

Vacant Parking Space Detection Based on Plane-based Bayesian Hierarchical Framework

Ching-Chun Huang, *Member, IEEE*, Yu-Shu Tai, and Sheng-Jyh Wang, *Member, IEEE*

Abstract—In this paper, we propose a vacant parking space detection system that operates day and night. In the daytime, the major challenges of the system include dramatic lighting variations, shadow effect, inter-object occlusion, and perspective distortion. In the nighttime, the major challenges include insufficient illumination and complicated lighting conditions. To overcome these problems, we propose a plane-based method which adopts a structural 3-D parking lot model consisting of plentiful planar surfaces. The plane-based 3-D scene model plays a key part in handling inter-object occlusion and perspective distortion. On the other hand, to alleviate the interference of unpredictable lighting changes and shadows, we propose a plane-based classification process. Moreover, by introducing a Bayesian hierarchical framework to integrate the 3-D model with the plane-based classification process, we systematically infer the parking status. Last, to overcome the insufficient illumination in the nighttime, we also introduce a preprocessing step to enhance image quality. The experimental results show that the proposed framework can achieve robust detection of vacant parking spaces in both daytime and nighttime.

Index Terms—Bayesian inference, histogram of oriented gradients, image classification, parking space detection.

I. INTRODUCTION

RECENTLY, video surveillance systems have become increasingly important in our daily life. With the noticeable progress of computer vision techniques, many video surveillance systems have been proposed to provide new kinds of intelligent functions, like object detection and tracking. Following the trend, vision-based systems for smart parking lot management have also attracted great attention in recent years. In general, these vision-based parking lot management systems can provide valuable information, like the location of vacant parking spaces, as well as some value-added services, like parking space guidance and vehicle finding. In this paper, we focus on a basic, yet crucial, function of vision-based

Manuscript received August 27, 2012; revised December 9, 2012, February 16, 2013; accepted February 18, 2013, February 20, 2013. Date of publication March 27, 2013; date of current version August 30, 2013. This work was supported in part by the National Science Council of Taiwan, under Grants 100-2218-E-151-007 and 101-2221-E-151-045. This paper was recommended by Associate Editor P. L. Correia.

C.-C. Huang is with the Department of Electrical Engineering, National Kaohsiung University of Applied Sciences, Kaohsiung 807, Taiwan, (e-mail: chingchun.huang5@gmail.com).

Y.-S. Tai and S.-J. Wang are with the Department of Electronics Engineering, Institute of Electronics, National Chiao Tung University, Hsinchu 30010, Taiwan (e-mail: kinn_92@hotmail.com; shengjyh@faculty.nctu.edu.tw).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TCSVT.2013.2254961



Fig. 1. Challenges for robust vacant parking space detection, including perspective distortion, inter-object occlusion, shadow effect, variations of lighting condition, and insufficient illumination at night.

parking lot management systems automatic detection of vacant parking spaces.

In Fig. 1, we show several parking lot images in our dataset. To robustly detect vacant parking spaces, we have to deal with a few challenges, including dramatic lighting variations, shadows cast on the scene, varying perspective distortion in the image, and inter-object occlusion among parked cars and the ground plane. Besides, insufficient illumination during the nighttime is another challenge. To overcome these problems, many novel methods have been proposed in the past. These methods can be roughly categorized into four major types: car-oriented methods, space-oriented methods, hybrid methods, and parking-lot-oriented methods.

Car-oriented methods [1]–[5] target car detection, and they determine the status of parking spaces based on the detection result. Space-oriented methods [6]–[13] model the appearance of the ground plane in advance. If the current appearance of a parking space is dissimilar to that model, they identify the parking space as occupied. Some hybrid methods [14]–[16], on the other hand, combine both space detection and car detection to find vacant parking spaces. These hybrid methods focus on the design of the fusion mechanism to achieve improved performance. Recently, unlike car-oriented or space-oriented methods which focus only on certain aspects of parking lots, parking-lot-oriented methods [17], [18] have been proposed to model the whole parking lot in unity and to integrate the 3-D scene model with the image observation for parking status inference.

For car-oriented methods, Tsai *et al.* [1] propose a global color-based model to efficiently detect vehicle candidates. In

their approach, a Bayesian classifier is trained to verify the detection of vehicles based on corners, edges, and wavelet features. In [2], Meji *et al.* propose a color-based texture segmentation process for vehicle detection based on color and texture features. Masaki [3] keeps tracking and recording the movement of vehicles in order to identify vacant parking spaces. On the other hand, many algorithms [4]–[5] adopt certain consistent texture features, such as histogram of oriented gradients (HOG) [19], to overcome lighting variations and geometric distortion. In general, these methods can achieve robust detection even under dramatic variations of lighting condition. However, for vacant space detection, these car-oriented methods do not take into account the inter-vehicle occlusion problem.

For space-oriented methods, the modeling of parking spaces is the key. Eigen-space representation [6] and many background modeling algorithms [7]–[9] provide pixel-based methods to provide ground models that can adapt to lighting variations. However, these pixel-based space modeling methods are usually sensitive to the shadows cast over the ground. To relieve the shadow effect, some texture-based methods assume that a vacant parking space possesses homogeneous appearance. Hence, they design certain measure of homogeneity to detect vacant parking spaces. For example, Yamada *et al.* [10] design a homogeneity measure by calculating the area of fragmental segments; Lee *et al.* [11] suggest an entropy-based homogeneity metric; and Fabian [12] uses a segment-based homogeneity measure similar to that in [10]. However, due to perspective distortion, a distant parking space may only occupy a small region in the captured image. This usually leads to unstable homogeneity measurement. To overcome the perspective distortion problem, López-Sastre *et al.* [13] suggest a method to rectify the perspective distortion and they use a Gabor filter bank to derive the homogeneity feature for vacant parking space detection. Basically, these space-oriented methods still suffer from the inter-object occlusion problem, which occurs when a parking space is partially or fully occluded by a car at an adjacent parking space.

Some researchers adopt hybrid methods to detect vacant parking spaces. For example, Dan [14] trains a general support vector machine (SVM) classifier to differentiate car regions from space regions by using image features made of the color vectors inside the parking space. However, this method cannot properly handle the inter-occlusion problem. To overcome the occlusion problem, Wu *et al.* [15] propose a method to group three neighboring spaces as a unit and they define the color histogram of the three-space unit as the feature in their SVM classifier. Even though these hybrid methods have considered both car model and space model, the classification performance of their algorithms is still affected by the environmental variations. In general, the lighting changes may dramatically degrade the detection accuracy. On the other hand, in [16], the authors propose an efficient method to combine static and dynamic information for vacant parking space detection. To extract static information, a histogram classification process is used to detect pavement regions while an edge counting process is used to identify vehicle regions. To extract dynamic information, they use blob analysis to track moving vehicles.

TABLE I

COMPARISONS OF VACANT SPACE DETECTION ALGORITHMS FOR FIVE TYPES OF CHALLENGES: PERSPECTIVE DISTORTION (*PD*), INTER-OBJECT OCCLUSION (*IO*), SHADOW EFFECT (*SE*), LIGHTING VARIATIONS (*LV*), AND INSUFFICIENT ILLUMINATION AT NIGHT (*IIN*). MEANING OF SYMBOLS: X: NOT GOOD ENOUGH; Δ : FAIR; O: GOOD

Type	Method	<i>PD</i>	<i>IO</i>	<i>SE</i>	<i>LV</i>	<i>IIN</i>
Car	Tsai [1]	X	X	Δ	Δ	X
Car	Mejía-Iñigo [2]	X	X	Δ	Δ	X
Car	Masaki [3]	X	X	Δ	Δ	X
Car	Schneiderman [4]	Δ	X	O	O	Δ
Car	Felzenszwalb[5]	Δ	X	O	O	Δ
Space	Funck [6]	X	X	X	O	X
Space	Background Modeling [7]–[9]	X	X	X	O	X
Space	Yamada [10]	X	X	Δ	Δ	X
Space	Lee[11]	X	X	Δ	Δ	X
Space	Fabian [12]	X	X	Δ	Δ	X
Space	López-Sastre [13]	O	X	Δ	Δ	X
Hybrid	Dan [14]	Δ	Δ	Δ	X	X
Hybrid	Wu [15]	Δ	O	Δ	Δ	X
Hybrid	Blumer [16]	Δ	Δ	Δ	Δ	X
Parking-lot	Huang [17]	O	O	Δ	Δ	X
Parking-lot	Huang [18]	O	O	O	Δ	X
Parking-lot	Proposed Method	O	O	O	O	Δ

In order to alleviate the inter-occlusion problem, however, their camera usually needs to be placed at a very high altitude.

Rather than focusing on the detection of individual cars or parking spaces, parking-lot-oriented methods model the geometric structure of the whole parking lot in order to properly handle the inter-occlusion situations. In [17]–[18], Huang *et al.* propose a Bayesian hierarchical framework (BHF) to integrate the 3-D scene knowledge and the classification of image pixels into a three-layer hierarchical framework. The structural scene properties of a parking lot, together with the pixel-based car model and parking space model, are well utilized to improve the performance of vacant space detection. Moreover, to conquer the variations of lighting condition and the shadows cast on the scene, Huang *et al.*'s method assumes that the parking lot scene is uniformly lighted by sunlight and they have made a lot of effort to dynamically estimate the lighting condition. However, although their method can produce robust detection results in the daytime, it fails in the nighttime due to the complicated lighting condition at night. Actually, so far as we know, very few systems have ever discussed the vacant parking space detection problem in the nighttime.

