Detecting Shadows of Moving Vehicles Based on HMM

Yuxuan Wang, H. D. Cheng, Juan Shan
Dept. of Computer Science, Utah State University, Logan, UT 84322
yuxuan.wang@aggiemail.usu.edu , hengda.cheng@usu.edu and Juan.Shan@aggiemail. usu.edu

Abstract

Vehicle detection and classification are invaluable in many transportation systems, such as traffic flow analysis, abnormal events detection, automotive driver assistant systems and so on. Many vehicle detection systems have been proposed so far. However, there are still some problems not well solved yet. One of the crucial problems is how to eliminate shadows cast by moving vehicles. In this paper, we propose a novel method based on combination of HMM and background subtraction. A public database of shadow: http://cvrr.ucsd.edu/aton/shadow is employed to test the performance of the new algorithm. Result shows that the proposed method is much better than other old methods.

1. Introduction

Intelligent Transportation Systems (ITS) is very important for traffic planners since it is too difficult to deal with a great number of data gathered from the roads manually. ITS mainly contains traffic flow analysis, events detection and a lot of other automatic jobs to help the planners to design new transportation systems, to modify the old systems, to monitor the current systems and so on. Vehicle detection and classification play a fundamental and important role in all of the intelligent systems [1]. Although a lot of tasks have been done to realize those techniques, they have still not been well solved yet.

This paper deals with the problem—shadows which obstacle the efficiency and robustness of ITS. The difficulty is due to the fact that the shadow always has the same motion as the objects and also is different from the background. Usually the shadows were misclassified as the foreground objects. Shadow will change the shapes of objects; merge two separated objects and even cause an object disappeared if the shadows cast over the objects.

A lot of efforts have been done to deal with shadows. The popular one is based on statistical approach and HSV color space [2]. It is not stable enough due to the different illumination conditions and some vehicles which have the similar colors as shadows will be classified as the shadows. Nowadays, another method based on physical models is proposed [3], and it had showed some advantages over previous algorithms, however, it is too complicated and still not very stable. To create more reliable system, some other properties of shadows should be proposed. [4] suggested a new way to detect shadows based on HMM. Not only considered the intensity of pixels, these two papers also concentrated on temporal information between frames and use this information to build up HMM. However, since the experiments are based on gray-level images and the HMM built in the papers are too coarse, the results are not very good.

In this paper, we will propose an improved HMM based shadow detection algorithm. Color images are employed instead of gray level images. A new way to initialize the parameters of HMM is suggested which is good for HMM to avoid a local maximum problem. Relevant background frame is built up to modify the results of Viterbi decoding algorithm of HMM.

2. Structure of HMM

Since the color distributions among background, foreground and shadow are overlapped with each other, only based on color domain information, it is hard to separate shadows from foreground pixels. However, there is another property that moving information can be employed. A good assumption can be made that once a pixel belongs to a group of background, foreground or shadow at current frame; it is most likely at next frame this pixel still belongs to this group.

Based on this assumption, HMM is good to model this kind of temporal relationship. Three states are created according to three different groups: background, foreground and shadow. They are denoted as B, F, S. Each state can transit to any other state include itself.

Differ from [4], we use color images. Input value

978-1-4244-2175-6/08/$25.00 ©2008 IEEE

of each pixel is a 1*3 vector indicating r, g, and b values. We can model the observations of states S and B as Gaussian distributions and for state F it satisfies the uniform distribution. Thus we can build up a three states, fully connected HMM structure.

To overcome the illumination variation problem, we use Mixture Gaussian Model instead of Single Gaussian Model. That is to say for state B and S, we not only can observe a Gaussian distribution but also a summation of several Gaussians. Each Gaussian Model represents a kind of illumination condition. And different Gaussian will be summed together due to different mixture rate. The formulas are showed as follows:

$$b_j(k) = \sum_{m=1}^{M} c_{jm} G[\mu_{jm}, \Sigma_{jm}, O_k] \quad j = B, S \tag{1}$$

$$G[\mu_{jm}, \Sigma_{jm}, O_k] = \frac{1}{\sqrt{2\pi}\,|\Sigma_{jm}|^{\frac{1}{2}}} \exp[-\frac{1}{2}(O_k - \mu_{jm})^T \Sigma_{jm}^{-1} (O_k - \mu_{jm})]$$

$b_j(k)$ means the probability to observe input vector O_k at state j. c_{jm} is the mixture rate of the number m Gaussian at state j. We have $\sum_{m=1}^{M} c_{jm} = 1$ $c_{jm} \geq 0$. μ_{jm} and Σ_{jm} indicate the mean vector and covariance matrix of number m Gaussian of state j.

The total number of mixture Gaussians is M. In this paper, M=2 stands for normal and dark.

For state F, since it satisfies uniform distribution. We can use formula (2) to calculate $b_F(k)$.

$$b_F(k) = \frac{1}{256^3} \tag{2}$$

3. Training of HMM

3.1. Training Method

Training sequences are collected according to the pixels indicated by the green crosses showed in Fig. 1. These two positions have good characteristic to represent the transition among three states.

