國 立 交 通 大 學

電 機 與 控 制 工 程 研 究 所

碩 士 論 文

利用基因演算法之Fuzzy ID3方法於階層式場景分析系統

A Hierarchical Approach for Scene Analysis Using Genetic Algorithm Based Fuzzy ID3 Method

研 究 生：陳 世 宗

指 導 教 授：張 志 永

中 華 民 國 九十三 年 六 月

利用基因演算法之Fuzzy ID3方法於
階層式場景分析系統

A Hierarchical Approach for Scene Analysis Using
Genetic Algorithm Based Fuzzy ID3 Method

研 究 生：陳世宗　　　　Student：Shih-Tsung Chen

指導教授：張志永　　　　Advisor：Jyh-Yeong Chang

國立交通大學

電機與控制工程學系

碩士論文

A Thesis

Submitted to Department of Electrical and Control Engineering

College of Electrical Engineering and Computer Science

National Chiao Tung University

in Partial Fulfillment of the Requirements

for the Degree of Master in

Electrical and Control Engineering

June 2004

Hsinchu, Taiwan, Republic of China

中華民國 九十三 年 六 月

利用基因演算法之Fuzzy ID3方法於階層式場景分析系統

學生：陳世宗　　　指導教授：張志永博士

國立交通大學電機與控制工程研究所

摘要

本篇論文應用機器學習演算法幫助我們分析影像之場景，並將場景中的物件區域辨識出來。在本篇論文中，我們提出一個階層式場景分析系統，它主要是利用以基因演算法為基礎的Fuzzy ID3理論方法。首先，我們提出一個基於基因演算法的Fuzzy ID3理論方法來產生模糊決策樹，並且從這顆建構的模糊決策樹，萃取出描述資料集的模糊規則。然後我們顯示出這些規則如何應用在車前的場景分析。並且利用自然界的法則，來加以改善場景中物件區塊的修正，以提高辨識的正確性。

本篇論文所提出的階層式場景分析方法，主要是應用在分析從駕駛者的角度往外看出去的道路場景。從測試的結果可以證明，我們所提出的這套系統，可以有效的分出場景中的物件，並且能夠提供於避免潛在車禍碰撞事件的應用。

A Hierarchical Approach for Scene Analysis Using Genetic Algorithm Based Fuzzy ID3 Method

STUDENT: SHIH-TSUNG CHEN ADVISOR: Dr. JYH-YEONG CHANG

Institute of Electrical and Control Engineering

National Chiao-Tung University

ABSTRACT

In this thesis, we consider to utilize machine learning algorithm to segment natural objects in outdoor scene images. First, we proposed a scene analysis system that is rooted from genetic algorithm based fuzzy ID3 method. We develop a genetic algorithm based fuzzy ID3 method, which is designed to generate a fuzzy decision tree and the decision tree can extract fuzzy rules to summarize the regularities existing in the data set. Afterward we show how the resultant rules can be used for object recognition and then apply image ground-truthing to further improve the rule-based object classification accuracy.

The proposed hierarchical scene analysis method is applied to analyze the forward-looking road scene from a car. The testing results have demonstrated the natural object segmentation accuracy is quite high and this method provides a potential application in automated car collision avoidance.

ACKNOWLEDGEMENTS

I would like to express my sincere appreciation to my advisor, Dr. Jyh-Yeong Chang. Without his patient guidance and inspiration during the two years, it is impossible for me to complete this thesis. In addition, I am thankful to all my Lab members for their discussion and suggestion.

Finally, I would like to express my deepest gratitude to my family. Without their strong support, I could not go through the two years.

Contents

List of Figures

List of Tables

Chapter 1. Introduction

1.1. Research Background

Computer has made a great contribution to the promotion of human lives. With the development of the computer, it is easier for people to communicate with each other. Since the mutual communication methods of mankind hinge on computers, they can identify and even can understand the data given to them. It is widely recognized that computers should be more humanized. In order to make it easier for people to use computers and even to communication with them, a lot of hand-written recognition and speech recognition systems have commercialized.

At present, images are regard as an essential variety of information. They are the most common media around us. One of the important problems in a vision system is identifying the image regions that represent objects. This operation, which is so natural and easy for people, is surprisingly difficult for computer. The first step toward a vision system for object identification is to partition or segment an image into meaningful regions. Ideally, a region represents an object or part of an object. But current progresses been made are limited, mostly because of the variation of natural scenes. By using some knowledge, we are likely able to let the system adapt itself to most situation that may be encounter on the road.

In recent years, a great deal of scientific efforts has already dedicated to image analysis problems. The paradigm of image analysis has been successfully utilized in various fields. For example, robot vision, medical diagnose, image segmentation, pattern recognition, and so on.

1.2. Overview

The purpose of image understanding is to enable a computer to understand its environment from visual information. There are diverse studies about image understanding using knowledge base system [1], [2]. In the study of image understanding, it is important to recognize the natural elements by image processing, and to understand their situations with knowledge base. But at present, a general purpose image understanding system is still very difficult for scientists to achieve. Because the input image data is very diverse and also sensitive to the environment, a scene in the real world may produce completely different data under different situation, for example, time of a day or year, wealth condition, and viewing angle.

In recent years, fuzzy set has provided an effective scheme of knowledge representation. One major feature of fuzzy set theory is its ability to express the ambiguity in human thinking, and uncertain information obtained from the real world. Furthermore, fuzzy set theory provides the inference that applies human reasoning ability to the knowledge base system. Therefore, fuzzy systems are suitable for various fields, e.g., decision making, pattern recognition, and control problems etc. In particular, fuzzy set theory has also been an appropriate framework for many problems encountered in the image scene analysis system, because various kinds of variation found in image scene can be resolved by fuzzy logic [3]. Due to the inherent complexity existing in the scene, the problem of image understanding has more developmental area. Keller *et al.* [4] has proposed the fuzzy-set-based aggregation networks for scene analysis by neural network structure using fuzzy set theory.

More recently, a number of authors have endeavored to use techniques from machine learning to increase the robustness and efficiency of scene understanding system. Foresti *et al*. [5] have proposed a new model of neural tree, called generalized neural tree (GNT). GNT joins together some characteristics of decision trees and some of classical feed-forward neural networks. In the GNT's learning rule, the whole tree structure is considered at each learning step, and the entire training set is used to update each node in order to provide a better classification. Bischof *et al*. [6] have proposed a conditional rule generation technique (CRG) that is designed to describe structures using part attribute and their relations. The idea of CRG is to generate classification rules that include structural pattern information to the extent that is required for classifying correctly a set if training patterns.

Knowledge-based image understanding and interpretation has been investigated by several researchers. The knowledge-guided segmentation and labeling (KGSL) [7] approach can be applied to segment the images, automatically extract cluster labeling rules and add them to the knowledge base to improve its performance. The promise of the system was demonstrated by processing a set of color images and the application of the approach to ocean satellite images. Fan *et al*. [8] have proposed a knowledge-based road scene understanding system KRUS. In this system, a road recognition information algorithm combining fusing of edge and region information and using of scene knowledge is presented, which makes the KRUS system intelligent enough to recognize road robustly and precisely.

1.3. Thesis Outline

In this thesis, we proposed a hierarchical scene analysis system. In this system, a supervised learning algorithm is applied to construct a decision tree and extract fuzzy rules from the decision tree. Finally, the fuzzy rules were used to classify the pixels of images.

The remainder of this paper is structured as follows. Chapter II introduces the genetic algorithm based Fuzzy ID3 method. Chapter Ⅲ discusses the overall structure of the hierarchical scene analysis approach and describes the experiment to a simple gray image. In Chapter Ⅳ, we will show the application of the approach to a set of color images of road scene taken on freeway by using digital camera. Chapter Ⅴ presents our conclusions.

Chapter 2. Fundamental Concept

Our proposed image analysis system relies heavily on the fuzzy set theory, genetic algorithm, and fuzzy ID3 algorithm. First, it is instructive to explain them in detail.

2.1. Introduction to Fuzzy ID3

Knowledge acquisition from data is very important in knowledge engineering. A popular and efficient method is ID3 algorithm [9]. The ID3 approach to pattern recognition and classification consists of a procedure for synthesizing an efficient decision tree for classifying pattern that have non-numeric feature values. The decision tree can also be expressed in the form of rules. Therefore, ID3 is often thought of as an inductive inference procedure for machine learning or rule acquisition.

Fuzzy ID3 (FID3) algorithm [10], [11] extended from ID3 to incorporate fuzzy notation. The decision tree using fuzzy ID3 algorithm is similar to that of ID3 algorithm. Fuzzy ID3 algorithm is extended to apply to a data set containing numeric feature values instead of symbolic feature and generates a fuzzy decision tree using fuzzy sets. A fuzzy decision tree consists of nodes for training features, edges for branching by given feature values of fuzzy sets, and leaf node for final decision classes with certainties.

The feature ranking step is optional as we can use any arbitrary order of the features, but it is a desirable step because it can reduce the size of the tree and hence produce an efficient and accurate decision tree. While construct the decision tree, we use genetic algorithm to tune the fuzzy set membership function of features and parameters of decision tree to improve the classification performance and reducing the rule number.

2.2. Feature Ranking

When we start to construct decision tree, we have to choose the most important feature from the whole features. The order of feature to construct decision tree is an important issue to be investigated. In order to construct a decision tree with high accuracy and small size, the order of feature is evaluated using information gain [12]. In the process of deciding the order of features is called feature ranking.

The information theory that underpins this information gain criterion can be given in one statement: The information conveyed by a message depends on its probability and can be measured in bits as minus the logarithm to base 2 of that probability. So, for example, if there are 8 equal probable messages, the information conveyed by any one of them is $-\log_2(1/8)$ or 3 bits. Therefore, the information gain criterion provides a mechanism for a ranking a set of features so that the most favorable order can be chosen.

Assume that we have a training data set D, where each training data has l features $A_1, A_2, ..., A_l$ and one classified class $C = \{C_1, C_2, ..., C_m\}$ and fuzzy sets $F_{i1}, F_{i2}, ..., F_{im}$ for the feature A_i. Let D^{C_k} to be a fuzzy subset in D whose class is C_k and $|D|$ is the sum of the membership values in a fuzzy set of training data. We use the information gain $G(A_i, D)$ to estimate the gain of the feature, and decide the order of feature, where A_i represents the *ith* feature.

$$G(A_i, D) = I(D) - E(A_i, D), \tag{2.1}$$

where

$$I(D) = -\sum_{k=1}^{n}(p_k \cdot \log_2 p_k), \tag{2.2}$$

$$E(A_i, D) = \sum_{j=1}^{m}(p_{ij} \cdot I(D_{F_{ij}})), \tag{2.3}$$

$$p_k = \frac{\left|D^{C_k}\right|}{|D|}, \tag{2.4}$$

$$p_{ij} = \frac{\left|D_{F_{ij}}\right|}{\sum_{j=1}^{m}\left|D_{F_{ij}}\right|}. \tag{2.5}$$

$I(D)$ stands for the entropy before branch, and $E(A_i, D)$ means the entropy after branch according to the feature A_i. We will the select the feature with maximum information gain for constructing the decision tree at root. So we will set higher order to the feature with higher $G(A_i, D)$. Because of the feature ranking procedure will influence the performance and size of the decision tree. Accordingly, we will obtain not only minimized rule number but also maximized accuracy.

2.3. Tree Construction

Assume that we have a training data set D, where each training data has l features $A_1, A_2, ..., A_l$ and one classified class $C = \{C_1, C_2, ..., C_m\}$ and fuzzy sets $F_{i1}, F_{i2}, ..., F_{im}$ for the feature A_i. Let D^{C_k} to be a fuzzy subset in D whose class is C_k a. Then the algorithm to generate a fuzzy decision tree is shown as follows:

1) Generate the root node that has a set of all training data, i.e., a fuzzy set of all training data with the unit membership value.

2) If a node t with a fuzzy set of data D satisfies the following conditions:

   2.1) The proportion of a data set of a class C_k is greater than or equal to a threshold θ_r, that is,

$$\frac{\left|D^{C_k}\right|}{|D|} \geq \theta_r, \tag{2.6}$$

   2.2) The number of a data set is greater than a threshold θ_n, that is,

$$|D| < \theta_n, \tag{2.7}$$

   2.3) There are no attributes for more classifications,

   then it is a leaf node and we assign the certainty $\frac{\left|D^{C_k}\right|}{|D|}$ with all classes at this node.

3) If it does not satisfy the above conditions, it is not a leaf node, and the internal node is generated as follows:

   3.1) Select the feature which has next large $G(A_i, D)$ value for test feature A_{test}.

   3.2) Divide D into fuzzy subsets $D_1, D_2, ..., D_m$ according to the test feature, where the membership value of data in D_j is the product of the membership value in D and the value of $F_{test,j}$ of the value of A_{test} in D.

   3.3) Generate new node $t_1, t_2, ..., t_m$ for fuzzy subsets $D_1, D_2, ..., D_m$ and label the fuzzy sets $F_{test,j}$ to edges that connect between the nodes t_j and t.

   3.4) Replace D by D_j and repeat from 2) recursively until the end of all path is leaf node.

2.4. Fuzzy Set Discretization

Continuous-valued features have to be discretized priori to selection, typically by partitioning the range of feature into subranges. In ID3-like algorithms, a threshold value for the continuous-valued feature partition into two subranges. We regard this threshold value as cut point. The objection may be raised is that the discretized schema will cause to produce “bad” cut point especially when there are more than two classes in the problem.

This drawback can be overcome by using a discretization algorithm, called class-attribute interdependence maximization (CAIM) [13]. The CAIM algorithm show that it generates discretization schemes with almost always the highest dependence between the class labels and discrete intervals, and always with significantly lower numbers of intervals. Nevertheless, this crisp set is unnatural in the real world. Therefore, a fuzzy set introduces vagueness by eliminating the sharp boundary that divides members from nonmembers in the group. Thus, the transition between full membership and nonmembership is gradual rather than abrupt. Hence, we introduce Gaussian-type membership functions to each feature in our fuzzy ID3 algorithm.

The fuzzy ID3 scheme is determined by the parameter which includes the thresholds θ_r, θ_n, and the membership functions of each feature fuzzy set. A good selection of fuzzy rule base, leaf node threshold, and membership functions would greatly improve the accuracy of decision trees. To this end, genetic algorithm (GA) based scheme is utilized because of the essential nature of nonlinear of decision trees which limits the feasibility of traditional gradient method. In our work, GA [14] is used to tune the leaf node thresholds θ_r, θ_n, and the parameter of the membership functions of each feature. The membership function for each feature adopts the Gaussian-type and is given by

$$m(x) = \exp(-\frac{(x-\mu)^2}{2\sigma^2}), \tag{2.8}$$

where x is the corresponding feature value of the data with mean μ and variance σ. Thus for each membership function, two parameters μ and σ must tune. To minimize the rule number and maximize the accuracy, the fitness function [15] is defined as

$$f = (A - A_0) + \frac{\eta}{L}, \tag{2.9}$$

where A is the accuracy of the classification, A_0 is the lowest accuracy of the classification in the current population, L is the average depth of the decision tree, and η is the influence of the average depth. While starting the tuning procedure, initially we set η to a value such that $\frac{\eta}{L}$ is greater than $A - A_0$. This means that reduction of average depth of decision tree obtains a higher priority over maximization of accuracy. Therefore, the thresholds are tuned that the data classifies at lower depth, so that nodes at higher depth of the tree becomes redundant. As GA evolves, we gradually continue to decrease the value of η so that maximization of the accuracy starts dominating. Subsequently, we reduce η to zero in k steps. After k steps, η is always zero. In other words, we focus on the improvement of the accuracy after η becomes zero. Thus we can decrease the rule number without losing classification performance.

Now, we will illustrate one cycle of the tuning process. Assume we have a data set with four features and three classes, such that there are twelve membership functions. Each membership function has two parameters μ and σ, and two thresholds θ_r and θ_n in addition. Thus we have to tune 26 parameters totally. Initially we set random number to these parameters. In our setting, μ is randomly chosen from the range between the maximum and minimum value of the corresponding feature among the data and σ is chosen between 0 and the standard deviation of that feature.

There are several encoding of GA which depends on the problem heavily. Binary encoding is the most common one, mainly because the first research of GA used this type of encoding and because of its relative simplicity. In binary encoding, every chromosome is a string of bits 0 or 1. Crossover and mutation are two basic operators of GA. Performance of GA depend on them very much. There are many ways how to perform crossover and mutation. We briefly describe how to perform these two operators.

In Fig. 2.1. Multi point crossover method selects two crossover points, binary string from the beginning of the chromosome to the first crossover point is copied from the first parent, the part from the first to the second crossover point is copied from the other parent and the rest is copied from the first parent again. We repeat this

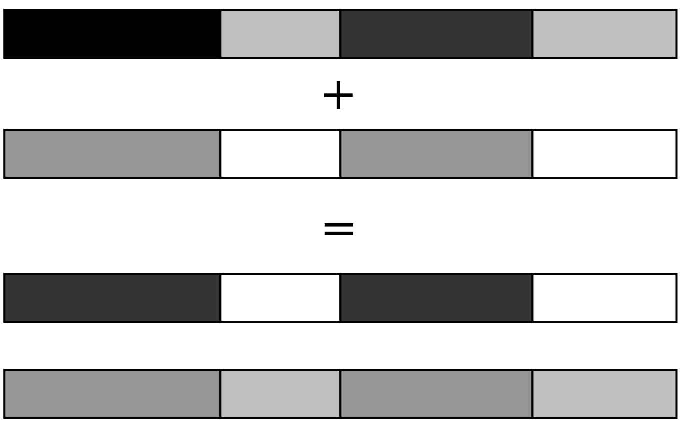


Fig. 2.1. Multi point crossover.

procedure until the end of those parent chromosomes. This process produces two new offspring chromosome, each of which is similar to both parent chromosomes. There are other ways to make crossover, for example we can choose more crossover points. Crossover can be quite complicated and depends mainly on the encoding of chromosomes. Specific crossover made for a specific problem can improve performance of the genetic algorithm.

