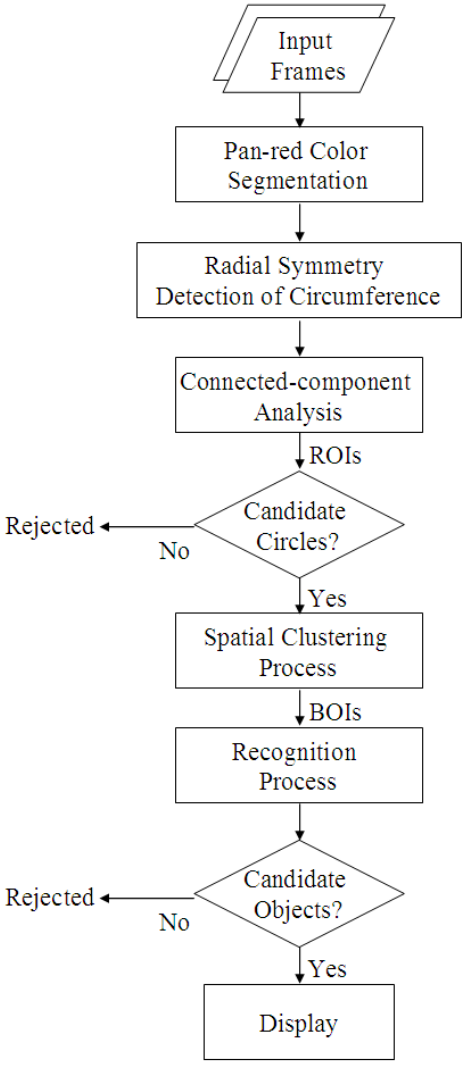


FIGURE 1. The overall structure of proposed speed sign recognition system

equations between the two color spaces can be computed as,

$$h = \cos^{-1}\left(\frac{0.5\left[(r-g)+(r-b)\right]}{\sqrt{(r-g)^2+(r-g)(r-b)}}\right) \tag{1}$$

$$s = 1 - \min(r, g, b) \tag{2}$$

$$i = (r+g+b)/3 \tag{3}$$

where r, g, and b are the normalizing values of RGB space in the range $[0, 1]$ and that

$$r + g + b = 1 \tag{4}$$

$h \in (-180°, 180°]$, $s \in [0, 1]$, and $i \in [0, 1]$ refer to corresponding components in the HSI space.

For color distribution of HSI color space, by analyzing the phenomenon when the component i was at a random value, red mainly lied at direction of 0° on the hue axis. Although the hue values may slightly vary with respect to the changes of ambient lighting conditions, from our previous experimental studies [28,29], their variations can be investigated during the statistical analysis and are mostly within the range $[-60°, 60°]$.

Therefore, the scope of the pan-red color is defined as

$$h \in [-60^\circ, 60^\circ] \tag{5}$$

To obtain the corresponding pan-red color space in RGB space, from (1) to (3), the condition (5) also yields the following conditions:

$$r > g \tag{6}$$

$$r > b \tag{7}$$

The analysis results specify the required condition for the pan-red color detection for candidate object regions of the speed signs. Figure 2 exhibits that the original image and the segmented image in which the possible pan-red object regions of the speed sign candidates can be effectively detected.

(a) (b)

FIGURE 2. Segmentation result from HSI color space. (a) Original image. (b) Result of pan-red segmentation.

3. **Speed Sign Detection.** The common characteristic of speed signs is that they have a white plate surrounded by a red border. Hence, the two essential features, including colors and shapes, can be adopted for detecting the position of a speed sign in a captured image in front of the host car. After the segmentation stage, image pixels belonging to the pan-red color category are extracted as connected regions. Here the circular-shaped regions are prospective object regions of speed signs among the object regions with the pan-red color features, which are obtained by the segmentation stage. However, most conventional shape detectors cannot offer sufficient robustness on handling non-closed object regions, and conventional robust detectors; for example, the generalized Hough circle detector (GHT), may suffer high computational costs and memory storage requirements in detecting circular regions in large-sized images. Several detection methods with the limited performance have also been discussed in recent literatures.

To overcome the mentioned problems above, this study adopts the radially symmetric characteristics of speed signs to efficiently detect the circular regions that frequently appear in several consecutive image frames and reflect a high count of pixels in proportion to the radius. For producing the edge map images, the image obtained from the pan-red segmentation need to be undergone a color-to-gray transformation to obtain a gray intensity image. This study used the luminosity equation, which forms a weighted average to

accommodate for human perception, as follows:

$$G = 0.21r + 0.71g + 0.07b \tag{8}$$

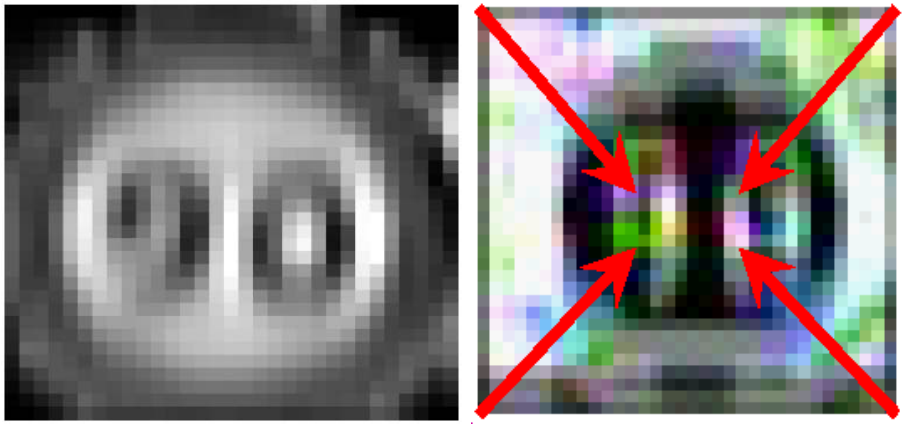
Subsequently, two 3×3 Sobel operators featuring first-order differential operation are applied to extract all horizontal and vertical edges of the gray intensity image. The edge image presents the gradient magnitudes and directions comprised by the potential objects, and its equations are

$$\begin{aligned} &G(i,j) = f(G_x(i,j), G_y(i,j)) \\ &\text{with} \quad \begin{array}{l} G_x(i,j) = G(i+1,j) - G(i,j) \\ G_y(i,j) = G(i,j+1) - G(i,j) \end{array} \end{aligned} \tag{9}$$

The gradient magnitude for a pixel (i, j) is computed from

$$G(i,j) = |G_x| + |G_y| \tag{10}$$

If the larger G values are obtained, the illumination gradient features of the gray intensities of adjacent pixels would be larger and reflect sufficiently significant edge features. Figure 3(a) is a sample speed sign with a red border being transformed into a gray intensity image. The inner round edge is one slope with higher gray intensities than those on the outside, as indicated by the red arrows in Figure 3(b). Those arrows point from the regions with lower gray intensities to those with higher gray intensities, in the following order: 1/4 inner round edge in upper-left corner, 1/4 inner round edge in bottom-left corner, 1/4 inner round edge in bottom-right corner, and 1/4 inner round edge in upper-right corner.



(a) (b)

FIGURE 3. Detected region of a candidate speed sign. (a) Grayscale diagram. (b) Gradient direction characteristics of edge.