The left cross indicates the transition between background and shadow and the right cross indicate the transitions among all three states. These two positions can represent the relationship among those states successfully. For each position we will collect two sequences at different time. And totally 4 training sequences are used.

Fig. 1 Collect training data according to the center of green crosses.

After gather the training data, we employed Baum-Welsh (BW) algorithm [5] to estimate the unknown parameters of HMM.

3.2. HMM Initialization

We know that BW algorithm is a special case of EM algorithm. EM algorithm has a problem so called local maximum optimization. EM algorithm can only generate local maximum solution instead of global solution. If we set wrong initial value to HMM, it may not produce the result as we expected. Thus the initialization of HMM is very important.

In this paper we use Maximum Likelihood Estimation (MLE) method to initialize the Gaussian Observation Model of HMM which is a good way to avoid local maximum.

For background and shadow which satisfy Gaussian distribution, we manually collect some pixel data according to each group. For each mixture Gaussian Model, a few data will be collected. And based on these data, we use MLE to get the mean value and covariance matrix for each single Gaussian Model.

$$\mu_{jm} = \frac{1}{N_{jm}} \sum_{k=1}^{N_{jm}} O_k \tag{3}$$

$$\Sigma_{jm} = \frac{1}{N_{jm}} \sum_{k=1}^{N_{jm}} (O_k - \mu_{jm})(O_k - \mu_{jm})' \tag{4}$$

N_{jm} represents the number of pixels collected for the mth Gaussian at state j. To initialize transition probability a_{ij}, we sign a big value to a_{ij} if i=j, otherwise we just sign a small value. For mixture rate, we initialize the same value to different Gaussian Model which is 0.5.

By using these mean vector, covariance matrix and transition probability to initialize HMM; we may avoid the local maximum and get a global solution.

4. Testing Strategy

4.1. Significant Zone

For all testing images, we will not pass all pixels into HMM but create a significant zone showed in cyan in Fig. 2. This zone can be seen in many tracking and detection system. The advantage of introducing this zone is that first not all parts of the image are important, only remain significant part can reduce the time complexity; second introduce this zone can overcome some overlapping problems.

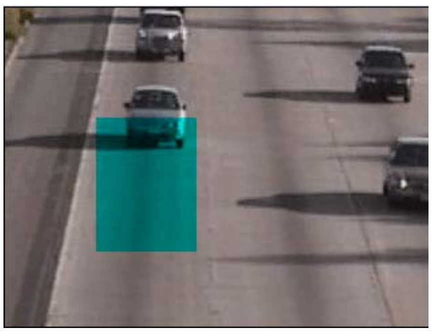

Fig. 2 Significant Zone

Each pixel in the zone will correspond to a trained HMM. This strategy maintains the resolution of image and dose not change the shape of the objects. Even it needs more calculations, it still acceptable.

4.2. Decoding

Traditional Viterbi algorithm [5] can be used to decode the optimum sequence of states according to input sequence of observations. We found that using Viterbi individually will cost some background pixels to be recognized as foreground pixels.

Based on [6] we can get a suitable background image for each input frame using adaptive method. We combine this background information with Viterbi algorithm to obtain the decoding results.

Let $\delta_t(i)$ indicates the optimum probability that the state is i at frame t.

$$\delta_{t+1}(j) = \{\max_i \quad \delta_t(i) a_{ij}\} b_j(O_{t+1}) \qquad (5)$$

Eq. (5) means if we know the probability of each state at frame t we can get value $\delta_{t+1}(j)$ for state j at frame t+1 based on this formula. This is the tradition Viterbi algorithm.

Now we introduce a new definition of $Backvalid(t)$ which verifies that at frame t the pixel is similar with reference background image. If it is similar enough (in each dimension of r, g and b the difference is less than 5), then $Backvalid(t) = 1$ otherwise, $Backvalid(t) = 0$.

Then we test whether the difference between $\delta_t(B)$ and $\delta_t(F)$ are significant. If so, we just follow Eq. (5). Otherwise, we test $Backvalid(t)$, if it is 1, then we use formula (6) to replace (9) in frame t+1.

$$\delta_{t+1}(j) = \delta_t(B) a_{Bj} b_j(O_{t+1}) \qquad (6)$$

Using this combined algorithm, we can reduce the misclassification from background to foreground significantly.

4.3. Result Enhancement

After applying Viterbi algorithm, we compose the states information of each pixel together to build a direct result from HMM. However, there are still some noises. To solve this problem, we first employ median filter to eliminate small noise point and then using area threshold upon connected component to remove larger noise blob. Finally perform the morphological operation to fill the hole in object. Fig. 3 shows an example of such enhancement.

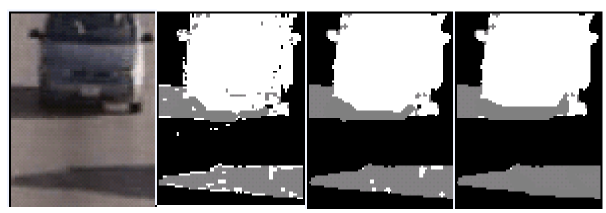

Fig. 3 Enhancement result. Left to Right: Original image, after passing through HMM, after passing through filters, after morphological operations.

