Proceedings of the
2007 IEEE Intelligent Vehicles Symposium
Istanbul, Turkey, June 13-15, 2007

Vision-Based Forward-Looking Traffic Scene Analysis Scheme

Jyh-Yeong Chang and Chien-Wen Cho

Abstract—**Vision-based driver assistant systems are very promising in Intelligent Transportation System (ITS); however, algorithms capable of describing traffic scene images are still very difficult to date. This paper proposes a system which can segment forward-looking road scene image into natural elements and detect front vehicles. First, the scene analysis system deals with scene segmentation and natural object labeling of forward-looking images. By the use of fuzzy Adaptive Resonance Theory (ART) and fuzzy inference techniques, the scene analysis task is accomplished with tolerance to uncertainty, ambiguity, irregularity, and noise existing in the traffic scene images. Secondly, the proposed system can detect the front vehicles and utilize a bounding box shape to further refine the segmentation result. Compared with conventional approaches, the proposed scheme can analyze forward-looking traffic scenes and yield reliable and efficient segmentation results. The validity of the proposed scheme in car detection was verified by field-test experiments. The traffic scene segmentation and front vehicle detection are successful.**

I. Introduction

IT is important to develop vision-based safety enhancement systems in the research of Intelligent Transportation System (ITS). There are various man-made and natural objects in a traffic scene, and they can be divided into two main categories: the foreground objects and the background objects. The foreground objects such as pedestrians, vehicles, motorcycles, and bicycles are frequently those which occupy the regions on the ground and are close to the host car in a forward-looking road scene image. On the other hand, the background objects such as the sky, trees, roads, and lane markings are often farther compared to foreground objects. Generally, the problems of recognizing and measuring those objects in a road scene focus on two main topics: the detection of roads [1], [2] and the detection of obstacles or vehicles [3], [4]. He *et al*. [1] used RGB color space in extracting road areas. They estimated the positions of left and right road boundaries and computed the mean and standard deviation vectors of the image pixels between these two boundaries. Using Gaussian model, they segmented road areas of urban traffic scenes by thresholding. Sotelo *et al*. [2] utilized the characteristic of cylindrical distribution in the Hue, Saturation, and Intensity (HSI) color space as features to segment the road region by calculating the distance of the vectors of input image pixels. Vision-based obstacle/vehicle detection systems have been widely studied. They take several hints of vehicles as criteria to distinguish vehicles from other objects. Kato and Ninomiya [3] reported their learning algorithm using a template matching method with modified quadratic discriminant functions. They exploited binocular vision images to remove the perspective effect from incoming images and to remap each pixel of these images toward new position. Sun *et al*. [4] utilized a Gabor filter bank for feature extraction in vehicle detection. To improve detection performance, they optimized these Gabor filters by genetic algorithms.

The detection of important objects in traffic images also has been reported. Chen [5] investigated highway overhead structure detection using horizontal edge projection model. After Sobel operator, the horizontal edges of traffic images were extracted and projected in each row of the images. He then detected peaks to determine overhead structures. In [6], Wu *et al*. further detected text on road signs: they utilized *K*-means algorithm to cluster the hue value of image pixels and, by assuming all text lies on planar surfaces, then localized road signs. To proceed, they detected text by integrating edge detection, searching and excluding candidate text areas, analyzing the difference of text color from its background's, and grouping using geometry properties.

Although a lot of research concentrated on forward-looking image processing for driver assistance has been reported to date, more studies need to be conducted to improve the analysis of traffic scenes for sophisticated ITS application. For a road scene image, we can divide it into two main parts: the upper part and the lower part. It is true that the lower part usually contains more important objects than the upper does. Conventionally, lane line tracking and obstacle detection algorithms [1]–[4] ignore the upper part directly to reduce searching area aiming for shortening its processing time. However, in some situations, like a lane on a slope, the road plane will extend from the lower part to the upper part. Besides lanes and vehicles, there are some important objects locating in the upper part of the image such as overhead structures [5], traffic signs and sign text [6], etc. Therefore, it is necessary to analyze the whole traffic image and segment it into different meaningful regions. Unfortunately, there are few algorithms [7] proposed yet to deal with the segmentation of the whole traffic image. In this

This research was supported in part by the National Science Council under grants NSC 95-2213-E-009-212 and NSC-95-2752-E-009-011-PAE, and by the Ministry of Economic Affairs under grant 95-EC-17-A-02-S1-032, Taiwan, ROC..