For the sake of clarification, we summarize in Table I the comparisons of several algorithms for vacant parking space detection. As indicated in this table, none of these existing methods can handle all the five types of challenges, including perspective distortion, inter-object occlusion, shadow effect, lighting variations, and insufficient illumination at night. In this paper, a new parking-lot-oriented method is presented to deal with all these challenges.

In the proposed method, we further improve the Bayesian hierarchical framework (BHF) in [18] to achieve robust detection of vacant parking spaces in both daytime and nighttime. In our method, we model the whole parking lot as a 3-D

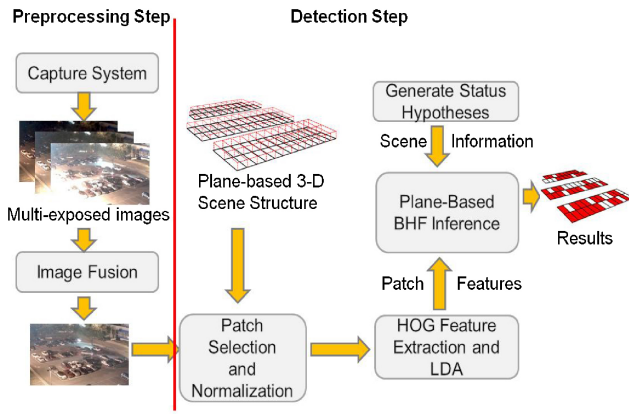


Fig. 2. System flow of the proposed algorithm.

structure consisting of plentiful planar surfaces. A plane-based classification process using robust texture features is proposed to replace the pixel-based classification in [18]. Furthermore, by using a modified BHF framework for inference, we can systematically model the relation between the 3-D planar surfaces and their image appearance. The inter-vehicle occlusion is well modeled in the modified BHF framework and illumination-insensitive object textures are well used for robust parking space detection. Furthermore, by introducing a multi-exposure pre-process to enhance the captured image sequence, we can perform vacant parking space detection days and nights under a unified framework.

The rest of this paper is organized as follows. In Section II, we briefly introduce the overview of the proposed system. In Section III, we illustrate the preprocessing stage for generating high-dynamic-range images in the nighttime. In Section IV, we present the proposed plane-based BHF inference framework for vacant space detection. Experimental results and discussions are presented in Section V. Last, Section VI concludes this paper.

II. OVERVIEW OF THE PROPOSED METHOD

In order to develop a vacant parking space detection system that can work all day, we focus on two major issues. The first issue is about how to obtain well exposed images for inference. In an outdoor scene, the lighting condition may have dramatic changes. Those variations greatly affect the appearance of image features, such as edges or colors, especially in the nighttime. To deal with this issue, we adopt a pre-process to enhance the visibility of image contents. On the other hand, the second issue is about how to improve the performance of vacant parking space detection and how to speed up the system for practical applications. To deal with this issue, a plane-based BHF framework is proposed for vacant parking space detection. By decomposing a parking lot into many 3-D planar surfaces, we can effectively exploit the texture information for vacant parking space detection and well represent the patterns of inter-vehicle occlusion.

In Fig. 2, we show the flowchart of the proposed method, which consists of a preprocessing step and a detection step. In the preprocessing step, we design a multi exposure system to



Fig. 3. (a) Image with a short exposure. (b) Image with a medium exposure. (c) Image with a long exposure. (d) Fusion result of the images in (a), (b), and (c).

capture images with different exposure settings. These images are then fused to obtain images with improved quality. In the detection step, a plane-based BHF inference framework is proposed. First, based on the proposed plane-based 3-D scene model, the normalized patches of interest, corresponding to the projection of 3-D surfaces onto the fused image, are identified. For each normalized patch, the histogram of oriented gradients (HOG) features are extracted and are further compressed via linear discriminant analysis (LDA) [20]. Finally, we use the proposed plane-based BHF framework to integrate 3-D scene information with plane-based classification results for the optimal inference of the status of the parking spaces. In the following sections, we will explain the details of the proposed system in steps.

III. PRE PROCESSING STEP

When capturing images in a dark environment, the color and texture information degrades. The degradation of image features may dramatically deteriorate the performance of vacant parking space detection. Hence, a pre-processing stage is used in our system to enhance the quality of nighttime images.

Up to now, plentiful methods have been proposed to enhance image contrast, like the Retinex-based algorithms in [21], [22], the histogram-equalization-based algorithms in [23]–[26], the Gray level grouping method in [27], [28], the discrete cosine transform DCT-based method in [29], the tone-mapping method in [30], and the Bayesian inference method in [31]. Although those methods can improve image quality impressively, some side-effects, like noise amplification and halo effects, may generate extra image features and harm the following detection process. Different from those approaches which are based on a single image, we enhance nighttime images based on multiple images under different exposure settings. In a dark environment, some image features, like colors or edges, may be missing if the exposure time is too short, as shown in Fig. 3(a). On the contrary, image color or intensity may get saturated if the exposure time is too long,

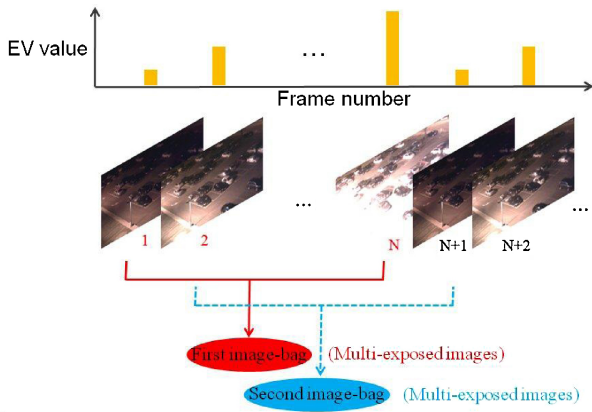


Fig. 4. Illustration of multi-exposure image capturing.

as shown in Fig. 3(c). The choice of exposure time is usually a trade-off. With the use of multiple images under different exposure settings, we are able to extract useful image features in both dark and bright areas. By fusing these images into a single image, we can obtain an image with improved details, as shown in Fig. 3(d).

To get multi exposure images, we use the AXIS M1114 IP camera which can adjust the exposure value (EV) during image capturing. By using the software development kit (SDK) provided by AXIS, we capture images from a short exposure period to a long exposure period in a cyclic manner with the period of N image frames, as illustrated in Fig. 4. In our system, the longest and shortest exposure period is 3 s and 0.33 s, respectively. By using the two-step exposure fusion method proposed in [32] to combine every N images, we get images of improved contrast.

IV. DETECTION STEP

A. Plane-Based Structure and Feature Extraction

In our system, we attempt to find a way to benefit from both car-oriented and space-oriented approaches. For car-oriented methods, they usually check a car area like that in Fig. 5(a); while for space-oriented methods, they check the ground area like that in Fig. 5(b). In our approach, we treat the parking spaces as a set of cuboids, as illustrated in Fig. 5(c). Each cuboid is composed of six patches, as illustrated in Fig. 5(d). Based on the 3-D cuboid model, we represent the structure of parking by a set of 3-D planar surfaces, as shown in Fig. 5(e). By projecting those 3-D surfaces onto the image, we get image patches of parallelogram shape. These patches are to be used for the status inference of parking spaces.

Due to the perspective projection in image formation, image patches may appear to be quite different in shape and size. To overcome perspective distortion, we normalize each image patch into a rectangle, with R_l pixels in length and R_w pixels in width. After that, we extract features from the normalized patches. For feature extraction, we adopt the HOG feature proposed in [19], which is less affected by shadows and the changes of illumination. To extract HOG features, a normalized image patch is regularly segmented into

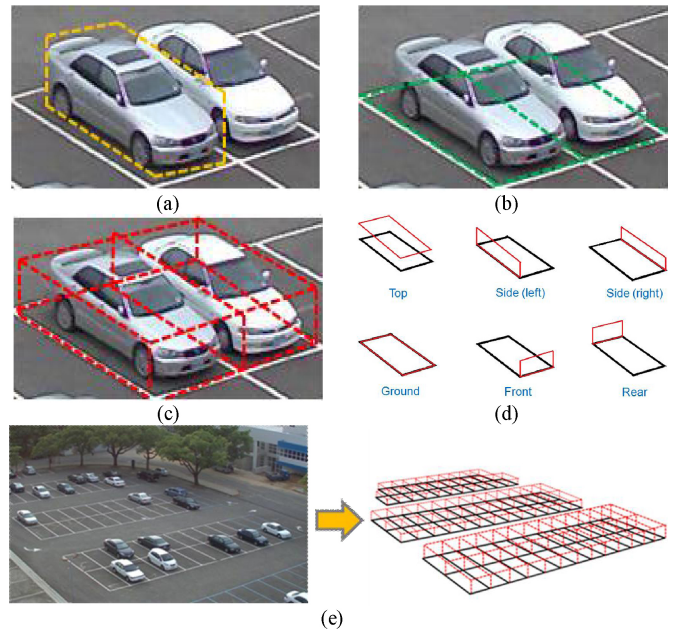


Fig. 5. (a) Image region for car-based inference. (b) Image region for space-based inference. (c) Cuboid modeling of parking spaces. (d) Planar surfaces in the cuboid model. (e) Parking lot model composed of planar surfaces.