After a crossover is performed, mutation takes place probably. Mutation is intended to prevent falling of all solutions in the population into a local optimum of the solved problem. In Fig. 2.2. Mutation operation randomly changes the offspring. In case of binary encoding we can switch a few randomly chosen bits from 1 to 0 or from 0 to 1. Mutation can be illustrated as follows:

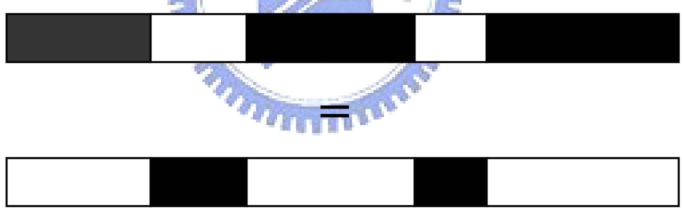


Fig. 2.2. Mutation.

In our GA scheme, assume we generate 50 chromosomes of these parameters, and use them to generate decision trees. After each decision tree is generated, for example, one individual has accuracy 87% and average depth is 3.5. The lowest accuracy in this population is 72%, and the fitness function of this individual is $f = (8.7 - 7.2) + \frac{\eta}{3.5}$. Accordingly, we perform the reproduction, crossover, and mutation operators to generate the new chromosomes and continue until the predetermined condition is achieved. Here we select the crossover probability

$p_c = 1$ and mutation probability $p_m = 0.0001$ for the GA evolving algorithm.

The advantage of GA is in their parallelism. GA is traveling in a search space using many individual trials in each generation so that they are less likely to get stuck in a local extreme like the other methods. The disadvantage of GA is in the computational time. GA can be slower than other methods. But since we can terminate the computation in any time, the longer run is acceptable. For some problems, choosing and implementation of encoding and fitness function can be difficult even though GA is powerful. To apply GA to fuzzy ID3 scheme, Fig. 2.4. is a flowchart of our genetic algorithm based fuzzy ID3 method.

2.5. Fuzzy Rule Inference

According the rule base, inference of the decision tree starts from the root node and iteratively tests each node indicated by the rule until reach at a leaf node. Note that we have recorded certainty values $\frac{|D^{C_k}|}{|D|}$ at leaf nodes as mentioned above and it represents the certainty of each class of the corresponding rule.

Since we obtain the $\frac{|D^{C_k}|}{|D|}$ values of each leaf node, the node is assigned by all class name with certainty value $\frac{|D^{C_k}|}{|D|}$. On the other hand, every leaf node has all class name with corresponding certainty value. The rule produced by each leaf node which can classify the data to every class with certainty value and does not directly classify the data to a specific class. For example, the fuzzy rule extracted from the leaf node as follows:

IF X_1 is F_{12} **AND** X_2 is F_{21}

THEN Class 1 with certainty 0.3 and Class 2 with certainty 0.7 (2.10)

In the pre-condition of the above rule, X_1 assumes the membership value in the second fuzzy set F_{12} of the first feature and X_2 assumes the membership value in the first fuzzy set F_{21} of the second feature. In the consequent part, there are two certainty values, 0.3 certainty value for class 1 and 0.7 certainty value for class 2. The steps of using the rule base to classify are as follows:

1) For each extracted fuzzy rule, we multiply the membership value of the corresponding fuzzy set of the testing data from the root to the leaf node sequentially. That is the firing strength, i.e. suppose the membership value of the testing data associated with the *ith* fuzzy set is m_i. The firing strength is given by

$$\prod_{i=1}^{l} m_i(x). \tag{2.11}$$

2) Multiply the certainty of the classes of the leaf node associated with the current fuzzy rule by the firing strength and denote the values as $J(n)$, $(n = 1, 2, ..., class\ \ number)$.
3) Repeat 1) and 2) until that all rules have been evaluated.
4) Sum the result in 3) of all the rules. Note that we must sum up the each class respectively.
5) Assign the testing data to the class with the maximum value in 4).

For example, a simple decision tree with 2 features, 3 subsets and 3 classes is shown in Fig. 2.3. This decision has four leaf nodes F_{11}, F_{21}, F_{22}, and F_{13} with their certainties C_1, C_2, and C_3 respectively. In addition, the membership values of the testing data are shown in the branch. Thus we can use these 4 fuzzy rules to classify the testing data as follows:

class 1: $\Sigma J(1)=0.1*0.1+0.5*0.8*0.3+0.5*0.3*0.3+0.9*0.8=0.895$

class 2: $\Sigma J(2)=0.1*0.2+0.5*0.8*0.7+0.5*0.3*0.4+0.9*0.1=0.45$

class 3: $\Sigma J(3)=0.1*0.7+0.5*0.8*0.0+0.5*0.3*0.3+0.9*0.1=0.205$

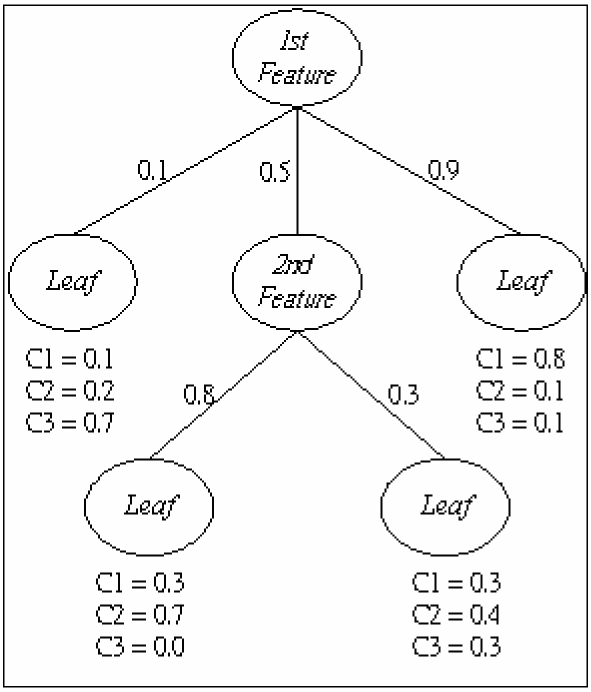


Fig. 2.3. An example of decision tree.

The testing data is assign to class 1 because $\Sigma J(1)$ is the maximum. Note that we classify one testing data need to evaluate all fuzzy rules but not just rely on a specific rule. In this way, we can use all rules together to decide the class of every testing data instead of generating a new specific rule only for some specific data. Therefore, we can reduce the size of rule base without losing performance.

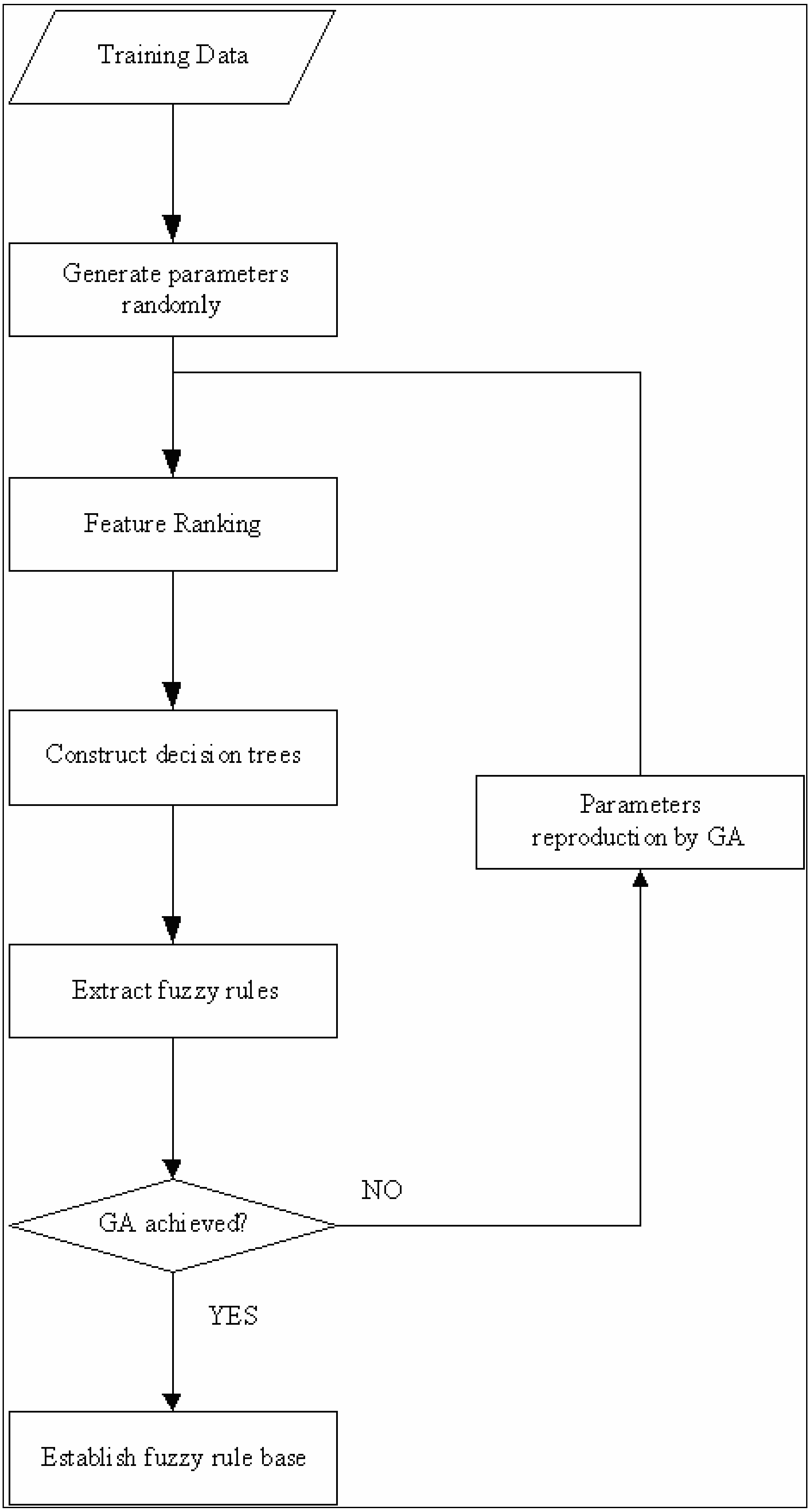


Fig. 2.4. Flowchart of genetic algorithm based fuzzy ID3 method.

Chapter 3. Scene Analysis System

In this chapter, we propose a system to analyze freeway scenes which usually contain sky, trees, road, vehicles, and so on. The functional module of the proposed system is shown in Fig. 3.1. The main capability of this system can automatically classify natural objects in the scene by the characteristics of image feature without too many subjective opinions. The details of the system are described in this chapter.

3.1. Feature Selection

Color is one of the most interesting characteristic of the natural world and can be computed in many different ways. Moreover, it is well known that chromatic characteristics of natural elements are not stable and highly dependent on color brilliance, reflections from the objects, illumination geometry, viewing geometry, and camera parameters. Unfortunately, up to now no single solution has been found to sufficiently characterize objects belonging to natural scenes. When given an image representing an outdoor scene, the first difficulty in describing it is to choose the most appropriate color space for object characterization. Selecting the best color space still is one of the difficulties in scene analysis [16].

3.1.1 RGB Color Space

Color is perceived by humans as a combination of tristimuli red, green, and blue

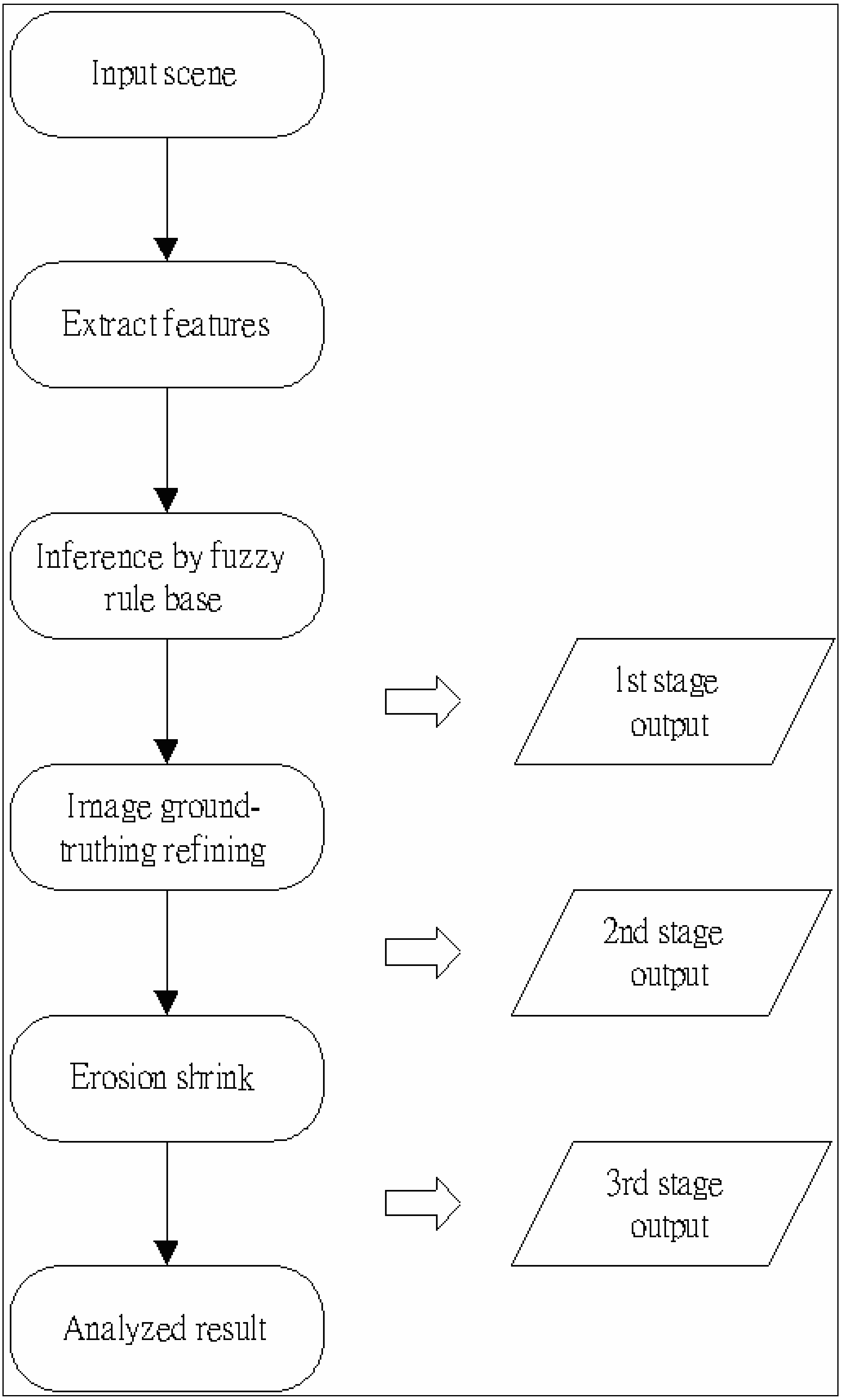


Fig. 3.1. The functional modules of the scene analysis system.

which are usually called three primary colors. From RGB representation, we can derive other kinds of color space by using either linear or nonlinear transformations. In the RGB model, each color appears in its primary spectral components of red, green, and blue. This model is based on a Cartesian coordinate system. The RGB color space can be geometrically represented in a 3 dimensional cub in Fig. 3.2. The coordinates of each point inside the cube represent the values of red, green, and blue components, respectively.

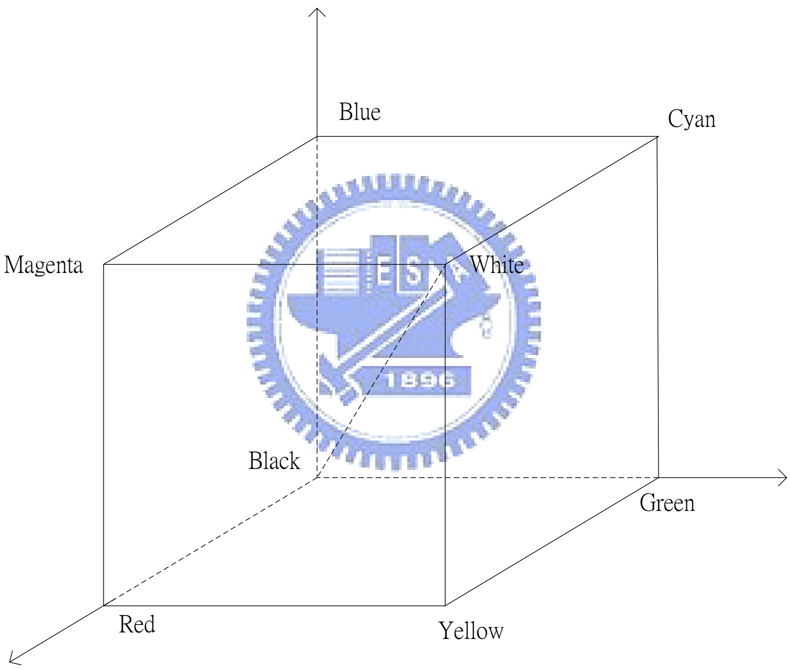


Fig. 3.2. RGB color space represented in a 3 dimensional cube.

