Comparing image features and machine learning algorithms for real-time parking space classification

M. Tschentscher[1], M. Neuhausen[1], C. Koch[2], M. König[2], J. Salmen[1], M. Schlipsing[1]

[1]Institute for Neural Computing, Ruhr-Universität Bochum,
Universitätsstraße 150, 44780 Bochum, Germany;
PH (+49) 234-32 27 977; FAX (+49) 234-32 14 210;
email: marc-philipp.tschentscher@ini.rub.de
[2]Chair of Computing in Engineering, Ruhr-Universität Bochum,
Universitätsstraße 150, 44780 Bochum, Germany;
PH (+49) 234-32 26 174; FAX (+49) 234-32 14 292;
email: koch@inf.bi.rub.de

ABSTRACT

Finding a vacant parking lot in urban areas is mostly time-consuming and not satisfying for potential visitors or customers. Efficient car-park routing systems could support drivers to find an unoccupied parking lot. Current systems detecting vacant parking lots are either very expensive due to the hardware requirement or do not provide a detailed occupancy map. In this paper, we propose a video-based system for low-cost parking space classification. A wide-angle lens camera is used in combination with a desktop computer. We evaluate image features and machine learning algorithms to determine the occupancy of parking lots. Each combination of feature set and classifier was trained and tested on our dataset containing approximately 10,000 samples. We assessed the performance of all combinations of feature extraction and classification methods. Our final system, incorporating temporal filtering, reached an accuracy of 99.8 %.

INTRODUCTION

In urban areas, finding a vacant parking lot in parking garages or parking lots is time-consuming and a tedious task for drivers. A system to detect available parking spaces to route drivers efficiently to proper lots is desirable. Some systems have reached the market or are under research promising to support the driver by locating a vacant parking lot. The oncoming section gives a detailed overview. A video-based system offers a proper alternative to deal with the classification problem. It is possible to combine low-cost hardware requirements with providing detailed occupancy maps for parking areas, which most of the current systems do not provide. As we show in this paper, several image processing and machine learning algorithms which can be employed to classify parking lots already exist. By using video-based systems several challenges occur especially on outdoor car-parks. Different weather and lighting conditions or objects occluding parking lots might influence the accuracy for the given task. In the implementation and the results section our experiments and the best performing feature combination and machine learning algorithm are shown. Finally, a conclusion about the developed system is given.

RELATED WORK

Sensor-based methods. A common type of pavement embedded sensors are inductive loop detectors (ILDs), which are wire loops either installed at the car park entrance to count entering and leaving vehicles or at each parking space leading to expensive and disruptive maintenance work (Ristola 1992). Another type of embedded systems uses magnetic field sensors, that measures changes in the magnetic flux to detect parking vehicles (Wolff et al. 2006). These kind of sensors need to be employed at each parking lot which can be very costly as each sensing unit is usually attached with a processing unit and a transceiver (Bong et al. 2008). Overhead occupancy detection methods are either based on light, sound or radar sensor systems. The drawback of infrared sensors is their sensitivity towards environmental conditions such as heavy rain, dense fog and blowing snow (Idris et al. 2009). Sound based sensors are insensitive to humidity, but large temperature changes and extreme air turbulence negatively affect their performance. Radar sensors perform well in inclement weather conditions, but sometimes need to be equipped with auxiliary sensors to detect stopped vehicles (Ichihashi et al. 2010). In general, overhead technologies are difficult to install in large outdoor car parks, which limits their applicability in such environments.

Video-based methods. Video-based systems have gathered great attention in recent years (Huang and Wang 2010), since they have the potential to provide a cost effective solution as they allow wide area detection and regular maintenance is possible without disturbing the traffic flow. Moreover, they can use existing visual surveillance infrastructure such as security cameras to capture images and videos (Nallamuthu and Lokala 2008).

True (2007) has combined car feature point detection and color histogram classification using the *k*-nearest neighbor algorithm and support vector machines to detect vacant parking spaces. The limitations of this work are the relatively low detection accuracy (94%). Ichihashi et al. (2012) have improved the detection performance using fuzzy c-means (FCM) clustering and hyperparameter tuning by particle swarm optimization (PSO). Their system has reached a detection rate of 99.6% for outdoor environments. However, they have not reported on the method's performance in terms of real-time applicability. Tsai et al. (2007) have presented a general approach to the detection of vehicles on highways, road intersections, and parking lots under different weather conditions and vehicle orientations. In this approach, a cascaded multichannel classifier based on corner features, edge features and wavelet coefficients was trained to verify the vehicle detection. However, the system is solely based on static images and has reached an average accuracy of merely 94.9%.