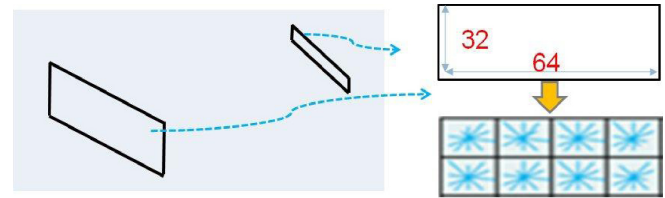


Fig. 6. Patch normalization and HOG feature extraction.

non overlapping cells, with each cell containing $C_l \times C_w$ pixels. In total, there would be $(R_l/C_l) \cdot (R_w/C_w)$ cells in each normalized patch. For each cell, a histogram of oriented gradients, as defined in [19], is built. Each histogram has H_b histogram bins. By combining the histograms of all cells in the normalized patch, we obtain the HOG feature. In our system, the parameters $(R_l \times R_w, C_l \times C_w, H_b)$ are empirically chosen to be $(64 \times 32, 16 \times 16, 10)$. That is, each normalized patch contains eight cells and the dimensions of its HOG feature is 80. In Fig. 6, we illustrate the processes of patch normalization and HOG feature extraction. As will be explained later, these high-dimensional HOG features will be converted into low-dimensional features via LDA so that the following inference process can be implemented in a more efficient way.

B. Patch Classification

In this section, we explain how to perform patch classification in the proposed plane-based BHF framework. As mentioned before, in our plane-based model, each parking space is approximated as a cuboid with six 3-D planar surfaces. We classify these surfaces into four different types: ground surface (G), side surface (S), front (or rear) surface (F), and top surface (T). Via perspective projection, these four types of planar surfaces are projected onto four types of image patches:

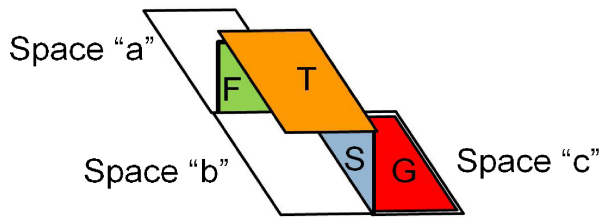


Fig. 7. Illustration for the subclass definitions of image patch.

G-patch, S-patch, F-patch, and T-patch. Due to the inter-object occlusion in the 3-D space, we further classify each type of image patch into a few sub-classes. In Fig. 7, we illustrate how we define the sub-classes for each kind of image surface. Here, without loss of generality, we use a camera configuration with a 45-degree view to explain the proposed patch classification. For the G-patch of the parking space “c” in Fig. 7, its image content is affected not only by the status of the parking space “c” but also by the status of the adjacent parking space “b.” Depending on whether these two parking spaces are occupied or vacant, there are four types of image patterns according to four different parking statuses: 1) “c” is occupied while “b” is vacant; 2) “c” is vacant while “b” is occupied; 3) both “c” and “b” are occupied; and 4) both “c” and “b” are vacant. Similarly, for the S-patch shared by “b” and “c” or the F-patch shared by “a” and “b,” there would be four major kinds of image patterns. For the T-patch of the parking space “b,” on the other hand, we may either classify its image patterns into four sub-classes that relate to the four status combinations of the spaces “b” and “c,” or into eight sub-classes that relate to the eight status combinations of “a,” “b,” and “c.” In our experiments, for the sake of simplification, we choose the four-subclass classification for T-patches

In Table II, based on the illustration in Fig. 7, we further define the indices of the four subclasses for each surface type according to the four status combinations of the present parking space and the most influential adjacent parking space. In this table, “o” means “occupied,” “v” means “vacant,” and “X” means “do not care.” In total, there are 16 kinds of image patches related to the four different types of planar surfaces and the four different combinations of parking statuses for each surface type. In the following paragraphs, we will use the notation $\text{Type}_{\text{Index}}$, where $\text{Type} \in \{T, G, S, F\}$ and $\text{Index} \in \{1, 2, 3, 4\}$, to label these 16 kinds of image patches. The whole set of these 16 patch labels is denoted as $\mathbf{L} \equiv \{T_{-1}, T_{-2}, T_{-3}, T_{-4}, G_{-1}, G_{-2}, G_{-3}, G_{-4}, S_{-1}, S_{-2}, S_{-3}, S_{-4}, F_{-1}, F_{-2}, F_{-3}, F_{-4}\}$.

In Fig. 8, we illustrate the 16 kinds of image patches for a parking space, together with some patch samples. Note that the front surface and the rear surface belong to the same surface type. Similarly, the surfaces on the two sides of the parking space belong to the same surface type. It can be observed in these samples that the image content inside a patch may reveal not only the information of the current parking space but also the information of the adjacent parking space. Moreover, for each surface type, the image contents for different combinations of parking statuses appear to be quite

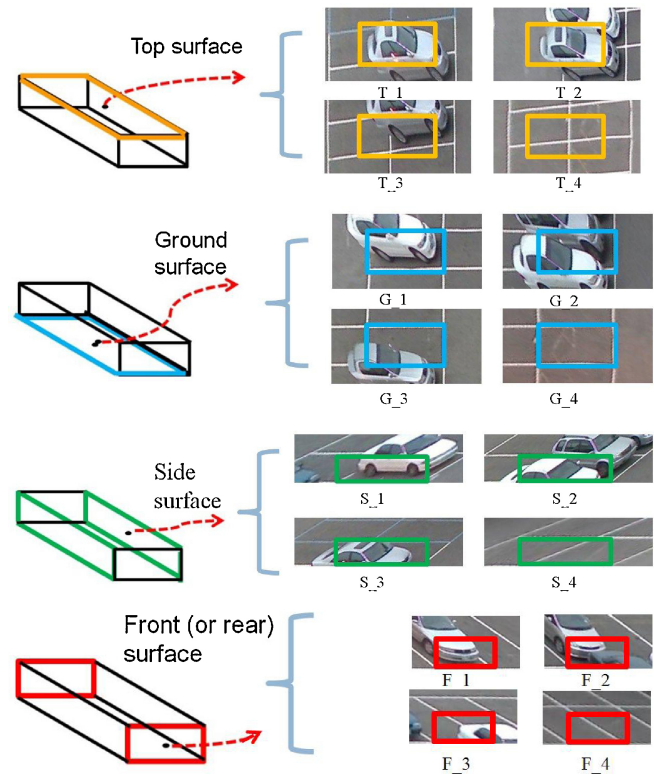


Fig. 8. Sixteen kinds of patch patterns and their classification labels. Each patch pattern is indicated by the rectangular region.

different. Hence, it would be possible for us to classify a given image patch into one of the four subclasses simply based on its image content. The classification result provides evidence to support not only the status inference of the current parking space but also the inference of the adjacent parking space. Even though the classification result at a single image patch may not be always correct, we can combine the classification results of several image patches around a parking space to achieve more robust inference.

Given a parking lot, we first set up an IP camera on the roof of a building near the parking lot. The camera is geometrically calibrated to obtain the 3-D to 2-D projection model and to construct the 3-D plane-based scene model. After that, we capture a few image sequences of the monitored parking lot and extract plentiful image patches for each type of planar surface. For each image patch, we manually collect its patch label and extract its HOG feature from the normalized patch. Based on the labeled patch type and the HOG feature, we learn the conditional probability function $p(\mathbf{o}|l)$, where \mathbf{o} denotes the observed feature of an image patch and $l \in \mathbf{L}$ denotes the label of the image patch.

Since the surface type of an image patch can always be determined based on the 3-D scene model and the 3-D to 2-D transformation, we simply construct the conditional probability model for each of the four surface types. Before classification, we apply the multi-class LDA over the training image patches of each surface type to reduce the high-dimensional HOG features down to a much lower dimension. Taking the learning process of the surface type T as an example, each image

TABLE II
SPECIFICATION OF SURFACE SUB-CLASSES BASED ON FIG. 7. MEANING OF SYMBOLS: V: VACANT; O: OCCUPIED; AND X: DO NOT CARE

Surface Type	T				G				S				F			
Subclass Index	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
Status of Space "a"	X	X	X	X	X	X	X	X	X	X	X	X	O	V	O	V
Status of Space "b"	O	V	O	V	V	O	O	V	V	O	O	V	V	O	O	V
Status of Space "c"	V	O	O	V	O	V	O	V	O	V	O	V	X	X	X	X

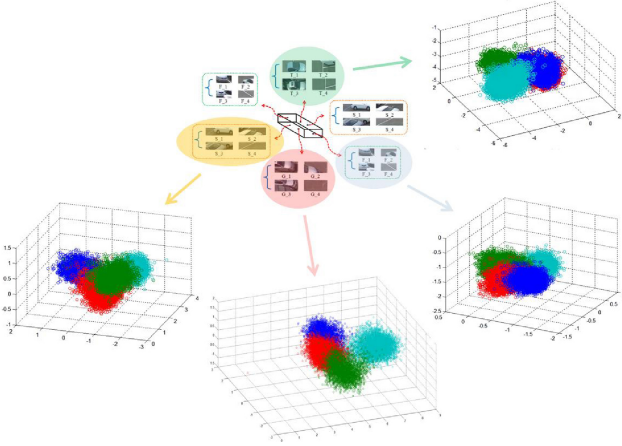


Fig. 9. Feature distributions of 16 subclass models.

patch is normalized and its HOG feature vector is extracted. Each image patch is manually labeled as one of the label set $\{T_{-1}, T_{-2}, T_{-3}, T_{-4}\}$ for a four-class LDA analysis. While adopting an LDA process, we performed eigenvalue decomposition. We found that for each surface type most training samples can be well approximated (up to 99%) by using only the first three eigenvectors that have the largest eigenvalues. Hence, after the LDA process, the dimension of the HOG features is reduced from 80 down to 3 and we obtain four subsets of 3-D features for the four subclass of the surface type T . For each of the four subclasses, say l , we model the conditional probability function of observing the 3-D feature \mathbf{x} given l as a Gaussian function represented by the following equation:

$$N(\mathbf{x}; \boldsymbol{\mu}_l, \boldsymbol{\Sigma}_l) = \frac{1}{(2\pi)^{3/2} |\boldsymbol{\Sigma}_l|^{1/2}} \exp\left\{-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_l)^T \boldsymbol{\Sigma}_l^{-1}(\mathbf{x} - \boldsymbol{\mu}_l)\right\}. \quad (1)$$

In (1), the mean vector $\boldsymbol{\mu}$ and the covariance matrix $\boldsymbol{\Sigma}$ are estimated from the 3-D training features of the corresponding subclass.