Based on these changing features in the intensity and direction of the gradients, a corner detection scheme can be applied to determine whether a pixel is a circular corner. Because some noise will appear in a real image in which a corner cannot be perfectly detected, our system uses the optimal corner detector [30], and can be computed as

$$O_w(i,j) = c \sin \frac{-\pi i}{W} \left[-(e^{zW} + e^{-zW}) + e^{jW} + e^{-jW} \right] \tag{11}$$

$$O_d(i,j) = c \sin \frac{\pi i}{W} \sin \frac{-\pi j}{W} \tag{12}$$

Equations (11) and (12) are applied for the correlation filters with the masks of the optimal corner detector; W reflects the size of the mask; c reveals the scaling factor of the mask; and z denotes the factor for determining the capacity on filtering the noise around the corner. The corners of all the contained objects can be obtained by performing the convolution computation on a particular image using the masks to improve the speed and the capacity of detecting circular corners [28,29]. To coincide with the gradient feature, as shown in Figure 3(b), the optimal corner detector was designed with four approximate masks that serve as circumference arcs in 45°, 135°, 225°, and 315° angles.

After obtaining the gradients of the corners that present possible portions of a circular object, each pixel is labeled by specific values through various direction combinations based on the horizontal gradient features G_x and the vertical gradient features G_y. For each pixel, its corresponding gradient direction value p_m for the four partitions of a circle can be determined as

$$p_m = \begin{cases} 1, & \text{if } G_x > 0 \text{ and } G_y > 0 \\ 2, & \text{if } G_x < 0 \text{ and } G_y > 0 \\ 3, & \text{if } G_x < 0 \text{ and } G_y < 0 \\ 4, & \text{if } G_x > 0 \text{ and } G_y < 0 \\ 0, & \text{otherwise.} \end{cases} \tag{13}$$

Accordingly, the direction features of each edge pixel, which are candidates of parts of circular objects, are efficiently obtained. A single edge pixel would vote the circle center in the direction being orthogonal to the edge direction of this pixel. In processing an edge image, a set of voting masks are established and adopted for associating with various radii features. When several edge pixels align around a boundary of a particular circular object, each edge pixel will vote along their orthogonal lines for the circle center to lead to a peak in this voting space. The radius space is a map with separate voting spaces for each radius to simplify the interpretation. A sample map of the radial symmetric voting space obtained from a circular object is shown in Figure 4.

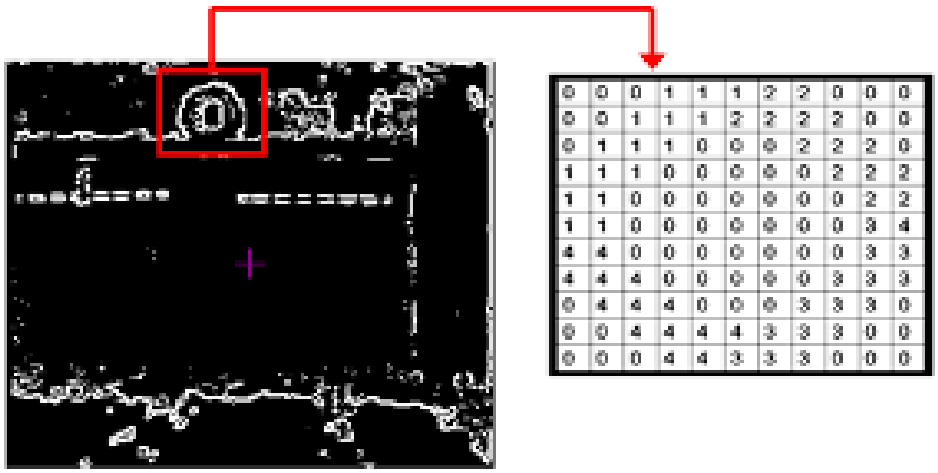


FIGURE 4. Image including a speed sign and its corresponding sample of radial symmetry voting space

Before introducing the vote process for a possible centre of circle, the required parameters are accordingly defined as follows:

(i). Edge image, denoted by E_{sub}: E_{sub} is an $N \times N$ matrix of the sub-image inside the edge image.

(ii) Phase image, denoted by P_{sub}: P_{sub} is an $N \times N$ matrix of the sub-image which reflects the phase variation of edge image E_{sub}. The value of each pixel is referred to (13).

(iii) Voting mask, denoted by V_{mask}: the center of a circle will be on the direction of the gradient of each edge pixel. The voting mask was designed as a discrete form that reveals the gradient values (defined as in (13)) of a circular object, and it was incremented for each pixel based on the results of the orientation votes. The location of a pixel of interest, which will be voted as a candidate center, is summarized as the following process:

(1) Gradient comparison process: compare the gradient value of each pixel between the corresponding edge image P_{sub} and phase image V_{mask} in $N \times N$, and the result is computed as

$$K_{phase}(i,j) = \begin{cases} 1, & P_{sub}(i,j) = V_{mask}(i,j) > 0 \\ 0, & \text{Otherwise} \end{cases} \tag{14}$$

That is, $K_{phase}(i,j) = 1$ indicates that only consistent gradient comparison results between P_{sub} and V_{mask} will be considered.

(2) Edge comparison process: compare edge image E_{sub} associated with K_{phase} in $N \times N$,

$$I_{mask} = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} E_{sub}(i,j) K_{phase}(i,j) \tag{15}$$

When the intensity I_{mask} is larger than a predefined threshold value, Thr, the central point of the voting mask V_{mask} can reveal a circle center; otherwise, it cannot. It is difficult to determine the quantity of Thr. From the histogram of I_{mask} computed over a large quantity of images, we defined a best-fitted selection interval for Thr as follows:

$$Thr = \kappa \cdot NV_{mask} \tag{16}$$

where NV_{mask} is the number of nonzero elements in V_{mask}, and $\kappa \in (0.6, 1]$.

Because our goal is to extract any potential circle centers, according to our experiments it is preferable for the appropriate range of κ to be close to the right boundary of this interval as values for Thr. We assumed that the large values of I_{mask} have high possibility to associate with the circular border of a speed sign.

After performing the voting process, the speed sign candidates can now be detected and located. Accordingly, the criteria of the geometrical properties of speed sign are applied to gradually removing false candidates. The first criterion requires that the actual speed sign should have a sufficient quantity of edge pixels. Let R denote a candidate speed sign. If the number of edge pixels is lower than 5% of the area of R, then this candidate region is eliminated. Next, the second criterion reflects that the dimension of a speed sign should be sufficiently large, that is,

$$W_R > 10 \text{ and } H_R > 10 \tag{17}$$

where W_R and H_R are the width and height of R, respectively, and their values can be measured by calculating the Euclidian distance between any two adjacent control points. Since the ratio between width W_R and height H_R of a traffic sign is mostly near to 1 due to its typical shape, it is required that a possible speed sign region R should satisfy the third criterion as follows,

$$\min\left(\frac{W_R}{H_R+1}, \frac{H_R}{W_R+1}\right) > 0.6 \tag{18}$$

The speed sign can be detected by searching the transformed images through circular corner extraction and symmetry vote process.