Other methods take advantage of the homogeneous appearance of vacant parking spaces. For example, Yamada and Mizuno (2001) have proposed a homogeneity measure by calculating the area of fragmental segments. Although their system has reached a detection rate of 98.7% with real-time performance, shadows and occlusions caused by adjacent cars are ignored. To reduce perspective effects, Lopez-Sastre et al. (2007) have suggested applying a Garbor filter bank on rectified images to derive the homogeneity feature for vacant lot detection. Their method has reached an overall classification rate of 99.7%, but it might fail in cases of strong shadows and over-exposure effects which violate homogeneity assumption. Wu et al. (2007) have presented a multi-class SVM to classify the state of three neighboring spaces as a unit. The reported performance is an error rate of 2.6% with real-time capability.

Huang and Wang (2010) have presented a Bayesian detection framework that takes into account both a ground plane model and a vehicle model to represent environmental changes. Although high accuracy of 99.4% has been reported, their system does not reach real-time performance. Recently, Kabak and Turgut (2010) and Seo and Urmson (2009) have presented methods to detect vacant parking spaces using aerial images. However, real-time applicability has not been in the focus due to the latency in image acquisition. A different approach is to use two cameras. For example, Jung et al. (2005) presented a method based on pixel structure classification and feature-based stereo matching to extract 3D information in real time. However, this approach requires sufficient overlapping of the two camera views, which prevents it from being practical on large car parks.

METHODOLOGY

In this section the methodology of the classification system as proposed in this paper is described. Figure 1 shows an overview of the involved modules.

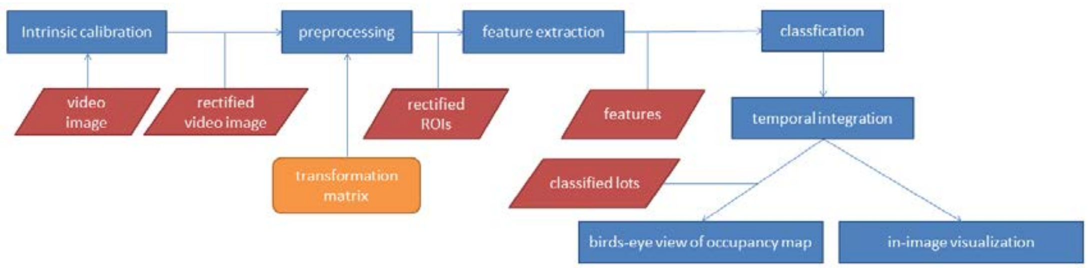

Figure 1. Overview about the different modules used for classification.

Intrinsic Calibration As a first step, one has to develop an intrinsic calibration for the camera. Due to the wide-angle lens of the camera we used in our experiments, the images are radially distorted. To get undistorted images, we used the *radial distortion model* (Weng et al. 1992).

Preprocessing In order to get exact information of the observed lots one has to provide a transformation from camera-image to the world-frame. Therefore, a direct linear transformation (Abdel-Aziz and Karara 1971) is employed which calculates the required transformation matrix. One has to provide at least four world coordinates and their corresponding points in the image manually.

As a next step, the parking lots to be observed have to be marked in the video image manually. It is sufficient to mark the four edges of each lot. These labeled *regions of interests* (ROIs) are extracted from the image and then rectified to fit a common size. This is a requirement of the classification. Figure 2(a) shows examples of such rectified ROIs.

Feature Extraction We implemented several image features to test their ability to describe vacant and occupied parking lots. We concentrate on four different features, namely color histograms, gradient histograms, difference-of-Gaussian (doG) histograms and finally Haar features and tested them separately and in combination.

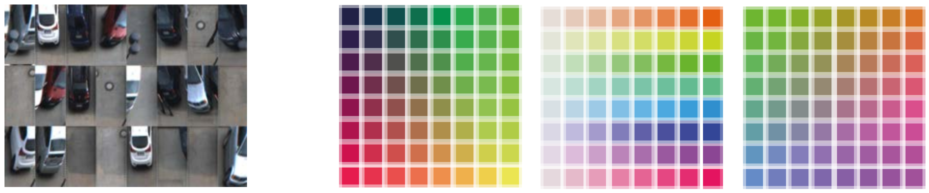

Figure 2(a). Rectified ROIs. (b) Example of RGB, HSV and YUV color spaces.