Similarly, for each of the surface types G , S , and F , we can obtain the 3-D training features for its four subclasses and construct the corresponding four Gaussian models, as shown in Fig. 9. In total, we obtain 16 low-dimensional training features for the 16 patch labels $\{T_{-1}, T_{-2}, T_{-3}, T_{-4}, G_{-1}, G_{-2}, G_{-3}, G_{-4}, S_{-1}, S_{-2}, S_{-3}, S_{-4}, F_{-1}, F_{-2}, F_{-3}, F_{-4}\}$.

For an image patch \mathbf{o} , if we denote the overall feature extraction process as a mapping from \mathbf{o} to \mathbf{x} , we can model the conditional probability function $p(\mathbf{o}|l)$ in terms of \mathbf{x} as

follows:

$$p(\mathbf{o}|l) \equiv p(\mathbf{x}|l) = \frac{1}{(2\pi)^{3/2} |\boldsymbol{\Sigma}_l|^{1/2}} \exp\left\{-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_l)^T \boldsymbol{\Sigma}_l^{-1}(\mathbf{x} - \boldsymbol{\mu}_l)\right\}. \quad (2)$$

C. Inference of Parking Status

Once we have learned the conditional probability function $p(\mathbf{o}|l)$, we can start to infer the status of parking spaces in new coming images. For example, given an image patch, we first determine its surface type based on the pre-established plane-based 3-D scene model and the pre-calibrated 3-D to 2-D transformation. After that, based on the surface type, the image patch \mathbf{o} , and the measured 3-D feature \mathbf{x} , we obtain the likelihood function $p(\mathbf{o}|l)$. By finding the label that maximizes the likelihood function, we can classify the image patch into one of the four subclasses and then deduce the status of the corresponding parking space. This maximum likelihood approach is pretty simple and efficient. However, the deduced classification label, which depends on a small image patch, provides only local decision of the parking status. The inevitable errors in subclass classification may lead to a wrong decision of the parking status.

As mentioned before, the classification results of relevant patches may provide extra supports for the status inference of parking spaces. However, from time to time, the inferred status from relevant patches may happen to conflict with one another. For example, after patch classification, the top patch of the parking space "c" is labeled as T_{-4} , which implies that the space "c" is vacant. On the other hand, the ground patch of space "c" is labeled as G_{-1} , which implies that "c" is occupied. The parking status inferred from these two labels conflict with each other. Hence, it is apparently not enough to infer the parking status by simply using the local classification results. We need to find a way to effectively fuse the supports from relevant image patches.

To unify the 3-D scene information and patch classification results for vacant parking space detection, we propose a three-layer hierarchical framework, named plane-based BHF, as shown in Fig. 10. Based on this framework, we build the connections between the 3-D scene model and 2-D image patches in a plane-based manner so that we can infer the statuses of parking spaces in a unified way. The proposed hierarchical framework consists of three layers: the scene layer (S_L), the label layer (L_L), and the observation layer (O_L). Each layer contains either nodes or parallelograms. In the scene layer, each node indicates the status of a parking space. For example, the node s_j in S_L denotes the status of the j th parking space, which can be either occupied or vacant. In

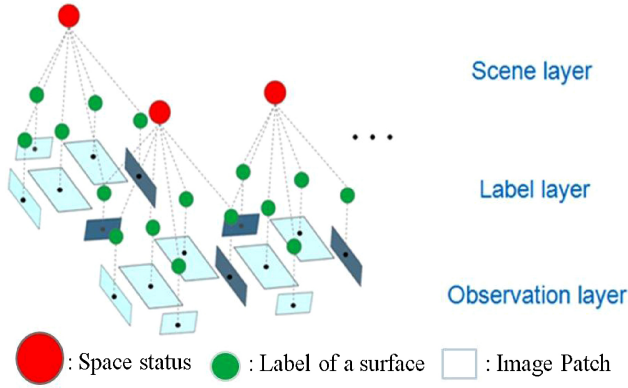


Fig. 10. Proposed plane-based BHF.

the observation layer, each parallelogram represents an image patch obtained by projecting the corresponding 3-D surface onto the image. These image patches in the observation layer provide image features for patch classification. Between the scene layer S_L and observation layer O_L , we design a label layer L_L . Each node in the label layer indicates the inferred label of the corresponding image patch. The label can be one of the 16 labels in the aforementioned label set L . The label layer L_L links to both the scene layer S_L and the observation layer O_L . These links play important roles in connecting the 3-D scene model and the 2-D image observation for the inference of vacant parking spaces. The links between S_L and L_L pass the information from the 3-D scene model to enforce consistent labels in the labeling layer. On the other hand, the links between O_L and L_L convey the likelihood messages from image patches. In the following paragraphs, we will explain how to utilize both image observations and 3-D scene information for parking status inference.

In our framework, we denote a status hypothesis S_t as a combination of the status of the parking spaces. For example, as shown in Fig. 11, if we consider only three parking spaces in the parking lot, the status hypothesis S_t can be one of the eight combinations: $\{(v, v, v), (v, v, o), (v, o, v), (v, o, o), (o, v, v), (o, v, o), (o, o, v), (o, o, o)\}$, where “v” denotes vacant and “o” denotes occupied. On the other hand, we denote our image observation as a set of image patches: $O = \{o_i\}_{i=1 \sim N}$, where i is the patch index and N denotes the total number of image patches in the parking lot. For the three parking spaces in Fig. 11, there are $N=16$ image patches in total. Based on the above definitions, the inference of the optimal parking status is equivalent to the maximization of $p(S_t|O)$. That is

$$S_t^* = \arg \max_{S_t} p(S_t|O). \quad (3)$$

To solve (3), we treat the set of classification labels L as a hidden layer that bridges the gap between S_t and O . Based on a status hypothesis, like the $S_t = \{v, o, o\}$ hypothesis in Fig. 11, we deduce the expected label set $L^{S_t} = \{l_i^{S_t}\}_{i=1 \sim N}$ for all N patches. Here, the superscript in L^{S_t} indicates that this label set is constrained by the status hypothesis S_t . Since there is a one-to-one mapping between S_t and L^{S_t} , we have $p(L|S_t)$

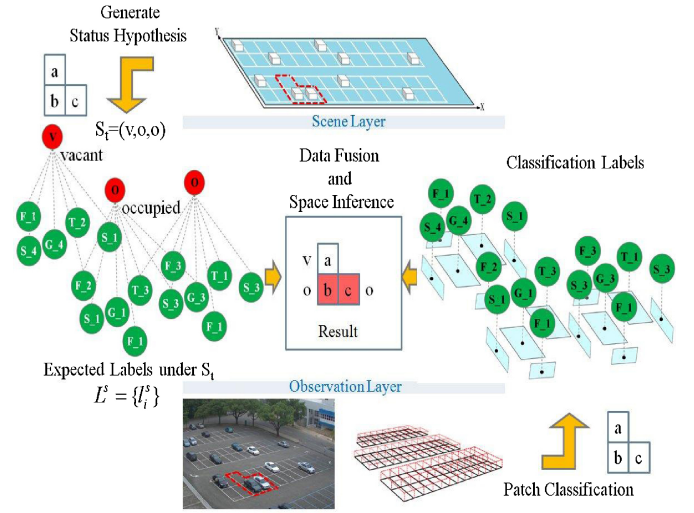


Fig. 11. Example of parking status inference in the proposed framework.

$$= \delta(L - L^{S_t}) \text{ and}$$

$$p(O|S_t) = \int p(O|L)p(L|S_t)dL = \int p(O|L)\delta(L - L^{S_t})dL = p(O|L^{S_t}). \quad (4)$$

Moreover, based on the Bayes' rule, we can rewrite (3) as

$$S_t^* = \arg \max_{S_t} [p(O|S_t)p(S_t)] = \arg \max_{S_t} [p(O|L^{S_t})p(S_t)]. \quad (5)$$

In our system, $p(O|L^{S_t})$ is formulated as

$$p(O|L^{S_t}) = \prod_i^N p(o_i|l_i^{S_t}) \quad (6)$$

where we assume that the observation nodes $\{o_i\}_{i=1 \sim N}$ are conditionally independent once if the labels of the labeling layer are given. In (6), $p(o_i|l_i^{S_t})$ represents how likely the image observation o_i on the i th image patch belongs to the patch label $l_i^{S_t}$. Note that $l_i^{S_t}$ is the expected label of the i th patch under the status hypothesis S_t . Moreover, the probability function $p(S_t)$ in (5) represents the prior knowledge about the parking statuses in the parking lot. In our system, we simply assume $p(S_t)$ to be uniformly distributed. Under this uniform-distribution assumption, the $p(S_t)$ term in (5) can be ignored and we get the final formulation as follows:

$$S_t^* = \arg \max_{S_t} [(\prod_i^N p(o_i|l_i^{S_t})) p(S_t)] = \arg \max_{S_t} [\prod_i^N p(o_i|l_i^{S_t})]. \quad (7)$$