Figure 5 demonstrates the processing results of the proposed traffic sign detection method. In Figure 5(b), the pan-red color segmentation results obtained from the sample image in Figure 5(a) are illustrated. Then Figure 5(c) depicts the results of radial symmetry vote space. Here a few local maxima as a high voting result will result in false

positives, as shown in Figure 5(c). Therefore, the appropriate candidates can be further validated by using the above-mentioned three verification stages, as shown in Figure 5(d). A number of small blobs and irregular blobs from other pan-red objects can be removed. By using a connected-component analysis [31], the extracted sign is shown in Figure 5(e), in which the approximation to the center and the radius of the circular is also provided. The inner regions with normalized size for the following recognition strategy are shown in Figure 5(f).

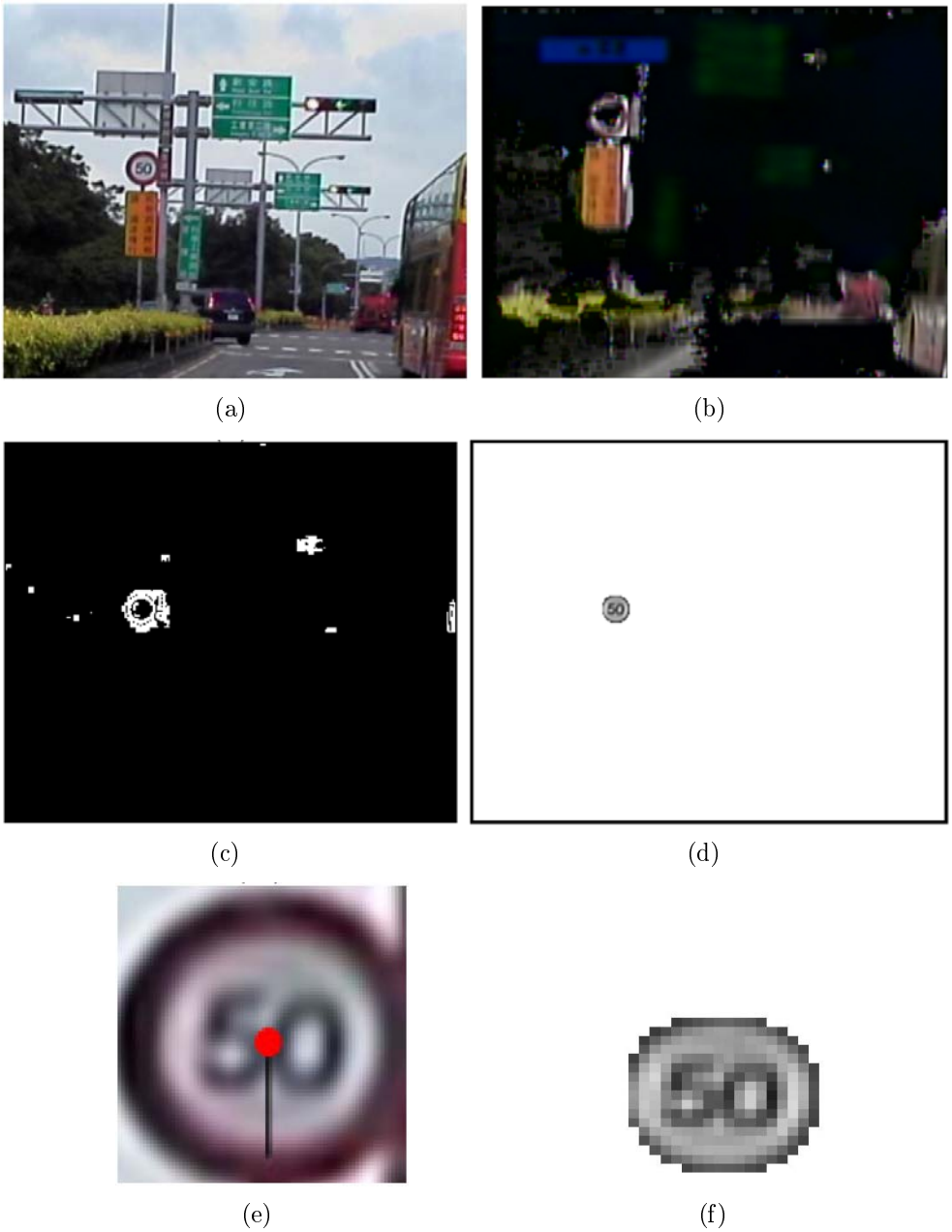


FIGURE 5. Results of traffic sign detection. (a) Original image. (b) Segmentation by pan-red color. (c) Sample of radial symmetry vote space. (d) Verified ROI. (e) Extracted sign with center and radius. (f) Inner regions of normalized size.

4. **Speed Sign Recognition.** The recognition process is activated once the candidate blobs are extracted by the detection stage. The knowledge-based object extraction and identification can be well applied to efficiently extract the digital contents with different characteristics from each candidate plane [32]. With the image of blob of interests (BOIs) and the extracted connected components of these BOIs, various candidates of the traffic sign numbers can subsequently be extracted in the raster-scanning order from the binary image, which are mapped either to a specific digit class or explicitly classified as noises. In the recognition strategy, this study presents the FAKLVQ with a batch version algorithm that advances the FLVQ approach by inducing a kernel-based metric of distance measure. In the following, we first briefly review the batch version of LVQ and FLVQ.

LVQ is the simplest case of SOM, which only updates the weight vector associated with the winning neuron for every data point [22,23,24,27]. The winning neuron w among all competitors is determined by the nearest matching condition:

$$||x_k - W_w(t)|| = \min_i ||x_k - W_i(t)|| \tag{19}$$

and then the LVQ updates its neuron as follows:

$$W_i(t+1) = W_i(t) + \alpha_i(t) h_{ik}(x_k - W_i(t)) \tag{20}$$

where W_i is the weighting vector of the neuron i the input pattern x_k appeared at time t, α_i is the learning rate of the neuron i in the range $[0, 1]$ and decays monotonically with time, and h_{ik} is an indicative function, the value of which is 1 if neuron i is the winning neuron (i.e., $i = w$); otherwise, it is 0.

Inspired by the success of FCM, the FLVQ introduces the fuzzy membership by representing the learning rate α_i as a number of monotone functions of u_{ik}. When the Euclidean distance metric is used, the update equation of fuzzy membership value between the input pattern and a weighting vector of neuron is calculated by

$$\begin{aligned} u_{ik}(t+1) &= \left(\sum_{j=1}^{c} \left(||x_k - W_i(t)|| / ||x_k - W_j(t)|| \right)^{2/(m'_t - 1)} \right)^{-1}, \text{ for } 1 \le i \le c. \\ &\text{with } m'_t = m'_0 + t(m'_f - m'_0)/t_{\max}. \end{aligned} \tag{21}$$

where c is the number of categories, m'_t denotes the degree of fuzziness and decreases with time, m'_0 and m'_f are the initial and final values of the fuzzifier, respectively (it was suggested that $1.1 < m'_f < m'_0 < 7$ [26]), and $t_{\max}$ is the iteration limit. The prototype update equation of each weighting vector is

$$W_i(t+1) = W_i(t) + \frac{\sum_{k=1}^{N} \alpha_{ik}(t)(x_k(t) - W_i(t))}{\sum_{k=1}^{N} \alpha_{ik}(t)} \tag{22}$$

where the fuzzified learning rate is defined as $\alpha_{ik}(t) = (u_{ik}(t))^{m'_t}$.

