Combined Spatial and Transform Domain Analysis for Rectangle Detection

Harish Bhaskar, Naoufel Werghi & Saeed Al. Mansoori
Khalifa University of Science, Technology and Research
Sharjah Campus, U.A.E.
{harish,bhaskar, naoufel.werghi, 1748}@kustar.ac.ae

Abstract – *We describe and compare methods for detecting rectangles in images using Hough and Radon transforms. Locating rectangles in a new image involves an alternating scheme of transform peak extraction and peak filtering. During the step of peak extraction, we apply the hough/radon transform on the target image and extract peaks (corresponding to line segments) from the transformed image. We filter these peaks based on certain geometric constraints in the transform and spatial constraints on the coordinate (or image) domain such that every set of 4 filtered peaks correspond to a rectangle in the image. We explore the effect of model parameters on system performance and show that the proposed methods achieves good accuracy for rectangle detection on several synthetic and real datasets.*

Keywords: rectangle detection, hough, radon, spatial constraints, transform domain, licence plate, aerial images.

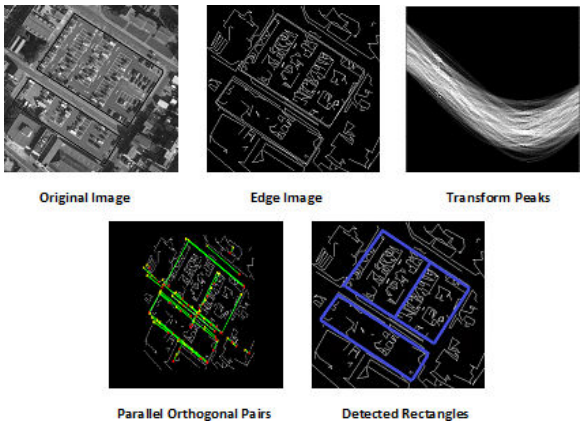


Figure 1: Schematic Illustration of the Proposed Rectangle Detection Technique

1 Introduction

Rectangle detection is a central problem in many applications including: microscopy for detecting particles [19], aerial imaging for detecting vehicle [13] and building structures [7], in surveillance for monitoring rectangular parking spaces [17] and even sometimes for recognizing number plates of vehicles [5]. However, these tasks are complicated by the presence of noise or clutter, illumination changes etc. A popular approach to addressing the problem of rectangle detection is to rely on low-level feature primitives such as edges [15] or lines [10]. Such methods extract spatial consistencies between a pairs of parallel and orthogonal line or edge segments for detecting rectangles.

However, there exist other competing approaches that explore certain geometric characteristics of rectangles in the Hough transform space to detect rectangles [8]. These techniques explore the Hough space for geometric consistencies between pairs of peaks in order to detect rectangles. But, a majority of these strategies cannot detect multiple rectangles simultaneously and also cannot handle inconsistencies that arise due to noise, clutter and variation in illumination conditions.

In this paper, we explore methods of detecting rectangles using steps of transform peak extraction and peak filtering with Hough and Radon transforms. Our model combines geometric constraints over the Hough/Radon domain with spatial constraints in the co-ordinate (or image) domain for accurate and simultaneous detection of multiple rectangles. The model is capable of detecting rectangles of unknown dimensions and orientation and is robust to noise, clutter and illumination changes. We demonstrate and compare the performances of the models on synthetic data and in addition, we present results of applying the models to real-time applications such as car parking space detection using aerial images and number plate detection.

2 Related Work

The goal of most rectangle detection techniques is to analyze if low level feature primitives satisfy certain geometric conditions in either the ***spatial*** or ***transform*** domain. Here in this section, we will briefly describe these methods and highlight their limitation to motivate our proposed methodology.