In (7), the calculation of $p(o_i|l_i^{S_t})$ is based on the pre-learned conditional probability model and the expected label $l_i^{S_t}$ for the i th image patch. To find the optimal status hypothesis S_t^* , an exhaustive search over all possible status hypotheses is used. However, for a parking lot containing tens of parking spaces, the number of status hypotheses would be extremely high. Hence, instead of generating the status hypotheses for all parking spaces in the parking lot at one time, we adopt a sequential block-wise approach. In Fig. 12(a), we illustrate how we perform the generation of status hypotheses. In this example, we sequentially infer the parking status from left to

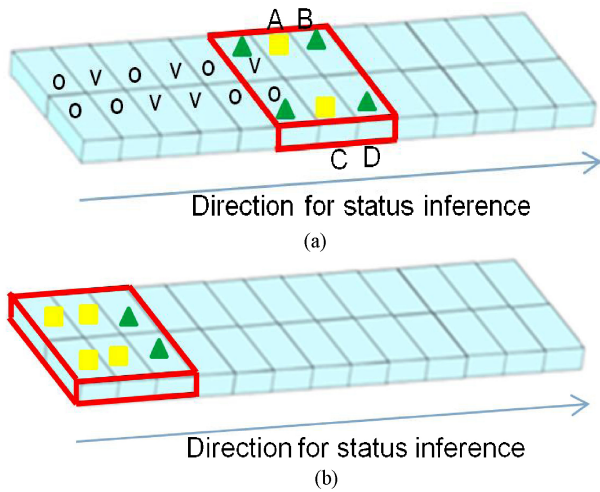


Fig. 12. (a) Illustration of status hypothesis generation and the order of parking status inference. (b) Status hypothesis generation on the border.

right. Note that there is no difference if we infer the status from right to left. Due to inter-object occlusion, we need to take into account the relevant parking spaces as we infer the status of a parking space. In our system, we consider six spaces at one time and determine the parking statuses of the two central parking spaces. In the example in Fig. 12(a), these parking spaces whose statuses have already been inferred are labeled as either vacant (*v*) or occupied (*o*). The region within the red boundary indicates the six parking spaces to be considered at this moment. Within the region, the parking spaces to be inferred are marked by yellow squares while the relevant parking spaces are marked by green triangles. For these six parking spaces, a plane-based BHF framework is built. Since the statuses of the two parking spaces in the left have already been inferred, we only need to generate 16 status hypotheses relating to the 16 status combinations of the parking spaces A, B, C, and D. By calculating the value of $\prod_{i=1}^N p(o_i|l_i^S)$ for each of the 16 hypotheses, the optimal combination can be picked and the status of the two central parking spaces is inferred. Based on this sequential way to generate status hypotheses, the system complexity grows only linearly as the number of parking spaces increases.

To detect the parking status on the border, the inference process is slightly modified. As shown in Fig. 12(b), the red boundary indicates the six parking spaces to be considered and there are four parking spaces marked by the yellow squares. These four parking spaces need to be inferred at the same time. Similar to the aforementioned inference process, a plane-based BHF framework for six parking spaces is built. However, without knowing the two parking spaces on the left, we need to test 64 status combinations of the six parking spaces. By finding the optimal status combination, the status of the four yellow-marked parking spaces is inferred.

In the proposed plane-based BHF framework, the inclusion of the 3-D scene information provides a few benefits. First, the image patches under analysis can be systematically selected via geometric projection and the surface type of each image patch can be easily determined. This allows us to apply LDA

TABLE III

CONDITIONS OF 12 TEST DATABASES. MEANING OF SYMBOLS: SSHOT: SNAPSHOT; ID: IDENTITY; FPM: FRAME PER MINUTE; NOF: NUMBER OF FRAME; AT: ACQUISITION TIME; WC: WEATHER CONDITION; CP: CAMERA POSITION

SShot						
ID	DS 1	DS 2	DS 3	DS 4	DS 5	DS 6
FPM	0.2	0.2	0.2	0.2	0.2	0.2
NoF	160	160	160	145	100	100
AT	5:00~18:00	5:00~18:00	5:00~18:00	6:00~18:00	18:00~3:00	18:00~3:00
Period	Daytime	Daytime	Daytime	Daytime	Nighttime	Nighttime
WC	 Typical	 Sunny	 Cloudy	 Rainy	 Rainless	 Rainless
CP	Camera Position 1	Camera Position 1	Camera Position 1	Camera Position 1	Camera Position 1	Camera Position 1
SShot						
ID	DS 7	DS 8	DS 9	DS 10	DS 11	DS 12
FPM	0.2	0.2	0.2	0.2	0.2	0.2
NoF	100	145	145	145	108	145
AT	18:00~3:00	6:00~18:00	6:00~18:00	6:00~18:00	9:00~17:00	6:00~18:00
Period	Nighttime	Daytime	Daytime	Daytime	Daytime	Daytime
WC	 Rainy	 Cloudy	 Rainy	 Rainy	 Sunny	 Cloudy
CP	Camera Position 1	Camera Position 2	Camera Position 3	Camera Position 4	Camera Position 5	Camera Position 6

to deal with four surface types separately. Second, the status hypotheses in the 3-D space can be systematically converted into expected classification labels that consistently relate relevant image patches. This automatically avoids the inconsistency problem mentioned in the aforementioned maximum-likelihood approach.

V. EXPERIMENTS AND DISCUSSIONS

A. Experiment Environment and Test Data

In our experiments, we evaluate our system at three outdoor parking lots. For each parking lot, we set up an IP camera to monitor the statuses of parking spaces day and night. The camera was geometrically calibrated beforehand and a few image sequences were captured for the learning of $p(\mathbf{o}|l)$ for 16 patch labels. Note that these training sequences were collected differently from our test datasets. For performance evaluation, we have collected 12 datasets in total, including nine daytime datasets under four different weather conditions and three nighttime datasets under rainy or rainless conditions. These 12 datasets were captured at 6 different locations with different capturing conditions, as listed in Table III. In this table, we assign each dataset an identity (ID), together with a snapshot of the dataset. Based on these datasets, we tested our system in daytime and nighttime periods, under different weather conditions, different viewing angles, different surrounding objects, different perspective distortion, and different levels of inter-vehicle occlusion. The 12 datasets, together with the ground truth of parking status and the detection results of our system, are available at our website [33]. Among these 12 datasets, the dataset DS_11 is originally released by Huang *et al.* in

TABLE IV
SPACE DETECTION RESULTS IN THE DAYTIME

Data Set ID	Number of tested spaces			Proposed method		
	vacant	parked	total	FPR	FNR	ACC
Typical (DS_1)	4123	7397	11520	0.280%	1.190%	99.39%
Sunny (DS_2)	8121	3399	11520	0.088%	0.860%	99.37%
Cloudy (DS_3)	5353	6167	11520	0.047%	0.260%	99.85%
Rainy (DS_4)	3544	6896	10440	0.640%	0.370%	99.45%

[18]. Since the parking statuses in these databases changed gradually, we performed vacant parking space detection for every five minutes.

B. System Evaluation

To quantitatively evaluate the performance of our system, we calculate false positive rate (FPR), false negative rate (FNR), and accuracy (ACC). Their definitions are expressed as follows:

$$\text{FPR} = \frac{\text{number of parked spaces being detected as vacant}}{\text{total number of parked spaces}}. \quad (8)$$

$$\text{FNR} = \frac{\text{number of vacant spaces being detected as parked}}{\text{total number of vacant spaces}}. \quad (9)$$

$$\text{ACC} = \frac{\text{number of correct detection in both parked and vacant spaces}}{\text{total number of tested spaces}}. \quad (10)$$

To test our system, we evaluate the performance in the day time under different weather conditions, the performance at nighttime, and the performance with different geometric settings.

1) *Daytime Performance Evaluation*: To evaluate our system under different conditions, we firstly divide a day into the daytime period and the nighttime period. For the daytime period, we tested four video sequences, including a typical day case (DS_1), a sunny day case (DS_2), a cloudy day case (DS_3), and a rainy day case (DS_4). Here, we use the term “typical day” to indicate the weather condition which is partially cloudy and partially sunny. These four test sequences contain different kinds of shadow effects and varying lighting conditions. For the sunny day case, there are plentiful overly exposed regions and strong shadows in the images. The variations of illumination cause apparent drifts in colors and brightness. For the typical day case, the lighting condition switches dramatically between sunny and cloudy. Sometimes, a shadowed region may suddenly disappear when the sunlight is blocked by a passing cloud. For the rainy day case, the raindrops may degrade the visibility of the camera. Moreover, moving pedestrians may affect the detection of vacant parking spaces.

