

# A BAYESIAN HIERARCHICAL DETECTION FRAMEWORK FOR PARKING SPACE DETECTION

Ching-Chun Huang<sup>1</sup>, Sheng-Jyh Wang<sup>1</sup>, Yao-Jen Chang<sup>2</sup>, and Tsuhan Chen<sup>2</sup>

<sup>1</sup> Department of Electronics Engineering, National Chiao Tung University, Hsinchu, Taiwan.

<sup>2</sup> Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, USA

## ABSTRACT

In this paper, a 3-layer Bayesian hierarchical detection framework (BHDF) is proposed for robust parking space detection. In practice, the challenges of the parking space detection problem come from luminance variations, inter-occlusions among cars, and occlusions caused by environmental obstacles. Instead of determining the status of parking spaces one by one, the proposed BHDF framework models the inter-occluded patterns as semantic knowledge and couple local classifiers with adjacency constraints to determine the status of parking spaces in a row-by-row manner. By applying the BHDF to the parking space detection problem, the available parking spaces and the labeling of parked cars can be achieved in a robust and efficient manner. Furthermore, this BHDF framework is generic enough to be used for various kinds of detection and segmentation applications.

**Index Terms**— Graphical models, Segmentation, Semantic Detection, Optimization, Bayesian framework.

## 1. INTRODUCTION

Using intelligent surveillance systems to monitor parking lots is becoming a practical application. To assist users in efficiently finding empty parking spaces, a parking lot management system may perform parking space detection to identify available parking spaces over time. Up to now, many parking space detection algorithms have already been proposed. For example, Wang and Hanson [1] extracted and analyzed the structured geometric information of parking lots from aerial images. Lee et al.[2] and Masaki [3] suggested keeping tracking and recording the movement of vehicles to locate empty parking spaces. Dan [4] trains a general SVM classifier to detect the status of a single parking space. Basically, these methods analyze the status of each parking space individually, without seriously considering luminance variations and occlusion problems.

In practice, for parking lot images, the condition of luminance changes over time, as shown in Fig. 5(a) and (c). The change of luminance may produce inconsistent image

features and seriously affect detection results. On the other hand, occlusions are very common in parking lot images. For example, due to the spatial arrangement of parking spaces, we expect each parked car may occlude some portions of its neighboring cars or parking spaces, as shown in Fig. 5(a) and (c). Hence, the status of a parking space may actually affect the inference of its neighboring parking spaces. On the other hand, there could be some other occlusions that are caused by obstacles in the 3-D scene, like the occlusion caused by the leaves in Fig. 5. These occlusions partially occlude the parking spaces and may cause incorrect inference. In this paper, we name these two types of occlusions as car occlusion and environmental occlusion. Here, we propose a new 3-layer Bayesian hierarchical detection framework, called BHDF, for parking space detection. By introducing a hidden labeling layer between an observation layer and a semantic layer, the proposed framework provides a systematic structure to deal with luminance variation, car occlusion, and environmental occlusion.

## 2. SYSTEM OVERVIEW

As mentioned above, since the status of a parking space may actually affect the inference of neighboring parking spaces, it is unsuitable to decide the status of each parking space individually. Hence, in our system, we analyze the status of parking spaces in a row by row manner. Here, we regard each parking space as an equal-sized cube on the ground of the 3D scene. As illustrated in Fig. 1(a), based on camera parameters, a row of parking spaces can be formed by projecting the 3D cubes to the image.



Fig. 1(a) An image partition for a row of parking spaces in the image. (b) The projection of 3-D cubes onto the 2-D image forms the top (green) and bottom (red) lines of the parking spaces.

This work was supported by the National Science Council, Taiwan, under the Grants No. NSC95-2221-E-009-106-MY2.

To analyze the statuses of each row, we regard detections as an optimal inference problem which is subject to three aspects. First, local regions should be correctly classified. Second, adjacent regions should follow neighboring constraints. Third, the parked cars in an image partition would match the inter-occluded patterns. In our system, an image partition and three trainable models are inputted to the BHDF framework to model this optimal problem. Moreover, belief propagation techniques [5] for speeding up the optimal inference on the BHDF framework are used to determine the output. In Fig. 2, we show the detail system flow of the proposed approach.