Jyh-Yeong Chang and Chien-Wen Cho are with the Dept. of Electrical and Control Eng., National Chiao Tung University, Hsinchu, Taiwan.

study, we will explore a framework to analyze a traffic scene and then detect the front car aiming for improving the driver's safety.

II. The Fuzzy ART Based Scene Analysis

The goal of this paper is to construct a scene segmentation system capable of automatically classifying and labeling the objects in images via image pixel features. To this end, we first construct a fuzzy rule base to analyze the scene and then utilize this rule base to classify pixels of the given traffic scene images. Fuzzy approach is adopted so that the module can be better tolerant of uncertainty, ambiguity, irregularity, and noise existent in the image. To obtain suitable fuzzy sets of features, we use fuzzy Adaptive Resonance Theory (fuzzy ART) to summarize the grouping nature among pixel features. Unlike other clustering approach, such as fuzzy *c*-means, fuzzy ART can produce appropriate number of clusters subject to the specified setting of vigilance value. In this paper, we will demonstrate our scene analysis system using road scene images as a test-bed. For safety improvement, we also focus on segmenting cars ahead. The details of the procedure are described as follows.

A. Feature Extraction and Fuzzy Clustering

The selection of meaningful features for scene analysis is vital to the success of image segmentation. We humans recognize the objects appearing in scenes according to many features like colors, intensity, boundaries, and their positions. For example, the sky often has a bright blue color and is always in the topmost space. Trees are mostly in green color, grow on the ground, stretch the branches and leaves toward the sky. The ground is in dark gray and always locates at the lowest space. To model scene objects, we adopt the intuition of human beings, i.e., the features of color, height, and shapes; these are the most useful ones to describe objects.

It is well known that colors can be described by many color spaces. Among them, RGB is mostly accepted color space when figures are to be displayed on the computer screen. The object is, however, more adherent to another color space, that is, the HSI system. According to the intrinsic component's feature, the components of Hue and Intensity are more efficient in classifying an object than Saturation. Hence, we choose only the hue and intensity components as our first two features selected. Moreover, we will also use the vertical position as the third feature to describe the objects because the vertical position contains many groundtruths for objects to appear while the horizontal position usually does not.

Accordingly, three inherent features, including hue, intensity, and height of image pixels, are extracted from a given $M \times N$ image in the proposed scene analysis module. For fuzzy ART [8] requirement, we encode the feature vector of the *i*-th pixel using the complement coding form as $\mathbf{I}_i = (a_i^1, a_i^2, a_i^3, a_i^{1c}, a_i^{2c}, a_i^{3c})$, where a_i^1, a_i^2, and a_i^3 are the piexel's hue, intensity, and height, respectively. Furthermore, a_i^{1c}, a_i^{2c}, and a_i^{3c} represent their corresponding complements. It is noted that a_i^1, a_i^2, and a_i^3 should be normalized in advance. We use fuzzy ART to cluster MN $\mathbf{I}_i$'s automatically. After the fuzzy clustering, the pixels of the image can be divided into *n* clusters. Value *n* depends on the vigilance parameter chosen in the fuzzy ART algorithm, in which a larger vigilance value produces larger cluster number *n*. The *i*-th pixel $\mathbf{I}_i$ is assigned to the category with the largest choice function by

$$\mathbf{O}_p = \{\mathbf{I}_k \mid T_p(\mathbf{I}_k) = \max\{T_1(\mathbf{I}_k), T_2(\mathbf{I}_k), \cdots, T_p(\mathbf{I}_k), \cdots, T_n(\mathbf{I}_k)\}\}, j = 1, 2, \cdots, n; k = 1, 2, \cdots, P, \quad (1)$$