5. Experiments

346 frames are collected from a public data set for shadow detection: http://cvrr.ucsd.edu/aton/shadow.

To show the result more clearly, we mark the significant zone with three kinds of colors. Blue represents the background. Green is for shadow. And red stands for foreground object. From Fig. 4 we can find that shadows and foreground moving vehicle are successfully separated from each other. The shape of the vehicle is kept without changing too much which shows a very good result.

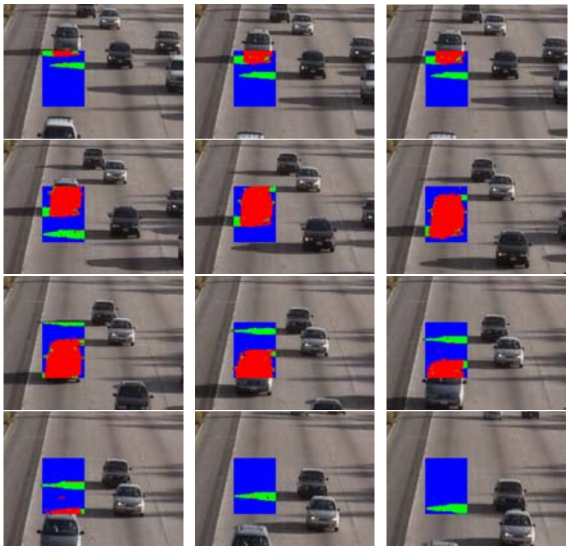

Fig. 4 A sequence of 12 continues frames result: red->vehicle, green->shadow.

In this experiment, the size of significant zone is 100*75. And we use pixel based algorithm which means 7500 HMMs will be assigned to each individual pixel. However, to generate the final result, each frame only takes less than 0.4 second which is a reasonable time cost. For training part, 4 sequences, each of which is 300 in length, will be trained. And the time cost by training is only less than 30 seconds.

To quantify the result of our algorithm, we introduce two metrics [7]: the shadow detection rate η and the shadow discrimination rate ξ.

$$\eta = \frac{TP_{shadow}}{TP_{shadow} + FN_{shadow}} \qquad (7)$$

$$\xi = \frac{\overline{TP}_{foreground}}{TP_{foreground} + FN_{foreground}} \qquad (8)$$

The $\overline{TP}_{foreground}$ is the number of ground-truth points of the foreground objects minus the number of points detected as shadows, but belonging to foreground objects. All the two values are the larger the better.

Table 1

Method	η	ξ
SNP	81.59%	63.76%
SP	59.59%	84.70%
DNM1	69.72%	76.93%
DNM2	75.49%	62.38%
OURS	92.1%	89.2%

Table1 shows the result of our proposed method compared with other 4 methods mentioned in the survey paper [7]. Result shows that our method is higher in both shadow detection rate and shadow discrimination rate than any other method.

6. Conclusion

In this paper, we proposed a novel method to detect shadows. A mixture Gaussian HMM is built to model three kinds of groups: background, shadow and foreground. Background validation is employed to justify Viterbi algorithm to decode the state sequences which are corresponding to three groups. Several image processing methods are used to enhance the final result. Experiments indicate that our proposed method is superior to other 4 existing algorithms. For further research, some other color domain can be considered as the input of HMM. Such as HSV which has been proved suitable for shadow detection. And also edge information can be employed as another feature of the shadow, since the shadow is smoother than foreground object. Also some potential properties of shadows will be further investigated to make systems more efficient and robust.

References

[1] S. Gupte, O. Masoud, R. F. K. Martin and N. P. Papanikolopoulos. "Detection and classification of vehicles", IEEE Trans. Intelligent Transportation Systems, pp. 37–47, 2002.

[2] R. Cucchiara, C. Grana, M. Piccardi, A. Prati, and S. Sirotti, "Improving Shadow Suppression in Moving Object Detection with HSV Color Information," Proc. IEEE Intelligent Transportation Systems Conf., pp. 334-339, Aug. 2001.

[3] S. Nadimi and B. Bhanu, "Physical Models for Moving Shadow and Object Detection in Video", IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 26, pp. 1079-1087, 2004.

[4] J. Kato, T. Watanabe, S. Joga, J. Rittscher, and A. Blake, "An HMM-based segmentation method for tracking monitoring movies", IEEE Trans. Pattern Anal. Machine Intell., vol. 24, pp. 1291–1296, Sept. 2002.

[5] LR Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition", Proceedings of the IEEE, 1989.

[6] Yuxuan Wang, H. D. Cheng and Chris Glazier, "An Improved Approach for Vehicle Detection and Classification", Joint Conference on Information Science, 2007.

[7] A. Prati, I. Mikic, M. Trivedi, and R. Cucchiara, "Detecting moving shadows: algorithms and evaluation," IEEE Transactions on Pattern Analysis and Machine Intelligence 25, pp. 918~923, July 2003.