Parking Lots Space Detection

Qi Wu
qwu@ece.cmu.edu

Yi Zhang
yiz1@ece.cmu.edu

Abstract

A problem faced in major metropolitan areas, is the search for parking space. In this paper, we proposed a novel way for automatic parking lots detection. In our approach, we first extract features by generating detection patchs and building Gaussian ground model from input video frames. Then this model is used to train an eight-calss multi-SVM classifier. Finally, the classfication is optimized globally via applying Markov Random Field.

1 Introduction

In recent years, parking has become a serious problem under the increasing of the private vehicles. Looking for parking space always wastes travelling time. For the driver's convenience, public parking lots should be responsible to inform the availability and location of parking space. However, maintaining such kind of parking space system manually needs lots of human resource. Therefore, unsupervised parking lots detection has been currently employed in many systems for counting the number of parking space, identifying the location, monitoring the changes of space status over the time.

Recent years, many researches have been done on improving parking lots detection systems. Some approaches use visual surveillance, which requires real-time interpretation of image sequences in order to automate the detection of parking spaces [1]. Some other approaches keeps tracking and recording the movement of vehicles in order to find the empty parking space [2, 3]. Nevertheless, these real-time detection methods require high computation and large storage. In this paper, we present a novel method via using only a few frames which are captured by a single camera for for unsupervised parking lots space detection. Because it is impossible to set the camera perpendicular to the parking lots, it is challenging for space detection under the influence of light variety, vehicles' shadow and occlusion. To obtain high detection accuracy under these critical conditions, we train nd recognized from the video frame by machine learning methods, instead of segment them directly to find out the available space. Our goal is to build a highly accurate automatic detection system which is stable and economic for industry application.

2 System Overview

The outline of our proposed parking lots space detection system is shown in Fig 1. this system consists four parts: Preprocessing, ground model feature extraction, multi-class SVM recognition and MRF based verification. In the first section, we rotate the raw input frames into the uniform axis and segment them into small patches which include 3 parking
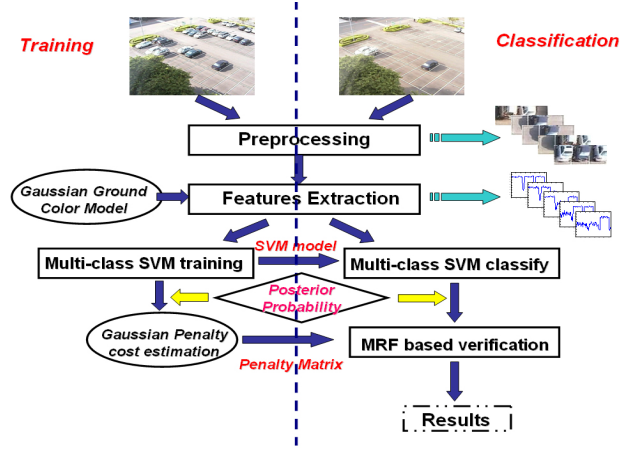
Figure 1: System Overview

spaces each. Then, the probability of ground color of these patches are extracted as the input features, via applying a ground model. Next, Multi-class SVM is trained to analyze and classify the patches into 8 classes (status). Finally, Markov Random Field (MRF) are built to solve the latent conflicts between two neighboring patches. In this step, we use the posterior probability generated by the SVM to improve the accuracy.
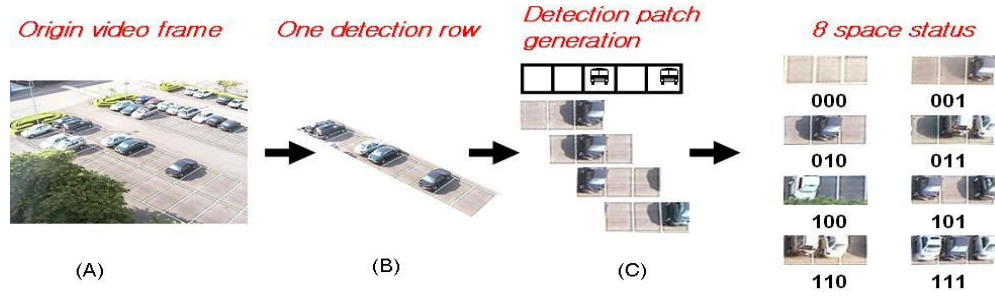
3   Preprocessing

