RECTANGULAR EMPTY PARKING SPACE DETECTION USING SIFT BASED CLASSIFICATION

Harish Bhaskar, Naoufel Werghi
Dept. of Computer Engineering, Khalifa University, Sharjah Campus, Sharjah, U.A.E.

Saeed Al-Mansoori
Dept. of Communication Engineering, Khalifa University, Sharjah Campus, Sharjah, U.A.E.

Keywords: Radon transform, Parking space detection, SIFT, Supervised classification, Peak extraction, Filtering, Geometric and spatial constraints.

Abstract: In this paper, we describe a method of combining rectangle detection and scale invariant feature transform (SIFT) analysis for empty parking space detection. A parking space in a parking lot is represented as a rectangular region of pixels in an image captured from an aerial camera. Detecting rectangular parking spaces in a new image involves an alternating scheme of extracting peaks from the Radon transform for the whole image and filtering them against specific geometric and spatial constraints. We then compute SIFT descriptors from these detected rectangular parking spaces and further apply supervised classification methods for detecting empty parking spaces. We demonstrate the performance of our model on several synthetic and real data.

1 INTRODUCTION

With increasing number of vehicles over years, parking has become an important issue particularly in commercial environments such as shopping malls and airports. Empty parking space detection system that can automatically identify empty parking spaces and guide users to it will save lot of time, money, effort and are highly desirable. Recently, a number of solutions that employ different sensors have been suggested, but they are either too expensive to implement or have failed to be effective. According to (Bong et al., 2008), we can conveniently categorize these car parking management or guidance system based on their technologies into: counter-based, wired sensor-based, wireless sensor-based and image-based.

The simplest is the counter-based system that counts the number of cars that enter and exit a car park area (Ristola, 1992). This system is capable of providing information on the total number empty parking spaces in an area but cannot guide the driver to the location of empty parking spaces. The wired sensor system rely on installing ultrasonic systems in each parking lot to detect the occupancy of parking spaces. These sensors are managed by a control unit to which they are wired and are capable of directing drivers to empty parking spaces. Wired systems differ by the type of sensor that is used. Different sensors including but not restricted to vortex berth detector (Yu and Liu, 2004), magnetic field sensor (Wolff et al., 2006) are currently available for parking guidance. These systems are limited due to the high installation and maintenance costs that results from the long and complicated wiring that is required to get them functional. A simple extension of wireless sensor systems to parking guidance have also been available in recent years. The sensors in most wireless systems are micro-controlled and typically include multi-sensor controls. They are more effective than wired sensor system but are very expensive than most other systems. Some common examples of these systems include: (Tang et. al, 2006) that uses the extended crossbow network architecture, (Benson et al., 2006) that adopts an anisotropic magneto-resistive magnetic field sensor together with a microprocessor transceiver.

Finally, image based techniques based on video capture sensors are also commonly used. These are particularly suitable where parking areas are already monitored by CCTV surveillance systems. Image based parking guidance systems are simple and inexpensive. However, transmitting image or video information through wireless sensors could be slightly costly. There has been some recent research efforts

Bhaskar H., Werghi N. and Al-Mansoori S..
RECTANGULAR EMPTY PARKING SPACE DETECTION USING SIFT BASED CLASSIFICATION.
DOI: 10.5220/0003358702140220
In *Proceedings of the International Conference on Computer Vision Theory and Applications* (VISAPP-2011), pages 214-220
ISBN: 978-989-8425-47-8

that has focused on the applications of car parking system using image sensor technologies as in (Funck et al., 2004). Other examples include: (Seo et al., 2009; Wang and Hanson, 1998) where empty parking spaces are estimated using an aerial image and semi-supervised learning mechanism and (Bong et al., 2008) where an integrated approach using image processing algorithms called Car Occupancy Information System (COINS) for accurate and robust monitoring of parking spaces. However, these models often need manual seeding and are restrictive for

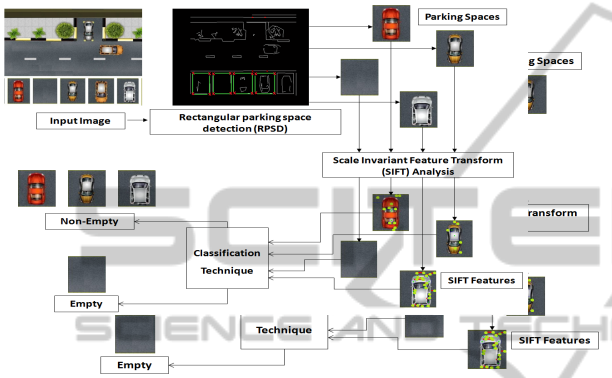


Figure 1: Schematic illustration of the proposed empty car parking space detection technique.