Under the category of spatial domain methods, a three step approach is commonly used. In the first step, conventional edge detection is performed using any edge operator such as the Canny, Robert [1], etc. Next, using the edges extracted from the previous step, lines are detected. The most common approach to line detection is using the Hough transform [3], although other methods of extracting linear elements from edges using splitting arithmetic [18] have also been tried. Finally, lines are clustered together forming groups of parallel and anti-parallel lines based on spatial constraints on their lengths and orientation. An extension to the classical spatial domain approaches is to use an 2D accumulator array to detect the center and orientation of the rectangles using a Rectangular Hough Transform assuming that the all rectangles have similar dimensions and that the dimensions of the rectangle are known.

On the other hand, the transform domain approaches analyze the peaks of the transform space for rectangle detection. Here, we begin with performing edge detection followed by Hough transform. However, we now detect rectangles directly by analyzing peaks against geometric constraints (including parallelism and orthogonality). The use of a sliding window method has been proven flexible and accurate to detect rectangles of known fixed dimensions. According to the windowed Hough transform framework [8], every pixel of the image is scanned using a sliding window, Hough transform of the edge image within this fixed neighborhood is computed and finally the peaks (that correspond to line segments) in the Hough space are checked against relevant geometric conditions for detecting rectangles directly. However, the method is inefficient for images containing rectangles of different dimensions. In addition, key system parameters that are specified empirically in most current methods such as the size of the neighborhood region for the windowed hough transform method control directly the efficiency and performance of the method. Moreover, these parameters are specific for different applications and have to be decided based on trial and error or using prior knowledge.

One further problem with both the spatial and transform domain approaches is that the line detection mechanism based on Hough transform is highly sensitive to the presence of noise, clutter and changes in illumination conditions. Therefore, much interest has been directed towards the use of Radon transform [21] that exploits integral projections for line detection in the presence of noise. Such methods have been successfully applied to detecting lanes [14], road centerlines [20], detecting internal rail cracks [2] etc.

We hypothesize that a Radon transform based peak extraction mechanism in conjunction with combined spatial and transform domain geometric constraints for peak filtering can allow detecting multiple rectangles simultaneously on the whole image with robustness against noise, clutter and illumination changes.

3 Contribution and Structure

The method proposed in the paper combines Hough transform based peak extraction with a geometric constrained peak filtering algorithm for rectangle detection. One novelty of the paper is that we perform these steps of peak extraction and filtering on the whole image allowing the model to account for detecting multiple rectangles simultaneously. Our results suggest that such an implementation can improve the accuracy and efficiency when compared to other methods.

Furthermore, we extend the model for peak extraction using the Radon transform. We prove quantitatively through extensive experimentation that a Radon-based rectangle detection model is more accurate and robust in comparison to the Hough-based counterpart. This is particularly proven when the images are influenced by the presence of noise, changing illumination conditions and clutter. Finally, we critically evaluate the impact of different system parameters on the performance of the model and draw analysis that would help improve the model.

We begin by describing our method, including the Hough and Radon transform models in Section 4. We then conduct systematic experiments on varied synthetic and real time data (in Section 5) that investigate the accuracy and robustness of the model in addition to investigating the effect of key model parameters and demonstrate the efficacy of the framework when compared to other developed techniques. In Section 6 we present some concluding remarks and directions of future work.

4 Method

In our proposed technique, we formulate the problem of rectangle detection as a alternating scheme of transform peak extraction and combined spatial and transform domain geometric constraints based peak filtering as depicted in Figure 1. We represent the rectangle as

two pairs of transform peaks that satisfy 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 $\mathbf{I}$, our aim is then to find rectangle(s) $\mathbf{R}$ 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.