In several cases, the input patterns are highly nonlinear and cannot be separated by a linear function, and the FLVQ using the Euclidean norm easily causes inferior performance. A solution is to transform the input data into a higher or infinite feature space by using a mapping function Θ, and replace the Euclidean distance as a kernel-induced one, as follows:

$$\begin{aligned} d^2(\Theta(W_i), \Theta(x_k)) &= ||\Theta(x_k) - \Theta(W_i)||^2 \\ &= \Theta(x_k)^T \Theta(x_k) + \Theta(W_i)^T \Theta(W_i) - 2\Theta(x_k)^T \Theta(W_i) \\ &= K(x_k, x_k) + K(W_i, W_i) - 2K(x_k, W_i) \end{aligned} \tag{23}$$

Because the training data are used through a dot product, the kernel function is subsequently defined as $K(x_i, x_j) = \Theta(x_i)^T \Theta(x_j)$ for x_i and x_j in input space. In this study, the Gaussian kernel was used as follows:

$$K(x_i, x_j) = \exp\left(\frac{-||x_i - x_j||^2}{2\sigma^2}\right) \tag{24}$$

where σ corresponds to the kernel width of the neurons and also determines the degree of neuron excitation.

By minimizing (23) by gradient descent, the batch update formula for W_i can be presented as

$$\begin{aligned} W_i(t+1) &= W_i(t) - \frac{\sum_{k=1}^{N} \alpha_{ik}(t) \left(\frac{\partial d^2(\Theta(W_i), \Theta(x_k))}{\partial W_i}\right)}{\sum_{k=1}^{N} \alpha_{ik}(t)} \\ &= W_i(t) - \frac{2}{\sigma^2} \frac{\sum_{k=1}^{N} \alpha_{ik}(t) K(x_k, W_i(t))(x_k - W_i(t))}{\sum_{k=1}^{N} \alpha_{ik}(t)} \end{aligned} \tag{25}$$

where the fuzzified learning rate is defined as

$$\alpha_{ik}(t) = (u_{ik}(t))^{m'_t}, \quad 1 \le i \le c \tag{26}$$

Based on the Gaussian kernel in (24), the fuzzy membership value of the input x_k with respect to the ith neuron can be derived as follows:

$$\begin{aligned} u_{ik}(t+1) &= \left(\sum_{j=1}^{c} \left(||\Theta(x_k) - \Theta(W_i)||/||\Theta(x_k) - \Theta(W_j)||\right)^{2/(m'_t - 1)}\right)^{-1} \\ &= \left(\sum_{j=1}^{c} \left((1 - K(x_k, W_i))/(1 - K(x_k, W_j))\right)^{1/(m'_t - 1)}\right)^{-1} \end{aligned} \tag{27}$$

with $m'_t = m'_0 + t(m'_f - m'_0)/t_{\max}$.

As in the batch version algorithm of LVQ, the expectation values of $W_i(t+1)$ and $W_i(t)$ in the kernel-based FLVQ will be equal and converge into a stationary weighting vector W_i^* while t goes infinity; that is,

$$E[\alpha_{ik}(t)(x_k(t) - W_i^*)] = 0, \quad t \to \infty \tag{28}$$

All weighting vectors of neurons are updated after the whole set of input patterns is presented.

Since the Gaussian kernel function (24) is used in this study, the kernel parameter also plays the important role of the clustering performance. Instead of using a constant value, in this study, the kernel parameter is updated during the learning process. Recall that in the FCM algorithms, a given data set $X = \{x_1, x_2, \ldots, x_N\}$ is partitioned into c fuzzy categories to minimize the following objective function:

$$J_N(X) = \sum_{k}^{N} \sum_{i}^{c} u_{ik} ||x_k - W_i||^2 \tag{29}$$

This objective function provides the fuzzy clustering with a well-known measurement of the clustering quality. By using the kernel function from the input data space to the

feature space, it needs to minimize the reformulated objective function as:

$$\begin{aligned} J_N(X) &= \sum_{k}^{N}\sum_{i}^{c} u_{ik}||\Theta(x_k) - \Theta(W_i)||^2 \\ &= \sum_{k}^{N}\sum_{i}^{c} u_{ik}(2 - 2K(x_k, W_i)) \end{aligned} \tag{30}$$

The kernel parameter σ is initialized and then can be updated as

$$\sigma(t+1) = \sigma(t) - \eta_\sigma \frac{\partial J_N(X)}{\partial \sigma} \tag{31}$$

where $\eta_\sigma > 0$ denotes the updating rate of the kernel parameter. Based on the above discussion, the algorithm of batch type of FAKLVQ is summarized in the following:

1. Fix the number of category c, and select a termination $\varepsilon > 0$.
2. Select $t_{\max}$, m_0', and m_f'. Initialize weighting vector for each neuron, $W_1(0)$, $W_2(0)$, ..., $W_c(0)$, fuzzy membership u_{ik}, and kernel parameter $\sigma(0)$.
3. For t= 1, 2, ..., $t_{\max}$.
   (a) For k = 1, 2, ..., N (the number of input patterns)
      (i) Calculate the fuzzy learning rate from (27).
      (ii) Update the kernel parameter from (31).
      (iii) If $k = N$ stop; otherwise, next k
   (b) Update all weighting vectors by using (25) and (26).
   (c) Calculate the object function $J_N(X)$.
   (d) Compute

$$E(t) = \sum_{i=1}^{c} ||W_i(t+1) - W_i(t)||^2 \tag{32}$$

   (e) If $E(t) < \varepsilon$ stop; otherwise, next t.

During the process of clustering, the values of u_{ik} and σ are continually updated, and this process stops until the condition of (32) is satisfied. Note that the updated rate σ should be always greater than zero in each iteration. The value of ε will be assigned to σ while it converges to zero or becomes a negative number.

The proposed FAKLVQ performs efficiently if the initial weighting vectors are close to their final values, and thus, the weighting initialization is required before the training process. First, we input all the samples of binary image for each number. The features of *hollow* and *concave* for each number digit were selected based on the region growing algorithm [33]. Figure 6 illustrates the features of number digit '6' after a conventional edging process. After obtaining the quantity and the prototypes of output neurons of FAKLVQ, the initial weighting vector of each neuron can be selected as the mean of features for each number digit. In the weighting update strategy, each training image of number is with size 25×25. An average of 50 training patterns, some including samples of noisy objects that may be confused, was used for each number digit of the speed sign. The recognition performance varied with the number of feature category. In Table 1, the comparison of FAKLVQ with batch LVQ and FLVQ is shown in regards to changing values of c. This experimental result indicates that the proposed FAKLVQ with the optimized kernel parameter provides superior recognition performance compared with those of the classical batch LVQ and FLVQ. When c is set to 6, the system can achieve optimal recognition results with a recognition rate of 97.62%.