Figure 2: Preprocess the input frame and generate detecting patches

Given an input video frame such as Fig 2, we focus on the parking regions. They can be easily obtained as we know the 3D points of real scene and the intrinsic and extrinsic parameters of camera. Intuitively, one may choose a whole row or a single parking space as a detection patch. However, the influence of light variety, vehicles' shadow and occlusion is accumulated in rows and cannot be attenuated if single space is chosen for detection. Thus, 3 parking spaces are used as a detection patch, which has two spaces in common with neighbor ones. After rotation and interpolation (Fig. 2), the original patches have been normalized into rectangular ones. Additionally, in the training process, we manually classify them into 8 (2^3) space status when label empty space with 0 and occupied one with 1.

4 Feature Extraction

The aim of Feature extraction sections is to obtain the probability which indicate the comparability between the predetermined ground color and the color of a certain pixel on the patches. Firstly, a vehicles color model is build up, which follows a similar way to build a skin color model. However, it is impossible to cluster the vehicles color simply into a color class as what happens in skin color model, because the vehicles's color various a lot in a large range. Therefore, a gaussian ground color model is trained in our approach, because of its stableness in the color space and little variance in different lighting situations. The probability of ground color in pixel x is defined as:

$$p(x) = \frac{1}{2\pi\sqrt{|\Sigma_g|}} exp(-d_g(x)) \tag{1}$$

$$d_g(x) = \frac{1}{2}(x - m_g)\Sigma_g^{-1}(x - m_g)^t \tag{2}$$

where m_g is the mean of ground color and Σ_g is the variance of color ground.

Three scanning lines are used to get the values of pixels from left to right. We computed the probability of every pixel by our ground color model. For a single patch, because there are 75 pixels along each scanning line, 215 ($75 \times 3sacnline$) features are extracted in one sample. Nevertheless, it's not necessary to use all the features, because unimportant features can be dismissed to decrease the training complexity. Thus, Principle Component Analysise (PCA) is applied to extract the critical features. After this demension decreasing, 50 critical features are picked out for further training, which contain over 99% energy of the complete feature set.

5 Recognition

After feature extraction, SVM(Support Vector Machine), which is a powerful tool for classification developed by Vapnik, can be trained quickly. Binary (dual-class) classifier are popular in practical application, which maps vectors with appropriate kernel function into high-dimension of the feature space and satisfy the linearly separable constraint:

$$min|w^T x_i + b| = 1, i = 1, \ldots, N \tag{3}$$

Unlike the classical SVM which use Signum function and has binary output (+1 or -1), we need to know the posterior probability of a single sample to be labelled as +1 or -1. In the general binary classification problem, the probability distribution for y can be define as:

$$p(y_i = \pm 1|x_i, w) = \frac{1}{1 + exp(-\eta f(x_i))} \tag{4}$$

$$f(x) = \sum_{i=1}^{N} \alpha_i y_i K(x, x_i) \tag{5}$$

where α_i denotes the Lagrange multipliers, $\{(x_i, y_i)|x_i \in R, y_i = \pm 1, i = 1, \ldots, N\}$ denotes a set of training samples, $K(x, x_i)$ is the kernel function and $\eta = 1/||w|| > 0$ is the distance between the hyperplane(w, b) and the Support vectors, the closest of the training points. Thus, the binary posterior probability distribution can be written as:

$$p(y_i = \pm 1|x_i, w) = \frac{1}{1 + exp(-1/||w||f(x_i))} \tag{6}$$

The SVM classifier, as introduced above, is a binary classifier, while our parking lots space classification is a multi-class problem, which will identify the detection patches into eight

parking status. Hence, we adapt the general binary SVM classifier by using one-against-one strategy which takes all possible two-class combinations. Therefore, $n(n-1)/2$ SVMs are trained and each SVM classifier separates a pair of classes. Here, n is the number of status, which is 8 in our case.