In Fig. 13, we show a few daytime images and their detection results. In Table IV, we list the evaluation results, including the detection results under various kinds of weather conditions. Even though the changes of weather condition may cause slight performance degradation in our system,

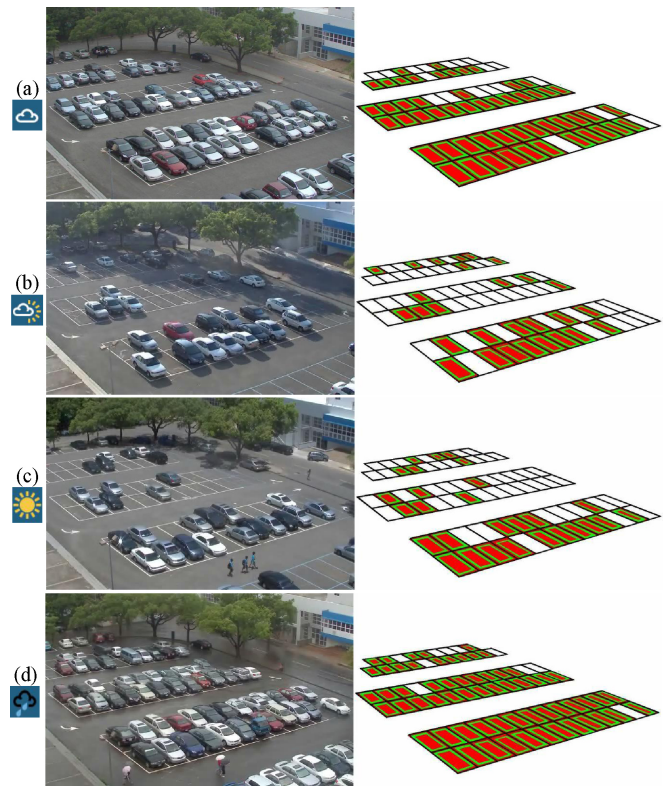


Fig. 13. Results of vacant space detection in the daytime. (a) Cloudy day. (b) Typical day. (c) Sunny day. (d) Rainy day. Green rectangles indicate the ground truth and red rectangles indicate the detection results.

these experimental results demonstrate that our system can effectively deal with the occlusion problem, shadow effect, and the changes of lighting condition in the daytime period.

2) *Nighttime Performance Evaluation*: For the nighttime period, we tested another three datasets DS_5, DS_6, and DS_7. To detect vacant parking spaces at night, we need to deal with different kinds of shadow effects caused by multiple lighting sources. In addition, the non-uniform illumination makes some parts of the image overly exposed while some other parts poorly exposed. Moreover, vehicles tend to dynamically move in the scene and the unpredictable lighting changes produced by car headlights causes another big challenge. In Fig. 14, we show some night images and their detection results. Here, we also show the nighttime images captured under different exposure settings. These images are further fused to generate a well-exposed image for vacant parking space detection.

In Table V, we list the evaluation results of these three nighttime datasets. Comparing with the detection in the day time, the performance is slightly degraded. We can further divide the nighttime datasets into static cases and dynamic cases. In a static case, as shown in Fig. 14(a), there is no moving car in the scene. The challenges come from insufficient illumination, heavy shadows caused by streetlamps, and inter-vehicle occlusion. These variations can be well learned beforehand and can be well handled by the proposed framework. Hence, the system performance is quite stable in static cases. On the other hand, in the dynamic case, moving cars with unpredictable lighting changes make our system less stable.

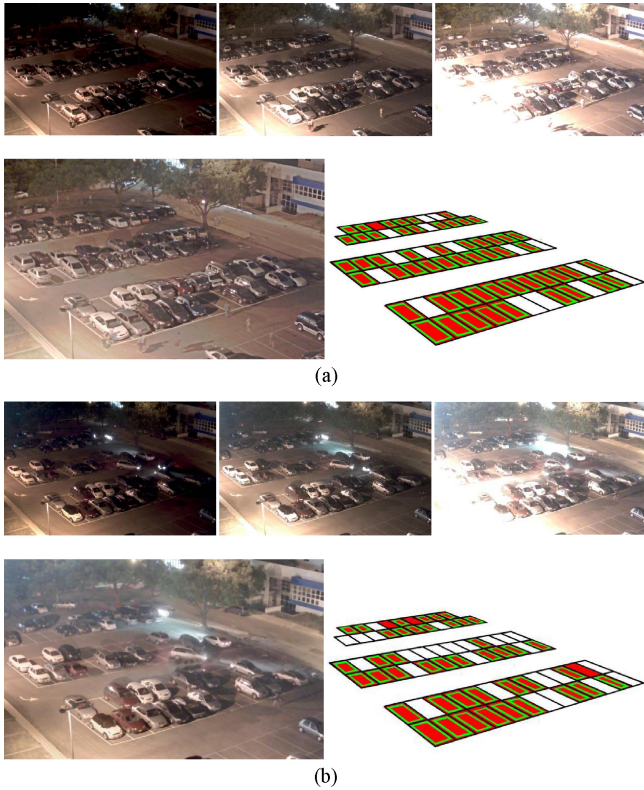


Fig. 14. Results of vacant space detection in the nighttime (a) without moving cars in the scene and (b) with moving cars in the scene. Row 1: image samples with EV=10, EV=50, and EV=90. Row 2: fused images and the corresponding detection results. Green rectangles indicate the ground truth and red rectangles indicate the detection results.

TABLE V
SPACE DETECTION RESULTS IN THE NIGHTTIME

	#of tested spaces			Proposed method		
	vacant	parked	total	FPR	FNR	ACC
Seq_1 (DS_5)	4165	3035	7200	1.610%	2.420%	97.92%
Seq_2 (DS_6)	4569	2631	7200	0.800%	1.730%	98.61%
Seq_3 (DS_7)	1428	2172	3600	0.46%	3.29%	98.42%

As shown in Fig. 14(b), there are three cars moving at the same time. Since the car is moving during the acquisition period, the car headlights introduce unexpected textures in the fused image. These unexpected textures may sometimes cause incorrect inference. Moreover, based on the evaluation results of the three nighttime datasets, we could find that the proposed system is able to achieve stable performance under both rainy and rainless conditions. Note that DS_5 and DS_6 were collected on days without rain while DS_7 was collected in a rainy day.

3) *Performance Under Different Geometric Settings:* To evaluate the influence of perspective distortion, we divide a parking lot into three blocks as shown in Fig. 15. The performance for each block is evaluated. The detection results are listed in Table VI. In general, the performance is getting better as the parking block is closer to the camera. This is because we have less perspective distortion and larger patches for vacant space detection. In this experiment, the

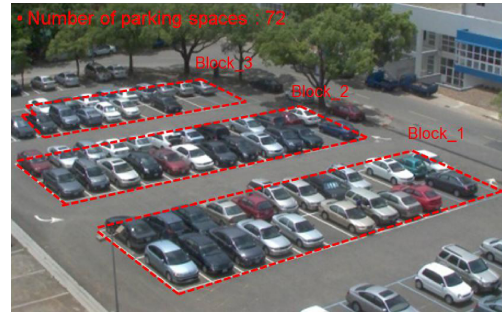


Fig. 15. Experiment environment for evaluating the influence of perspective distortion.

TABLE VI
SPACE DETECTION RESULTS IN DIFFERENT BLOCKS

Dataset Snapshot	Block number	#of tested spaces			Proposed method		
		vacant	parked	total	FPR	FNR	ACC
Datasets: DS_1~DS_4							
	Block_1	6309	9941	16250	0.260%	0.03%	99.83%
	Block_2	8168	8082	16250	0.540%	0.69%	99.38%
	Block_3	6664	5836	12500	0.240%	1.13%	99.29%
Datasets: DS_5~DS_6							
	Block_1	4333	3467	7800	1.640%	1.15%	98.63%
	Block_2	4930	2870	7800	0.870%	0.73%	99.22%
	Block_3	4077	1923	6000	3.380%	3.56%	96.50%

datasets DS_1~DS_4 are integrated together to evaluate the performance in the daytime. On the other hand, the datasets DS_5~DS_6 are integrated to evaluate the performance in the nighttime. These results show that the proposed system can handle the perspective distortion well.

On the other hand, to evaluate the robustness against different camera settings, we also tested our system by using datasets collected with different camera heights and different viewing angles. In this experiments, six datasets (DS_1 and DS_8~DS_12), which were captured at six different camera positions, were used. To illustrate the field of view of each dataset, six snapshots corresponding to the six datasets are shown in Fig. 16. These six datasets also include the variations of the surrounding objects, like trees and buildings. The detection results are shown in Table VII. In the results, we may find that the height and the viewing angle of a camera are critical for vacant parking space detection. For a camera setting at a lower altitude, such as the case in Camera Position 6 (DS_12), the inter-vehicle occlusion is very severe and the system performance degrades. On the other hand, If we compare DS_8 with DS_11, we can find that a wider coverage of the parking lot can monitor more parking spaces at one time but with the degradation in accuracy.

4) *Evaluation of Patch Classification:* In the proposed method, the patch classification is quite sensitive to the setting of camera height and viewing angle. In our experiment, for each camera setting, we collected the training sequence in one day while collected the test sequences in other days. We also manually labeled the ground truth of the parking status of each parking space and determined the relationship between patch labels and parking status. Based on the 3-D cuboid model and the status of each parking space, we can automatically crop

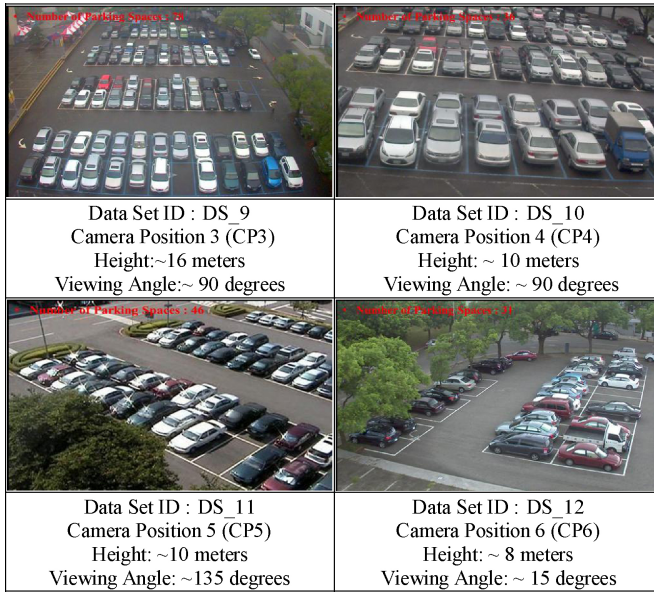


Fig. 16. Datasets captured at six different positions.