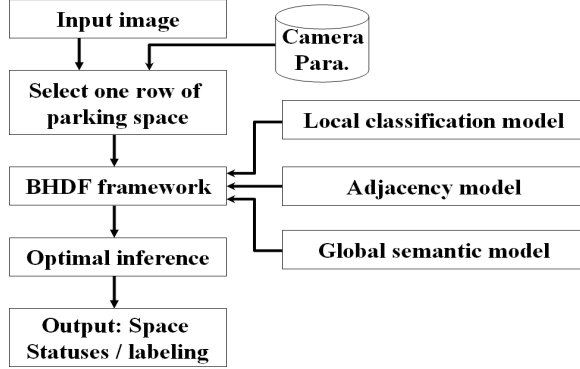


Fig. 2. System flow of the proposed algorithm.

### 3. BAYESIAN HIERARCHICAL DETECTION FRAMEWORK (BHDF)

The proposed Bayesian hierarchical detection framework for parking space detection can be represented by a 3-layer graphical model, as shown in Fig. 3. Without loss of generality, we show the case of four parking spaces. The top layer is an observation layer ( $D_L$ ), where each node indicates a local feature. The local feature can be either texture-based or pixel-based. The middle layer is a hidden labeling layer ( $H_L$ ), where each node represents the categorization of a local region or an image pixel. The bottom layer is a semantic layer ( $S_L$ ), which includes the semantic hypotheses of the parking spaces. Moreover, in our BHDF framework, the topology of the inter-layer connections is flexible and application-oriented.

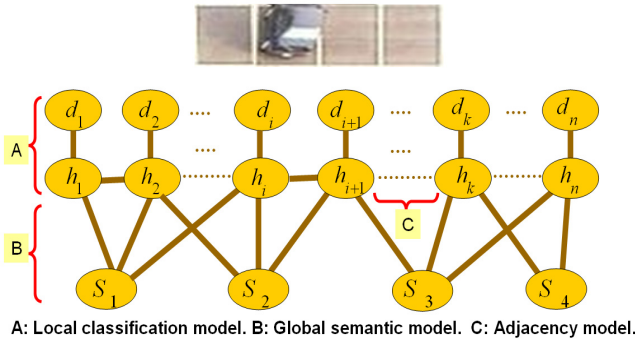


Fig. 3. Detection of four parking spaces based on BHDF.

In our approach, we model the node  $d_i$  in  $D_L = \{d_1, \dots, d_M\}$  as the RGB color value at the  $i$ th pixel of the image partition. We define each labeling node  $h_i$  in  $H_L = \{h_1, \dots, h_M\}$  to be one of the labels in the set  $\{C, G, O\}$ , which  $C$  indicates a “car” pixel,  $G$  indicates a “ground” pixel, and  $O$  indicates an “otherwise” pixel. Here,  $M$  is the total number of pixels in the image partition. In the “four parking spaces” case, the semantic layer  $S_L = \{s_1, s_2, s_3, s_4\}$  corresponds to the status of four parking spaces, with each semantic node being either occupied (1) or available (0).

In BHDF, the inter-layer and intra-layer message propagations are defined by using three trainable models. As illustrated in Fig. 3, we assume observation nodes are conditionally independent when the status of the labeling layer is given. This implies no connection among observation nodes. Here, we assume there are one-to-one connections between the observation layer and the labeling layer. These connections are well defined by the local classification model,  $p(D_L|H_L)$ . On the other hand, since the local decisions of two adjacent labeling nodes are usually highly correlated, we build the adjacency model  $p(H_L)$  based on a two-clique random field [6]. Moreover, to model the interactions between the labeling layer and the semantic layer, each semantic node is connected to related labeling nodes. Through those connections, global semantic knowledge can be used to influence the classification of local labeling nodes. Those topological connections are automatically learned and characterized by the global semantic model,  $p(D_L|S_L)$ . In the parking space detection problem, the inter-occluded patterns between cars are treated as semantic prior and are modeled in  $p(D_L|S_L)$ . Note that since a car can only park at a well specified parking slot and occlude only its neighbors, each semantic node only connects to its relative nodes on the labeling layer.