where T_j is the choice function produced after the fuzzy ART learning procedure and $P = MN$ is the total number of image pixels. Any cluster containing pixels fewer than a small fraction, 5% in this study, of the number of the image pixels, *MN*, is considered not important enough to be a category and would be ignored by augmenting each pixel to the category of the second largest choice function. Then, the remained *j*-th cluster's mean vector $\mathbf{m}_j$ can be calculated. Let $\mathbf{m}_j = (m_j^1, m_j^2, m_j^3)$, where m_j^1, m_j^2, and m_j^3 are the mean of hue, intensity, and height of the *j*-th cluster, respectively. Value m_j^k is computed by

$$m_j^k = \frac{\sum_{\mathbf{I}_i \in \mathbf{O}_j} a_i^k}{N_j}, \quad i = 1, \cdots, P; j = 1, \cdots, n; k = 1, 2, 3, \quad (2)$$

where N_j is the total number of pixels of the *j*-th cluster. The standard deviation vector may also be computed for each cluster in a similar manner. Let $\boldsymbol{\sigma}_j = (\sigma_j^1, \sigma_j^2, \sigma_j^3)$, where σ_j^1, σ_j^2, and σ_j^3 are the standard deviation of the hue, intensity, and height of the *j*-th cluster given by

$$\sigma_j^k = \frac{\sum_{\mathbf{I}_i \in \mathbf{O}_j} (a_i^k - m_j^k)^2}{N_j - 1}, \quad (3)$$
$$i = 1, \cdots, P; j = 1, \cdots, n; k = 1, 2, 3.$$

B. The Construction and Merging of Membership Functions

Fuzzy rules provide us a convenient way to deal with imprecise, noise, or ambiguous data. We make use of the membership functions to represent the features' possibility of each cluster. Many types of membership functions, e.g., bell-shaped, triangular, and trapezoid ones, are frequently used in a fuzzy system. We choose the Gaussian type membership function to represent the features because the Gaussian type membership function can reflect the first order and second order statistics of clusters and is differentiable. As stated earlier, after the fuzzy ART clustering, the image pixels are assigned into different clusters. The corresponding mean and standard deviation of each cluster can be computed. We can thus obtain associated Gaussian type membership functions for the features in a cluster. However, very often some membership functions are too close and will generate inefficient and/or redundant rules in the fuzzy rule base. These rules not only increase the size of rule base but also decrease the generalization ability of the rule base. Suppose feature k, $k = 1, \ldots, K$, has l_k linguistic labels. The maximum number of fuzzy rules that can be learned is $\prod_{k=1}^{K} l_k$. In this image analysis application, we usually have a large number of training pixels so that all $\prod_{k=1}^{K} l_k$ fuzzy rules can possibly be learned. By reducing the number of linguistic labels for each feature, the number of the produced fuzzy rules can also be reduced. A simple way is to combine two neighboring membership functions which have a great degree of overlap. It is known that the degree of overlap of two membership functions depends on the common portion of these two functions. To integrate their common part in an interactive fashion, it seems reasonable to combine these two functions. Two neighboring membership functions of the j-th and (j+1)-th clusters with means m_j^k and m_{j+1}^k and standard deviations σ_j^k and σ_{j+1}^k can be merged owing to their closeness if

$$| \; m_j^k - m_{j+1}^k \; | \le M_F \cdot \min(\sigma_j^k, \sigma_{j+1}^k), \tag{4}$$

where M_F is a pre-specified *merging factor*. The merged membership function $\mu_j^{k^*}(a_j^k)$ is re-constructed using, by recomputing (2) and (3), the mean and standard deviation of the new cluster which combines the j-th and the (j + 1)-th clusters.