RGB is the most commonly used model for the television system and picture acquired by digital cameras. RGB is suitable for color display, but not good for color scene analysis because the high correlation among the R, G, and B components. By high correlation, we mean that if the intensity changes, all the three components will change accordingly. Also, the measurement of a color in RGB

space does not represent color differences in a uniform scale, hence, it is impossible to evaluate the similarity of two colors form their distance in RGB color space.

3.1.2. HSI Color Space

The HSI color space is another commonly used color space in image processing, which is more intuitive to human vision. The HSI color space separates color information of an image from its intensity information. Color information is represented by hue and saturation values, while intensity which describes the brightness of an image, is determined by the amount of the light. Hue represents basic colors. Saturation is a measure of the purity of the color, and signifies the amount of white light mixed with the hue.

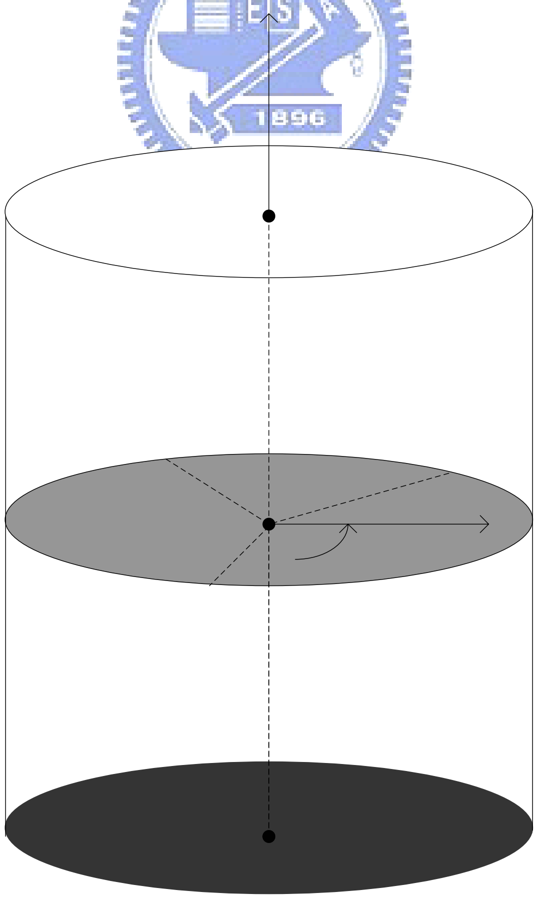


Fig. 3.3. HSI color space represented in a cylindrical coordinates.

The HSI color model can be described geometrically as in Fig. 3.3. Generally hue is considered as an angle between a reference line and the color point in RGB color space. The range of hue value is from 0^{o} to 360^{o} , for example, yellow is 60^{o}, green is 120^{o} , blue is 240^{o}, and magenta is 300^{o}. The saturation component represents the radial distance from the cylinder center. The nearer the point is to the center, the lighter is the color. Intensity is the height in the axis direction. The axis of the cylinder describes the gray levels, for example, minimum intensity is black, maximum intensity is white. Each slice of the cylinder perpendicular to the intensity axis is a plane with the same intensity.

The HSI color space has a good capability of representing the colors of human perception [17], because human vision can distinguish different hues easily, whereas the perception of different intensity or saturation does not imply the recognition of different color. The HSI color space can be transformed from the RGB [18]. The formulation for hue, saturation, and intensity are

$$H = \begin{cases} \theta, & \text{if} \quad B \le G \\ 360 - \theta, & \text{if} \quad B > G \end{cases}, \tag{3.1}$$

where

$$\theta = \cos^{-1}\left\{ \frac{\frac{1}{2}[(R-G)+(R-B)]}{\left[(R-G)^2+(R-B)(G-B)\right]^{\frac{1}{2}}} \right\} \tag{3.2}$$

$$S = 1 - \frac{3[min(R,G,B)]}{(R+G+B)} \tag{3.3}$$

$$I = \frac{(R+G+B)}{3}. \tag{3.4}$$

Hue reflects the predominant color of an object and has a great capability in subjective color perception. Hue is also the most useful feature in color image processing since it is less influenced by the nonuniform illumination such as shade, shadow, or reflect lights.

3.1.3. Gray Level

The gray level Y is a measurement of the luminance of the color, and is a likely candidate for edge in a color image. The formulation of gray level Y from RGB components is given by

$$Y = 0.299R + 0.587G + 0.114B \tag{3.5}$$

3.1.4. Spatial Information

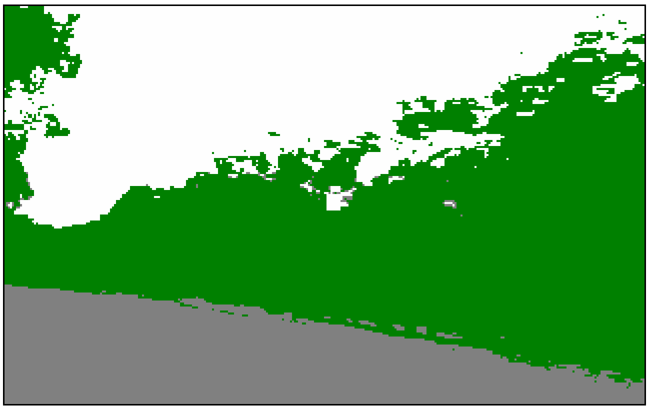
One of the drawbacks of color space clustering is that the cluster analysis does not utilize any spatial information. Therefore, the spatial information, which involves vertical position and horizontal position, is suitable for our scene analysis system. But the horizontal position has less unique information than that of the vertical position. Because of the natural elements, for example, the house or tree, maybe locate from left to right but can not locate from up to down in the scene. For this reason, we only choose the vertical position as our spatial information feature.

3.2. Establish Fuzzy Rule Base

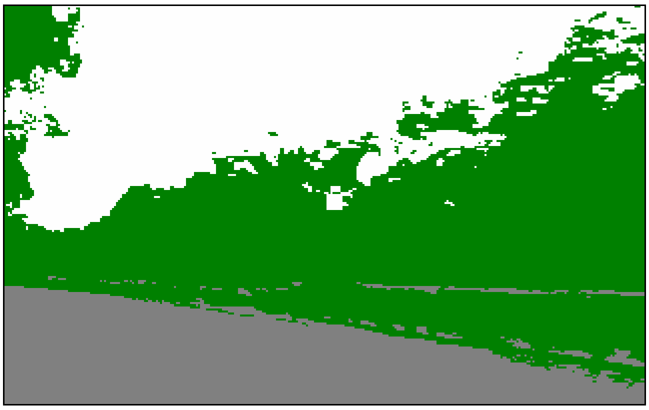
If we want to build a system that is able to understand a natural element in complex scenes, we can generate a fuzzy rule base system to describe the natural element. However, scene analysis in the computer vision research has been known to be one of the most difficult fields. Consequently, we proposed the genetic algorithm based fuzzy ID3 method to solving this problem.

Provided here is a gray image as Fig. 3.4(a); we can see this scene with our eyes and understand it with our brain. Nevertheless, the scene analysis system can not recognize anything initially. Therefore, we have to provide a desired recognition by our effort. In Fig. 3.4(b), the desired recognition image has three colors which represent three distinct objects in Fig. 3.4(a), in which the white represents sky, green represents tree, and gray represents road. Then, we use the genetic algorithm based fuzzy ID3 method to generate fuzzy rules to analyze the image. In order to let the scene analysis system makes sense, we expect these fuzzy rules generated by the genetic algorithm based fuzzy ID3 method is reasonable and accurate enough.

(a)



(b)



(c)

Fig. 3.4. Results of the proposed approach: (a) original image, (b) desired output image, and (c) resulting image by fuzzy rule base.

In this case, we use Fig. 3.4(a) as the input image and 3.4(b) as the desired output image. And we choose gray level and vertical position as our input features. After running the genetic algorithm based fuzzy ID3 method, the generated decision tree is

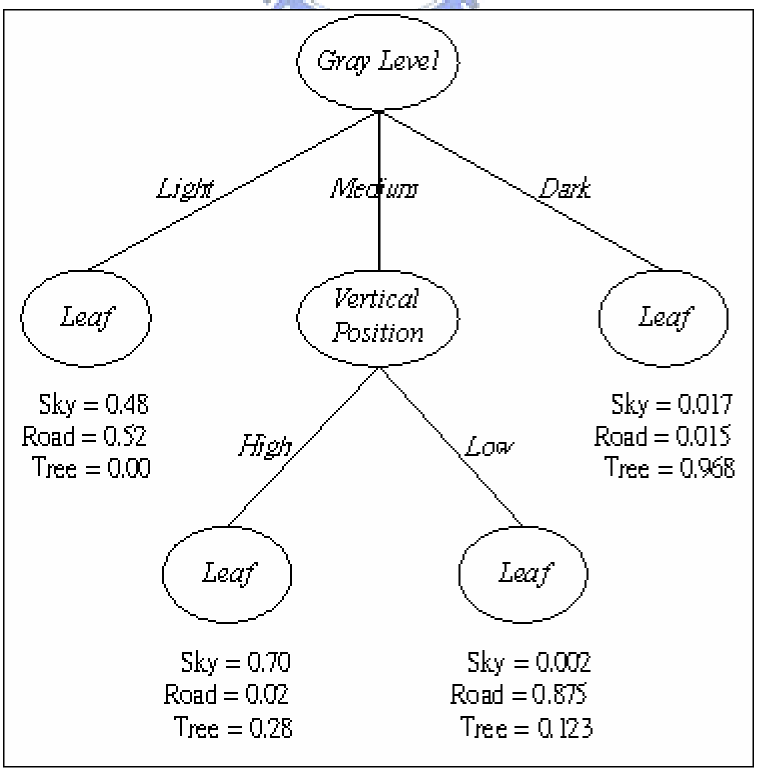


Fig. 3.5. The decision tree generated by the scene analysis system on Fig. 3.4(a).

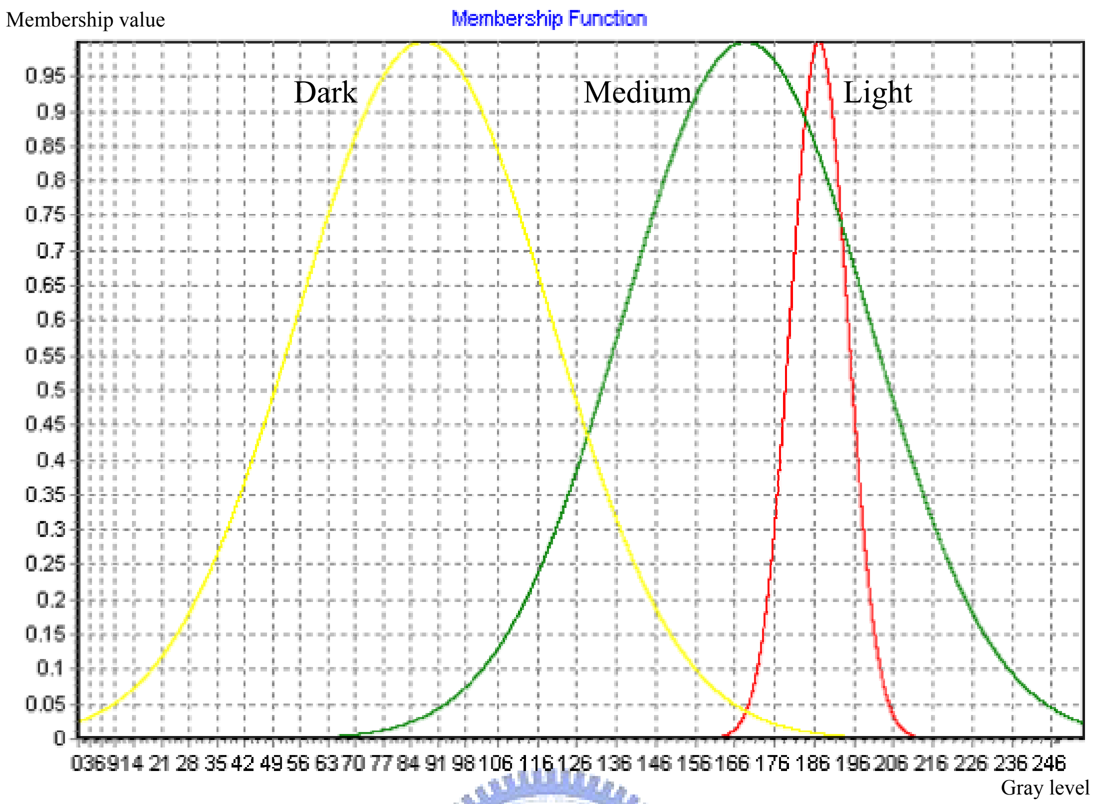


Fig. 3.6. Membership function of gray level.

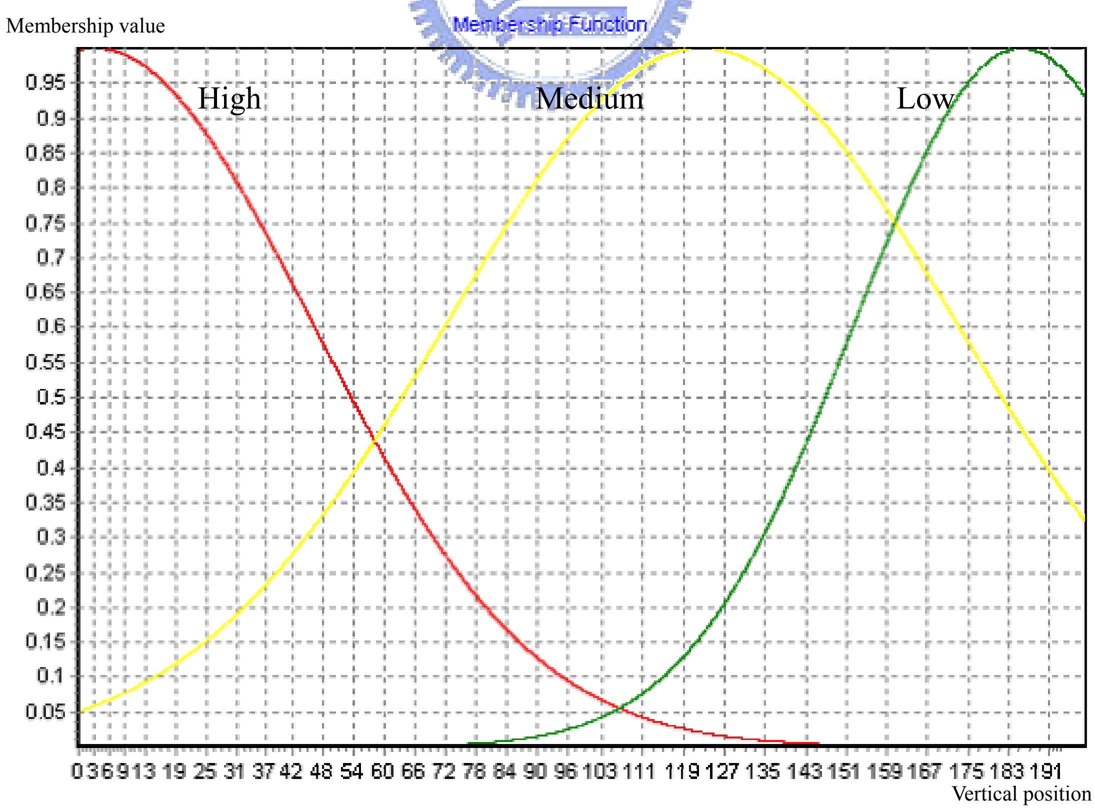


Fig. 3.7. Membership function of vertical position.

shown in Fig. 3.5. In addition, the membership function of gray level is shown in Fig. 3.6 and the membership function of vertical position is shown in Fig. 3.7. Consequently, we can infer fuzzy rules from the decision tree shown in Fig. 3.5. These fuzzy rules are as follows.

IF *Gray level* **is** *Light*

THEN *Sky* with certainty 0.480, *Road* with certainty 0.520,

and *Tree* with certainty 0.000.

IF *Gray level* **is** *Dark*

THEN *Sky* with certainty 0.017, *Road* with certainty 0.015,

and *Tree* with certainty 0.968.

IF *Gray level* **is** *Medium* **and** *Vertical position* **is** *High*

THEN *Sky* with certainty 0.700, *Road* with certainty 0.002,

and *Tree* with certainty 0.280.

IF *Gray level* **is** *Medium* **and** *Vertical position* **is** *Low*

THEN *Sky* with certainty 0.002, *Road* with certainty 0.875,

and *Tree* with certainty 0.123.

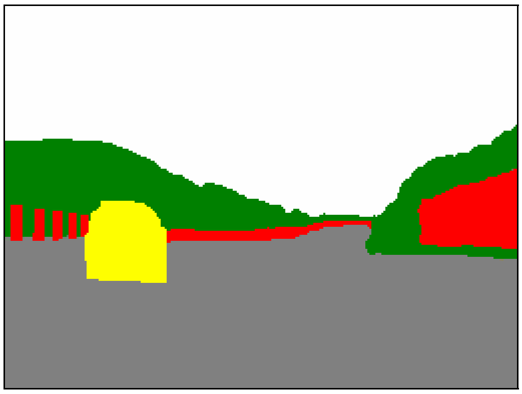
These fuzzy rules evaluate gray level first and vertical position next, because gray level has lager feature rank than vertical position. It means that gray level is more discriminative than vertical position in this case. This is because gray level can separate sky, tree, and road better than vertical position. The training result which evaluated by these four fuzzy rules is shown in Fig. 3.4(c). We obtain a region pixel accuracy of 97.7% in comparison Fig. 3.4(c) with Fig. 3.4(b). It is evident that these four fuzzy rules do locating the pixel to the appropriate natural element well. Therefore, in our proposed approach, we can say this scene analysis is an efficient and reasonable system.