In this paper, we explore an unified framework for detecting rectangular parking space using combined Radon transform and SIFT descriptor based classification for image based empty parking space detection. Our model combines geometric constraints over the Radon transform domain with spatial constraints in the co-ordinate (or image) domain for accurate and simultaneous detection of multiple rectangular parking spaces. Further, we integrate SIFT descriptor analysis and classification for efficient empty parking space detection. An illustration of the framework is presented in Figure 1.

Contributions & Structure: The method proposed in this paper combines Radon transform based rectangle detection for parking space extraction with SIFT analysis (Lowe, 2004) based empty parking space classification. One novelty of the method is the integrated three stage framework for empty parking space detection. In the first step, we use Radon transform based rectangle detection that allows detecting parking spaces in real image and is robust to the presence of noise and illumination changes. In the second step, we classify the extracted parking spaces as empty spaces using SIFT features and supervised classification techniques. Our results indicate that using the proposed framework presents good performance at high accuracy rates.

2 METHOD

We formulate our proposed method as a three stage process of: rectangular parking space extraction using Radon transform based rectangle detection algorithm, SIFT feature detection on extracted parking spaces and finally supervised classification. We represent a parking space in the image as a rectangular region $\mathbf{R}$. Given an image $\mathbf{I}$ of the parking space, our first step is to extract n parking spaces $\mathbf{R}_n$ from $\mathbf{I}$. Once, we obtain the rectangular parking space regions, we perform SIFT analysis on each of the n parking spaces to obtain $\mathbf{S}_n$, where $\mathbf{S}_i$ is a feature vector containing descriptors of that region $\mathbf{R}_i$. These feature descriptors are further classified into either empty or full classes representing whether the parking space is empty or occupied using a supervised classifier. We summarize our approach as follows:

1. Input any given image $\mathbf{I}$ of the parking area.
2. Parking space extraction: We adopt a rectangle detection algorithm based on Radon transform for parking space extraction. i.e $\mathbf{R}_n = \Gamma(\mathbf{I})$, where Γ is the rectangle detection function and $\mathbf{R}_n$ are the n rectangle regions extracted from $\mathbf{I}$
3. SIFT Feature Detection: For every parking space region $i = 1 \ldots n$,
   - Compute SIFT descriptors: $\mathbf{S}_i = \psi(\mathbf{R}_i)$, where $\psi(.)$ is the SIFT operator
4. Perform binary Classification of descriptors to obtain class label $\ell = -1 or 1$, where -1 represents empty parking space and 1 represents an occupied parking space for each of n SIFT descriptors $\mathbf{S}_n$. i.e. $\ell = \chi(\mathbf{S}_n)$, where $\chi(.)$ represents the classifier.

In the following subsections, we describe the $\Gamma(.)$, $psi(.)$ and $\chi(.)$ functions in more detail. We formulate the problem of rectangular parking space detection as a alternating scheme of transform peak extraction and combined spatial and transform domain geometric constraints based peak filtering as depicted in Figure 2.

Figure 2: Schematic illustration of the proposed rectangular parking space detection methodology.

We represent the rectangular parking space as two pairs of transform peaks that satisfy certain specific geometric constraints in the transform domain in addition to the spatial constraints of the corresponding line segments in spatial co-ordinate (image) domain. Given an image **I**, our aim is then to find rectangle(s) **R** each represented as two pairs of peaks PP_1 and PP_2, where PP represents a **P**eak **P**air that satisfy certain conditions. We summarize our approach as follows:

1. ***Edge Detection and Enhancement:*** Get the edge image using: $\mathbf{E} = \epsilon(\mathbf{I})$, where $\epsilon(.)$ is the edge detection and enhancement operator. We perform conventional edge detection using the Canny operator and further enhance the edge image using classical denoising, edge linking and edge cleaning operations. We use this edge image **E** for further processing.
2. ***Peak Extraction:*** Apply the Radon transform on the edge image: $\mathbf{A} = \mathrm{T}(\mathbf{E})$, where T(.) represents the Radon transform. The Radon transform of an image pixel $\mathbf{I}(x,y)$ as the integral projection along a straight line defined by its distance ρ from the origin and its angle of inclination θ, a definition that is similar to that of the Hough transform. The Radon transform is considered to be the generalization of the Hough transform and has been proven to be more robust to noise for rectangle detection in (Bhaskar et al., 2010). Mathematically,

$$\mathbf{A}(\rho,\theta) = \int_x \int_y \mathbf{I}\delta(xcos\theta + ysin\theta - \rho)dxdy \quad (1)$$

   where the δ is the dirac delta function defines integration only over the line. The range of θ limited to $[0,\pi]$. Similar to the Hough transform, each point in the Radon transform space correspond to a straight line in the spatial co-ordinate domain. Conversely, each point in the spatial domain becomes a sine curve in the projection domain.Select n peaks from **A** i.e. $\mathbf{P} = \psi(\mathbf{A})$, where **P** are n peaks extracted from the **A** and ψ function which finds n global maxima from **A**
3. ***Peak Filtering:*** We filter peaks at two levels:
   - We filter all n peaks **P** in the transform domain to obtain intermediate peaks pairs **IP** that satisfy certain geometrical constraints. i.e. $\mathbf{IP} = C_{\mathbf{A}}(\mathbf{P})$, where $C_{\mathbf{A}}$ represents the parallel and orthogonal constraints in the transform space.
   - We further filter all intermediate peaks pairs **IP**, in the form of line segments in the spatial co-ordinate domain to get the selected peak pairs **PP**, i.e. $\mathbf{PP} = C_{\mathbf{I}}(\mathbf{IP})$ that form potential rectangles. where $C_{\mathbf{I}}$ are spatial constraints in the image domain.

Having transformed the original edge image **E** using one of the transformations mentioned above, we have **A**. Our next step is to identify n peaks in the transform space that correspond to lines in the image. Since **A** is a type of accumulator, the simplest mechanism of extracting peaks is using *thresholding*. That is, $\mathbf{A} > t_p$, where t_p is a pre-defined threshold value, which if needed can be learnt from training.
It is possible that there can be more than one peak in the transform space that all satisfy the threhsolding constraint within any reasonable neighborhood. These multiple peaks of the transform space within a small neighborhood region will correspond to duplicate lines in the spatial domain. Duplicated peaks will also be discarded in the Transform Domain Peak Filtering stage. We therefore, discard all other peaks within any specified neighborhood region.

We extend the algorithm of (Jung and Schramm, 2004), to detect rectangle patterns in the transform space. We denote the n extracted peaks as $P_1, P_2, \ldots P_n$ of $\mathbf{A}(\rho,\theta)$. We iteratively compare peaks P_i and P_j in order to identify those that are parallel to one another, i.e. those that satisfy the following conditions:

$$\begin{aligned} \triangle\theta &= |\theta_i - \theta_j| < t_\theta \\ |A(\rho_i,\theta_i) - A(\rho_j,\theta_j)| &< t_l \frac{A(\rho_i,\theta_i) + A(\rho_j,\theta_j)}{2} \end{aligned} \quad (2)$$

where, t_θ is an angular threshold that determines if P_i and P_j correspond to parallel lines. t_l is the threshold that determines if lines corresponding to P_i and P_j are of the same lengths. Each pair of peaks that satisfy 2 denote extended peaks $\mathbf{EP} = (\beta_k, \mathcal{A}_k)$, where

$$\beta_k = \frac{1}{2}(\theta_i + \theta_j) \; and \; \mathcal{A}_k = \frac{1}{2}(A(\rho_i,\theta_i) + A(\rho_j,\theta_j)) \quad (3)$$