In mathematics, the BHDF framework is a well defined graphical model. It models the parking space detection problem as an MAP optimization problem, as expressed in (1). Basically, we aim to find the most reasonable space status  $S_L$  and pixel labeling  $H_L$  based on the image partition  $D_L$ .

$$H_L^o, S_L^o = \arg \max_{H_L, S_L} p(H_L, S_L | D_L) \quad (1)$$

Moreover, the graphical model reveals the whole statistical properties embedded in the detection problem. These properties include that  $D_L$  are conditionally independent of  $S_L$ ,  $d_i$  at pixel  $i$  only depends on its corresponding labeling node  $h_i$ , and different  $s_i$ 's are independent of each other. Additionally,  $p(s_i)$  is assumed to be a uniform distribution. “Occupied” status and “available” status are equally possible for every parking slot. Based on these statistical assumptions, (1) could be rewritten as (2). In (2), the statistical distributions of  $\ln(p(d_i|h_i))$ ,  $\ln(p(h_i, h_j))$ , and  $\ln(p(h_i|R_s(h_i)))$  are defined by the local classification model, the adjacency model, and the global semantic model,

respectively. The details of these three models are to be explained later.

$$\begin{aligned} & \arg \max_{H_L, S_L} p(D_L | H_L, S_L) p(H_L | S_L) p(S_L) \\ &= \arg \max_{H_L, S_L} \left[ \sum_{i \in M} \ln(p(d_i | h_i)) + \ln(p(H_L | S_L)) \right] \\ &= \arg \max_{H_L, S_L} \left[ \sum_{i \in M} \ln(p(d_i | h_i)) + \ln(p(h_i | R_s(h_i))) \right] \\ & \quad + \sum_{i \in M} \sum_{j \in N_i} \ln(p(h_i, h_j)) \end{aligned} \quad (2)$$

Here, due to the uniform distribution assumption, we ignore the effect of  $\ln(p(S_L))$ . Note that  $R_s(h_i)$  is the set of semantic nodes connecting to  $h_i$  in the graph.  $N_i$  is the neighbors of the  $i$ th pixel. To determine the output, we adopt the belief propagation technique [5] to derive the optimal inference of the graphical model.

#### 4. THREE TRAINABLE MODELS

##### 4.1. Local classification model

To train  $p(D_L | H_L)$ , we first assume each image pixel has the identical local classification model and we aim to calculate  $p(d_i | h_i)$ . For  $p(d_i | h_i = G)$  and  $p(d_i | h_i = O)$ , we use (3) to estimate the distributions.

$$\begin{cases} p(d_i | h_i = G) = \frac{1}{N_G} \sum_{j=1}^{T_n} \phi(d - d_{Gj}) \\ p(d_i | h_i = O) = \frac{1}{N_D} \sum_{j=1}^{T_m} \phi(d - d_{Dj}) \end{cases} \quad (3)$$

where  $N_G$  and  $N_D$  denote the normalization terms, and  $\phi(\cdot)$  is a symmetric kernel function. Moreover,  $d_{Gj}$  denotes one of the  $T_n$  “ground” pixel samples collected from the “ground” region, while  $d_{Dj}$  denotes one of the  $T_m$  “otherwise” pixel samples from the “otherwise” regions. The “ground” regions and the “otherwise” regions are pre-defined, as illustrated in Fig. 4(b). To overcome the luminance variation problem, we train different probability models for  $p(d|h=G)$  and  $p(d|h=O)$  under different luminance conditions. Since the camera is basically stable in the parking lot, some image regions can be selected in advance. Within those image regions, the first and second moments of the image data under different luminance conditions are calculated and stored. By comparing to those statistics, we may select a suitable classification model dynamically.