3.3. Image Ground-Truthing

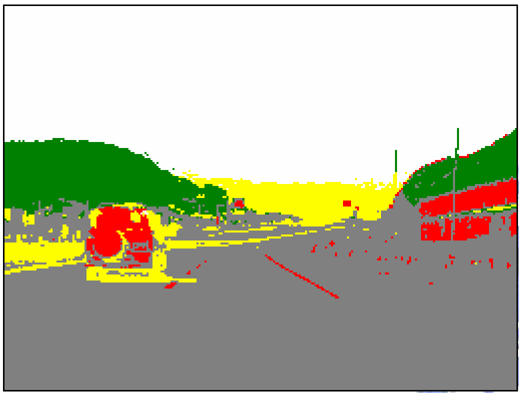
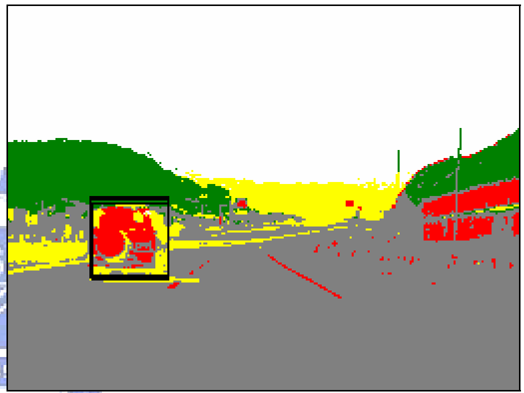
The vehicles consist of many inhomogeneous components, such as glass, license plate, tire, lamp, steel plate, and so on. By using fuzzy rules inference, it encounters great difficulty in describing the vehicle class. The vehicle class is not pure as sky, tree, and road classes. Therefore, we introduce image ground-truthing into our scene analysis system to improve the vehicle region accuracy. The ground-truthing useful for this proposed scene analysis system are summarized below.

1) The vehicles must run on the road.
2) Any vehicle highly probable has a shadow area, and this area is the darkest pixels in the scene.
3) All kinds of vehicles have a fixed height/width ratio.

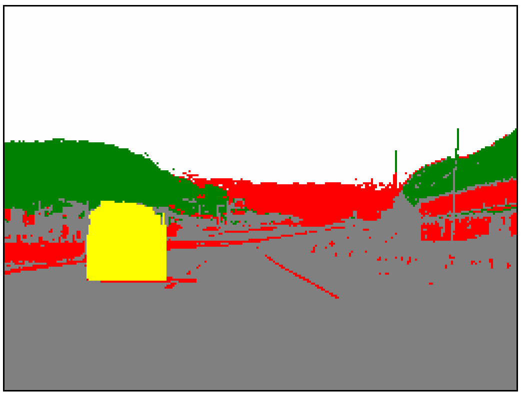
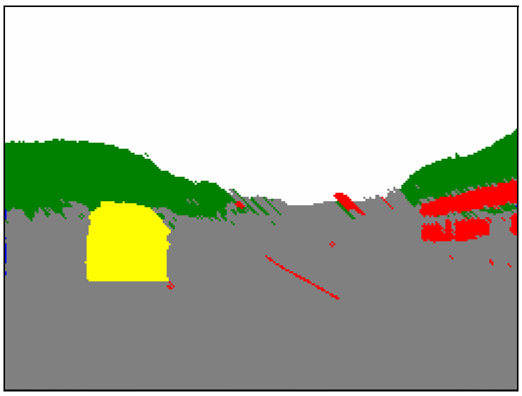
We describe our scene analysis system in Figs. 3.8(a)–(f). We establish fuzzy rule base in advance and test the input image Fig. 3.8(a) which is a 256×192 chromatic image. And Fig 3.8(b) is the desired output to verify the object region analysis accuracy which includes five colors: white denotes sky, green denotes tree, gray denotes road, yellow denotes vehicle, and red denotes the others. After fuzzy rule inference on each pixel of the image, the testing result is shown in Fig. 3.8(c). We can see the image recognition accuracy and vehicle recognition accuracy are not good enough. Therefore, we use three image ground-truthing rules to find the vehicle possible region, and sketch the vehicle possible region in black line square. The result is shown in Fig. 3.8(d). In the vehicle possible region, we use Sobel operator [19]

(a) (b)

(c) (d)

(e) (f)

Fig. 3.8. (a) original image, (b) desired output image, (c) resulting image by fuzzy rule base inferring, (d) possible vehicle region finding by image ground-truthing, (e) vehicle region refining by edge detection, and (f) final scene image obtained by image

erosion.

to detect the vehicle upper contour and replace the wrong classes in our detected vehicle region by vehicle class. Then, we assign the wrong vehicle class at the outside of the vehicle region to the others class. In Fig. 3.8(e), the vehicle was improved by our approach, but the wrong vehicle class is still not correct and needs changing. Hence, we apply erosion method [19] to remove the pixels in the wrong class. The final experimental result is shown in Fig. 3.8(f). In comparison the final result of Fig. 3.8(f) with the resulting image by fuzzy rule base inferring of Fig 3.8(c), we can find out Fig. 3.8(f) is more similar to desired output of Fig. 3.8(b) than Fig. 3.8(c). In the intelligent transportation system, the vehicle recognition is the most important figure to be considered, and we can recognize vehicle correctly in this case.

Chapter 4. Simulation and Experiment

In this chapter, we illustrate our proposed scene analysis system by simulating the outdoor images, which are chromatic image with size 256×192. These images are taken on freeway by using digital camera. In order to confirm the validity of the proposed system, we provide the desired output image by manual recognition of segmented natural objects. Then, we compare the simulated results with the desired output image. This simulation was done on Pentium4 2.4G personal computer.

4.1. Fuzzy Rule Base for Freeway Scene

As mentioned in Chapters 2 and 3, we design a scene analysis system based on fuzzy ID3 algorithm. With this algorithm, we can establish a fuzzy rule base to analyze the freeway image taken from a driver view in a driving car.

As developed by Wang and Mendel [20], fuzzy rules were generated by learning from examples. A training data with feature vector and is associated with the desired output of corresponding objects. Such image pixel constitutes an input-output pair. These fuzzy rules are a series of associations of the form "if antecedent conditions hold, then consequent conditions hold." For our system, we consider the consequent conditions are the name of the natural objects in the scene such as sky, tree, road, vehicle, and others. And the number of antecedent conditions equals the number of features.

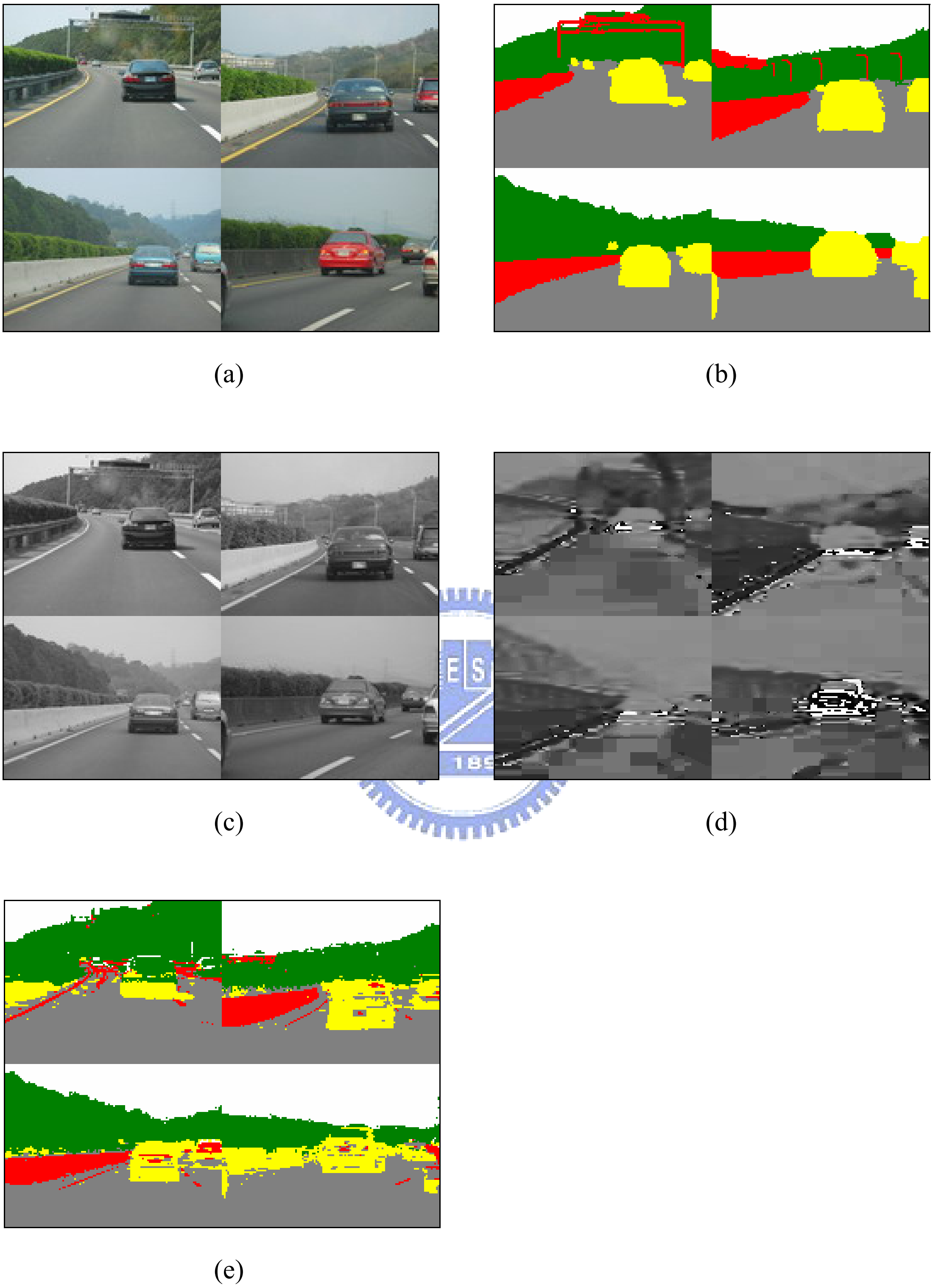


Fig. 4.1. (a) four 128×96 training images, (b) desired output image, (c) gray level of the training images, (d) hue of the training images, and (e) training result by our proposed scene analysis system.

In order to construct the fuzzy ID3 decision tree to specify the fuzzy rule base to analyze the image scene, we use four typical 128×96 outdoor images as our training data. These outdoor images are shown in Fig. 4.1(a). Then these features used for the training data are gray level, hue, and vertical position. Gray level and hue are shown in Figs. 4.1(c) 4.1(d), respectively. As to these images, the following five classes will be assigned: the sky is denoted by white, the tree is denoted by green, the road is denoted by gray, the vehicle is denoted by yellow, and the others is denoted by red. Fig. 4.1(b) is the desired output image, obtained by manual segmentation. Subsequently, we use these data as our training data set. After running the genetic algorithm based fuzzy ID3 method, the scene analysis system generates 30 fuzzy rules and provides a training result which is shown in Fig. 4.1(e). After recognizing by this fuzzy rule base, the training data set has obtained an 87.5% accuracy of region segmentation. In the following, we utilize this fuzzy rule base obtained as the default fuzzy rule base in the consequent freeway scene analysis

4.2. Simulation Results

After establishing the fuzzy rule base, we apply our scene analysis system to analyze 20 chromatic images taken from a driving car on freeway by using digital camera. Here, the first stage output represents the output after fuzzy rule base inferring, the second stage output represents the output after image ground-truthing refining, and the third stage output represents the output after image erosion shrinking. These analyzed results are shown in Figs. 4.2–4.21. In the intelligent transportation system, the vehicle recognition rate is the most important figure to be considered. Therefore, in addition to the image accuracy of segmented natural objects, we use another criterion A_V defined by

$$A_V = \frac{\begin{array}{c}\textit{Vehicle pixels in the segmented image that is classified correctly} - \\ \textit{Non-vehicle pixels in the segmented image that is classified to vehicle class}\end{array}}{\textit{Total vehicle pixels}}$$

to evaluate the vehicle recognition accuracy of our proposed scene analysis system.

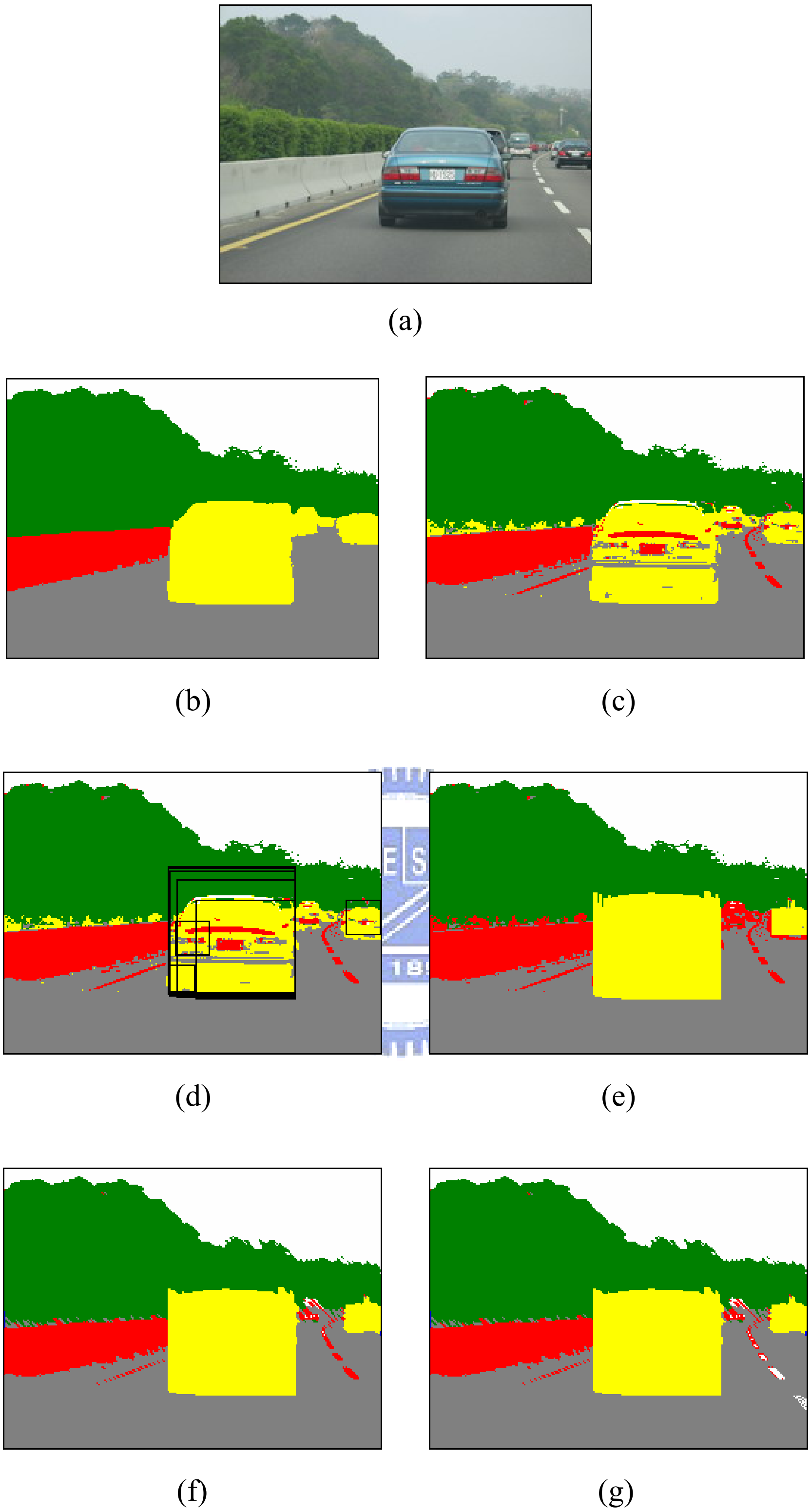
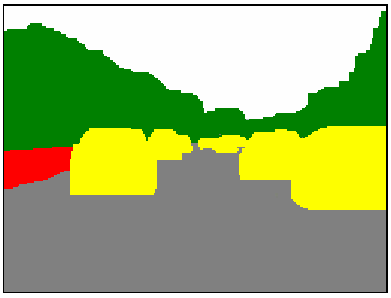
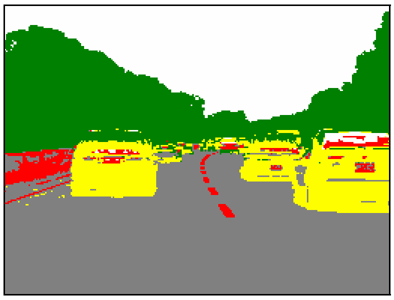


Fig. 4.2. Testing on Image 1. (a) original image, (b) desired output image, (c) resulting image by fuzzy rule base inferring, (d) possible vehicle region finding by image ground-truthing, (e) vehicle region refining by edge detection, (f) final scene image obtained by image erosion, and (g) lane line recognition.