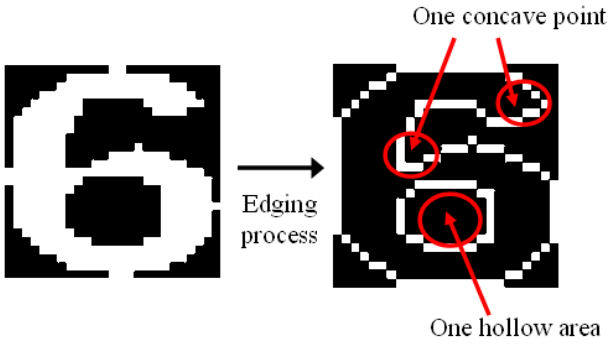


FIGURE 6. The binary image of "6" (left) and the exampled features of inner and outer shapes (right)

TABLE 1. Recognition accuracy in selected category

Value of c	LVQ	FLVQ	FAKLVQ
4	81.52%	88.23%	94.84%
5	83.29%	90.52%	95.08%
6	86.02%	91.97%	97.62%
7	86.33%	91.71%	96.22%
8	84.91%	90.30%	94.77%

5. **Experimental Results.** To examine the performances of our approach, test sequences were recorded by a CCD camera (with focal length 40mm and lens 1/3") fixed onto the front windshield of a vehicle under a normal driving speed on several public roads. The standard size of speed signs with a round shape is 65cm in diameter. Based on Gaussian imaging formula, the speed sign images ranging from 33×33 to 13×13 in size correspond to the detection distance at 50 to 100 meters ahead. In the detection procedure, only the upper half of each image was managed because most traffic signs appear in the region alongside the road and below the sky. In this way the required detection time can be further reduced by half. The proposed system was implemented on a Pentium 1.8GHz platform with 512 RAM. The frame rate of the vision system was 30 frames per second, and the size of each frame of the captured image sequences was 320 pixels by 240 pixels per frame.

Various experimental sets under various ambient lighting and weather conditions (such as sunny, cloudy, rainy, and nighttime) were used for comparative performance evaluation on speed sign detection and recognition. Two representative examples of speed signs with the 40 and 50 (KM/h), which usually appear on urban roads, are shown in Figures 7 and 8, respectively. The sub-images present the main result in each procedure, where (a) represents the original image, (b) is the image through pan-red color segmentation, (c) is the edge image formed by pan-red pixels, (d) indicates the center location of speed sign by using the proposed detection approach and the verification stages, (e) is the extracted ROI, (f) is the BOI for recognition process, (g) presents the division of decimal and unit number, and (h) shows the recognition result. In addition to inferior illumination on cloudy days, these two examples demonstrate that our system also works when the signs are not perpendicular to the movement of the vehicle. The whole images were searched, and the invariance of translation was supported by the fact that the speed signs appearing in the left and right side were successfully detected. In Figure 7, the number digits can be effectively recognized, even if it appears oriented with no ideal position. Regardless of whether the decimal and the unit number are consecutive in the BOI image, such

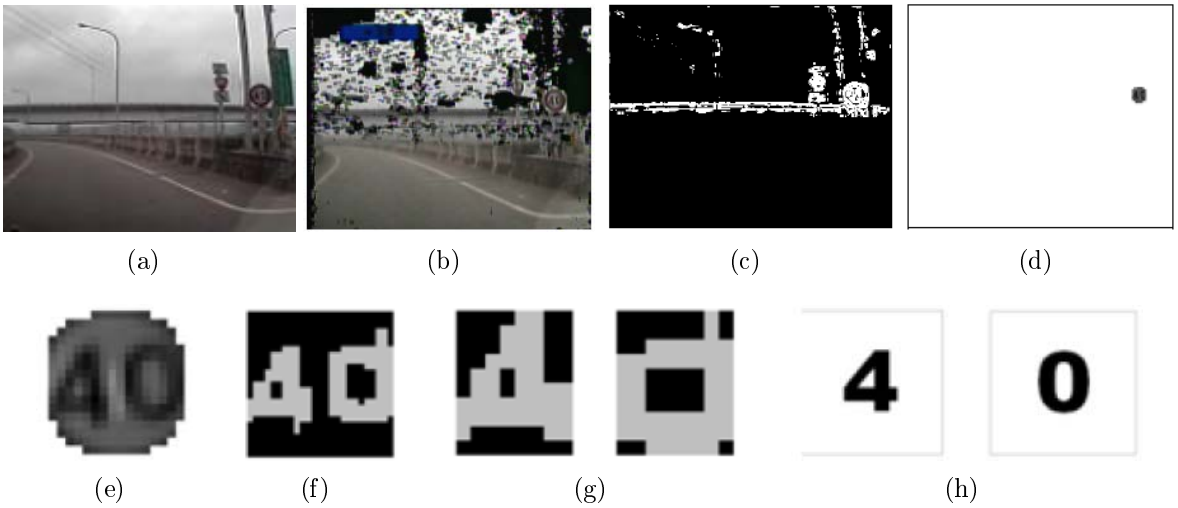


FIGURE 7. Experimental results in the entrance of freeway. (a)-(d) Speed sign detection process. (e)-(h) Speed-sign recognition process.

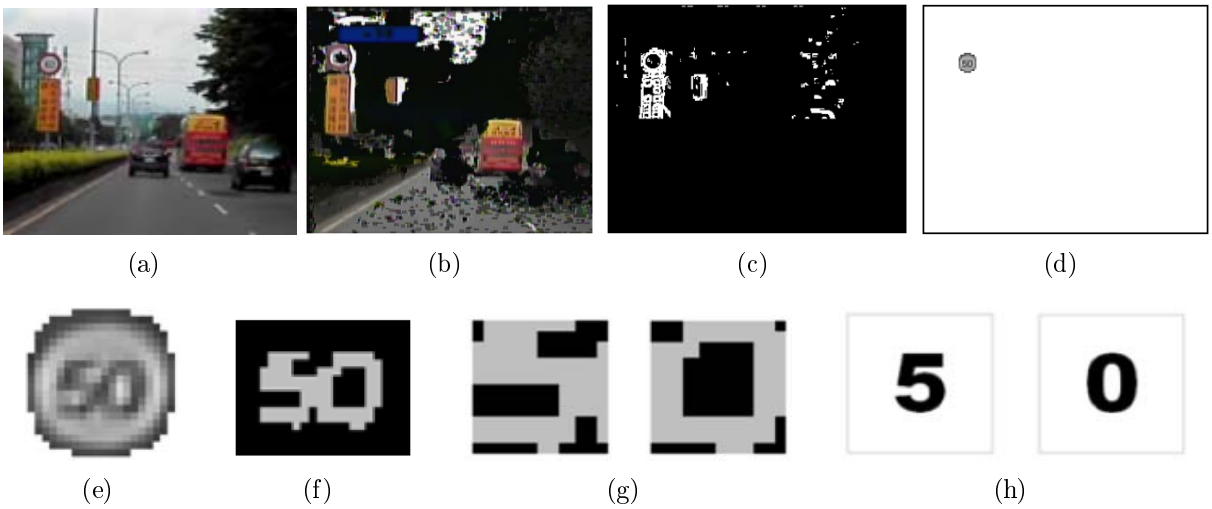


FIGURE 8. Experimental results in the urban road. (a)-(d) Speed sign detection process. (e)-(h) Speed-sign recognition process.

numbers, even in deformed binary image, are separated for successful recognition. By inspecting the images in Figures 7 and 8, the system can recognize contents with various scales as standard speed signs.

Figure 9 demonstrates that our system also works when two limitation signs occur in the same image. Both signs were appropriately detected with the same features of pan-red color edges and circular shapes. The correct speed sign candidate was verified through the proposed recognition scheme because its content was recognized as the digital numbers 5 and 0. In Figure 10, the system is able to detect and recognize the speed sign that are occluded by some steel frames.