In our approach, SVM with RBF(Radial Basis Function) kernels is adopted to be the component of the classifier (in Equation (7)).

$$R_i(P) = exp[-\frac{\|P - C_i\|^2}{\sigma_i^2}] \tag{7}$$

After trained by 28 $(8 \times (8-1)/2)$ binary classifiers, the probability that a set of samples belong to the ith class is shown in Equation (8):

$$p(y_i|x) = \frac{1}{2 - n + \sum_{j=1, i \neq j}^{8} \frac{1}{p_{ij}(y_i|x)}}, i = 1, \ldots, 8 \tag{8}$$

where $p_{ij}(y_i|x_i)$ is obtained by Equation (6). Knowing these probabilities, we can say a sample is in the ith status if $p(y_i|x) = \max\limits_{j=1,\ldots,n} p(y_j|x)$.
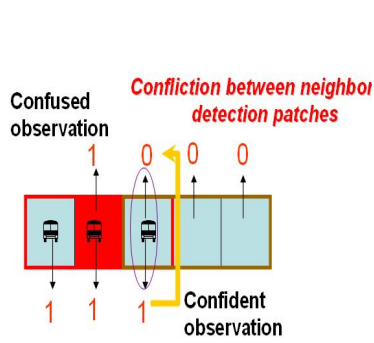
6 Verification and Conflict Correction

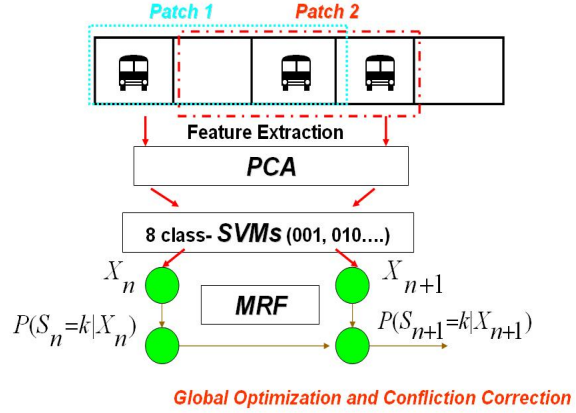Figure 3: Conflict between two neighboring patches

Figure 4: Markov Random Field based Verification

As any other Machine Learning algorithms, SVM cannot guarantee 100 percents accuracy in classification. In our parking lots detection, because the neighboring patches have 2 shared parking spaces, confliction may occur when one or both of the two patches are classified into wrong status. For example. in Fig. 3, we have the results of neighboring patches labelled as 001 and 110, or 110 and 010. To correct these conflictions, Markov Random Field (MRF) is applied to correct the conflicts and optimize the results.

6.1 Markov Random Field Model

As shown in Fig. 4, the status label S of each patch is independent and identically distributed when the posterior probability X of this patch is given. This can be shwon as

$p(S_i = k|X_i) \perp p(S_{i+1} = k|X_{i+1})$. So we can maximize the likeliy hood function:

$$
\begin{aligned}
p(S|X) &= \arg\max p(S_n = i, S_{n+1} = j|X_n, X_{n+1}) & (9)\\
&\propto \arg\max p(X_n, X_{n+1}|S_n = i, S_{n+1} = j) & (10)\\
&= \arg\max p(X_n|S_n)p(X_{n+1}|S_{n+1})p(S_{n+1}, S_n) & (11)
\end{aligned}
$$

$$
\log p(S|X) = \arg\max(\log p(X_n|S_n) + \log p(X_{n+1}|S_{n+1}) + \log p(S_{n+1}, S_n)) \quad (12)
$$