Color histograms provide detailed information about the distribution of colors in certain images. Therefore we developed an extractor, which is able to create multi-dimensional histograms with different resolutions per channel. We denote (4, 3, 2) as the number of bits for each channel. Figure 2(b) shows an example of the three most popular color spaces (RGB, HSV and YUV). These color spaces are combined with three different edge detection algorithms.

We used a Prewitt operator (Prewitt 1970) that extract edge information in an image and compute a histogram of their distribution (angle histogram). It is possible to define the resolution of these angles by parameter α and defining to use the whole (360) or the semicircle (180) by setting γ to 0 or 1.

The doG method is an alternative filter detecting edges. To extract edge information from a camera image, at first a highly smoothed image (with Gauss-filter) is generated. This image is then subtracted from the original. The histogram represents the distribution of the gray scale values of the resulting image. Figure 3(a) shows thresholded example of doG images.

Haar features (Viola and Jones 2001) are also useful edge-filters which are calculated efficiently with the help of integral images. Haar features can code different orientations of edges. Vertical edges can be coded with features like the first and third whereas the second and fourth detect horizontal edges (Fig. 3(b)).

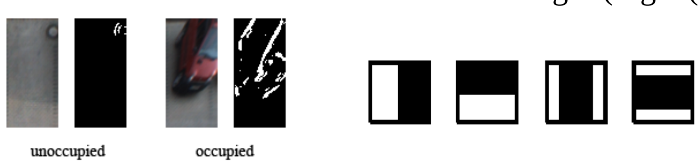

Figure 3. (a) DoG images. (b) Example of Haar features

Classification The system has to perform a binary classification (occupied/vacant). Therefor we tested three different learning algorithms such as *k*-nearest neighbor (*k*-NN), linear discriminant analysis (LDA) (Hastie and Tibshirani 2001) and support vector machine (SVM) (Vapnik 2008).

In order to further improve the classification results we implemented a temporal integration via *exponential smoothing*. The term $f(t)$ shows the filters equation where α represents the learning rate and Y_t represents the particular class label assigned in time step t for each parking lot. If $\alpha = 1$, no new classification will affect the older for a specific parking space. In opposite, $\alpha = 0$ means no influence of older results.

$$f(t) = \alpha * Y_{t-1} + (1 - \alpha) * Y_t$$

We build a dataset of approximately 10,000 samples in order to get significant results. To avoid overfitting in the experiments we divided this dataset into two disjoint sets for training and testing.

IMPLEMENTATION AND RESULTS

In this section the implementation of the whole system is explained. Therefore, an overview of the system and the setup is given. The result of each feature combination and machine learning algorithm are shown afterwards.

Hardware and environmental Setup The software runs on a standard desktop PC. In order to cover a possibly large amount of parking spaces, a of-the-shelf wide-angle lens camera is used for recording. It is positioned in the back of the parking lots with a slight top view (see Fig. 4). In our configuration this leads to a monitored area of 15 parking lots per camera. This keeps the costs rather low.

The implemented modules are lined up in a feed-forward pipeline such that the output of a preceding module acts as input of the succeeding. The implemented visualization is shown in Fig. 4.

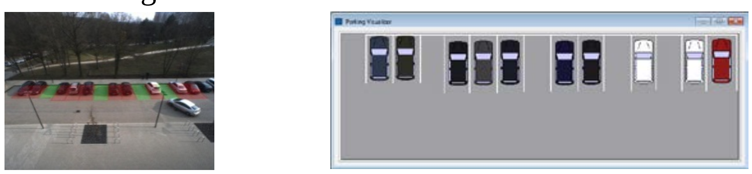

Figure 4. Camera image and corresponding visualization.

In the following sections, the results of each feature combination are presented. We tested each of the before mentioned classification algorithms with each feature combination. The following tables 1 - 7 show the results (test error) in percent on single images without temporal filtering. The minimal test error is highlighted.