After the merging possible procedure, linguistic labels can be used to describe the features' characteristics. To illustrate in a comprehensive way, we show in Fig. 1 a two-feature and six-cluster case. We can consider each cluster as a region with two axes which represent two feature spaces. In turn, the region projects onto either one of the feature axes respectively, and then produces a corresponding Gaussian fuzzy membership function with the peak being the mean of each feature. In Fig. 1, clusters 3 and 4 are so close, i.e., satisfying (4) for both axes, that they can be combined into one cluster. Accordingly, membership functions on these two axes, derived from these two clusters, are also merged. However, for the case of clusters 1 and 2 and the case of clusters 5 and 6, they are merged only on one axis (the horizontal axis and the vertical axis, respectively).

C. Fuzzy Rule Base Extraction and Classification

As developed by Wang and Mendel [9], fuzzy rules can be generated by learning from examples. An image pixel with feature vector (a^1, a^2, a^3) is associated with its desired output of corresponding natural elements. Such image pixel constitutes an input-output pair to be learned in the fuzzy rule base. In this setting, the rules generated are a series of associations of the form

"**IF** *antecedent conditions hold,* **THEN** *consequent conditions hold.*"

The number of antecedent conditions equals the number of features. Note that antecedent conditions are connected by "**AND.**" For illustrative purpose, assume now we have three linguistic labels, HIGH, MIDDLE, and LOW to describe the pixel's hue; three labels, BRIGHT, GRAY, and DARK to describe the pixel's intensity; and three labels, UP, MIDDLE, and DOWN to describe the pixel's height. For example, pixel i invokes the feature-target vector:

$$[a_i^1, a_i^2, a_i^3; D_i] = (0.80, 0.75, 0.93; \text{SKY}), \tag{5}$$

where a_i^1, a_i^2, and a_i^3 denote the normalized hue, intensity, and height of the pixel, respectively, and D_1 is the corresponding object of the pixel. First, we have to determine the membership values of a_i^1, a_i^2, and a_i^3 for different linguistic labels. Suppose, say, a_i^1 maps membership function HIGH, MIDDLE, and LOW to values 0.90, 0.32, and 0.15, respectively. a_i^2 maps membership function BRIGHT, GRAY, and DARK to values 0.73, 0.22, and 0.13, respectively. a_i^3 maps membership function UP, MIDDLE, and DOWN to values 0.56, 0.33, and 0.49, respectively. Next we assign the given inputs to the labels with maximum membership values. Thus, a_i^1 is specified by fuzzy set HIGH, a_i^2 by fuzzy set BRIGHT, and a_i^3 by fuzzy set UP. Hence, pixel i supports the rule of

> **IF** the pixel's hue is HIGH **AND** its intensity is BRIGHT **AND** its height is UP, **THEN** the pixel is SKY. (6)

with firing strength 0.56. Due to a large number of training pixels, some conflicting rules may be generated. The conflicting rules have the same antecedent conditions but lead to different consequent conditions (for example, the pixel is ROAD or the pixel is TREE). For a set of antecedent conditions, we can have only one rule to reflect it. Therefore, we have to choose one from the two or more conflicting rules from each qualified cluster. To this end, we choose the rule that is supported by a maximum number of examples. Furthermore, to prune redundant fuzzy rules, if the supporting pixels of a rule are less than 5% of the total pixels of the image, the rule is excluded from defining an **IF-THEN** rule. After the fuzzy rule base is established, we use the max-min inference to classify the elements of images.

C. Fuzzy K-Nearest Neighbor Classifier

Although fuzzy inference using fuzzy rule base is very promising in scene image segmentation, there are still erroneously classified pixels existing after rule-base classification. These wrongly classified pixels occur most frequently across the boundary of two natural element categories. To reduce these mis-classified pixels, we adopt the modified fuzzy *K*-Nearest Neighbor (*K*-NN) algorithm [10], which can provide more reliable differentiation than traditional *K*-NN method owing to the introduction of fuzzy notion. Similar to the traditional *K*-NN, the first step of this algorithm is to choose the *K*-nearest neighbors of an input sample x_p: a window mask of size $n_w \times n_w$ (n_w is an odd number), centered around the input pixel x_p, is used to define the *K*-nearest neighbors of the pixel. The second step is to assign a membership grade of this input sample x_p according to its distance from it. Let S_m, $0 \le m \le (n_w - 1)/2$, be the set of pixels in the window different from the input pixel x_p in positional index by m. The membership degree of the belongingness to class j of a pixel x in S_m (say x in class l), u_j is assigned by