To estimate the color model  $p(d_i | h_i = C)$  for “car” pixels, we apply the method introduced in [6]. The color data of many “car” pixels are collected and converted to a new color domain based on the following transformation.

$$\begin{aligned} Z &= (R+G+B)/3 \\ u &= (2Z-G-B)/Z \\ p &= \text{Max} \{ (Z-G)/Z, (Z-B)/Z \} \end{aligned} \quad (4)$$

To combat the luminance variation problem, only the chromatic information  $(u, p)$  is used. Based on [7], the chromatic values of the “car” pixels concentrate compactly in the  $u$ - $p$  color space. The distribution can be approximated by a Gaussian function:

$$p(x | \text{vehicle}) = \frac{1}{2\pi\sqrt{|\Sigma_v|}} \exp\left(-\frac{1}{2}(x-m_v)\Sigma_v^{-1}(x-m_v)^t\right) \quad (5)$$

where  $x=(u, p)$ ,  $m_v$  is the chromatic mean, and  $\Sigma_v$  is the chromatic covariance matrix. If we define  $x = T(d)$  as the color transformation in (4), the “car” pixel model is defined as

$$p(d|h=C) \equiv \frac{1}{N_C} p(T(d)|\text{vehicle}) \quad (6)$$

where  $N_C$  denotes the normalization term.

##### 4.2. Global semantic model

To model  $p(h_i | R_s(h_i))$ , we build up a probability table for each pixel  $i$ . The probability model is approximated by an accumulated histogram and is trained by using lots of pseudo training samples. A pseudo training sample representing the inter-occluded pattern can be automatically generated as follows. Here, we assume each vehicle is a cube in the 3-D world. The length ( $l$ ), width ( $w$ ), and height ( $h$ ) of the cube are modeled as three independent Gaussian random variables with pre-trained means and variances. We also assume the status of a parking space  $s_i$  is a random variable with uniform distribution  $p(s_i)$ . The statuses of different parking spaces are considered to be independent of each another. In addition, inside a parking space, the location  $(x, y)$  where the center of a car is parked at is random. We model the location as two independent pre-trained Gaussian random variables as well. Based on those models and the camera projection parameters, we randomly generate 5000 pseudo training samples. For example, a pseudo training sample is shown in Fig. 4(a), where the white regions represent “car” pixels and black regions represent “ground” pixels or “otherwise” pixels. In Fig. 4(b), we further illustrate the definitions of “ground” regions and “otherwise” regions.

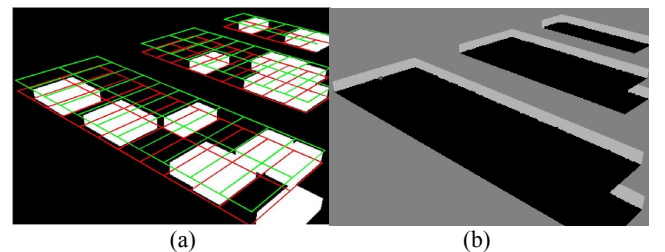


Fig. 4 (a) A training sample of parking lot status for global semantic model. (b) Pre-defined “ground” regions in the black color and “otherwise” regions in the light-gray color. Here, the dark-gray area indicates the region that we do not care. This “don’t care” region is not considered in parking space detection.

### 4.3. Adjacency model

As in the global semantic model, we build up an accumulated histogram to approximate the probability table  $p(h_i, h_j)$  for each neighboring pair based on 5000 pseudo training samples. Basically, most of the neighboring pairs follow the smoothness constraints. That is, neighboring labels tend to belong to the same categorization. However, some regions, such as the boundary between “car” pixels and “ground” pixels, have quite different statistical distributions. By using the adjacency model, our system may characterize adjacent constraints for various locations.

## 5. EXPERIMENTS AND RESULTS

In our experiments, we set up an IP camera on the roof of a building near a parking lot and monitor the status of parking spaces from morning to evening. To evaluate the performance of our detection algorithm, we test different parking lot images captured in different luminance conditions. The test images include 1300 available parking spaces and 1500 occupied spaces. The final false-acceptance rate (FAR) and false-rejection rate (FRR) are 0.032 and 0.02, respectively. In comparison, Dan’s method [4], which is a single-space detection method based on a SVM classifier, has  $FDR = 0.048$  and  $FRR = 0.071$ . In addition, we revealed the segmentation results using the proposed BHDF framework. As expected, the semantic knowledge is helpful in the detection of parking spaces. The world projection information is well-learned and embedded in the semantic knowledge so that the inter-object occlusion problem can be properly handled. Moreover, the partially occlusions caused by environmental obstacles, like the leaves in Fig. 5(a) and (c), do not affect the detection results. This is due to the fact that the other parts of the parking spaces provide enough supports to make global and correct decisions. Up to now, the whole system has been implemented in the Visual C++ environment on a PC with a 1.6GHz Pentium-4 CPU. It takes about 3 seconds to perform the detection and segmentation of parking spaces for a 320x240 color image. The speed is reasonably fast for practical usage.