In Figures 11 and 12, two examples show the limited sign detection and recognition at night. Speed signs may specularly reflect light at night; however, the correctness of detection and recognition is not affected. Similarly, only the content of speed sign is recognized. Our system was able to detect and recognize the speed sign at night if the objects were illuminated by vehicle headlights or street lights. However, the speed sign

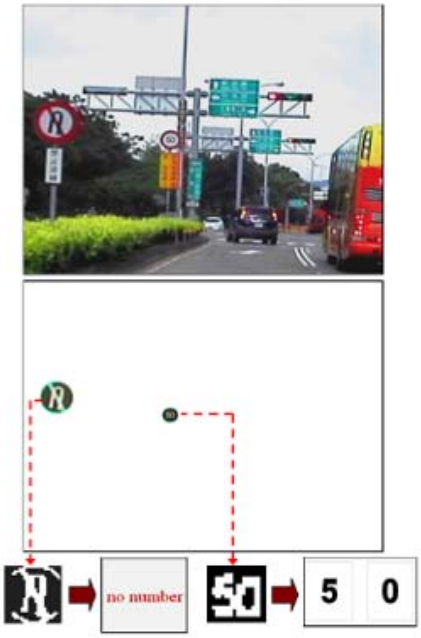


FIGURE 9. Experimental results with two limited signs at one scene

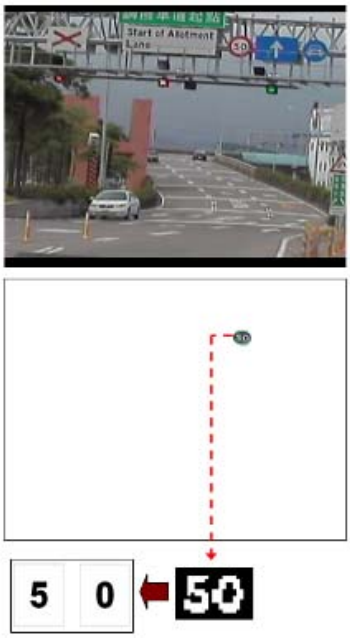


FIGURE 10. Experimental results with the occluded speed sign and its detection and recognition

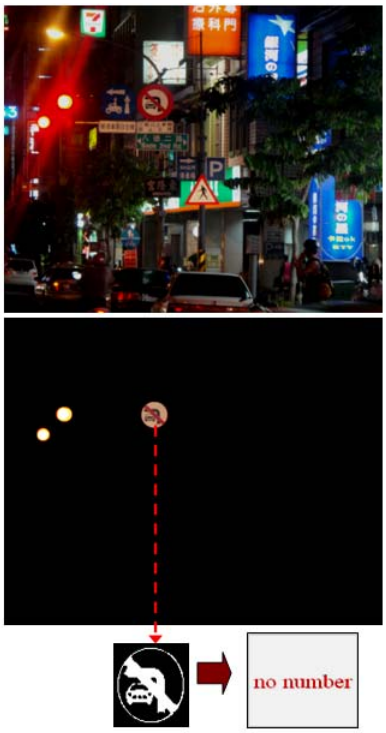


FIGURE 11. Experimental results at nights. The limited sign can be detected and its content is recognized as invalid speed-limited information.

FIGURE 12. Experimental results at nights. The speed sign can be detected and recognized correctly.

may be partially lost at the detection stage when the environmental illumination is too dim to detect the pan-red object.

For the quantitative evaluation of speed sign detection and recognition performance, the measures of the Jaccard coefficient [34] were applied to evaluate performance, and are defined as

$$J = \frac{T_P}{T_P + F_P + F_N} \tag{33}$$

where T_P (true positives) denotes the number of correctly detected and recognized speed signs, F_P (false positives) denotes the number of falsely detected and recognized speed sigs, and F_N (false negatives) is the number of missed detections and recognitions of speed signs.

In each frame, the value of Jaccard coefficient J for speed sign detection and recognition results can be manually counted, and its average value can subsequently be obtained from all frames of the video sequences, as computed from

$$\bar{J} = \sum\nolimits_{N_f} J \Big/ N_f \tag{34}$$

where N_f is the total number of frames of the video sequences.

Quantitative evaluation was performed on the testing video of 5821 frames with a total of 252 speed signs. Table 2 shows the results of evaluation between the proposed approach, support vector machines (SVM) based approach by [19], and fuzzy based approach by [35]. As stated in [19], a number of signs cannot be detected because of the color similarity between the sign and the background; thus, it results in more F_N in the detection phase of speed signs. In addition, because their algorithm mainly uses four distance-to-borders (DtBs) features for recognizing the inner area of each candidate, it easily recognizes several F_P that have similar appearance to digital numbers of speed signs, especially in a small quantity of pixels. In [35], since only the color segmentation algorithm is carried in the detection strategy, it easily misses candidates with faded red color which causes more FN, and also brings more non-candidates with similar color range which leads to more FP in the detection phase of speed signs. Consequently, the recognition performance also degrades. Besides, the membership functions for the inputs of segmented color and shape measures need large numbers of trials. In this comparison, our approach demonstrates considerably superior processing results in the recognition of speed signs.

TABLE 2. Experimental quantitative evaluation

	Our Approach	Approach in [19]	Approach in [35]
Detection (%)	96.13	92.28	88.42
Recognition (%)	95.82	91.92	86.68
Total # of Frames	5821		
Total # of Speed Signs	252		

If the vehicle speed is high, it may occasionally cause camcorder vibration when passing through uneven roads. Thus, the quality of input video sequences is, as shown in the two examples in Figure 13, which were extracted from the testing sequences. The speed signs were detected correctly and the recognition of these two objects were also correct, even though they appear jerky. This indicates that the proposed recognition strategy using FAKLVQ remains robust against deformation of objects. In addition, the repetition of our approach can be exhibited in Figure 14, where four frames of the route are shown. As can be seen, the developed algorithm is able to repeatedly detect and recognize the same sign in four images.

(a) (b)

FIGURE 13. Two examples of recognition results under camcorder vibration. The right images are the detection results and the recognition results.

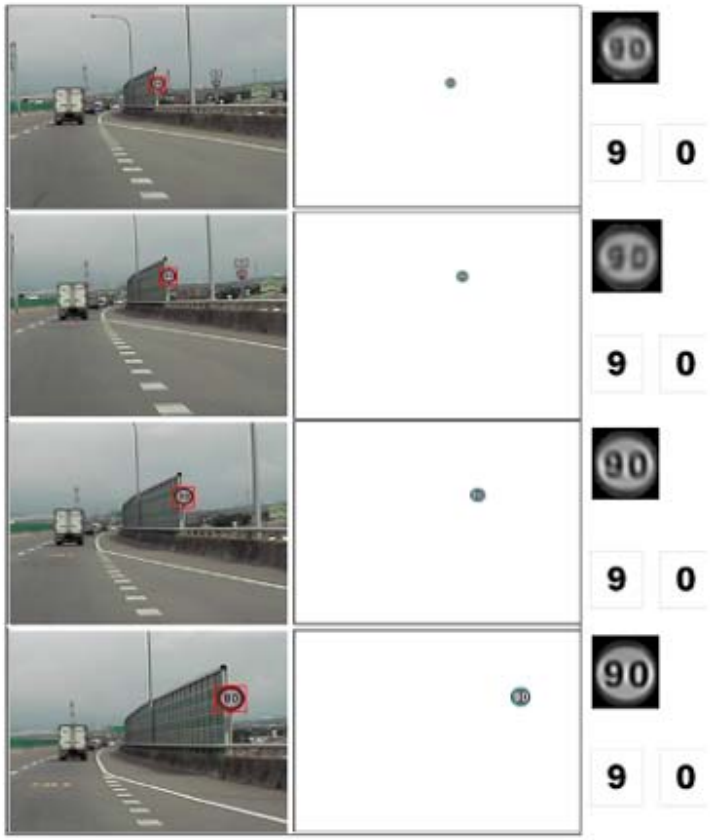


FIGURE 14. Speed sign detection and recognition in a sequence. The size of each image is 320 × 240 pixels, and the time between two successive images is 0.2s.