$$u_j(x \in S_m) = \begin{cases} 1 - \dfrac{2m}{n+1}, & \text{if } j = l \\ 0, & \text{if } j \neq l. \end{cases} \tag{7}$$

Finally, among these *K*-nearest neighbors, the input sample x_p is ascribed to the class j^* that accumulates the maximal class membership degree of belongingness among the pixels in the mask window:

$$j^* = \arg\max \sum_{x \in S_m} u_j(x). \tag{8}$$

D. The Road Scene Analysis Algorithm

Vehicles, which can have various colors, are difficult to detect using only color features, but they often can be better detected by some ground-truth cues described below.

1) Vehicles must be on the road.
2) Vehicles have shadows under the car bodies.
3) The height-width ratios of vehicles vary in a certain range, and are suitably represented by circumscribing rectangles.

For illustration purpose, as shown in Fig. 2(a), a forward-looking traffic scene image with some vehicles on the road. The road region in gray color is shown in Fig. 2(b). For car detection in a faster and reliable manner, we can limit the searching region of interest to a smaller area, i.e., the road, rather than the whole image. The shadow of a vehicle is shown in Fig. 2(c). Finally, as shown in Fig. 2(d), the preceding car is circumscribed with a rectangle. In the above, Cues 1) and 2) can be helpful to assure the car edges by some edge finding routines. Cue 3) is useful for car region refinement.

With the above cues in mind, the road scene analysis and car detection procedures are summarized as follows:

1) Extract three features containing the hue, intensity, and vertical position of each pixel of the training images.
2) Use fuzzy ART algorithm to cluster the pixels of training images.
3) Construct the membership functions of corresponding clusters and merge the obtained membership functions.
4) Extract fuzzy rule base and classify test images:
   i) Infer each pixel of test images and record the output scores of each class according to the trained fuzzy rule base and label it into SKY, PLANT, ROAD, VEHICLE, or BARRIER, respectively.
   ii) Detect vehicles of the test images using the cues above and circumscribe the vehicles by rectangles. Car region pixels will be improved according to the car edges detected as follows. For any pixel labeled as VEHICLE but not circumscribed inside the rectangle, change their label to the class with the second high score. Label any non-VEHICLE pixel circumscribed inside the rectangle as VEHICLE.
   iii) Use fuzzy *K*-NN algorithm to further remove the mis-classified pixels of test images.

III. Experimental Results

To evaluate the scene analysis module, we first trained a fuzzy rule base for this purpose. Next, we segmented 20 traffic scene images by inferring and labeling the image pixels using the fuzzy rule base obtained. Afterwards, the scene segmentation accuracy was also computed to test our scene analysis scheme. The details are given as follows.

1) Preprocessing

Four representative 256×192 color scene images (as shown in Figs. 3), which were taken on a highway were used to train the fuzzy rule base for scene analysis. Their associated labeled images were done manually. To analyze these traffic scenes, we had five objects to segment: they were the *sky* drawn in cyan, *roads* in dark gray, *trees* in green, *barriers* in light gray, and *vehicles* in red. To proceed, the three-dimensional features, hue, intensity, and height, were extracted from each pixel of the traffic images. Each feature value was first normalized to be within the interval [0, 1] then encoded as feature vector in the complement coding format.