In the our MRF framework, we define a node labelling problem as assigning to every node n in one parking row a label(see Fig 2(B)), which is written as S_n. The energy function E, which can be viewed as the log likelihood $\log p(S|X)$, is composed of a data energy E_d and smoothness energy E_s. The data energy is simply the sum of a set of per-node data costs $d_n(S)$, which is the negative log SVM posterior probability result:$-\log p(X_n|S_n)$, that is $E_d = \sum_n d_n(S_n)$. The smoothness energy, which is also called penalty value, is defined as the sum of spatially varying horizontal neighboring penalty cost,$V_{n,n+1}(S_n, S_{n+1})$, that is $E_s = \sum_n V_{n,n+1}(S_n, S_{n+1})$. Therefore, in order to solve our problem which has been changed into the energy minimization and global optimization, we should to train and estimate the appropriate penalty cost.

6.2 Penalty Cost Estimation

Since there are 2 overlapping parking spaces between two neighboring patches, we basically have 3 kinds of relationship between them.

- **No conflict** eg. the SVM results of two neighboring patches are (100 & 000),(101 & 011)...

- **One conflict** eg. the SVM results of two neighboring patches are (000 & 100),(111 & 011)...

- **Two conflict** eg. the SVM results of two neighboring patches are (110 & 010),(001 & 101)...

Therefore, based on the neighborhood relation on the set of nodes, we roughly define the penalty cost as:

$$
V_{n,n+1}(S_n, S_{n+1}) = \begin{cases} -\log p(S_n, S_{n+1}) = 0 & if.n = abc, n+1 = bcd, a, b, c, d \in [0,1] \\ -\log p(S_n, S_{n+1}) = C & elsewise, C \ is \ a constant \end{cases}
$$

$$(13)$$

However, our experiment shown that the penalty costs, which are although constant numbers, are different in the conditions of classifying a empty parking space to be 1 and labelling a occupied one as 0. Thus, to minimize the total energy, the penalty cost which solves the conflict can be computed by:

$$
V(S_n, S_{n+1}) = (d_n(S_n = i') - d_n(S_n = i)) + (d_{n+1}(S_{n+1} = j) - d_{n+1}(S_{n+1} = j'))
$$

$$(14)$$

where we assume $S_n = i, S_{n+1} = j$ are the results after SVM. $S_n = i', S_{n+1} = j'$ are the true label known from the ground truth. Because these pre-computed penalty costs on different kinds of neighboring patches follows normal distribution, we trained 64 (8×8) gaussian models for every situations. Hence, the estimated penalty values (Table. 1) are the means of these gaussian models. Using this penalty matrix and MRF framework, we can easily solve the conflict and improve our recognition accuracy.

	000	001	010	011	100	101	110	111
000	0.00	1.18	1.32	2.10	0.00	1.31	1.41	2.22
001	0.00	1.21	1.28	1.90	0.00	1.25	1.32	1.97
010	1.36	0.00	1.87	1.17	1.02	0.00	1.88	1.52
011	1.28	0.00	1.98	1.22	0.99	0.00	1.95	1.61
100	1.57	2.13	0.00	1.31	1.13	2.05	0.00	1.32
101	1.56	2.05	0.00	1.27	1.20	1.93	0.00	1.44
110	1.92	1.43	1.37	0.00	2.21	1.42	1.28	0.00
111	2.03	1.52	1.34	0.00	2.11	1.47	1.26	0.00