(a)

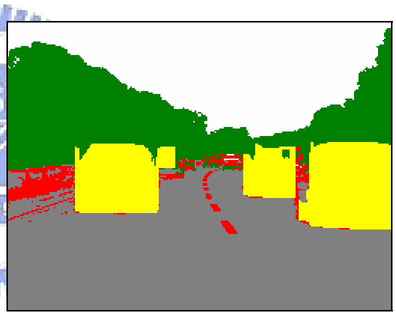


(b)

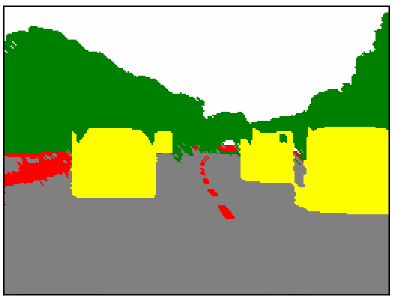


(c)

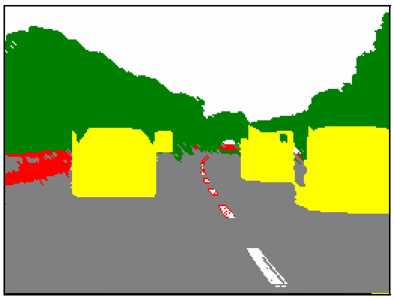
(d)



(e)



(f)



(g)

Fig. 4.3. Testing on Image 2. (a) original image, (b) desired output image, (c) resulting image by fuzzy rule base inferring, (d) possible vehicle region finding by image ground-truthing, (e) vehicle region refining by edge detection, (f) final scene image obtained by image erosion, and (g) lane line recognition.

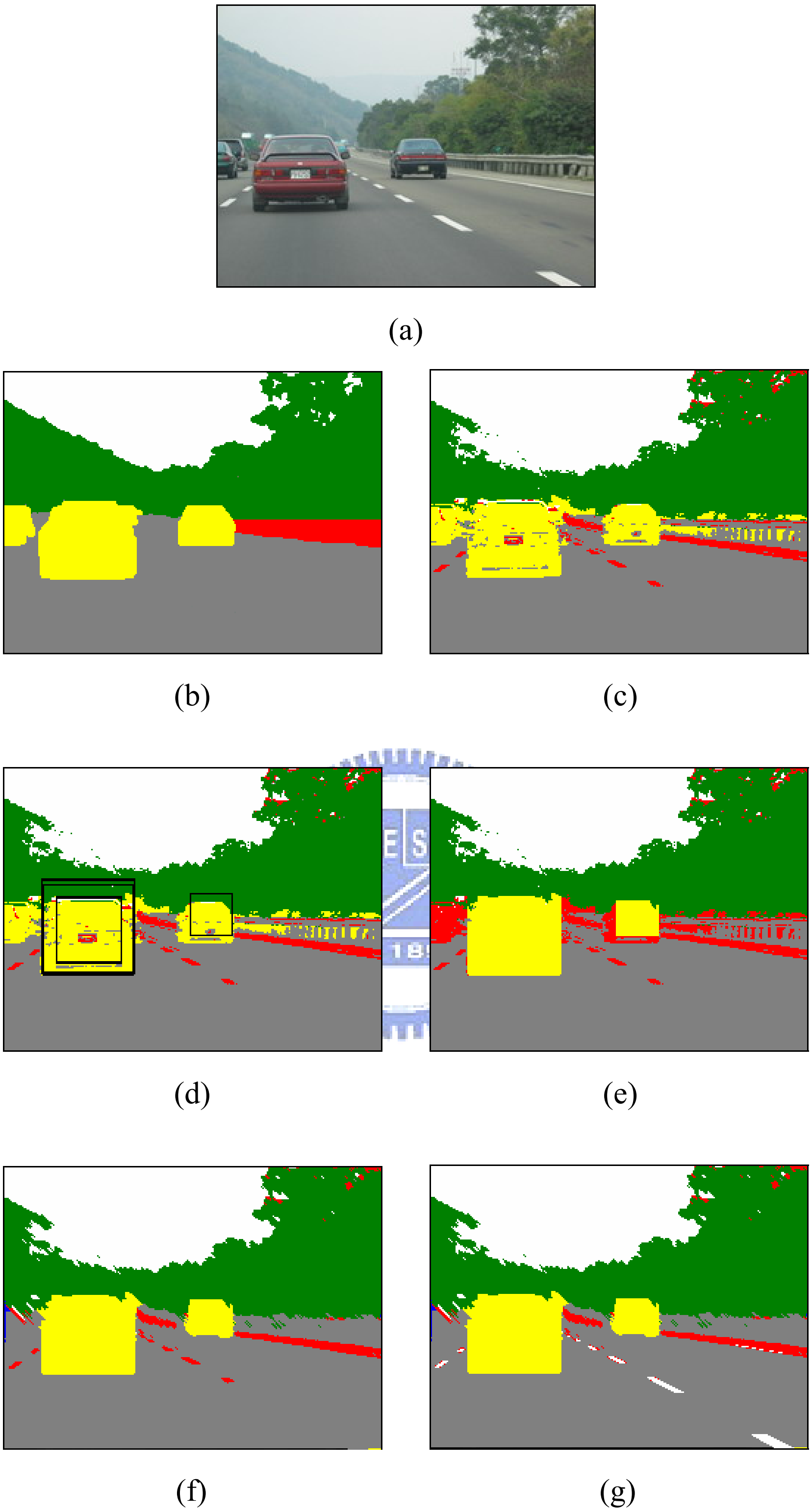
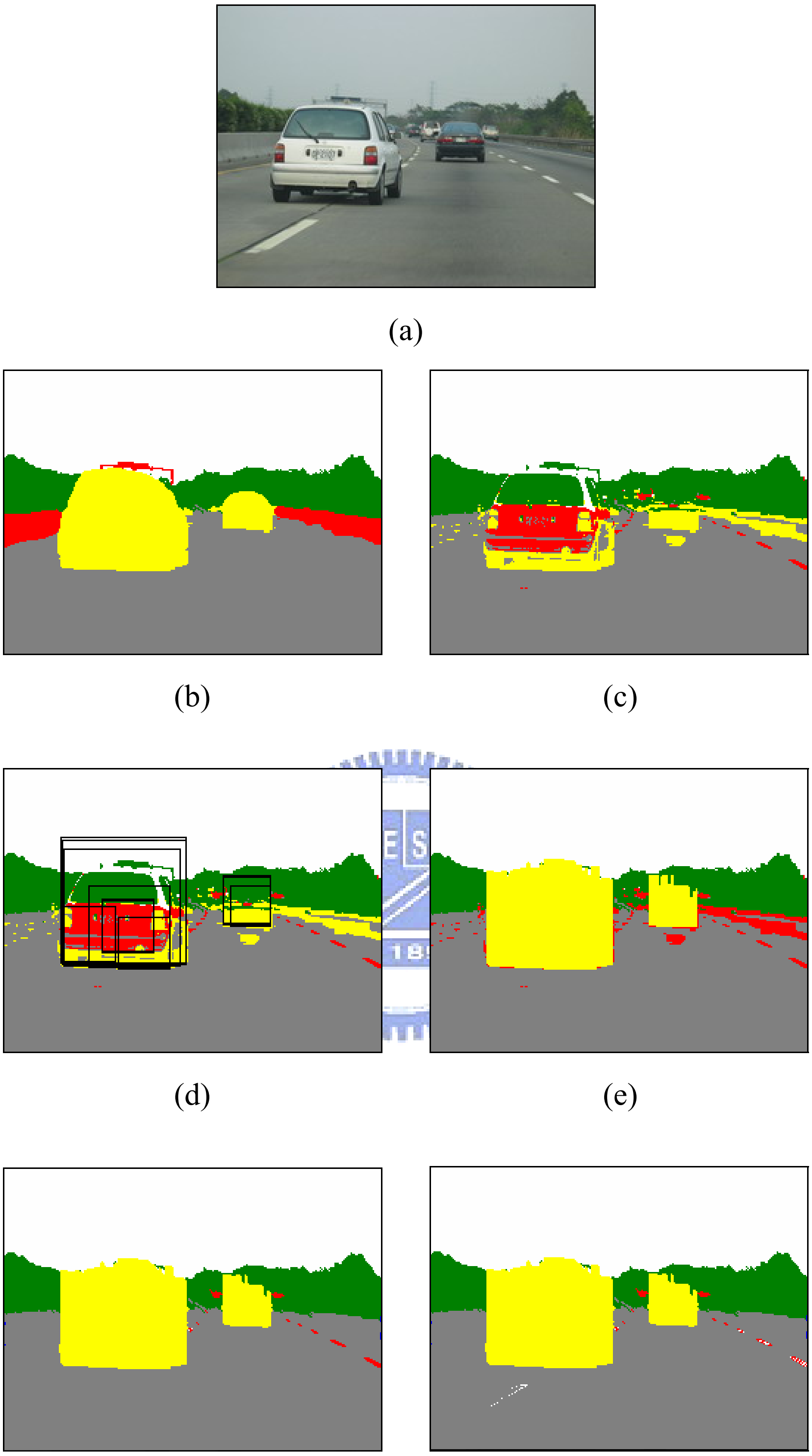


Fig. 4.4. Testing on image 3. (a) original image, (b) desired output image, (c) resulting image by fuzzy rule base inferring, (d) possible vehicle region finding by image ground-truthing, (e) vehicle region refining by edge detection, (f) final scene image obtained by image erosion, and (g) lane line recognition.



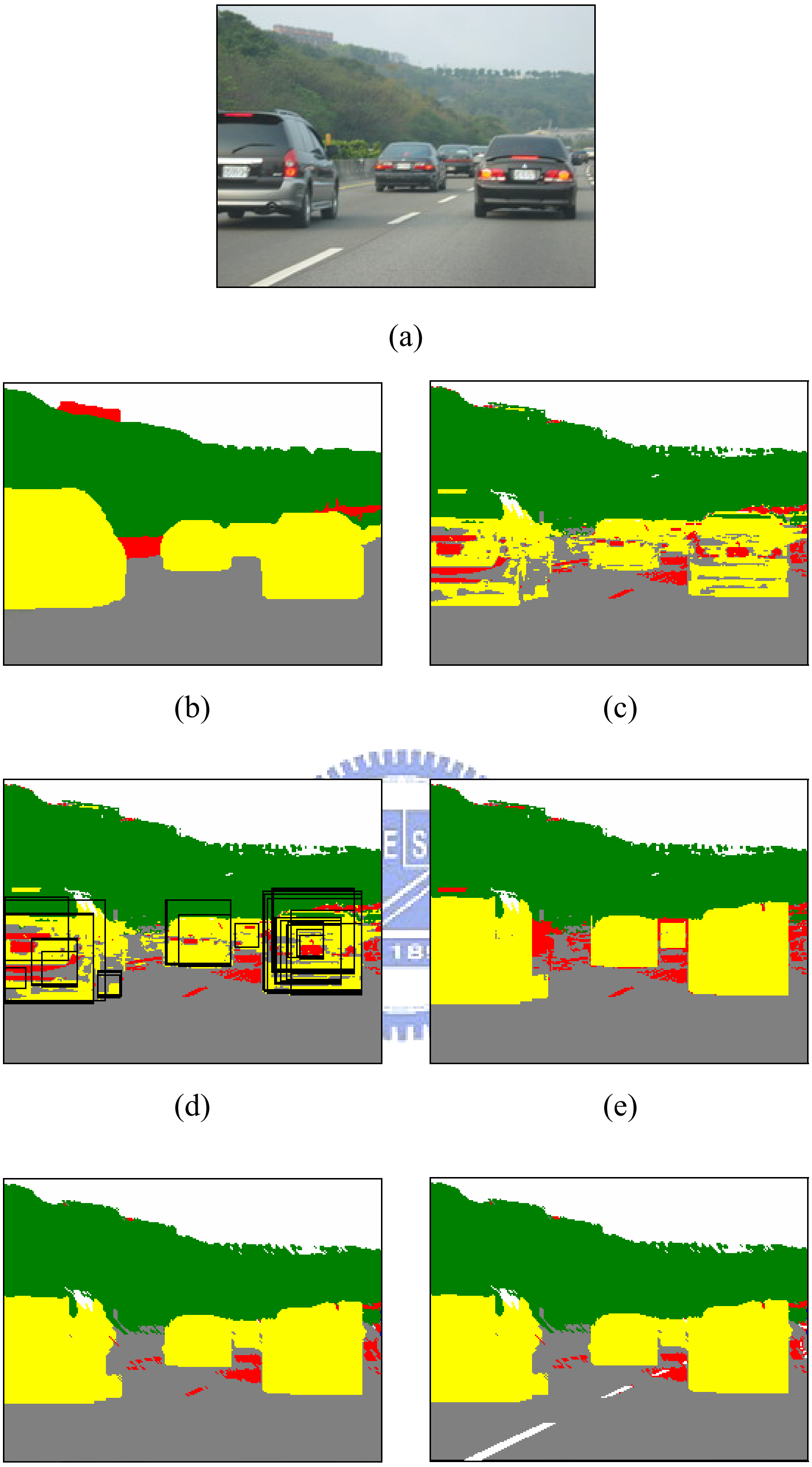
(a)

(b) (c)

(d) (e)

(f) (g)

Fig. 4.5. Testing on Image 4. (a) original image, (b) desired output image, (c) resulting image by fuzzy rule base inferring, (d) possible vehicle region finding by image ground-truthing, (e) vehicle region refining by edge detection, (f) final scene image obtained by image erosion, and (g) lane line recognition.



(a)

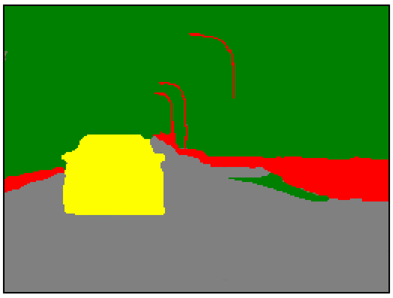
(b) (c)

(d) (e)

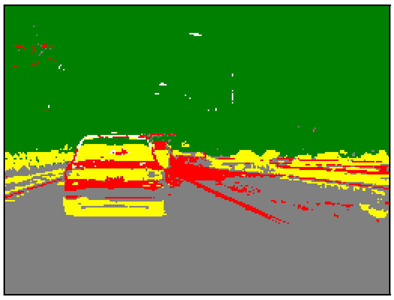
(f) (g)

Fig. 4.6. Testing on Image 5. (a) original image, (b) desired output image, (c) resulting image by fuzzy rule base inferring, (d) possible vehicle region finding by image ground-truthing, (e) vehicle region refining by edge detection, (f) final scene image obtained by image erosion, and (g) lane line recognition.

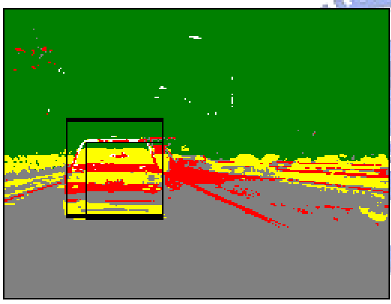
(a)



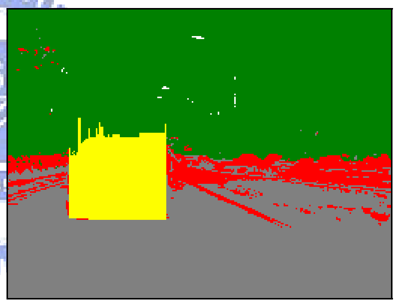
(b)



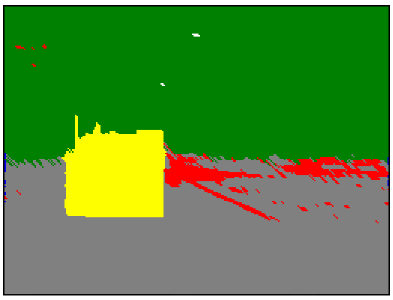
(c)



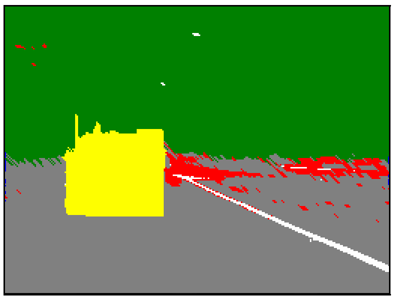
(d)



(e)



(f)

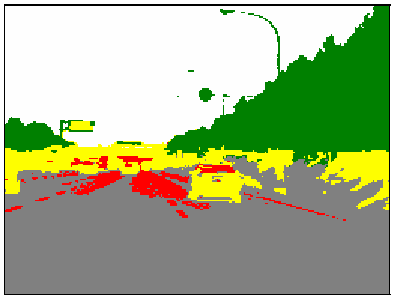


(g)

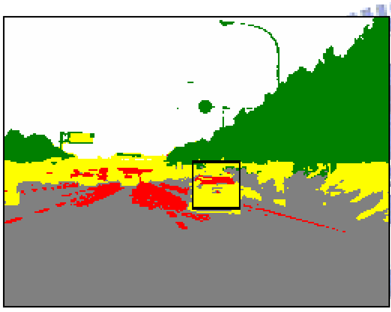
Fig. 4.7. Testing on Image 6. (a) original image, (b) desired output image, (c) resulting image by fuzzy rule base inferring, (d) possible vehicle region finding by image ground-truthing, (e) vehicle region refining by edge detection, (f) final scene image obtained by image erosion, and (g) lane line recognition.

(a)

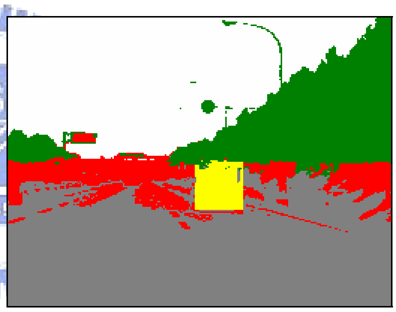
(b)



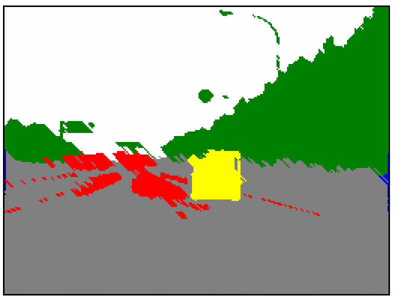
(c)



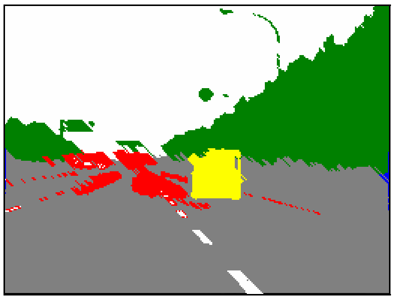
(d)



(e)



(f)



(g)

Fig. 4.8. Testing on Image 7. (a) original image, (b) desired output image, (c) resulting image by fuzzy rule base inferring, (d) possible vehicle region finding by image ground-truthing, (e) vehicle region refining by edge detection, (f) final scene image obtained by image erosion, and (g) lane line recognition.