These peaks are further checked for orthogonality. The lines that correspond to these extended peaks are orthogonal if the following condition is satisfied:

$$\triangle\beta = ||\beta_r - \beta_s| - 90^o| < t_\beta \quad (4)$$

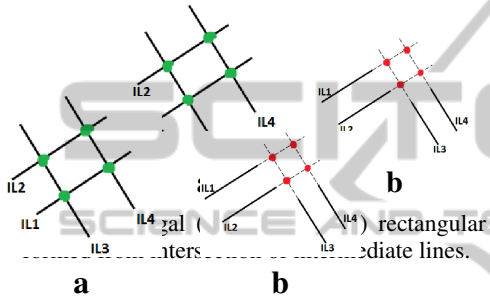
We call those extended peak pairs that satisfy condition 4 as the intermediate peak pairs **IP**. These **IP**'s, representing potential candidate rectangles, that are further filtered in the spatial domain, in order to extract the actual ones.
Let us consider one intermediate peak pair $\mathbf{IP}_i$ (containing 4 peaks, pair of 2 parallel peaks that are orthogonal to one another) and the corresponding line segments $\mathbf{IL}_i$ in the spatial domain. We use a line

intersection mechanism for detecting genuine rectangular structures subtended by these line segments $\mathbf{IL}_i$. According to our technique,

- We compute the point of intersection J for each pair of orthogonal line segments from the set $\mathbf{IL}$.
- For each intersection point J, we check if:

$$\left|\frac{\overrightarrow{\mathbf{IL}_i^s J}.\overrightarrow{\mathbf{IL}_i^f J}}{\|\overrightarrow{\mathbf{IL}_i^s}.\overrightarrow{\mathbf{IL}_i^f}\|^2}\right| < \xi \quad (5)$$

where, $\mathbf{IL}_i^s$ and $\mathbf{IL}_i^f$ are the end points of line segment $\mathbf{IL}_i$.



gal ( ) rectangular structures
diate lines.

The satisfaction of Equation 5 guarantees that the point of intersection lies in between the end points of the corresponding line segments. This constraint allows us distinguish between genuine rectangular patters as in Figure 3a from incorrect ones as in Figure 3b. The choice of the value of our threshold ξ introduces additional relaxations on the constraint. We have chosen the value of ξ to be 10 in all our experiments. Within this framework, it is also possible introduce application based size constraints on rectangles being detected. This technique of spatial filtering of peaks for detecting rectangles allow robust detections of contiguous blocks of rectangles simultaneously.

2.1 SIFT Descriptors & Classification

Scale Invariant Feature Transform (SIFT) is an approach for detecting and extracting local feature descriptors that are reasonably invariant to changes in illumination, image noise, rotation, scaling, and small changes in viewpoint. In our framework, we use the SIFT interest point detector proposed by (Lowe, 2004). From every parking space extracted using the rectangular parking space detection algorithm mentioned above, we extract key points and their corresponding SIFT descriptors. We acknowledge that other descriptors can be used, however, the performance of the system will depend on the robustness of the chosen descriptor. The SIFT descriptor is a 128-dimensional vector containing a set of gradient orientation histograms for every key point extracted for that image. We notice that for images containing occupied parking spaces, the number of set of key points that can be extracted is much larger than the number of key points that could be located on an empty parking space. We will exploit this characteristic during the classification. We illustrate the locations of some of these extracted descriptors in images as shown in Figure4.

Figure 4: Locations of SIFT descriptors extracted from the rectangular parking spaces.

For the purposes of image classification, using the high dimensional SIFT descriptors can often by an expensive procedure. Therefore, it is common to perform dimensionality reduction so that higher order features can be computed from these SIFT descriptors. Here, we cluster a large training set of descriptors sampled from our data set using the k-means clustering algorithm and quantize these original 128-dimensional SIFT descriptors by assigning a label of the closest center as described in the work of (Yang and Newsam, 2008). As in (Yang and Newsam, 2008), we associate the frequency count of these labels to each higher order features. In our work, we use a set of 50 higher order features extracted from the SIFT descriptors.

In the final step, we perform classification of the SIFT based features using supervised classification algorithms. We classify the detected parking spaces and their corresponding SIFT features using the Support Vector Machine (SVM) classifier. In addition, we also compare the results of the SVM classifier to a simple *thresholding* operator for empty parking space detection. SVM is a supervised learning algorithm that determines a hyperplane that separates classes by maximizing the margins between them (Yang and Newsam, 2008). SVM is easy to implement and use and have been proved to be very useful in handling high-dimensional feature vectors as in our case. In order to train our SVM, we use manually categorized, sizeable collection of positive (empty parking spaces) and negative (occupied parking spaces) examples. We apply a 3-fold cross validation to estimate the optimal parameters.