2. ***Peak Extraction:*** Apply the transform on the edge image: $\mathbf{A} = \mathrm{T}(\mathbf{E})$, where T(.) represents the transform (either Hough or Radon). Select n peaks from $\mathbf{A}$ i.e. $\mathbf{P} = \psi(\mathbf{A})$, where $\mathbf{P}$ are n peaks extracted from the $\mathbf{A}$ and ψ function which finds n global maxima from $\mathbf{A}$

3. ***Peak Filtering:*** We filter peaks at two levels:

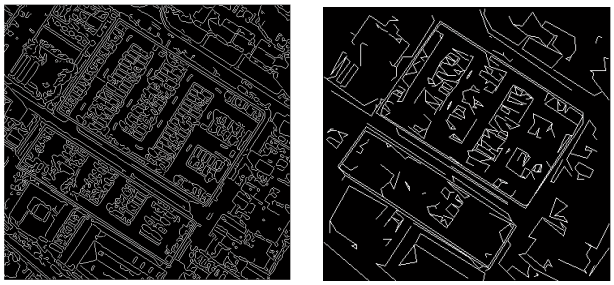
   - We filter all n peaks $\mathbf{P}$ in the transform domain to obtain intermediate peaks pairs $\mathbf{IP}$ that satisfy certain geometrical constraints. i.e. $\mathbf{IP} = \mathcal{C}_{\mathbf{A}}(\mathbf{P})$, where $\mathcal{C}_{\mathbf{A}}$ represents the parallel and orthogonal constraints in the transform space.
   - We further filter all intermediate peaks pairs $\mathbf{IP}$, in the form of line segments in the spatial co-ordinate domain to get the selected peak pairs $\mathbf{PP}$, i.e. $\mathbf{PP} = \mathcal{C}_{\mathbf{I}}(\mathbf{IP})$ that form potential rectangles. where $\mathcal{C}_{\mathbf{I}}$ are spatial constraints in the image domain.

4.1 Edge Detection and Enhancement

An important part of any rectangle detection procedure is edge detection. Our selected baseline method [8] use a combination of denoising method and Canny operator for generating the edge image. Though this is effective in simple real images, we have found through initial experimentation as in 5.2 that this is rather ineffective in a majority of real images containing a large number of edges, as illustrated in our example Figure 2a. Therefore, in addition to conventional edge detection based on the *Canny operator*, we perform edge enhancement. As a part of edge enhancement, in addition to classical denoising, we also link disconnected edges and clean edges that are shorter than a predefined threshold to obtain the final edge image as illustrated in Figure 2b. We use this edge image $\mathbf{E}$ for further processing.

4.2 Transform and Peak Extraction

The first step involves extracting a fixed number of n peaks from the transform space. So, we start by applying the transform T on the edge image $\mathbf{E}$ to obtain $\mathbf{A}$. Our method compares two transforms in the form of Hough and Radon for peak extraction.



a **b**

Figure 2: Canny Edge Detected Image (a) that are enhanced using Edge Enhancement Procedure (b)

We begin by briefly describing the Hough and Radon transforms.

4.2.1 Hough Transform

The Hough transform is an established technique for detecting linear structures in images. According to Hough transform, any line in the spatial co-ordinate domain can be described using the relationship $\rho = xcos\theta + ysin\theta$, where ρ is the normal distance from the center to the line and θ is the normal angle that the line subtends with the center. Edge points that all lie in a straight line are expected to satisfy this relationship and can be re-represented in the Hough space using parameters ρ and θ. Detecting line segments using the Hough transform has been well researched and different kinds of Hough transforms have been proposed including but not restricted to standard Hough transform [3], adaptive Hough Transform [6], probabilistic Hough transform [4], randomized Hough transform [12], etc. [9].

4.2.2 Radon Transform

We define 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. 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-orindate domain. Conversely, each point in the spatial domain becomes a sine curve in the projection domain.

In [11], the difference between the Hough and Radon transform is emphasized. The Hough transform is considered a discretisation of the continuous Radon transform where the mapping corresponds to how a data point in the spatial co-ordinate domain maps on to the transform space.

4.2.3 Peak Extraction

Having transformed the original edge image $\mathbf{E}$ using one of the transformations mentioned above, we have $\mathbf{A}$. Our next step is to identify n peaks in the transform space that correspond to lines in the image. Since $\mathbf{A}$ is a type of accumulator, the simplest mechanism of extracting peaks is using *thresholding*. That is, $\mathbf{A} > t_p$, where t_p is an user defined threshold value.