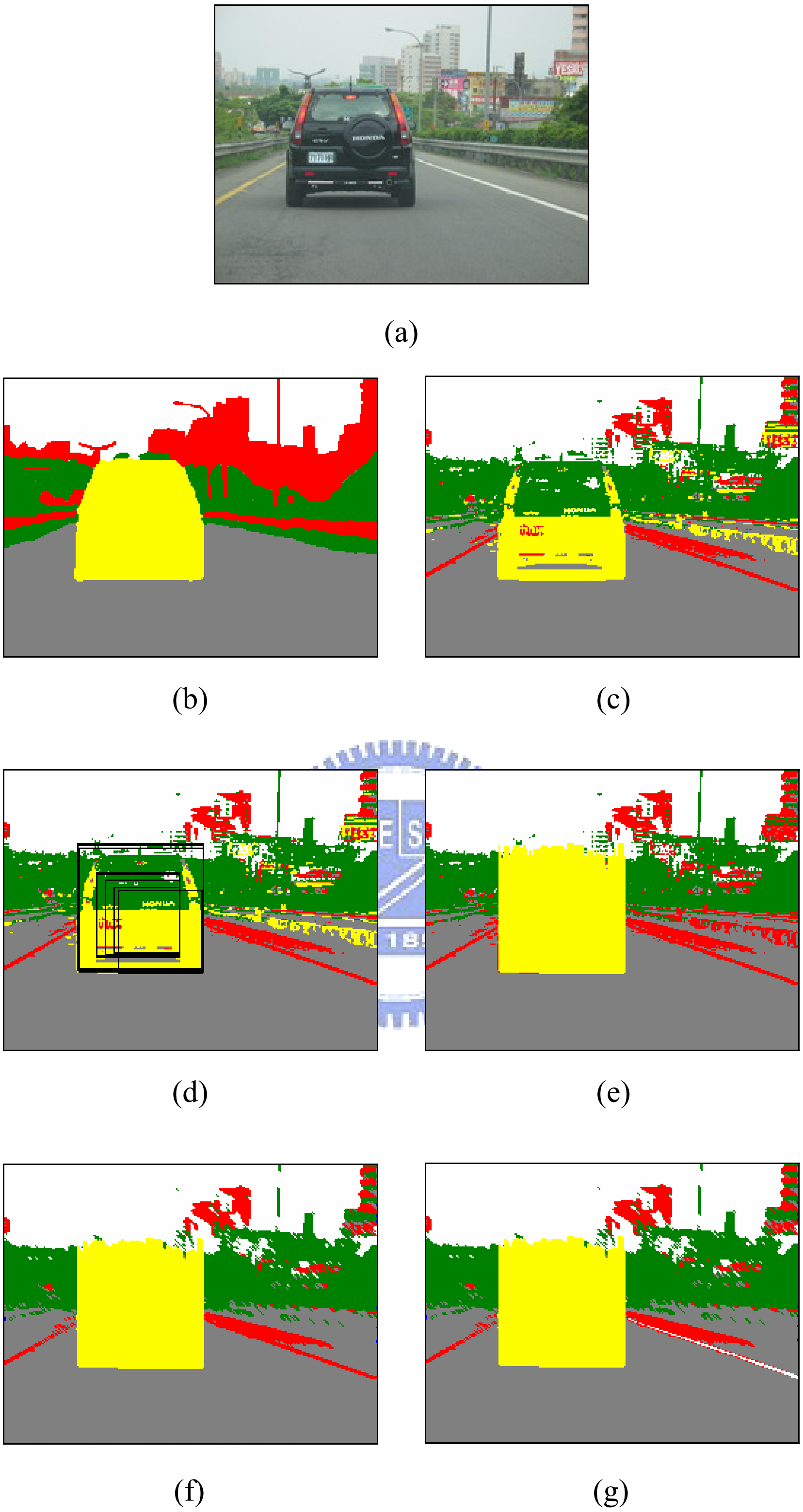


Fig. 4.9. Testing on Image 8. (a) original image, (b) desired output image, (c) resulting image by fuzzy rule base inferring, (d) possible vehicle region finding by image ground-truthing, (e) vehicle region refining by edge detection, (f) final scene image obtained by image erosion, and (g) lane line recognition.

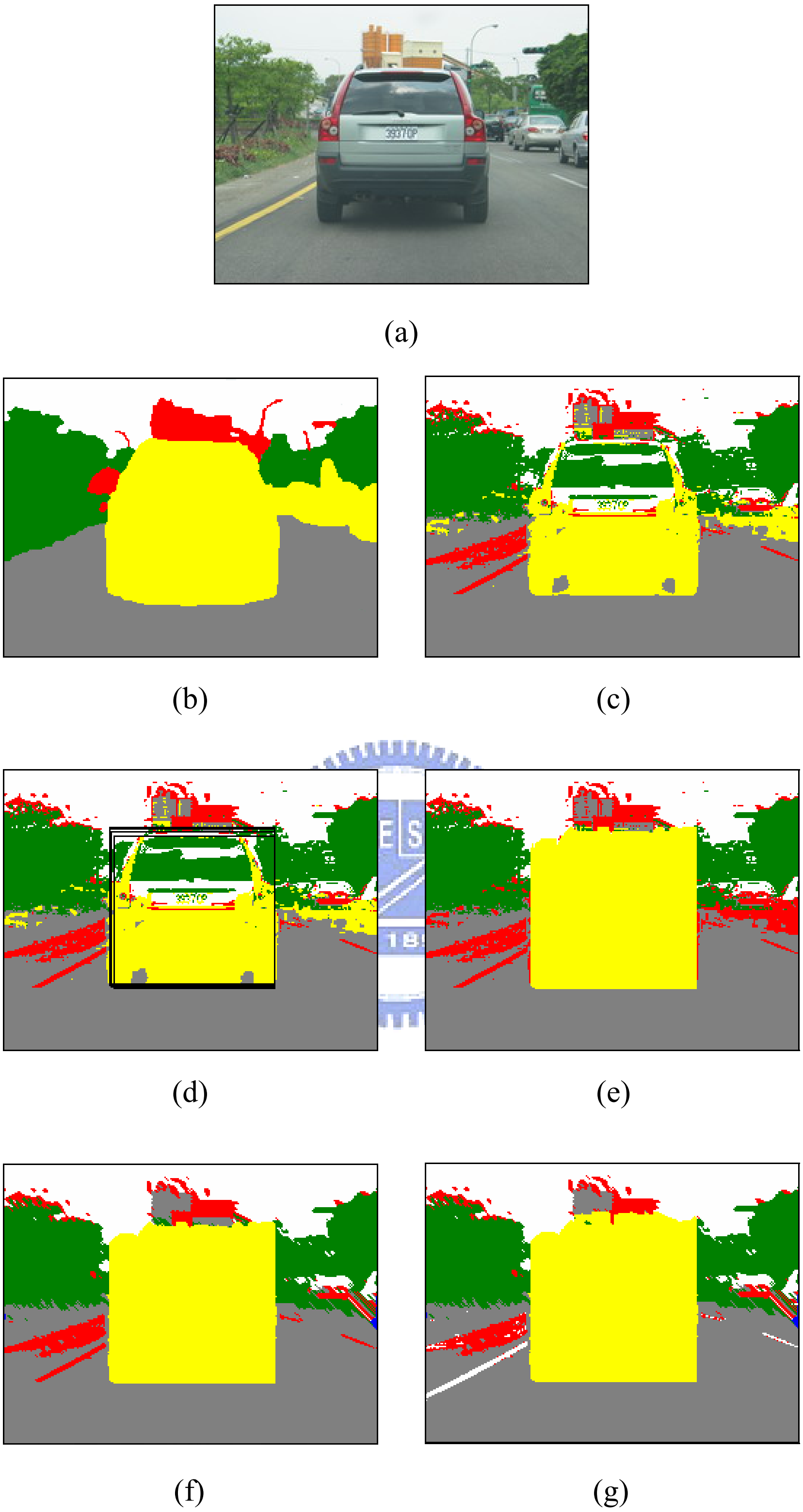
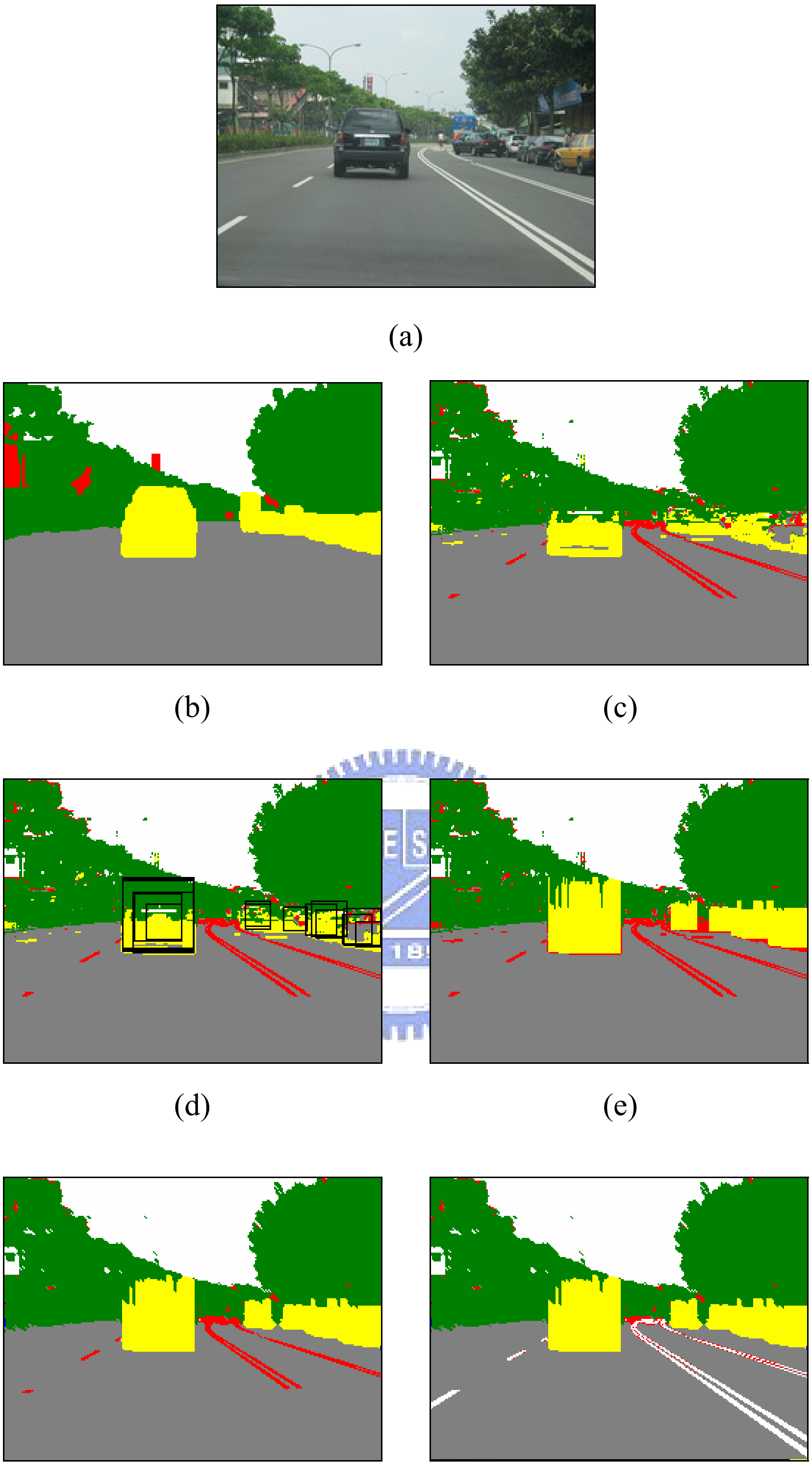


Fig. 4.10. Testing on Image 9. (a) original image, (b) desired output image, (c) resulting image by fuzzy rule base inferring, (d) possible vehicle region finding by image ground-truthing, (e) vehicle region refining by edge detection, (f) final scene image obtained by image erosion, and (g) lane line recognition.



(a)

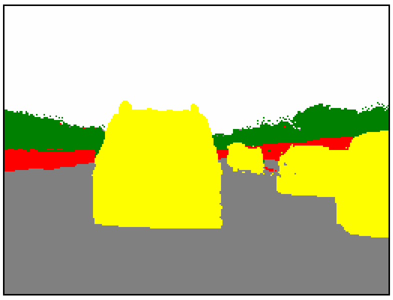
(b) (c)

(d) (e)

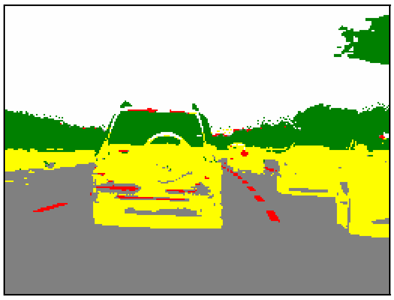
(f) (g)

Fig. 4.11. Testing on Image 10. (a) original image, (b) desired output image, (c) resulting image by fuzzy rule base inferring, (d) possible vehicle region finding by image ground-truthing, (e) vehicle region refining by edge detection, (f) final scene image obtained by image erosion, and (g) lane line recognition.

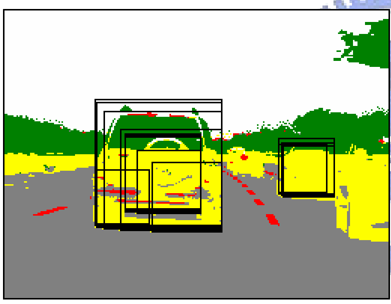
(a)



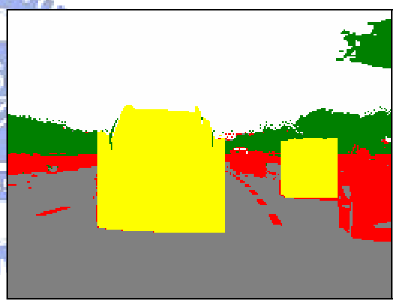
(b)



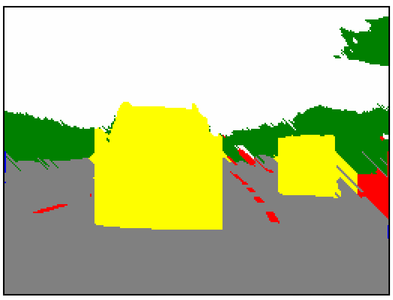
(c)



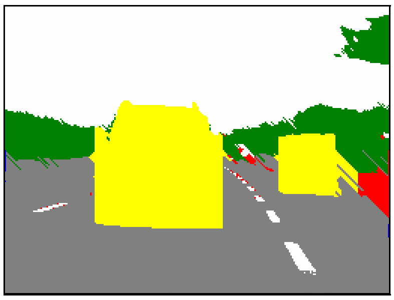
(d)



(e)



(f)



(g)

Fig. 4.12. Testing on Image 11. (a) original image, (b) desired output image, (c) resulting image by fuzzy rule base inferring, (d) possible vehicle region finding by image ground-truthing, (e) vehicle region refining by edge detection, (f) final scene image obtained by image erosion, and (g) lane line recognition.