In comparison to the SVM type supervised classification, we also implement a simple thresholding operator that categorizes parking spaces as empty if the number of keypoints extracted from the SIFT analysis is less than a pre-defined threshold. We perform a number of initial experimentations to determine an optimal threshold for this binary classification task.

3 EXPERIMENTATION & RESULTS

Dataset Description: A large collection of both synthetic and real data is used for evaluating our model. In the synthetic set, we use a collection of 60 images at varying resolutions starting from 256 x 256 pixels to 493 x 403 pixels. For real images, we again use a collection of 60 images captured aerially from the top of a building using a CANON G7 10 Mega pixel camera. We manually ground truth all the data in the synthetic set containing a total of 272 occupied parking spaces and 136 empty parking spaces. Similarly, we ground truth the real data set to contain a total of 285 occupied parking spaces and 161 empty parking spaces.

Performance Metrics: To evaluate the performance of classification, we compute the following performance metrics using a two-way contingency table as in (Seo and Urmson, 2009).

- Precision, $p = \frac{a}{a+b}$, if $a+b>0$, otherwise undefined
- Recall, $r = \frac{a}{a+c}$, if $a+c>0$, otherwise undefined
- False Positives, $fp = \frac{b}{b+d}$, if $b+d>0$, otherwise undefined
- False Negatives, $fn = \frac{c}{a+c}$, if $a+c>0$, otherwise undefined
- Accuracy, $acc = \frac{a+d}{a+b+c+d}$, if $a+d>0$, otherwise undefined

We begin by dividing the synthetic and real data into 3 sets containing 20 images each for training, validation and testing. During the training process, we manually extract rectangular parking spaces, perform SIFT analysis thus computing the 128 dimension local SIFT descriptors. Furthermore, we extract 50 dimensional higher order features from the 128 dimension descriptors using the k-means clustering approach described in section 2.1. Finally, we associate these features to the respective ground-truth class label and thus train the SVM classifier. We evaluate the feature and classifier combination using a manually intensive 3 fold cross validation.

Table 1: Compared mean (3 fold) metric values between SVM and Thresholding classification on synthetic dataset.

Metric	SVM	Thresholding
Precision	0.969	0.99
Recall	0.942	0.936
False Positives	0.032	0.020
False Negatives	0.031	0.021
Accuracy	0.969	0.975

Table 2: Compared mean (3 fold) metric values between SVM and Thresholding classification on real dataset.

Metric	SVM	Thresholding
Precision	0.952	0.965
Recall	0.932	0.947
False Positive	0.093	0.063
False Negatives	0.061	0.043
Accuracy	0.929	0.969

Synthetic Data: We perform experiments on synthetic data containing clear rectangular parking spaces created under controlled conditions. The images are free from noise variations and will simulate the effect of various parking scenarios. The images in the test sets are taken through steps of edge detection using a Canny operator with threshold 0.1 followed by edge enhancement. We then perform RPSD to detect and extract rectangular parking spaces. The mean error in detecting rectangular parking spaces using the proposed algorithm on the synthetic dataset was found to be around 4.85%.

The detected parking spaces are then subjected to SIFT analysis and classified using both the SVM and Thresholding approaches. We tabulate the quantitative evaluation of the performance of the classifiers in Table 1.

Real Data: We test the same proposed model on the set of real data. The images in the real data are more challenging with variations due to noise, varying illumination conditions etc. Table 2 shows the accuracy of the classifiers. We would like to highlight that the proportion of missed detections of the RPSD model on real images was found to be 12%. This is an increase of 7.15% in comparison to the detections on the synthetic data. However, the proposed model maintains an accuracy greater than 95%.

We notice from the results on both synthetic and real data that the simple threshold based classifier can produce better classification accuracy than the SVM classifier. However the threshold value is fixed on

a trial and error basis. On the other hand, the SVM classifier is more systematic approach that gives additional flexibility of easily extending the current binary classification problem into a multi-class problem. One of the other main problems of the proposed framework is the proportion of missed detections of the RPSD model. The following experiment examines the impact of the RPSD algorithm on the accuracy of the model. We summarize the results of the RPSD algorithm on some example synthetic and real-time images in Figure 5.