Our system was run over sequences with varying lighting conditions and visual quality of speed signs. As demonstrated in the experimental examples in which the scene is difficult for color-based approaches, the speed sign can be correctly detected against substantial lighting variation and inferior contrast. A number of false alarms in detection occurred because of the similar geometric appearance and physical property with limited signs; however, these can be reduced by our proposed recognition strategy. The detector may not reliably detect small signs because the number of possible pixels is too small and

these signs are usually too small to be recognized; however, they can be detected and recognized in subsequent images. Although the method of tracking speed signs was not integrated in this work, a candidate sign can be considered as valid in a sequence when it is detected and recognized in at least two frames of a sequence.

Table 3 shows the comparative analysis on the methods of Miura et al. [21], Saturnino et al. [19], Fleyeh [35], Baro et al. [36], and the proposed approach. The system in [21] used two cameras, one for candidate detection and one for capturing candidates in larger regions. Other approaches required only one camera mounted on the windscreen of the vehicle. For the detection strategy, Miura et al. [21] used color and shape information, and screening process in which the Hough transform was applied to extract the candidate edges from regions of interest. A normalized correlation-based pattern matching was also proposed in the speed signs recognition strategy, in which the normalized value between candidates and templates must be calculated to find the best candidate for identification. Thus, this method experiences high computational costs, and the improved reliability of recognition algorithm must also be guaranteed by using a robust tracking method. Saturnino et al. [19] used linear and Gaussian-kernel SVM for detection and recognition of traffic signs, respectively. Their approach performed effectively in shape classification and content recognition; however, a larger dimension of frames is required because only four feature vectors are used in the recognition strategy. This also results in higher computational costs. Fleyeh [35] presented a fuzzy system for detection and recognition of traffic signs. The construction of fuzzy inference systems in both the color segmentation and shape classification strategies involved complicated computations, such as fuzzy input pre-calculation, fuzzification, inference, and defuzzification. Moreover, the proper membership functions must be manually tuned from various trial-and-errors. Baro et al. [36] proposed a boosted detectors cascade, which are trained with a novel evolutionary version of Adaboost. Besides, the recognition strategy applied a multiclass learning technique. This system performed the robustness against a variability of sign appearance; however, the training times and the possible features for training examples demand a great quantity and limit the feasibility of this system. Besides, their detection stage using appearance-based information needed an extra computation cost so that the real-time

TABLE 3. State-of-the-art comparison

Method	Detection strategy	Recognition strategy	Rotation/ vibration invariance	Occlusion handling	Computation cost	Illumination condition adapting
Miura et al. [21]	Edge extraction and Hough transform	Normalized correlation-based pattern matching	Good	Fair	High	Fair
Saturnino et al. [19]	Color segmentation + linear SVMs	Gaussian-kernel SVMs	Fair	Fair	Medium	Fair
Fleyeh [35]	Fuzzy color segmentation	Fuzzy shape recognizer	Fair	Poor	High	Poor
Baro et al. [36]	Evolutionary adaboost detectors	Model-fitting classification	Good	Good	High	Good
Our approach	Pan-red segmentation + radial symmetry voting	FAKLVQ neural network	Good	Good	Low	Good

applications in driver-support was restricted. The proposed system in this study mainly uses radial symmetry detection combined with FAKLVQ neural network recognizer, and the algorithm is efficient and has high accuracy and robustness under the main difficulties, such as image scaling and vibration, occlusion, and illumination variations. Moreover, an efficient FAKLVQ neural network is presented to realize the rapid and reliable content recognition because of its parallel and distributed computation features. The computation time of the proposed system depends on the road-scene complexity of the frame. Most computation time is spent on the radial symmetry vote of the detection stage. For an input video sequence with 320×240 pixels per frame, the proposed system requires an average of 0.431s processing time per frame. Moreover, with regard to memory usage, at least the integral image of the processed image must be in memory during detection. The restriction in memory usage is not the system itself; that is, the parameters used in the detection and the recognition algorithm must also be available in memory. By inspecting the result, the proposed system demonstrated effectiveness in speed sign detection and recognition.

6. **Conclusions.** This paper described a complete approach to detect and recognize speed signs from video sequences, which considers the existing difficulties in outdoor environments regarding object detection and recognition. The detection stage, radial symmetry detection based on the circular diagram was highly effective for speed sign detection. The recognition module was developed based on the batch type of FAKLVQ, which uses the kernel-based approach with the updated fuzzy memberships and kernel parameter, and accurately recognized the digital numbers of speed signs. The major contributions of this paper are summarized as follows:

(a) The segmentation method using pan-red is proposed to extract speed sign pixels from video frames. Even though the speed signs are illuminated under different lighting conditions, our system can effectively detect them from the background.

(b) The optimal corner detector with the fast radial symmetry is proposed for the circular signs, which offers high robustness to image noise. In addition, a coarse-to-fine verification scheme is proposed to verify each speed sign candidate. Most of impossible candidates can be filtered out in this stage, and thus the further recognition process can be performed very efficiently to identify the content of each circular sign.

(c) The proposed FAKLVQ algorithm can achieve efficient training process and real-time capability to recognize the speed-limit information. The speed signs with rotations and occlusions can be accurately recognized. Moreover, this approach is invariant to small vibrations and changes of scale.

The average accuracy of speed sign recognition using our scheme achieves about 96%. Experimental results have proved that the presented methodology offers robustness, effectiveness, and efficiency in practical environments.

Acknowledgements. The authors gratefully thank Reviewers for the helpful comments and suggestions for improving the paper. This work was supported by the National Science Council of Taiwan under Contract Nos. NSC-100-2221-E-030-004, NSC-100-2221-E-027-033, NSC-101-2218-E-033-003, NSC-101-2219-E-027-006, and NSC-101-2219-E-027-007.

REFERENCES

[1] H. Kurihata, T. Takahashi, I. Ide, Y. Mekada, H. Murase, Y. Tamatsu and T. Miyahara, Detection of raindrops on a windowshield from an in-vehicle video camera, *International Journal of Innovative Computing, Information and Control*, vol.3, no.6(B), pp.1583-1591, 2007.