Angle Histogram We denote α as the resolution of the edge direction (1 means every full degree, 36 means every 10th degree is used) and $\gamma = 1$ if using a semicircle as edge direction and $\gamma = 0$ if using the full 360° direction.

We employed *k*-NN with *k* = 5 and 50 prototypes per class (Tab. 1). The *LDA* was regularized with $\sigma^2 \in [10^{-6}, 1]$ (Tab. 2). The *SVM* was parameterized during training by setting $\gamma \in [0.2, 1]$ and $C \in [10^2, 10^4]$ using a radial basis function as kernel (Tab. 3).

Table 1. Results for angle histogram features in percent for *k*-NN

	RGB (2,2,2)	RGB (3,3,3)	HSV (4,0,0)	HSV (3,3,0)	HSV (4,4,0)	HSV (3,3,2)	HSV (0,3,3)	YUV (2,2,2)	YUV (3,1,1)	YUV 0,4,4
$\alpha = 1, \gamma = 0$	3.38	3.33	3.65	2.40	2.58	3.53	4.07	3.18	4.10	5.38
$\alpha = 2, \gamma = 0$	3.53	3.77	3.74	**2.38**	2.47	3.36	3.56	3.45	3.92	5.35
$\alpha = 18, \gamma = 0$	3.17	3.59	3.42	3.03	3.15	3.00	4.04	3.86	4.60	5.70
$\alpha = 1, \gamma = 1$	4.27	4.51	3.96	3.00	3.39	3.80	3.95	4.07	6.68	6.39
$\alpha = 2, \gamma = 1$	4.48	4.51	4.99	3.42	3.15	4.07	4.93	4.87	4.16	4.43
$\alpha = 18, \gamma = 1$	4.31	4.90	4.72	4.01	4.78	4.31	5.26	4.48	5.67	6.39

Table 2. Results for angle histogram features in percent for *LDA*

	RGB (2,2,2)	RGB (3,3,3)	HSV (4,0,0)	HSV (3,3,0)	HSV (4,4,0)	HSV (3,3,2)	HSV (0,3,3)	YUV (2,2,2)	YUV (3,1,1)	YUV 0,4,4
$\alpha = 1, \gamma = 0$	7.63	5.20	8.46	6.33	5.02	5.85	8.52	8.26	6.95	11.82
$\alpha = 2, \gamma = 0$	7.22	6.95	9.39	7.16	4.90	6.71	7.87	6.80	6.03	12.15
$\alpha = 18, \gamma = 0$	5.20	3.77	6.50	4.69	4.34	5.44	4.46	4.81	5.14	9.27
$\alpha = 1, \gamma = 1$	6.59	4.93	11.73	6.71	4.48	4.43	7.31	6.65	6.77	10.04
$\alpha = 2, \gamma = 1$	5.73	5.32	8.67	6.42	4.31	4.28	6.21	6.21	6.06	9.27
$\alpha = 18, \gamma = 1$	5.17	4.37	5.76	**3.95**	**3.95**	4.81	5.11	4.90	4.34	7.72

Table 3. Results for angle histogram features in percent for *SVM*

	RGB (2,2,2)	RGB (3,3,3)	HSV (4,0,0)	HSV (3,3,0)	HSV (4,4,0)	HSV (3,3,2)	HSV (0,3,3)	YUV (2,2,2)	YUV (3,1,1)	YUV 0,4,4
$\alpha = 1, \gamma = 0$	2.29	2.49	2.73	1.93	**1.75**	2.47	2.72	2.35	2.29	3.50
$\alpha = 2, \gamma = 0$	2.94	2.79	3.18	2.47	1.84	2.14	3.03	3.18	2.88	3.39
$\alpha = 18, \gamma = 0$	3.83	3.03	6.24	3.09	2.94	3.18	3.50	4.48	4.78	5.97
$\alpha = 1, \gamma = 1$	3.24	2.41	2.64	2.14	1.93	2.44	2.97	2.85	3.30	3.03
$\alpha = 2, \gamma = 1$	3.00	3.15	2.49	2.20	2.29	2.73	2.85	3.21	3.21	3.53
$\alpha = 18, \gamma = 1$	4.07	2.76	3.89	2.38	2.29	2.55	3.65	3.83	4.43	5.88

DoG Histogram We used four different filter sizes. The classification algorithm is parameterized the same as in angle histograms because feature vectors have similar size. Table 4 – 6 show the results for each used classification algorithm.