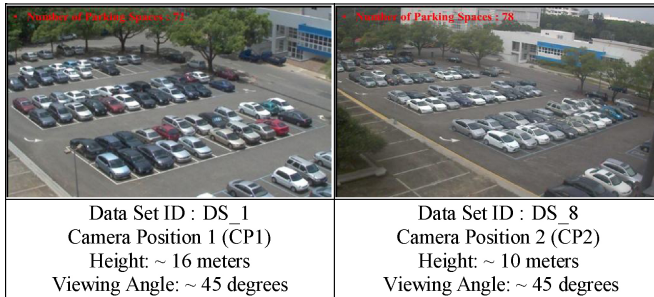


Fig. 17. Detection Results at Different Locations with Different Viewing Angles

the image patches and generate the classification label for each image patch. Those cropped patches are used for training and testing. Note that since the perspective projection process is highly relative to the camera setting, the patch classification models need to be re-trained for different camera settings. In Table VII, we summarize the performance results of the patch classification process under four different camera settings by using the datasets DS_1, DS_5, DS_8, and DS_9. For each dataset, the number of training patches (Train Num.), the number of test patches (Test Num.), and the accuracy (ACC) of patch classification for the S-patch, T-patch, F-patch, and G-patch are reported. In these results, it is noticeable that the performance of patch classification is highly relative to the viewing angle and the height of the camera position. Moreover, even though the classification performance of individual image patch may not be satisfactory, by using the proposed plane-based BHF to provide a unified framework for information fusion, we still obtain a reliable strong classifier to achieve robust vacant parking space detection

C. Comparison of System Performance

For comparison, the receiver operating characteristic (ROC) curves are plotted. To plot an ROC curve, we define the prior

TABLE VII
PERFORMANCE OF PATCH CLASSIFICATION

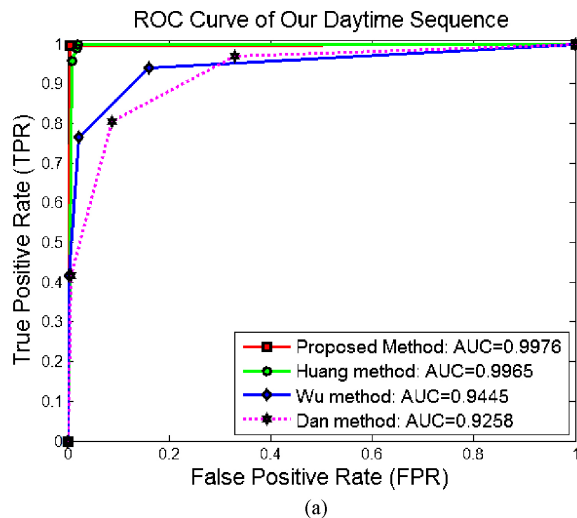
Patch	Dataset DS_1 Camera Position 1 (CP1) (Daytime, 45 Degrees, High)			Dataset DS_5 Camera Position 1 (CP1) Nighttime, 45 Degrees, High)		
	Train Num.	Test Num.	ACC	Train Num.	Test Num.	ACC
S-patch	12110	25408	86.86%	7490	20020	86.01%
T-patch	12453	26131	96.82%	7701	20589	94.60%
F-patch	18511	38840	86.43%	11449	30602	81.70%
G-patch	12110	25408	90.49%	7490	20020	89.70%
Patch	Dataset DS_8 Camera Position 2 (CP2) (Daytime, 45 Degrees, Low)			Dataset DS_9 Camera Position 3 (CP3) (Daytime, 90 Degrees, High)		
	Train Num.	Test Num.	ACC	Train Num.	Test Num.	ACC
S-patch	10920	12324	73.81%	11154	13386	52.82%
T-patch	10920	12324	93.38%	11136	13416	97.32%
F-patch	16380	18486	72.23%	16731	20124	86.64%
G-patch	10920	12324	70.41%	11154	13416	92.45%

probability $p(o_s)$ as a tunable parameter. Here, $p(o_s)$ indicates how likely a parking space is being occupied. In our experiment, we assume the prior probabilities of different parking spaces are independent. If S_i indicates the status of N parking spaces, with K of the N spaces being occupied, then the probability $p(S_i)$ is calculated by $p(S_i) = p(o_s)^k(1 - p(o_s))^{N-k}$. By changing the value of $p(o_s)$, we adjust the prior belief of the status of N parking spaces, $p(S_i)$, and an ROC curve can be plotted.

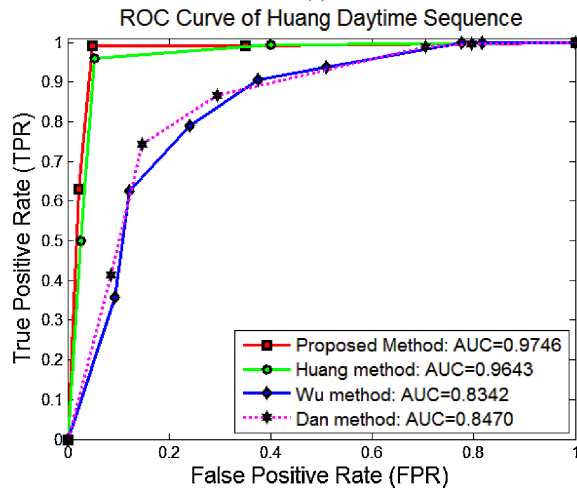
Moreover, we tested two daytime datasets (DS_1 and DS_11) and one nighttime dataset (DS_6) for performance comparison with some other systems. In this simulation, the methods proposed by Dan [14], Wu *et al.* [15], and Huang *et al.* [18] were tested for comparison. The ROC curves are plotted in Figs. 17 and 18 for comparison. The area under the ROC curve (AUC) is also computed for reference. Compared to with Huang *et al.*'s work in [18], the proposed method achieves comparable performance in the daytime period. However, as mentioned before, The method in [18] is much more complicated and needs to dynamically model the lighting condition for pixel-based classification. In addition, the method is mainly based on color intensity to classify the parking status. However, when the images are captured in the nighttime period, the color information degrades. This caused serious performance degradation of Huang *et al.*'s method. In comparison, the proposed method can well handle the change of lighting condition by using the texture information in image patches. In Fig. 18, we plot the ROC curves of the methods in [14], [15], and [18], and the proposed method for comparison. This figure shows that the proposed method can achieve better accuracy during the nighttime period.

D. System Complexity

The whole system has been implemented in the Visual C++ environment on a PC with a 2.4GHz Pentium-4 CPU. In the daytime, it takes about 2 s to infer the statuses of 72 parking spaces in a 320×240 color image. In the nighttime, due to the use of multi-exposure images, it takes about 5 s to perform the detection. In comparison, Huang *et al.*'s work in [18] takes about 30 s to perform vacant parking space detection.



(a)



(b)

Fig. 18. ROC curve comparison in daytime. (a) Performance comparison by using the dataset DS_1. (b) Performance comparison by using the dataset DS_11.

So far, we have tested our system at six different locations. In average, it takes two days to install each system. The first day is used for hardware setup, camera calibration, and training data collection. The second day is used to label the training data and to learn models for patch classification. Within the whole process, the most time-consuming steps are training data collection and training data labeling. For the data labeling step, we have designed a user-friendly interface to help the users label the status of each parking space. Based on the 3-D cuboid model and the status of each parking space, our system automatically crops the image patches and generates the classification label for each image patch. The whole system has been designed to make the system setup process simple and practical. For a new comer, he/she only needs to learn the camera calibration process.

E. Discussion and Future Work

Even though our system works pretty well in an outdoor parking area, it is still a challenge to detect vacant parking spaces in an indoor parking lot. For an indoor parking lot, the proposed method might not be suitable. The bottleneck is not the technical part but the efficiency and cost. Owing to

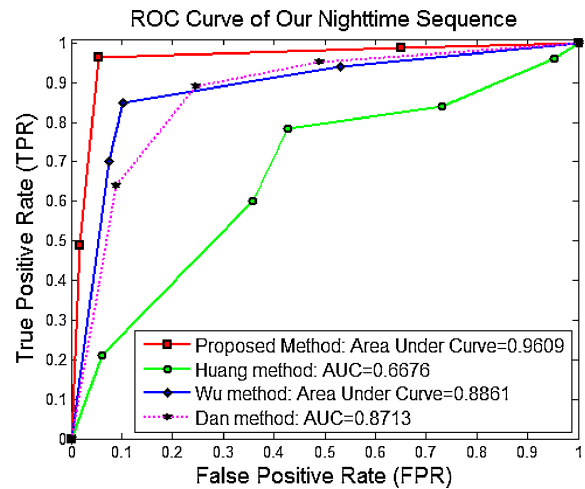


Fig. 19. ROC curve comparison by using the nighttime dataset DS_6.

the severe occlusion in an indoor environment and the limited field of view, we may need dozens of cameras to monitor the whole indoor parking lot.

On the other hand, the system performance in the nighttime can be further improved. Most of the failure cases are caused by the headlight of moving cars. In Fig. 14, we have shown a challenging case in which cars are moving in the parking lot. We will require a new mechanism to handle these unpredictable lighting changes. A possible solution is to include temporal information or to adopt moving vehicle tracking technique. These could be the future works for our vacant parking space detection system.