[2] D. Deguchi, K. Doman, I. Ide and H. Murase, Improvement of a traffic sign detector by retrospective gathering of training samples from in-vehicle camera image sequences, *Proc. of the 10th Asian Conference Computer Vision*, Queenstown, New Zealand, pp.204-213, 2010.
[3] B. F. Wu, C. T. Lin and Y. L. Chen, Dynamic calibration and occlusion handling algorithms for lane tracking, *IEEE Trans. Ind. Electron.*, vol.56, pp.1757-1773, 2009.
[4] C. P. Young, B. R. Chang, H. F. Tsai, R. Y. Fang and J. J. Lin, Vehicle collision avoidance system using embedded hybrid intelligent prediction based on vision/GPS sensing, *International Journal of Innovative Computing, Information and Control*, vol.5, no.12, pp.4453-4468, 2009.
[5] J. A. Jang, H. S. Kim and H. B. Cho, Smart roadside system for driver assistance and safety warnings: Framework and applications, *Sensors*, vol.11, pp.7420-7436, 2011.
[6] C. N. Kloeden, A. J. McLean, V. M. Moore and G. Ponte, Travelling speed and the risk of crash involvement, *Tech. Rep. CR 204*, NHMRC Road Accident Res. Unit, Univ. Adelaide Adelaide, Australia, 1997.
[7] H. Ishikawa, C. Miyajima, N. Kitaoka and K. Takeda, Detection of distracted driving using a Bayesian network, *ICIC Express Letters, Part B: Applications*, vol.2, no.3, pp.627-633, 2011.
[8] L. Bai, Y. Wang and M. Fairhurst, An extended hyperbola model for road tracking for video-based personal navigation, *Knowledge-Based Systems*, vol.21, pp.265-272, 2008.
[9] B.-F. Wu, C.-T. Lin and Y.-L. Chen, Robust position estimation and analysis based on a camera model with dynamic calibration for driving assistance systems, *IEICE Transactions on Information and Systems*, vol.E92-D, no.9, pp.1725-1735, 2009.
[10] W. C. Huang and C. H. Wu, Adaptive color image processing and recognition for varying backgrounds and illumination, *IEEE Trans. Industrial Electronics*, vol.45, no.2, pp.351-357, 1998.
[11] M. Benallal and J. Meunier, Real-time color segmentation of road signs, *Proc. of IEEE Canadian Conf. Electrical and Computer Engineering*, Canadian, vol.3, pp.1823-1826, 2003.
[12] H. Fleyeh, Color detection and segmentation for road and traffic signs, *Proc. of IEEE Conf. Cybernetics and Intelligent Systems*, vol.2, pp.809-814, 2004.
[13] L. W. Tsai et al., Road sign detection using eigen colour, *IET Computer Vision*, vol.2, no.3, pp.164-177, 2008.
[14] A. de la Escalera, J. M. Armingol, J. M. Pastor and F. J. Rodriguez, Visual sign information extraction and identification by deformable models for intelligent vehicles, *IEEE Trans. Intelligent Transportation Systems*, vol.5, no.2, pp.57-68, 2004.
[15] S. Xu, Robust traffic sign shape recognition using geometric matching, *IET Intelligent Transportation Systems*, vol.3, no.1, pp.10-18, 2009.
[16] C. Y. Fang, S. W. Chen and C. S. Fuh, Road-sign detection and tracking, *IEEE Trans. Vehicular Technology*, vol.52, no.5, pp.1329-1341, 2003.
[17] Y. X. Liu, T. Ikenaga and S. Goto, Geometrical, physical and text/symbol analysis based approach of traffic sign detection system, *IEICE Trans. Inf. & Syst.*, vol.E90-D, no.1, pp.208-216, 2007.
[18] A. de la Escalera, J. M. Armingol and M. Mata, Traffic sign recognition and analysis for intelligent vehicles, *Image Vis. Comput.*, vol.21, no.3, pp.247-258, 2003.
[19] S. Maldonado-bascon et al., Road-sign detection and recognition based on support vector machines, *IEEE Trans. Intelligent Transportation Systems*, vol.8, no.2, pp.264-278, 2007.
[20] C. Y. Fang, C. S. Fuh, P. S. Yen, S. Cherng and S. W. Chen, An automatic road sign recognition system based on a computational model of human recognition processing, *Computer Vision and Image Understanding*, vol.96, pp.237-268, 2004.
[21] J. Miura, T. Kanda, S. Nakatani and Y. Shirai, An active vision system for on-line traffic sign recognition, *IEICE Transactions on Information and Systems*, vol.E85-D, no.11, pp.1784-1792, 2002.
[22] T. Kohonen, The self-organizing map, *Proc. of IEEE*, vol.78, no.9, pp.1464-1480, 1990.
[23] V. P. S. Rallabandi and S. K. Sett, Knowledge-based image retrieval system, *Knowledge-Based Systems*, vol.21, pp.89-100, 2008.
[24] P. C. Chang, C. H. Liu and C. Y. Fan, Data clustering and fuzzy neural network for sales forecasting: A case study in printed circuit board industry, *Knowledge-Based Systems*, vol.22, pp.344-355, 2009.
[25] J. C. Bezdek and N. R. Pal, Two soft relatives of learning vector quantization, *Neural Networks*, vol.8, pp.729-743, 1995.
[26] K. L. Wu and M. S. Yang, A fuzzy-soft learning vector quantization, *Neurocomputing*, vol.55, pp.681-697, 2003.
[27] H. Zhang and J. Lu, Semi-supervised fuzzy clustering: A kernel-based approach, *Knowledge-Based Systems*, vol.22, pp.477-481, 2009.

[28] H. H. Chiang, Y. L. Chen and T. T. Lee, On-road speed sign recognition using fuzzy kernel-based learning vector quantization, *Proc. of Int. Conf. System Science and Engineering*, Macao, pp.49-54, 2011.

[29] H. H. Chiang, Y. L. Chen, W. Q. Wang and T. T. Lee, Road speed sign recognition using edge-voting principle and learning vector quantization network, *Proc. of Int. Computer Symposium*, Tainan, Taiwan, pp.246-251, 2010.

[30] A. de la Escalera, L. E. Moreno, M. A. Salichs and J. M. Armingol, Road traffic sign detection and classification, *IEEE Trans. Industrial Electronics*, vol.44, no.6, pp.848-859, 1997.

[31] K. Suzuki, I. Horiba and N. Sugie, Linear-time connected-component labeling based on sequential local operations, *Computer Vision & Image Understand.*, vol.89, pp.1-23, 2003.

[32] Y.-L. Chen, Automatic text extraction, removal and inpainting of complex document images, *International Journal of Innovative Computing, Information and Control*, vol.8, no.1(A), pp.303-327, 2012.

[33] J. Garcia-Consuegra, G. Cisneros and E. Navarro, A sequential echo algorithm based on the integration of clustering and region growing techniques, *Proc. of IEEE Geoscience and Remote Sensing Symposium*, vol.2, pp.648-650, 2000.

[34] P. Sneath and R. Sokal, *Numerical Taxonomy, the Principle and Practice of Numerical Classification*, W. H. Freeman, New York, 1973.

[35] H. Fleyeh, Traffic sign recognition by fuzzy sets, *Proc. of IEEE Intelligent Vehicles Symposium*, pp.422-427, 2008.

[36] X. Baro, S. Escalera, J. Vitria, O. Pujol and P. Radeva, Traffic sign recognition using evolutionary adaboost detection and forest-ECOC classification, *IEEE Trans. Intelligent Transportation Systems*, vol.10, no.1, pp.113-126, 2009.