Table 4. Results for doG histogram features in percent for *k*-NN

	RGB (2,2,2)	RGB (3,3,3)	HSV (4,0,0)	HSV (3,3,0)	HSV (4,4,0)	HSV (3,3,2)	HSV (0,3,3)	YUV (2,2,2)	YUV (3,1,1)	YUV 0,4,4
3 x 3	3.83	2.76	2.88	3.42	3.03	2.20	3.92	3.09	3.80	5.58
5 x 5	3.39	2.70	2.64	3.47	3.15	2.44	3.50	2.91	3.53	4.96
7 x 7	3.45	2.23	2.23	3.09	**2.14**	3.06	3.86	2.52	3.62	4.31
9 x 9	3.56	2.35	2.70	3.09	**2.14**	2.64	3.68	2.55	3.65	4.51

Table 5. Results for doG histogram features in percent for *LDA*

	RGB (2,2,2)	RGB (3,3,3)	HSV (4,0,0)	HSV (3,3,0)	HSV (4,4,0)	HSV (3,3,2)	HSV (0,3,3)	YUV (2,2,2)	YUV (3,1,1)	YUV 0,4,4
3 x 3	5.20	4.66	4.99	4.13	**3.03**	3.80	5.05	4.46	4.96	8.08
5 x 5	5.38	5.73	4.81	3.83	3.18	4.16	5.20	4.66	4.63	7.34
7 x 7	5.29	5.23	3.98	4.48	4.04	4.34	4.25	4.87	4.63	8.32
9 x 9	4.84	6.27	4.13	4.96	3.83	4.60	5.49	4.01	5.20	9.21

Table 6. Results for doG histogram features in percent for *SVM*

	RGB (2,2,2)	RGB (3,3,3)	HSV (4,0,0)	HSV (3,3,0)	HSV (4,4,0)	HSV (3,3,2)	HSV (0,3,3)	YUV (2,2,2)	YUV (3,1,1)	YUV 0,4,4
3 x 3	3.18	2.55	3.09	3.12	2.47	3.53	3.30	2.97	3.74	4.75
5 x 5	3.06	2.61	1.99	1.96	**1.69**	2.67	3.15	2.49	3.06	3.56
7 x 7	2.82	2.73	2.32	1.96	1.96	3.39	3.33	2.26	3.00	4.10
9 x 9	3.03	2.64	1.72	2.82	2.02	3.24	3.33	2.94	3.00	4.40

Haar Features This section shows the results of Haar features (Fig. 5) combined with color histograms in Table 7. For parameterizing the SVM, we used different parameters. We set $\gamma \in [0.0003, 0.0015]$ and $C \in [0.01,1]$.

Table 7. Results for Haar features in percent

	RGB (2,2,2)	RGB (3,3,3)	HSV (4,0,0)	HSV (3,3,0)	HSV (4,4,0)	HSV (3,3,2)	HSV (0,3,3)	YUV (2,2,2)	YUV (3,1,1)	YUV 0,4,4
k-NN	15.74	11.73	12.83	13.15	**10.13**	12.80	14.41	17.02	14.91	18.15
LDA	7.13	6.86	7.01	5.73	**4.25**	5.49	7.81	6.59	6.36	12.00
SVM	5.05	4.69	4.60	4.93	**2.85**	3.74	5.17	4.19	4.40	8.91

It is noteworthy that all of the feature combinations perform best either on HSV(3,3,0) or HSV(4,4,0) color histogram. In preliminary experiments these two color histograms reached the highest accuracy.

Temporal Filtering We employed temporal filtering on whole video sequences in order to improve the results. Varying $\alpha \in [0.7, 0.8]$ and choosing a threshold above 0.8 considering *f(t)*, we achieved classification results of 99.8%.

CONCLUSION

In this paper we developed a promising video-based system for vacant parking space classification using image features and machine learning algorithms. We compared four different image features (color, angle, doG histograms and Haar features) and three different classification algorithms (*k*-NN, LDA, SVM) and showed their performance combining color histograms with one of the others.

The final system relies on color and difference of Gaussian histograms, an SVM classifier and *exponential smoothing* for temporal filtering. This system reached an accuracy of 99.8% and achieved real-time speed.