2) Fuzzy Rule Base Construction

From (4), it can be expected that the segmentation accuracy of a fuzzy rule base will highly depends on the selection of the merging factor M_F. Therefore an iterative procedure to determine a best M_F value will be described below.

i) In the firs step, we construct the membership functions for the fuzzy rule base by fuzzy ART clustering. The feature vectors of the pixels of the training images were clustered by the fuzzy ART algorithm with vigilance parameter ρ = 0.5. After clustering, we obtained 20 three-dimensional feature vector categories of the training images. The means, standard deviations, and resultant membership functions of three features defined for each category were computed accordingly.

ii) In this step, we search a best M_F value for merging membership functions defined above. Initially, the merging factor M_F was set to 1 and the membership functions were merged if two neighboring membership functions satisfied (4).

iii) By the learning process described in Section II, every image pixel of the twenty training images has been used to train the fuzzy rule base for traffic scene analysis.

iv) To evaluate the effectiveness of the M_F, we segmented these twenty training images by fuzzy inference using the fuzzy rule base obtained. The segmentation accuracy, with respect to the groundtruth images obtained manually, was computed accordingly.

v) The merging factor M_F was decreased from 1 down to 0.1 with step size = 0.1. Then, iii), iv), and v) were repeated.

With these M_F learning trials, we had found that the highest accuracy occured at M_F = 0.1 and the segmentation accuracy was monotonically decreasing with respect to M_F value. With this trend in mind, we searched the best M_F between [0, 0.1] with step size = 0.02 in a similar manner and found that the highest accuracy still occurred at M_F = 0.1. Accordingly, we could conclude that 0.1 is the best merging factor M_F. With this M_F value, the numbers of membership functions of hue, intensity, and height were reduced, all are from 20, to six, seven, and seven, respectively. The final fuzzy rule base generated consisted of 216 fuzzy rules, some of which were listed in Table I, where the six fuzzy sets of hue attribute were ordered from small value to high value as U_1–U_6 according to the order of the values assuming full membership function degree. Similarly, the seven fuzzy sets of intensity attribute were ordered as I_1–I_7 and those of height attribute were ordered as H_1–H_7. For example, Rule 1 appears as

IF the pixel's hue is U_3 **AND** its intensity is I_7 **AND** its height is H_1, **THEN** the pixel is SKY.

The average segmentation accuracy of the training images was 88.59%, which were computed according to the groundtruth images obtained manually. We show the segmentation result of these four images in Figs. 4(a)–4(d).

3) Segmentation Accuracy Evaluation

To evaluate the segmentation accuracy of the fuzzy rule base obtained, 20 traffic scene images were utilized to test the proposed segmentation algorithm. For brevity, we only show one examples in Fig. 5. Among them, Fig. 5(a) is one of the traffic scene image. Fig. 5(b) is the segmented output image by the fuzzy rule base learned. Most pixels are correctly classified but there are still some mis-classified pixels. In Fig. 5(c), the detected vehicles are circumscribed by rectangles using edge detection algorithm. In Fig. 5(d), the pixels inside the rectangles are all assigned to VEHICLE class, whereas each pixel outside those rectangles is assigned to the class having the second high score if it is mis-classified to VEHICLE class through fuzzy rule inference. To further correct the erroneous pixels, the image is processed by the modified fuzzy *K*-NN algorithm as shown in Fig. 5(e). It is obvious that, by fuzzy *K*-NN algorithm, the noisy pixels are eroded and thus the segmentation result is greatly improved. On average, the segmentation accuracy for these 20 classified test images is 86.56%.

IV. Conclusion

In this paper, we have proposed a framework which can analyze traffic scenes and provide a practical solution to the detection of preceding vehicles. Our system deals with scene segmentation and object labeling of input image frames. Fuzzy inference based approach is adopted so that the system is tolerant to the uncertainty, ambiguity, irregularity, and noise exist in an image. Furthermore, to obtain suitable fuzzy sets of features, we use fuzzy ART to reflect the nature of feature space. We demonstrate our scene analysis system using road scene images as a test-bed. Moreover, for enhance safety driving, we also focus on segmenting the car ahead. A few pixels, however, may still be mis-classified after the scene segmented through fuzzy inference. We further adopt fuzzy *K*-NN algorithm to further refine the labeling result. The scene object segmentation accuracy are tested to be very reliable and promising in our experiment. Compared with conventional approach, the proposed scheme is capable of better understanding forward-looking traffic scenes and is effective and robust in detecting cars ahead.