Table 1: The matrix of penalty values on all the situations

7 Experiments

7.1 Data Acquisition

We captured the video frames from the real scene. Firstly, in order to build a model that is flexible enough to cover most of the variation of ground color, we have collected more than 200 ground patches in different light environments for training. Then from 500 frames which were captured from variant time slot, we generated and selected 2400 patches (300 for each status)as training data. Finally, using the trained Gaussian ground model and 8 class-SVMs, we evaluated the performances of various kernel functions of SVM. Moreover, we also compare the results with and without using MRF conflict correction, to evaluate the performance of MRF.
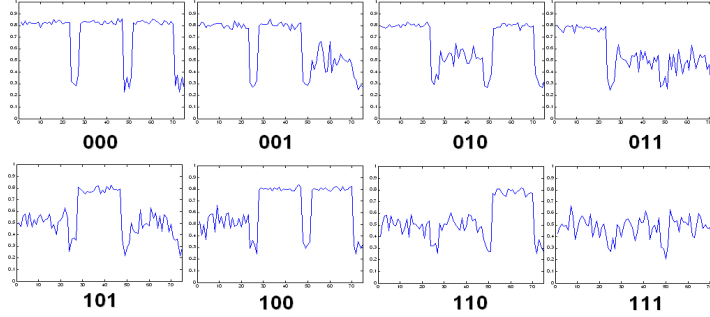
7.2 Features Preprocessing

Figure 5: Profiles on 8 patch status

The gaussian ground color model is used to extract the features from patches. The features are the profile constructed by the continues probabilities of ground color in pixels. Fig 5 shows these profiles in 8 status. Obviously, the empty spaces always have a higher probability than the occupied ones, because their color are closer to ground. Then, PCA is used to reduce the feature dimensions. The experiments demonstrate that first 50 eigen-vectors contribute over 99% information (99.0123%) and we project all the features to these 50 principle components. Fig 6 sketches the projection on the first 3 principle components, which shows our features are basically cluster into small regions and good for classification.
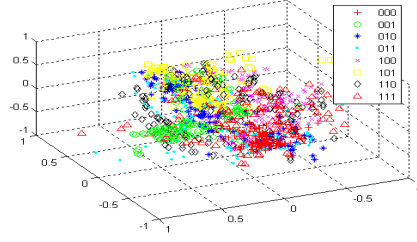
Figure 6: Features project on the first 3 principle components

7.3 Classification and Comparison

After obtaining the features, we classify our patches using Multi-class SVM. Using the kernel functions of linear, polynomial and gaussian radius basis function separately, Fig 7(Classification accuracy) and Fig 8(Conflict rate) demonstrate the comparisons among them. The x-axis of these plots shows the number of training samples. Note that we use separate patches of 8 parking status in the training processing, that x-axis just denote the number of patches for the training. However, in the test process, the input data are the frames of the whole parking lots. Therefore, the y-axis of Fig 7 denotes the classification accuracy of the frame test while that of Fig 8 denotes the average conflict rate of neighboring patches in the frame test(in our case, there are 34 patches in a frame, hence, we have total 29 potential conflicts between two neighboring patches) .
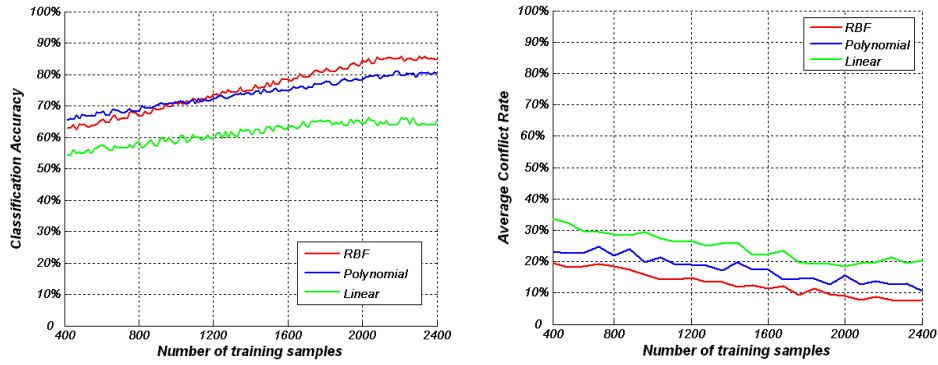