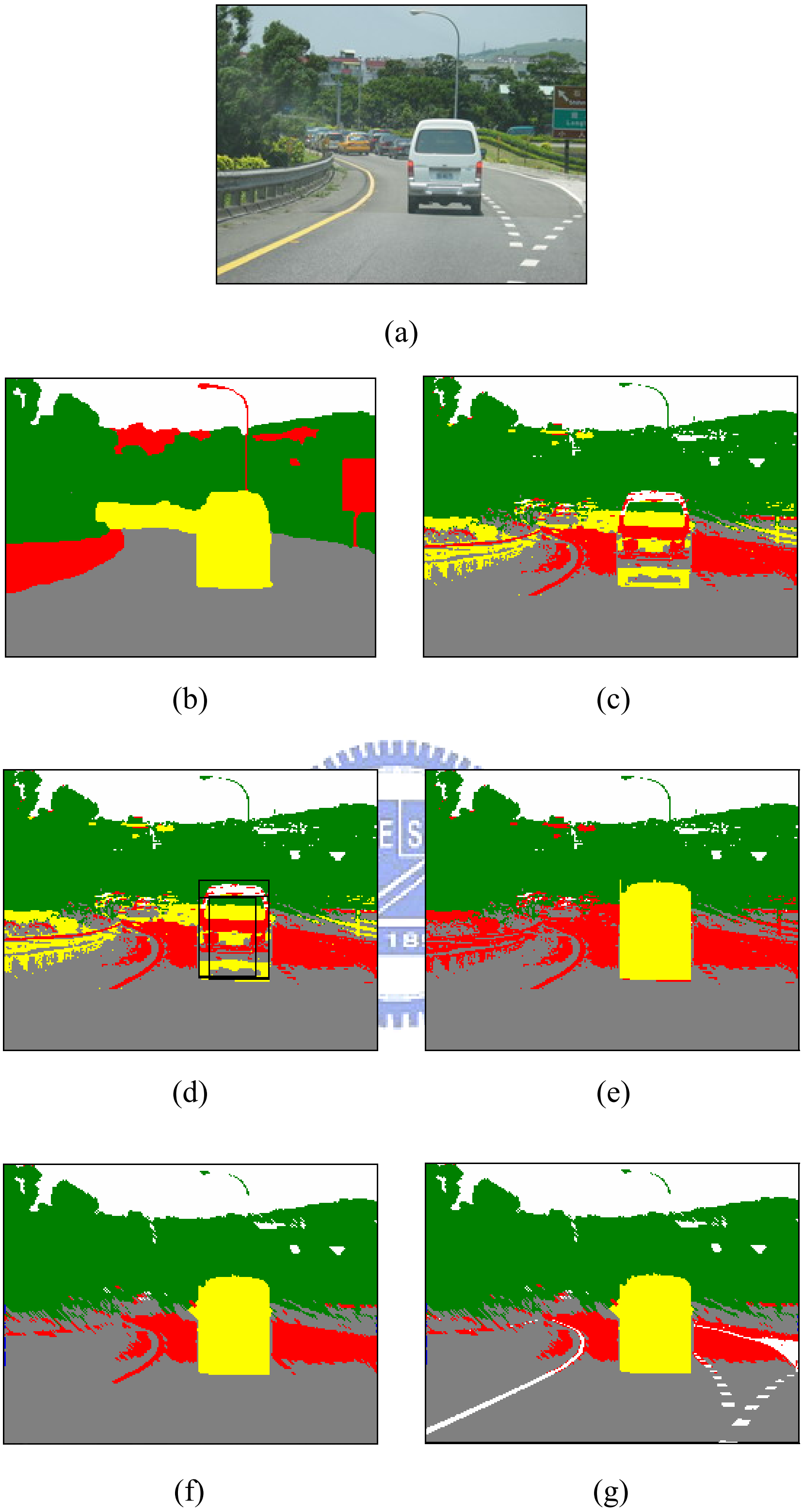
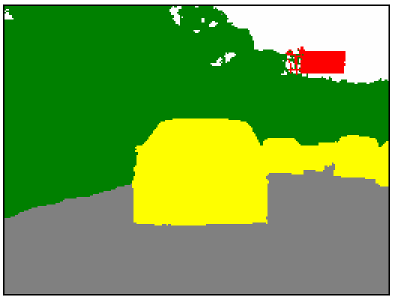
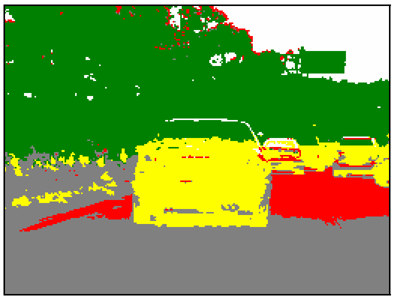


Fig. 4.13. Testing on Image 12. (a) original image, (b) desired output image, (c) resulting image by fuzzy rule base inferring, (d) possible vehicle region finding by image ground-truthing, (e) vehicle region refining by edge detection, (f) final scene image obtained by image erosion, and (g) lane line recognition.

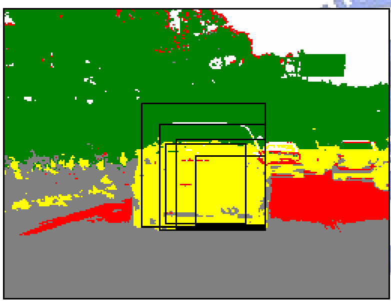
(a)



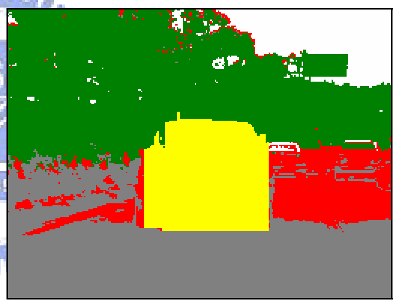
(b)



(c)



(d)



(e)

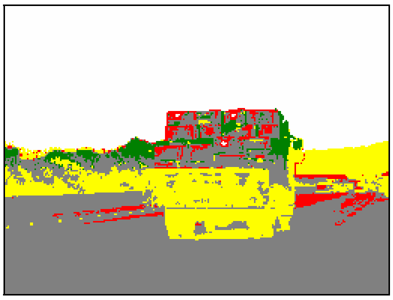
(f)

(g)

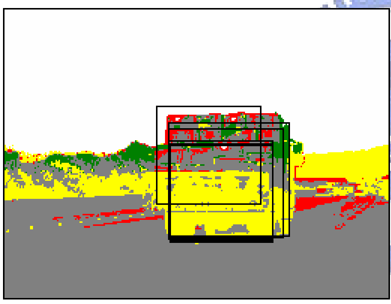
Fig. 4.14. Testing on Image 13. (a) original image, (b) desired output image, (c) resulting image by fuzzy rule base inferring, (d) possible vehicle region finding by image ground-truthing, (e) vehicle region refining by edge detection, (f) final scene image obtained by image erosion, and (g) lane line recognition.

(a)

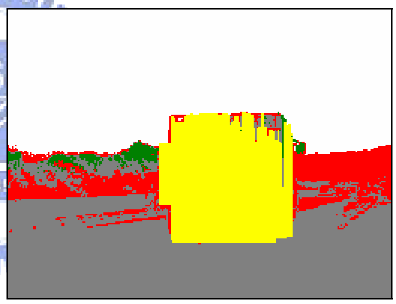
(b)



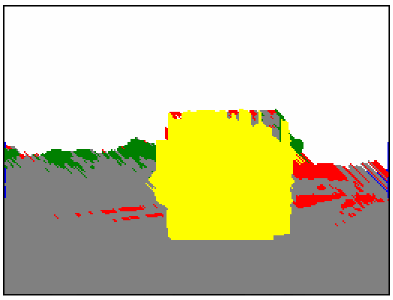
(c)



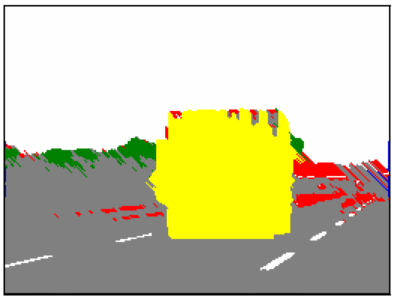
(d)



(e)

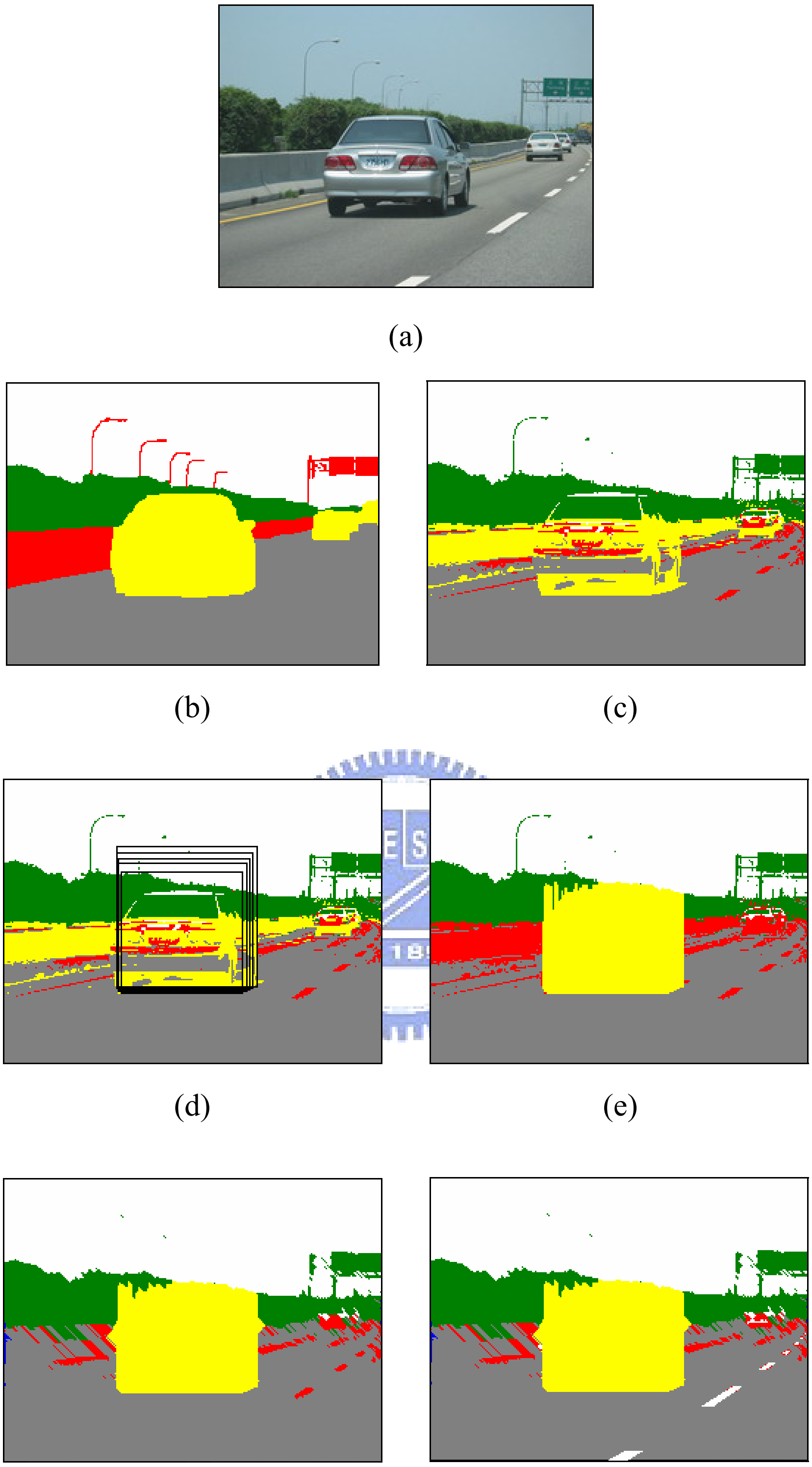


(f)



(g)

Fig. 4.15. Testing on Image 14. (a) original image, (b) desired output image, (c) resulting image by fuzzy rule base inferring, (d) possible vehicle region finding by image ground-truthing, (e) vehicle region refining by edge detection, (f) final scene image obtained by image erosion, and (g) lane line recognition.

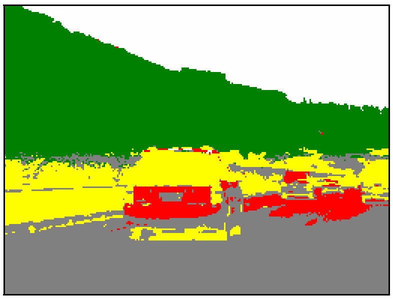


(a) (b) (c) (d) (e) (f) (g)

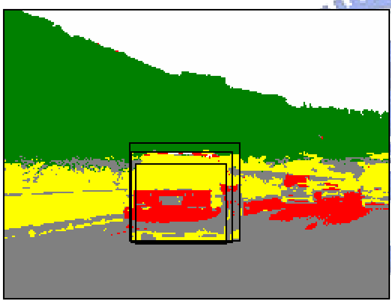
Fig. 4.16. Testing on Image 15. (a) original image, (b) desired output image, (c) resulting image by fuzzy rule base inferring, (d) possible vehicle region finding by image ground-truthing, (e) vehicle region refining by edge detection, (f) final scene image obtained by image erosion, and (g) lane line recognition.

(a)

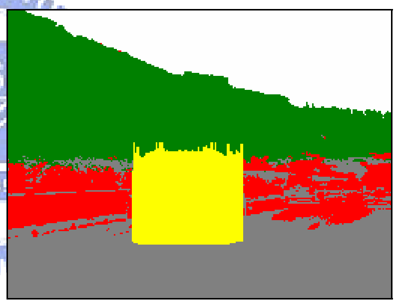
(b)



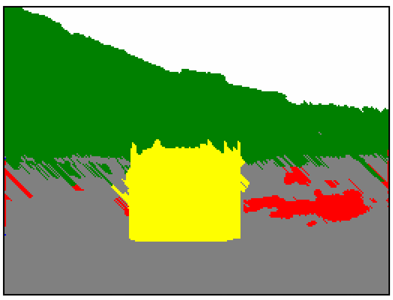
(c)



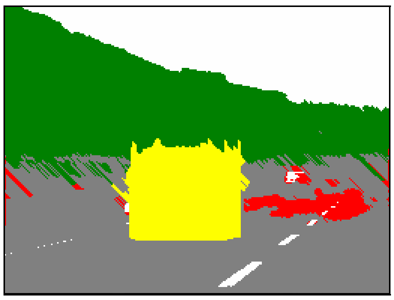
(d)



(e)

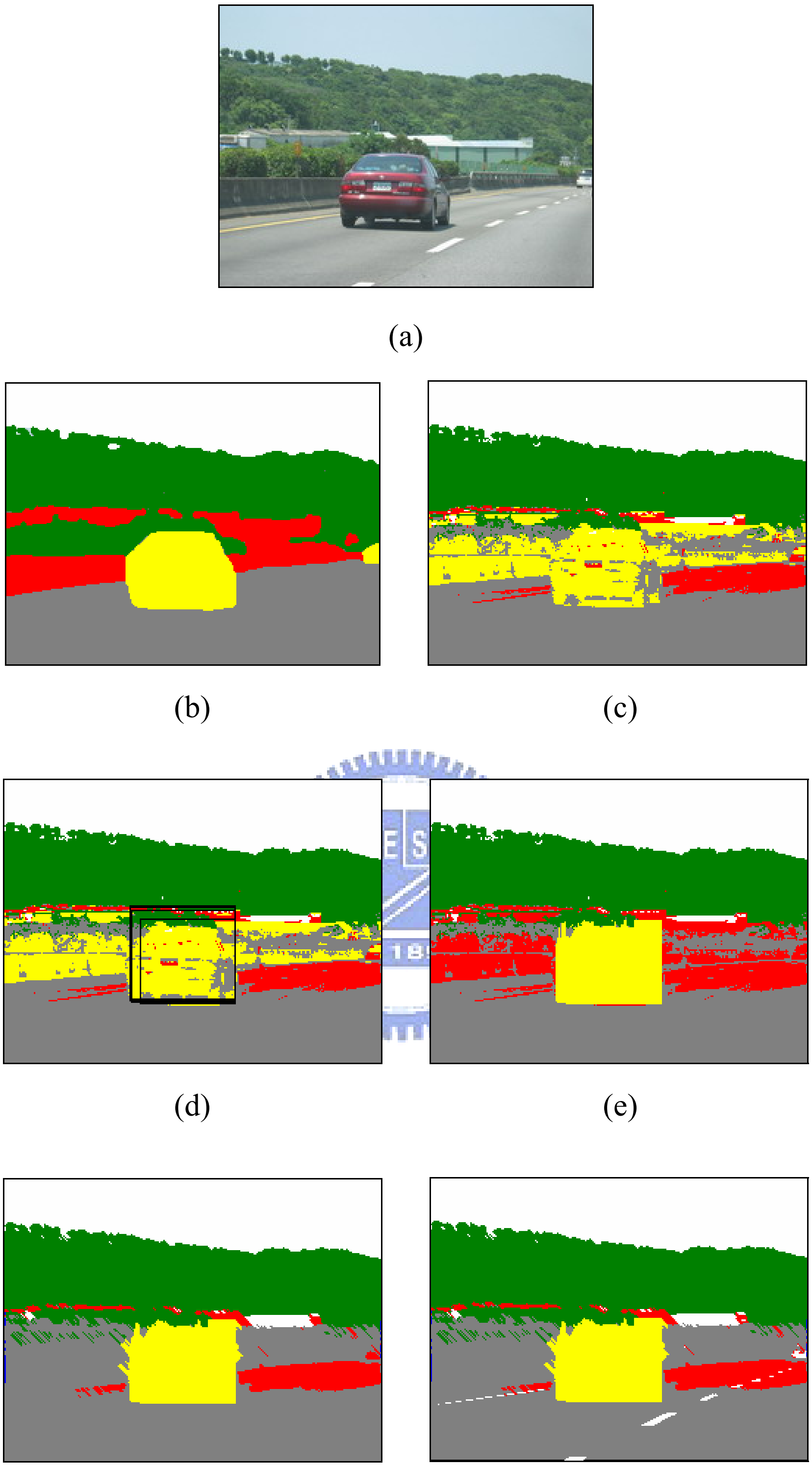


(f)



(g)

Fig. 4.17. Testing on Image 16. (a) original image, (b) desired output image, (c) resulting image by fuzzy rule base inferring, (d) possible vehicle region finding by image ground-truthing, (e) vehicle region refining by edge detection, (f) final scene image obtained by image erosion, and (g) lane line recognition.



(a) (b) (c) (d) (e) (f) (g)

Fig. 4.18. Testing on Image 17. (a) original image, (b) desired output image, (c) resulting image by fuzzy rule base inferring, (d) possible vehicle region finding by image ground-truthing, (e) vehicle region refining by edge detection, (f) final scene image obtained by image erosion, and (g) lane line recognition.