F sult ular ace n
sy rea

We also plot the precision-recall curves of model both on synthetic and real data. This is illustrated in Figure 6. We have computed the break-even-point to be at 0.947 for the real data and 0.963 for the synthetic data,which is indicative of a robust and accurate system.

Figure 6: Precision versus Recall curves for Real and Synthetic Data.

We attribute the reduced detection rate of the rectangular parking space detection algorithm to two main reasons. The RPSD algorithm depends on the parking lines to detect parking spaces and therefore partial or full occlusion of these parking lines can affect the detection process. In addition, the RPSD algorithm is aimed at detecting rectangular regions alone. However, in most real images, the parking regions are not all rectangular. This limitation is fairly easy to address by incorporating changes into the condition in Equation 2, such that small variations in the angular orientations of the parking lines is made acceptable.

4 CONCLUSIONS

We have proposed a framework for detecting empty parking spaces in images. Our method combines Radon transform based rectangular parking space detection scheme with SIFT analysis for classification of empty parking spaces. The results of applying our technique to synthetic and real data suggests that the proposed framework is more accurate and robust to the presence of noise, clutter and illumination changes in images. In our future work we plan to extend the proposed technique to detect parking spaces of other arbitrary shapes and compare them with other state-of-the-art methods.

REFERENCES

Benson, J. P., Donovan, T. O., Sullivan, P. O., Roedig, U., and Sreenan, C. (2006). Car-park management using wireless sensor networks. *Proc. 31st IEEE Conf. Local Computer Networks*, pages 588595.

Bhaskar, H., Werghi, N., and Mansoori, S. A. (2010). Combined spatial and transform domain analysis for rectangle detection. *Proc. of the 13th IET Conference on Information FUSION*.

Bong, D. B. L., Ting, K. C., and Lai, K. C. (2008). Integrated approach in the design of car park occupancy information system(coins). *IAENG International Journal of Computer Science*, 1(1):714.

Funck, S., Mohler, N., and Oertel, W. (2004). Determining car-park occupancy from single images. *IEEE Intelligent Vehicles Symposium*, pages 325328.

Jung, C. R. and Schramm, R. (2004). Rectangle detection based on a windowed hough transform. *SIBGRAPI 04: Proc. of the Computer Graphics and Image Processing*, XVII Brazilian Symposium, pages 113120.

Lowe, D. (2004). Distinctive image features from scaleinvariant keypoints. *International Journal of Computer Vision*, 60(2):91110.

Ristola, T. (1992). Parking guidance system in tapiola. *Proc. IEE Conf. Road Traffic Monitoring*, page 195.

Seo, Y.-W., Ratliff, N., and Urmson, C. (2009). Self-supervised aerial image analysis for extracting parking lot structure. Proc. of International Joint Conference on Artificial Intelligence.

Seo, Y.-W. and Urmson, C. (2009). A hierarchical image analysis for extracting parking lot structures from aerial images. *Robotics Institute, Carnegie Mellon University*: CMU-RI-TR-09-03.

Tang, V. W. S., Zheng, Y., and Cao, J. (2006). An intelligent car park management system based on wireless sensor networks. Proc. Int. Sym. Pervasive Computing and Applications, pages 6570.

Wang, X. and Hanson, A. (1998). Parking lot analysis and visualization from aerial images. Proc. of the 4th IEEE Workshop on Applications of Computer Vision, page 36.

Wolff, J., Heuer, T., Gao, H., Weinmann, M., Voit, S., and Hartmann, U. (2006). Parking monitor system based on magnetic field sensors. *Proc. IEEE Conf. Intelligent Transportation Systems*, pages 12751279.

Yang, Y. and Newsam, S. (2008). Comparing sift descriptors and gabor texture features for classification of remote sensed imagery. *Proc.IEEE International Conference on Image Processing*, pages 18521855.

Yu, C. and Liu, J. (2004). A type of sensor to detect occupancy of vehicle berth in carpark. *Proc. 7th Int. Conf. Signal Processing*, pages 27082711.