References

[1] Y. He, H. Wang, and B. Zhang, "Color-based road detection in urban traffic scenes," *IEEE Trans. Intell. Transport. Syst.*, vol. 5, no. 4, pp. 57–68, 2004.

[2] M. A. Sotelo *et al.*, "VIRTUOUS: Vision-based road transportation for unmanned operation on urban-like scenarios," *IEEE Trans. Intell. Trans- port. Syst.*, vol. 5, no. 2, pp. 69–83, 2004.

[3] T. Kato and Y. Ninomiya, "An approach to vehicle recognition using supervised learining," *IEICE Trans. Inform. Syst.*, vol. E83-D, no. 7, pp.1475–1479, 2000.

[4] Z. Sun, G. Bebis, and R. Miller, "On-road vehicle detection using evolutionary Gabor filter optimization," *IEEE Trans. Intell. Transport. Syst.*, vol. 6, no. 2, pp. 125–137, 2005.

[5] Y. Chen, "Highway overhead structure detection using video image sequences," *IEEE Trans. Intell. Transport. Syst.*, vol. 4, no. 2, pp. 67–77, 2003.

[6] W. Wu, X. Chen, and J. Yang, "Detection of text on road signs from video," *IEEE Trans. Intell. Transport. Syst.*, vol. 6, no. 4, pp. 378–390, 2005.

[7] J. D. Crisman and C. E. Thorpe, "UNSCARF–A color vision system for the detection of unstructuredroads," in *Proc. 1991 IEEE Int. Conf. Robotics and Automation*, pp. 2496-2501, Apr. 1991.

[8] C. T. Lin and C. S. G. Lee, *Neural Fuzzy Systems: A Neuro-Fuzzy Synergism to Intelligent Systems*,Upper Saddle River, New Jersey: Prentice-Hall,1996, pp. 626-630.

[9] L. X. Wang and J. M. Mendel, "Generating fuzzy rules by learning from examples," *IEEE Trans. Syst. Man Cybern.*, vol. 22, no. 6, pp. 1414–1427, 1992.

[10] J. Y. Chang and C. W. Cho, "Scene analysis system using a combined fuzzy logic-based technique," *J. Chinese Instit. Eng.*, vol. 25, no. 3, pp. 297–307, 2002.

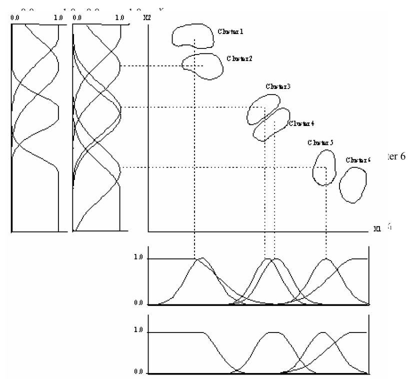

Fig. 1. Demonstration of deriving membership functions from the fuzzy ART clusters and the merged membership functions.

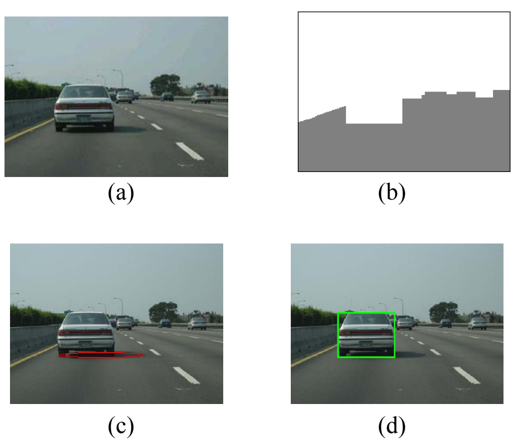

Fig. 2. (a) A forward-looking traffic scene image. (b) The road region is in gray color. (c) The shadow of the preceding car is marked. (d) The preceding car is circumscribed with a rectangle.