Figure 7: Comparison of The Classification Accuracy in Different Kernel Function of SVM

Figure 8: Comparison of The Average Conflict Rate in Different Kernel Function of SVM

From this two figures, it can be pointed out that the best kernel function for our case is the Gaussian Radius Basis Function(RBF). The classification accuracy is increasing when more samples are used for training. Thus, less conflicts will be occurred. However, it is obvious from the figure that the best result for SVM is 83.57% and we still have at least 7.32% average conflict rate. To solve this problem, Markov Random Field(MRF) is used for reducing conflict. Fig 9 and Fig 10 shows the result after MRF. Obviously,MRF process improve the accuracy significantly (increase around 10% from purly SVM result). Also, average conflict rate is reduced sharply to 2.57% from 7.32% of purely SVM. This improvement can be attributed to the fact that the posterior probabilities of true labels be-

tween two neighboring patches are always close to those of pervious false labels predicted by the SVM, which are easily corrected by the MRF.
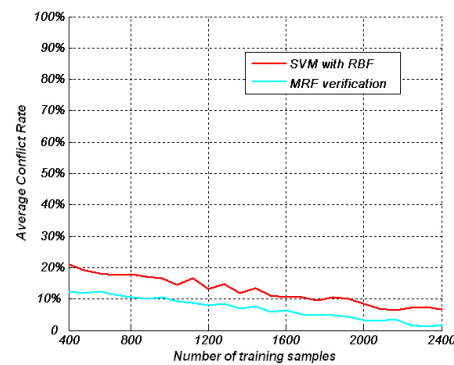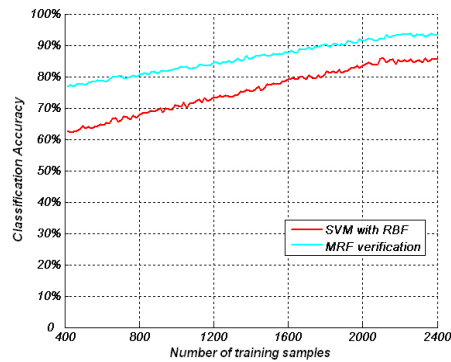

Figure 9: Comparison of The Classification Accuracy before and after MRF verification

Figure 10: Comparison of The Average Conflict Rate before and after MRF verification

8 Conclusion

In this paper, we propose a system for unsupervised parking lots space detection. Comparing with other pervious methods, we just use one camera and take a few frames over seconds. Thus, the workload and computation complexity are much lower than others. In addition, in order to avoid the high hardware requirement for realtime car segmentation, Gaussian Ground Model are presented to obtain the probabilities of ground color. Then, We use 3 parking space, instead of one space or the whole row space, to compose a patch for Multi-class SVM classification. The posterior probabilities which are obtained by SVM classifications are further optimized by applying Markov Random Field to solve the potential conflict between two neighboring patches. This reduces the error rate of SVM classification and improve the recognition accuracy. The experiments results show that our approach is rather robust with high precision.

References

[1] G.L.Foresti, C.Micheloni, and L.Snidaro, "Event Classification for Automatic Visual-based surveillance of Parking Lots" *Proceeding of the 17th International Confrence on Pattern Recognition*, pp. 1051-1081.

[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* 11-14 Oct. 2005 Page(s):271 - 274

[3] Masaki, I., "Machine-vision systems for intelligent transportation systems" *IEEE Conference on Intelligent Transportation System* 13-6 P:24 - 31 Nov. 1998

[4] Y.Boykov, O. Veksler, and R.Zabih, "Efficient Approximate Energy Minimization via Graph Cuts" *IEEE transactions on PAMI*,Vol.20,no.12,p.1222-1239,Nov. 2001

[5] Q. Tao, G.W. Wu, F.Y Wang, and J. Wang, "Posterior probability support vector Machines for unbalanced data" *IEEE transactions on Neural Networks*, Vol.16, Issue 6, P.1561-1573 Nov. 2005