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 described in section 4.3. We therefore, discard all other peaks within any specified neighborhood region.

4.3 Transform Domain Peak Filtering

We extend the algorithm of [8], 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 $\mathbf{IP}$. These $\mathbf{IP}$'s, representing potential candidate rectangles, that are further filtered in the spatial domain as described in section 4.4, in order to extract the actual ones.

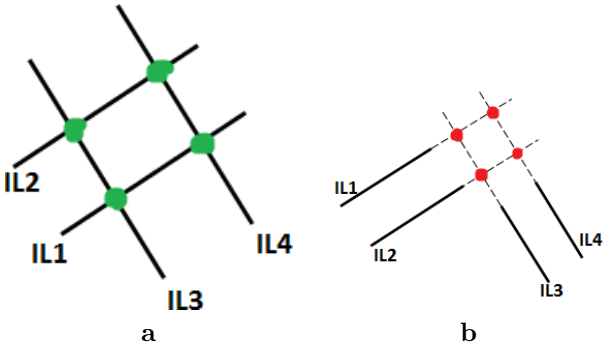


Figure 3: Legal (a) and False (b) Rectangular Structures formed from Intersection of Intermediate Lines

4.4 Spatial Domain Peak Filtering

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

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

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

The satisfaction of Equation 5 guarantees that the point of intersection $\jmath$ 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.

5 Results and Analysis

In this section, we perform systematic experiments evaluating the performance of the proposed models. We then compare performance of the two proposed methods and the windowed Hough transform method for rectangle detection [8] demonstrating them on image data. To evaluate the performance for a given image, we

compute a two-way contingency table as in [16] through visual inspections of the input and output images. It summarizes the outputs of binary classification in terms of: true positive (a) - frequency of accurate detection of rectangles, false positive (b) - frequency of unfiltered duplication of rectangles in the output, false negative (c) - frequency of misdetections and true negative (d) - will be zero for our case.

		Input	
		Found	Not Found
Output	Found	a	b
	Not Found	c	d

Table 1: Contingency Table for Performance Measurement

We compute the following performance metrics based on the contingency table.

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

5.1 Comparison of Performance on Synthetic data

To test the performance of the proposed Radon method, compare it with its Hough transform counterpart we use a dataset of 15 synthetic images. Each image starting from 1 to 15 has increasing complexity, in terms of the number of rectangles to be detected and the clutter in them. All models are initialized with the Canny edge detection method at an edge detection threshold of 0.1. We visually inspect the outputs of all techniques and build our contingency table as described in section 5. The mean metric values listed in Table 2 suggest the Radon transform method compares favorably than the Hough transform method and the windowed Hough transform baseline.

One of the important limitation of the windowed Hough transform technique [8], is its inability to cope with rectangles or even noisy edge lines that are aligned too close to each other. However, our spatial domain constraints based on line intersection principle adds to the robustness to the rectangle detection procedure allowing contiguous blocks of rectangles to be detected simultaneously with high resistance to noisy edges.

Metric	Hough	Radon	Windowed Hough
Precision	0.9435	1	0.8579
Recall	0.6667	0.7333	0.8388
False Negatives	0.3333	0.1667	0.1611
Accuracy	0.6167	0.7222	0.6989

Table 2: Compared Mean Metric Values between Hough and Radon Methods on Synthetic Dataset

Metric	Canny	Sobel	Roberts	Prewitt
Precision	0.8435	0.6579	0.7873	0.6129
Recall	0.6267	0.5333	0.6388	0.5234
Accuracy	0.4323	0.3982	0.4198	0.3872

Table 3: Comparison of the Precision and Recall metrics for model with different edge operators

In terms of computational demand, our proposed methods are slightly more efficient (3270ms average per image compared to 4540ms for the windowed hough transform method) though one should note that our proposed methods capture all rectangles simultaneously without any human intervention that often is the cause for high computational demand as in the case of the baseline. Both algorithms are implemented in Matlab and run on a Pentium Duo Core processor machine with 4GB RAM.