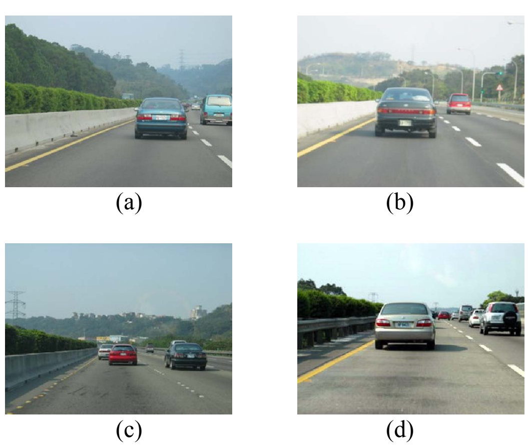

Fig. 3. Four representative 256×192 training images.

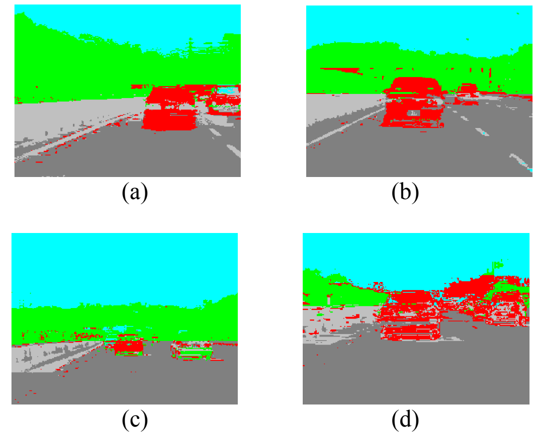

Fig. 4. The classification result of the four training images by inferring using established fuzzy rule base.

TABLE I
SOME OF THE FUZZY RULE BASE OBTAINED

Number	HUe	Intensity	Height	Class
1	U_3	I_7	H_1	SKY
2	U_3	I_7	H_2	SKY
⋮	⋮	⋮	⋮	⋮
17	U_1	I_3	H_7	ROAD
⋮	⋮	⋮	⋮	⋮
70	U_6	I_1	H_6	PLANT
⋮	⋮	⋮	⋮	⋮
167	U_1	I_1	H_4	VEHICLE
⋮	⋮	⋮	⋮	⋮
215	U_5	I_3	H_6	BARRIER
216	U_6	I_7	H_4	BARRIER

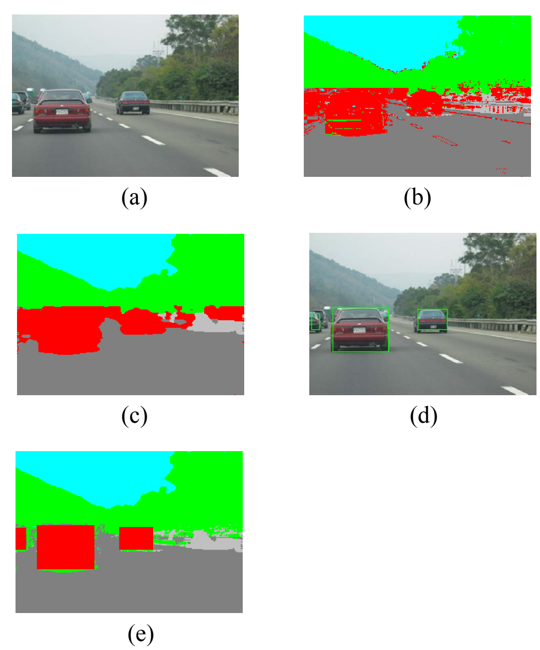

Fig. 5. An example selected from 20 test images and its segmentation results. (a) The original image. (b) The fuzzy ART based classification output image. (c) The segmentation result after fuzzy *K*-NN classification. (d) The detected vehicles. (e) The segmentation result combining vehicle detection.