## 6. CONCLUSION

In this paper, we propose a 3-layer Bayesian hierarchical detection framework for robust parking space detection. The practical problems caused by luminance variations and occlusions are properly handled. The luminance variation problem is tackled in the classification process between the observation layer and the labeling layer, while the occlusion problem is handled by the relationships between the labeling layer and the semantic layer. Since the parking space problem has special geometric structure, we further divide occlusions into car occlusions and environmental occlusions. Car occlusions are learnable and can be modeled

into the semantic prior knowledge to restrict the possible labeling configurations. On the other hand, incorrect inference caused by environmental occlusions can be suppressed by incorporating the global hypotheses passed from the semantic layer. By applying the BHDF to the parking space detection problem, the available parking spaces and the labeling of parked cars can be achieved in a robust and efficient manner. Furthermore, this BHDF framework is generic enough to be used for various kinds of detection and segmentation applications.

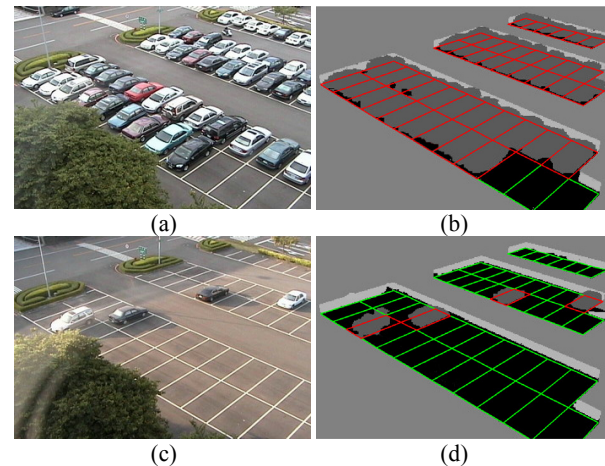


Fig. 5. Left column: two images for parking space detection under different luminance conditions. A parked car may occlude some portions of neighboring cars or adjacent parking spaces. Some parking spaces may get occluded by the environmental obstacles like the leaves. Right column: the corresponding detection results using the proposed BHDF framework. An occupied parking space is marked in red, while an empty parking space is marked in green. “Ground” pixels, “car” pixels, and “otherwise” pixels are represented in black, darker gray, and light gray, respectively.

## 7. REFERENCES

- [1] X. G. Wang, A. R. Hanson “Parking lot analysis and visualization from aerial images,” *Proc. in Fourth IEEE Workshop Applications of Computer Vision*, pp.36-41, 1998.
- [2] C. H. Lee, M. G. Wen, C. C. Han and D. C. Kou, “An automatic monitoring approach for unsupervised parking lots in outdoors,” *Security Technology, 2005. CCST '05. 39th Annual 2005 International Carnahan Conference*, pp.271 - 274, 2005.
- [3] I. Masaki, “Machine-vision systems for intelligent transportation systems,” *IEEE Conference on Intelligent Transportation System*, Vol.13(6), pp.24-31, 1998.
- [4] Noah Dan, “Parking management system and method,” *US patent*, Pub. No.: 20030144890A1, Jul 2003.
- [5] Christopher M. Bishop, “Pattern recognition and machine learning,” Springer, 2006.
- [6] T. Boykov, O. Veksler, R. Zabih, “Markov Random Fields with Efficient Approximations,” *CVPR'98*, pp. 648-655, June 1998.
- [7] Luo-Wei Tsai, Jun-Wei Hsieh, Kao-Chin Fan, “Vehicle Detection Using Normalized Color and Edge Map,” *ICIP'05*, pp. 598-601, Sept. 2005.