5.2 Effect of Edge Enhancement

As mentioned earlier, edge detection is crucial for accurate rectangle detection both in our proposed model and other baseline strategies. In order to probe the question of finding the best possible edge image further, we performed experiments comparing different edge detection operators including Canny, Sobel, Roberts and Prewitt on the several synthetic and real images. Though in majority of synthetic images these methods were found to be competitive, in all real images, pure edge detection and denoising was ineffective in producing faithful edge images. Our experiments studying the effect of edge enhancement over pure edge detection in Table 3 shows significant increase in the effectiveness of the method for detecting rectangles.

5.3 Effect of Increasing Noise

One of the novelties of the proposed Radon transform method is its ability to handle effectively noise and clutter in images. The following experiment examines the impact of increasing levels of noise on the proposed methods. The algorithms were run on all images of the synthetic dataset with 10 levels of increasing noise. The results in Figure 4, show that the Radon based

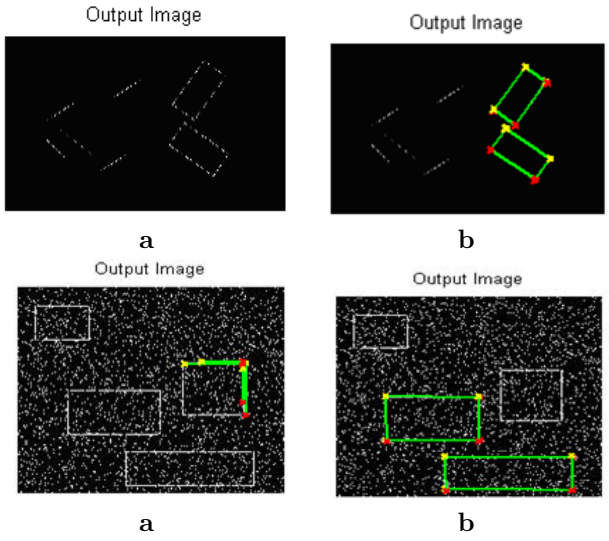


Figure 4: Failed Rectangle Detection of Hough (a) that are Correctly Detected by Radon (b)

Figure 5: Output on Real Images of Parking Lot (a) and License Plate Detection (b)

rectangle detection demonstrates more robustness than the Hough based model to noise and clutter.

5.4 Real-Data Application

Finally, we now present results showing that our Radon model on real data. We run the models on two different applications: a) parking lot detection from aerial images and b) number plate detection. The aerial images of the parking lot were captured at high resolution. For the purposes of the experiment we have considered a region of the aerial image at a resolution of 482 x 464. The car number plate images were captured using a 2M pixel mobile phone camera and processed using Matlab on a Pentium Duo-Core machine. We have measured the complexity of our proposed strategy as 23s on the number plate images. Two images of the number plate at varying illumination condition is considered. We apply our method on the real data and the results of the rectangle detection are illustrated in Figure 5. It is evident that the proposed method can also cope well with slight changes in the illumination conditions.

6 Conclusion

We have developed and compared novel techniques for rectangle detection based on Hough and Radon transforms. Our implementation of the transform peak extraction and peak filtering steps allows detecting multiple rectangles simultaneously over the whole image thus resulting in a method that outperforms other baseline models on different synthetic and real-time data. The results of experiments comparing the two transforms suggests that the Radon transform based rectangle detection algorithm is more robust and accurate in the presence of noise when compared to its Hough transform based counterpart. One of the open question concerns the various threshold that are specified empirically in the system. In our future work, we aim to address this issue by applying learning methods to extract possible threshold information from the data.

References

[1] S. Tabbone D. Ziou. Edge detection techniques - an overview. *Intl. Journal of Pattern Recognition and Image Analysis*, 8(1):537–559, 1998.