It is imaginable to use this system in parking garages even though it was only tested on outside parking lots. Because of less space, more cameras are needed to cover a large range of lots. Furthermore the lighting conditions have to be examined and adjusted to get sufficient results. Improvements can be achieved by minimizing the influence of adjacent cars parking left and right from the labeled area. Concerning other car parks or parking garages further experiments should be conducted adapting the system to different lighting conditions.

REFERENCES

Abdel-Aziz, Y. and Karara, H. (1971). Direct linear transformation from comparator coordinates into object-space coordinates in close-range photogrammetry. In *Proceedings of the ASP/UI Symposium on Close-Range Photogrammetry*, pages 1–18. American Society of Photogrammetry.

Bong, D. B. L., Ting, K. C., and Lai, K. C. (2008). Integrated approach in the design of a car park occupancy information system (COINS). *IAENG International Journal of Computer Science*, 35(1):7–14

Hastie, T., Friedman, J., and Tibshirani, R. (2001). *The elements of statistical learning*. Springer Series in Statistics.

Huang, C.-C. and Wang, S.-J. (2010). A hierarchical bayesian generation framework for vacant parking space detection. *IEEE Transactions on Circuits and Systems for Video Technology*, 20(12):1770–1785.

Ichihashi, H., Katada, T., Fujiyoshi, M., Notsu, A., and Honda, K. (2010). Improvement in the performance of camera based vehicle detector for parking lot. In *Proceedings of the IEEE International Conference on Fuzzy Systems*, pages 1950–1956

Idris, M. Y. I., Leng, Y. Y., Tamil, E. M., Noor, N. M., and Razak, Z. (2009). Car park system: A review of smart parking system and its technology. I*nformation Technology Journal*, 8(2):101–113.

Jung, H. G., Kim, D. S., and Yoon, P. J. und Kim, J. H. (2005). Stereo vision based localization of free parking site. In *Computer Analysis of Images and Patterns*, pages 231–239. Springer

Kabak, M. O. and Turgut, O. (2010). Parking spot detection from aerial images. Final project autumn 2010, machine learning class, Stanford University.

Lopez-Sastre, R. J., Jimenez, P. G., Acevedo, F. J., and Bascon, S. M. (2007). Computer algebra algorithms applied to computer vision in a parking management system. In *Proceedings of the IEEE International Symposium on Industrial Electronics*, pages 1675–1680.

Nallamuthu, A. and Lokala, S. (2008). Vision based parking space classification. Report, Clemson University, Department of ECE.

Prewitt, J.M.S (1970), Object enhancement and extraction, Lipkin, Rosenfeld (Eds.), Picture Processing and Psychopictorics, Academic Press, pages 75–149

Ristola, T. (1992). Parking guidance system in tapiola. In *Proceedings of the IEEE Conference on Road Traffic Monitoring*, page pp. 195.

Seo, Y.-W. and Urmson, C. (2009). Utilizing prior information to enhance selfsupervised aerial image analysis for extracting parking lot structures. In *Proceedingsof the IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages339–344.

True, N. (2007). Vacant parking space detection in static images. Report, University of California San Diego.

Tsai, L.-W. and Hsieh, J.-W.and Fan, K.-C. (2007). Vehicle detection using normalized color and edge map. *IEEE Transactions on Image Processing*, 16(3):850–864.

Vapnik, V. N. (2008). *Statistical Learning Theory*. Wiley-Interscience.

Viola, P. and Jones, M. (2001). Rapid object detection using a boosted cascade of simple features. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 511–518.

Weng, J., Cohen, P., and Herniou, M. (1992). Camera calibration with distortion models and accuracy evaluation. *IEEE Transactions on pattern analysis and machine intelligence*, 14(10):965–980.

Wolff, J., Heuer, T., Gao, H., Weinmann, M., Voit, S., and Hartmann, U. (2006). Parkingmonitor system based onmagnetic field sensors. In *Proceedings of the IEEE Conference on Intelligent Transportation Systems*, pages 1275–1279

Wu, Q., Huang, C., Wang, S.-Y., Chiu, W.-C., and Chen, T. (2007). Robust parking space detection considering inter-space correlation. In *Proceedings of the IEEE International Conference on Multimedia and Expo*, pages 659–662

Yamada, K. and Mizuno, M. (2001). A vehicle parking detection method using image segmentation. *Electronics and Communication*, 84(10):25–34.