VI. CONCLUSION

In this paper, our goal was to find suitable modeling so that the performance of texture-based detection can be leveraged for vacant parking space detection. In the survey of previous works, we found that decomposing a parking lot into a composition of individual cars or parking spaces made it difficult to model the inter-object occlusion. On the other hand, we also found that pixel-based detection was quite sensitive to environmental variations if compared with texture-based detection. Hence, to improve the system performance of vacant parking space detection, we proposed a plane-based modeling that regarded the whole parking lot as a structure containing plentiful planar surfaces. With the proposed structure, we systematically combined the texture information in image patches with the 3-D scene information to obtain robust detection results. We also introduced an image processing flow to improve the visual quality of nighttime images. Experiments demonstrated that our vacant parking space detection system performed well under various kinds of weather conditions in both daytime and nighttime.

ACKNOWLEDGMENT

The authors would like to thank all the editors and the anonymous reviewers for their comments. We would also like to thank B.-C. Tsai, Z.-Y. Hsu, Y.-C. Chen, T.-M. Su, and J.-H. Hsieh for collecting the datasets used in the experiments.

REFERENCES

- [1] L.-W. Tsai, J.-W. Hsieh, and K.-C. Fan, "Vehicle detection using normalized color and edge map," *IEEE Trans. Image Process.*, vol. 16, no. 3, pp. 850–864, Mar. 2007.
- [2] R. Mejía-Iñigo, M. E. Barilla-Pérez, and H. A. Montes-Venegas, "Color-based texture image segmentation for vehicle detection," *Image Segmentation*, Rijeka, Croatia: InTech, pp. 273–290, 2011.
- [3] I. Masaki, "Machine-vision systems for intelligent transportation systems," in *Proc. IEEE Int. Conf. Intell. Transp. Syst.*, vol. 13, no. 3, Dec. 1998, pp. 24–31.
- [4] H. Schneiderman and T. Kanade, "Object detection using the statistics of parts," *Int. J. Comput. Vision*, vol. 56, no. 3, pp. 151–177, Feb. 2004.
- [5] P. Felzenszwalb, R. Girshick, D. McAllester, and D. Ramanan, "Object detection with discriminatively trained part based models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 9, pp. 1627–1645, Sep. 2010.
- [6] S. Funck, N. Mohler, and W. Oertel, "Determining car-park occupancy from single images," in *Proc. IEEE Intell. Veh. Symp.*, Jun. 2004, pp. 325–328.
- [7] T. Horprasert, D. Harwood, and L. A. Davis, "A statistical approach for real-time robust background subtraction and shadow detection," in *Proc. IEEE Int. Conf. Comput. Vision*, Sep. 1999, pp. 1–19.
- [8] C. Stauffer and W. E. L. Grimson, "Adaptive background mixture models for real-time tracking," in *Proc. IEEE Int. Conf. Comput. Vision Pattern Recognit.*, vol. 2, Jun. 1999, pp. 246–252.
- [9] P. Power and J. Schoonees, "Understanding background mixture model for foreground segmentation," in *Proc. Image Vision Comput.*, New Zealand, Nov. 2002, pp. 267–271.
- [10] K. Yamada and M. Mizuno, "A vehicle parking detection method using image segmentation," *Electron. Commun.*, vol. 84, no. 10, pp. 25–34, Oct. 2001.
- [11] C. H. Lee, M. G. Wen, C. C. Han, and D. C. Kuo, "An automatic monitoring approach for unsupervised parking lots in outdoor," in *Proc. IEEE Int. Conf. Security Technol.*, Spain, Oct. 2005, pp. 271–274.
- [12] T. Fabian, "An algorithm for parking lot occupation detection," *IEEE Comput. Inf. Syst. Ind. Manage. Appl.*, pp. 165–170, Jun. 2008.
- [13] R. J. L. Sastre, P. Gil Jimenez, F. J. Acevedo, and S. Maldonado Bascon, "Computer algebra algorithms applied to computer vision in a parking management system," in *Proc. IEEE Int. Symp. Ind. Electron.*, Jun. 2007, pp. 1675–1680.
- [14] N. Dan, "Parking management system and method," *U.S. patent*, Pub. No.: 20030164890A1, Jul. 2003.
- [15] Q. Wu, C. C. Huang, S. Y. Wang, W. C. Chiu, and T. H. Chen, "Robust parking space detection considering inter-space correlation," in *Proc. IEEE Int. Conf. Multimedia Expo*, Jul. 2007, pp. 659–662.
- [16] K. Blumer, H. Halaseh, M. Ahsan, H. Dong, and N. Mavridis, "Cost-effective single-camera multi-car parking monitoring and vacancy detection toward real-world parking statistics and real-time reporting," in *Proc. Int. Conf. Neural Inf. Process.*, Nov. 2012, pp. 506–515.
- [17] C. C. Huang, S. J. Wang, Y. J. Chang, and T. Chen, "A Bayesian hierarchical detection framework for parking space detection," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, Apr. 2008, pp. 2097–2100.
- [18] C. C. Huang and S. J. Wang, "A hierarchical bayesian generation framework for vacant parking space detection," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 20, no. 12, pp. 1770–1785, Dec. 2010.
- [19] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proc. IEEE Conf. Comput. Vision Pattern Recognit.*, vol. 1, Jun. 2005, pp. 886–893.
- [20] A. M. Martínez and A. C. Kak, "PCA versus LDA," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 2, pp. 228–233, Feb. 2001.
- [21] J. Wu, Z. Wang, and Z. Fang, "Application of retinex in color restoration of image enhancement to night image," in *Proc. Int. Cong. Image Signal Process.*, Oct. 2009, pp. 1–4.
- [22] A. Yamasaki, H. Takauji, S. Kaneko, T. Kanade, and H. Ohki, "Denighting: Enhancement of nighttime images for a surveillance camera," in *Proc. Int. Conf. Pattern Recognit.*, Dec. 2008, pp. 1–4.
- [23] H. Ibrahim and Nicholas Sia Pik Kong, "Brightness preserving dynamic histogram equalization for image contrast enhancement," *IEEE Trans. Consum. Electron.*, vol. 53, no. 4, pp. 1752–1758, Nov. 2007.
- [24] Q. Wang and R. K. Ward, "Fast image/video contrast enhancement based on weighed threshold histogram equalization," *IEEE Trans. Consum. Electron.*, vol. 53, no. 2, pp. 757–764, May 2007.
- [25] M. Kim and M. G. Chung, "Recursively separated and weighed histogram equalization for brightness preservation and contrast enhancement," *IEEE Trans. Consum. Electron.*, vol. 54, no. 3, pp. 1389–1397, Aug. 2008.
- [26] T. Arici, S. Dikbas, and Y. Altunbasak, "A histogram modification framework and its application for image contrast enhancement," *IEEE Trans. Image Process.*, vol. 18, no. 9, pp. 1921–1935, Sep. 2009.
- [27] Z. Y. Chen, B. R. Abidi, D. L. Page, and M. A. Abidi, "Gray level grouping (GLG): An automatic method for optimized image contrast enhancement—Part 1: The basic method," *IEEE Trans. Image Process.*, vol. 15, no. 8, pp. 2290–2302, Aug. 2006.
- [28] Z. Y. Chen, B. R. Abidi, D. L. Page, and M. A. Abidi, "Gray level grouping (GLG): An automatic method for optimized image contrast enhancement—Part 2: The variations," *IEEE Trans. Image Process.*, vol. 15, no. 8, pp. 2303–2314, Aug. 2006.
- [29] J. Mukherjee and S. K. Mitra, "Enhancement of color images by scaling the DCT coefficients," *IEEE Trans. Image Process.*, vol. 17, no. 10, pp. 1783–1794, Oct. 2008.
- [30] Q. Shan, J. Jia, and M. S. Brown, "Globally optimized linear windowed tone-mapping," *IEEE Trans. Vis. Comput. Graphics*, vol. 16, no. 4, pp. 663–675, July–Aug. 2010.
- [31] T.-C. Jen and S.-J. Wang, "An efficient Bayesian framework for image enhancement with spatial consideration," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2010, pp. 3285–3288.
- [32] I. Mertens, J. Kautz, and F. V. Reeth, "Exposure fusion," in *Proc. 15th Pacific Conf. Comput. Graphics Appl.*, Mar. 2007, pp. 382–390.
- [33] C.-C. Huang (2012), *Huang's Projects* [Online]. Available: <http://140.113.238.220/~chingchun/Lotprojects.html>



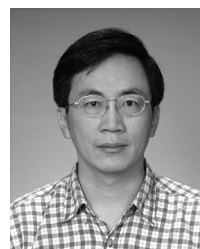
Ching-Chun Huang (M'09) received the B.S., M.S., and Ph.D. degrees in electrical engineering from National Chiao Tung University, Hsinchu, Taiwan, in 2000, 2002, and 2010, respectively.

He is currently an Assistant Professor with the Department of Electrical Engineering, National Kaohsiung University of Applied Sciences, Taiwan. His current research interests include image/video processing, computer vision, and computational photography.



Yu-Shu Tai received the B.S. degree in engineering science from National Cheng Kung University, Tainan, Taiwan, in 2009 and the M.S. degree in electronics engineering from National Chiao-Tung University, Hsinchu, Taiwan, in 2011.

His expertise area is in image processing and computer vision.



Sheng-Jyh Wang (M'95) received the B.S. degree in electronics engineering from National Chiao-Tung University (NCTU), Hsinchu, Taiwan, in 1984, and the M.S. and Ph.D. degrees in electrical engineering from Stanford University, Stanford, CA, USA, in 1990 and 1995, respectively.

He is currently a Professor with the Department of Electronics Engineering, NCTU. His current research interests include image processing, video processing, and image analysis.