[2] M. Djeddi and P. Aknin. Detection of internal rail defects based on the smoothed randon transform in the presence of ultrasonic grain noise. In *World Conf. on Nondestructive Testing*, 2008.

[3] R. O. Duda and P. E. Hart. Use of the hough transformation to detect lines and curves in pictures. *ACM Communication*, 15(1):11–15, 1972.

[4] C. Galambos, J. Kittler, and J. Matas. Progressive probabilistic hough transform for line detection. volume 1, page 1554, 1999.

[5] D-S. Gao and J. Zhou. Car license plates detection from complex scene. *Proc. Intl. Conf. on Signal Processing*, 1(1):1409–1414, 2000.

[6] J. Illingworth and J. Kittler. The adaptive hough transform. *IEEE Trans. Pattern Analysis Machine Intelligence*, 9(5):690–698, 1987.

[7] C. O. Jaynes, A. Hanson, E. Riseman, and H. Schultz. Building reconstruction from optical and range images. In *Proc. of Intl. Conf. on Computer Vision and Pattern Recognition*, volume 0, page 380, 1997.

[8] C. R. Jung and R. Schramm. Rectangle detection based on a windowed hough transform. In *SIBGRAPI '04: Proc. of the Computer Graphics and Image Processing, XVII Brazilian Symposium*, pages 113–120, 2004.

[9] C. R. Jung and R. Schramm. Parallelogram detection using the tiled hough transform. In *Proc. of Intl. Conf. on Systems, Signals and Image Processing*, pages 177–180, 2006.

[10] D. Lagunovsky and S. Ablameyko. Straight-line-based primitive extraction in grey-scale object recognition. *Pattern Recognition Letters*, 20(10):1005–1014, 1999.

[11] C.L. Luengo Hendriks M. van Ginkel and L.J. van Vliet. A short introduction to the radon and hough transforms and how they relate to each other. Technical Report QI-2004-01, Delft University of Technology.

[12] Robert A. Mclaughlin. Randomized hough transform: Improved ellipse detection with comparison. Technical report, Univ. of Western Australia, 1998.

[13] H. Moon, R. Chellappa, and A. Rosenfeld. Performance analysis of a simple vehicle detection algorithm. *Image and Vision Computing*, 20(1):1–13, 2002.

[14] R. Nourine, M. E. Boudihir, and S. F. Khelifi. Application of radon transform to lane boundaries tracking. In *Proc. of Intl. Conf. on LNCS Image Analysis and Recognition*, pages II: 563–571, 2004.

[15] A. Pujol and L. Chen. Line segment based edge feature using hough transform. In *VIIP '07: The Seventh IASTED Intl. Conf. on Visualization, Imaging and Image Processing*, pages 201–206, 2007.

[16] Y-W. Seo and C. Urmson. A hierarchical image analysis for extracting parking lot structures from aerial images. Technical Report CMU-RI-TR-09-03, Robotics Institute, Carnegie Mellon University.

[17] J. K. Suhr, K. Bae, J. Kim, and H. G. Jung. Free parking space detection using optical flow-based euclidean 3d reconstruction. In *Intl. Conf. on Machine Vision Applications*, pages 563–566, 2007.

[18] W-B Tao, J-W. Tian, and J. Liu. A new approach to extract rectangular building from aerial urban images. *Proc. Intl. Conf. on Signal Processing*, 1(1):143–146, 2002.

[19] Z. Yu and R. Bajaj. Detecting circular and rectangular particles based. *Journal of Structural Biology*, 145:2004, 2004.

[20] Q. Zhang and I. Couloigner. Comparing different localization approaches of the radon transform for road centerline extraction from classified satellite imagery. *Proc. of Intl. Conf. on Pattern Recognition*, 2:138–141, 2006.

[21] Q. Zhang and I. Couloigner. Accurate centerline detection and line width estimation of thick lines using the radon transform. 16(2):310–316, 2007.