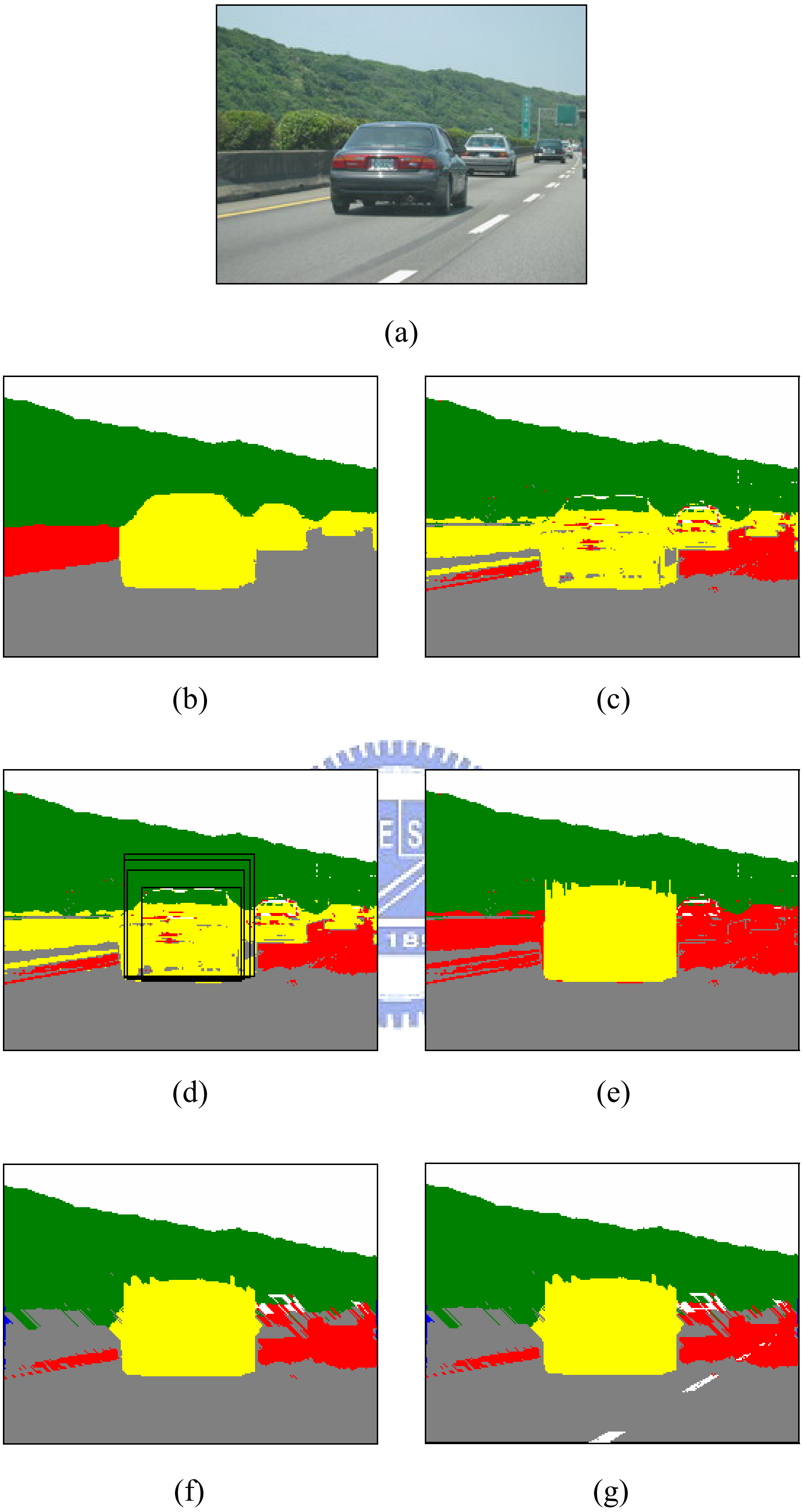


Fig. 4.19. Testing on Image 18. (a) original image, (b) desired output image, (c) resulting image by fuzzy rule base inferring, (d) possible vehicle region finding by image ground-truthing, (e) vehicle region refining by edge detection, (f) final scene image obtained by image erosion, and (g) lane line recognition.

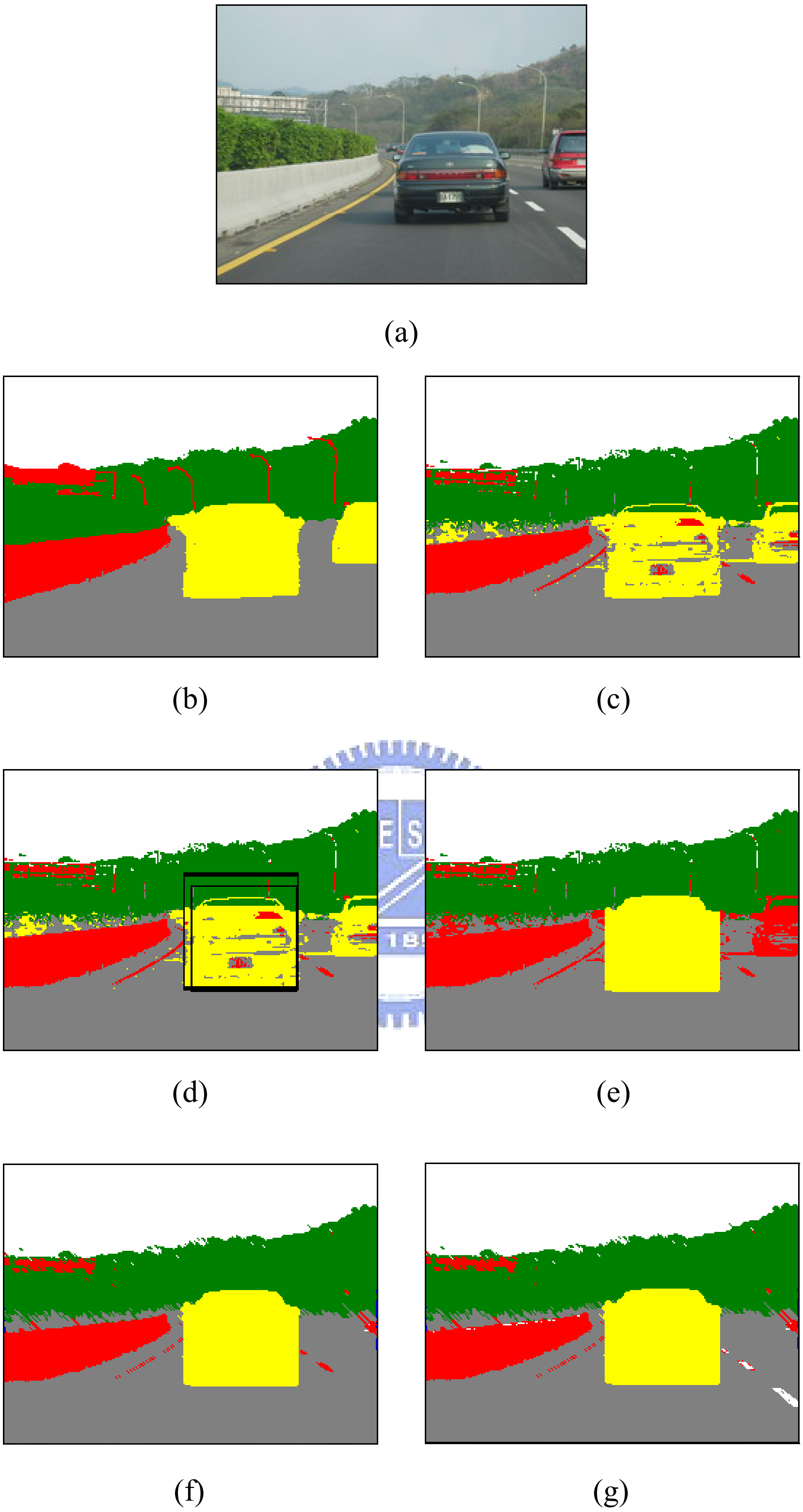


Fig. 4.20. Testing on Image 19. (a) original image, (b) desired output image, (c) resulting image by fuzzy rule base inferring, (d) possible vehicle region finding by image ground-truthing, (e) vehicle region refining by edge detection, (f) final scene image obtained by image erosion, and (g) lane line recognition.

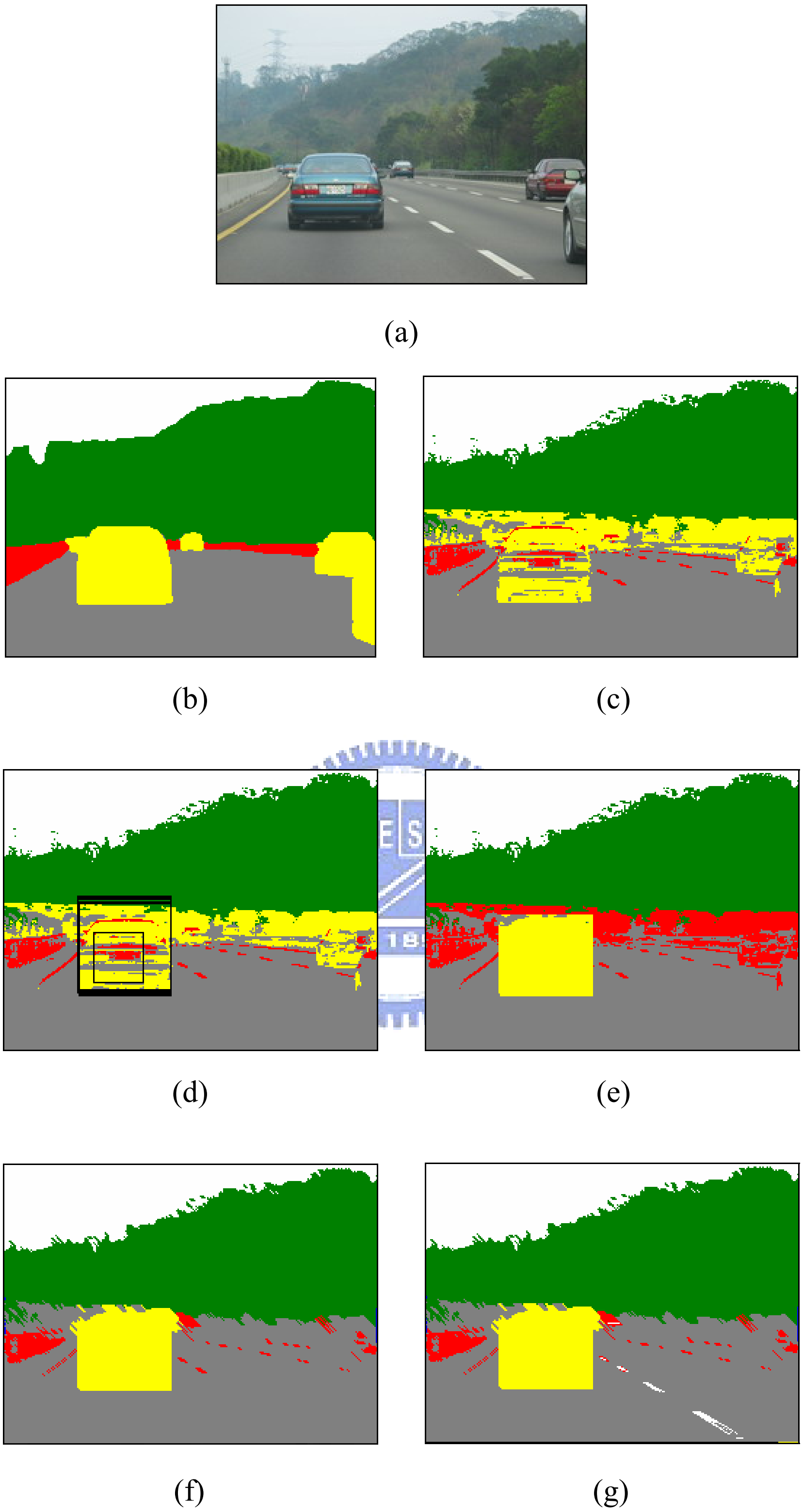


Fig. 4.21. Testing on Image 20. (a) original image, (b) desired output image, (c) resulting image by fuzzy rule base inferring, (d) possible vehicle region finding by image ground-truthing, (e) vehicle region refining by edge detection, (f) final scene image obtained by image erosion, and (g) lane line recognition.

In summary, the overall accuracy comparison of the above 20 images between each stage output are illustrated in TABLE I. After testing 20 chromatic images, we obtained 87.79% accuracy of the overall scene images and 84.12% accuracy of vehicles in the scene images. As demonstrated in the successful application on freeway images, the proposed scene analysis system is general and robust.

TABLE I

ACCURACY COMPARISON BETWEEN EACH STAGE OUTPUT

	Considering all vehicles in the image		Ignoring far away vehicles and specify far away vehicles as others	
	Image accuracy	Vehicle accuracy	Image accuracy	Vehicle accuracy
1st stage output	84.91%	35.57%	84.02%	28.35%
2nd stage output	87.93%	73.08%	88.88%	84.06%
3rd stage output	87.66%	73.87%	87.79%	84.12%

Chapter 5. Conclusion

This thesis has presented an automatic interpretation system of outdoor scene image by using a set of fuzzy rules. The scene analysis system based on the fuzzy rules derived from the GA based fuzzy ID3 technique has been developed in this thesis. The system is trained by a set of training images. Then, we construct the decision tree, and extract fuzzy rules from this decision tree. Finally, we apply this scene analysis system to analyze outdoor images.

Based on the knowledge bases of the image and supervised learning algorithm, the scene analysis use fuzzy IF-THEN rules to interpret image scenes can be generated. If more different objects in the images are analyzed, the more fuzzy rules will generate by the scene analysis system. After testing on a set of chromatic images, we obtained 87.79% accuracy of the scene images and 84.12% accuracy of the vehicles in the scene images. As demonstrated in the successful application on freeway images, the proposed scene analysis system is general and robust.

This proposed system develops an automatic scene analysis scheme by computer and we can extent this system to various fields, e.g. geographic image analysis, medical image analysis, robotic vision, and so on. But in each different application, we must add the specific domain knowledge to select the feature and refine the segmented image..

Knowledge plays a critical role in an intelligent vision system. More attention should be paid to knowledge structure and knowledge processing. Conventional image processing techniques are necessary, but not enough. In the field related to knowledge processing, fuzzy ID3 algorithm is a powerful tool.

References

[1] W. Zhang and M. Sugeno, "A fuzzy approach to scene understanding," in *Proc. Second IEEE Int. Conf. Fuzzy Systems*, 1993, pp. 564–569.

[2] Y. Nakagawa and K. Hirota, "Fundamentals of fuzzy knowledge base for image understanding," in *Proc. IEEE Int. Conf. Fuzzy Systems*, March 1995, pp. 1137–1142.

[3] K. Miyajima and A. Ralescu, "Modeling of natural objects including fuzziness and application to image understanding," in *Proc. Second IEEE Int. Conf. Fuzzy Systems*, 1993, pp. 1049–1054.

[4] J. Keller, R. Krishnapuram, R. Yager, and L. Zadeh, "Fuzzy decision models in computer vision," *Fuzzy Sets, Neural Network, and Soft Computing*, pp. 213–232, 1994.

[5] G.L. Foresti and W. Vanzella, "Generalized neural trees for outdoor scene understanding," in *Proc. 2000 Int. Conf. Image Processing*, 2000, pp. 336–339.

[6] W. F. Bischof and T. Caelli, "Scene understanding by rule evaluation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 19, pp. 1284–1288, 1997.

[7] M. Zhang, L. O. Hall and D. B. Goldgof, "A generic knowledge-guided image segmentation and labeling system using fuzzy clustering algorithm," *IEEE Trans. Syst., Man, Cybern. B*, vol. 32, no.5, pp.571–582, October 2002.

[8] C. Fan and X. Ye, W. Gu, "KRUS: A knowledge-based road scene understanding system," in *Proc. Fourteenth Int. Conf. Pattern Recognition*, 1998, pp.731–733.

[9] J. R. Quinlan, "Induction of decision trees," *Machine Learning*, vol. 1, pp. 81–106, 1986.

[10] M. Umano, H. Okamoto, I. Hatono, and H.Tamura, "Fuzzy decision trees by fuzzy ID3 algorithm and its application to diagnosis systems," in *Proc. Third IEEE Conf. Fuzzy Systems*, 1994, pp. 2113–2118.

[11] J. Rives, "FID3: fuzzy induction decision tree," in *Proc. First Int. Symposium Uncertainty Modeling and Analysis*, Dec. 1990, pp. 457–162.

[12] J. R. Quinlan, *C4.5 : Programs for Machine Learning*. San Mateo, CA: Morgan Kauffman, 1993.

[13] L. Kurgan and K. J. Cios, "CAIM discretization algorithm," *IEEE Trans. Knowledge Data Eng.,* vol. 16, pp. 145–153, Feb. 2004.

[14] C. T. Lin and C. S. G. Lee, *Neural fuzzy systems A neuro-fuzzy synergism to intelligent systems*, Prentice-Hall, Upper Saddle River, New Jersey, 1996.

[15] N. R. Pal and S. Chakraborty, "Fuzzy rule extraction from ID3-type decision trees for real data," *IEEE Trans. Syst., Man, Cybern. B*, vol. 31, no. 5, pp. 745–754, October 2001.

[16] J. Gauch and C. W. Hsia, "A comparison of three color image segmentation algorithm in four color spaces," *Visual Communication and Image Processing*, vol.1818, pp. 1168–1181, 1992.

[17] J. Freixenet, X. Llado, J. Marti, and X. Cufi, "Use of decision trees in colour feature selection. Application to object recognition in outdoor scenes," *in Proc. 2000 Int. Conf. Image Processing*, 2000, pp. 496–499.

[18] R. C. Gonzalez and R. E. Woods, *Digital Image Processing,* New Jersey, Prentice-Hall, 2002.

[19] I. Pitas, *Digital Image Processing Algorithm and Applications,* Canada, John Wiley & Sons, 2000.

[20] L. X. Wang and J. M. Mendel, "Generating fuzzy rules by learning from example," *IEEE Trans. Syst., Man, Cybern.*, vol. 22, pp. 1414–1427, 1992.