Nonparametric Background Modeling Using The CONDENSATION Algorithm

Xingzhi Luo, Suchendra M. Bhandarkar
University of Georgia, Athens, GA, USA
xingzhi@cs.uga.edu, suchi@cs.uga.edu

Wei Hua, Haisong Gu
Vident Systems Inc., Sunnyvale, CA, USA
wei@vident.com, haisonggu@vident.com

Abstract

Background modeling for dynamic scenes is an important problem in the context of real time video surveillance systems. Several nonparametric background models have been proposed to model dynamic scenes and promising results have been reported. However, a critical problem with existing nonparametric models is their high computational requirement because a large set of background samples is usually needed to model the background. In this paper, a nonparametric background model that uses an importance sampling method is proposed to overcome the problem of high computational complexity of conventional nonparametric background models. Instead of using a large number of samples to model the background probability densities, much fewer background samples are maintained and updated using the CONDENSATION algorithm. A Markov Random Field model is used to enhance the foreground detection results by imposing spatial constraints. Experimental results show that the proposed method is much faster and computationally more efficient than existing nonparametric background models. The proposed technique is observed to match the capabilities of existing nonparametric background models in terms of being able to effectively model dynamic backgrounds but with greatly reduced computational complexity.

1 Background and Previous Work

Background modeling of dynamic scenes is an important issue in automated video-based surveillance systems. In typical dynamic outdoor scenes, events such as a combination of random noise, trees swaying in the breeze, waves on water surfaces and camera jitter result in a dynamically changing background where multiple background colors can be observed at some pixel locations. Another common phenomenon in dynamic scenes is that the color observed at a pixel location could move to neighboring locations. Background models that use a single color value or single Gaussian distribution [7, 10] at each pixel location are no longer viable for modeling dynamic scenes. Stauffer and Grimson [9] have proposed an adaptive on-line parametric color model in which the background color of each pixel is modeled as a multiple Gaussian mixture (MGM). Computationally more efficient versions of the MGM background model have also been proposed in the literature [1, 3, 5]. The pixel-based MGM background model, though computationally efficient, is limited in its capability by the upper bound placed on the number of Gaussian distributions for each pixel. In the nonparametric model proposed by Elgammal *et al.* [2] the background pixel samples are directly used to represent the background color distribution. Sheikh and Shah [8] improve the above nonparametric model by integrating the color and location information derived from the background samples, thus exploiting both temporal and spatial information. The further incorporation of a foreground model and a Markov Random Field (MRF) to impose spatial constraints on the background and foreground, have been observed to reduce the effects of random noise [11].

There are certain key limitations to the conventional nonparametric background models [2, 8]. In a typical nonparametric background model, background samples $Y = \{y_i\}, i \in [1, N]$ and foreground samples $Z = \{z_j\}, j \in [1, M]$ are maintained to compute the background and foreground distributions using a kernel function (typically a Gaussian function). Each pixel sample is encoded using its color and location, that is, $y_i = (x, y, r, g, b)$. The kernel-based background probability density is given by $p(x|B) = \frac{1}{N}\sum_{i=1}^{N}\varphi_H(x - y_i)$ and the foreground probability is given by $p(x|F) = \alpha c + (1-\alpha)\frac{1}{M}\sum_{j=1}^{M}\varphi_H(x - z_j)$, where c is a constant (i.e., a uniform distribution), φ_H is the kernel function and α is a parameter which controls the relative con-

Proceedings of the IEEE International Conference
on Video and Signal Based Surveillance (AVSS'06)
0-7695-2688-8/06 $20.00 © 2006 IEEE

tributions of the uniform distribution and the kernel function-based distribution.

The first limitation of the above nonparametric dynamic background model is its high computational complexity. Suppose the background sample set uses all the pixels from K most recent images as the background samples where K usually ranges from $100 - 1000$. If the image size is $W \cdot H$, then the total number of samples for the background is $N = K \cdot W \cdot H$. If $K = 1000$, $W = 320$, $H = 240$, then the total number of samples in the background set is $N = 7.68 \times 10^7$. The total number of pixels in an image of size $W = 320$ and $H = 240$ is given by $p = W \cdot H = 7.68 \times 10^4$. Thus, the total time taken to compute the background probability for all the pixels in the image is given by $pN \times \text{Time}(\varphi_H(x)) = 5.9 \times 10^{12} \times \text{Time}(\varphi_H(x))$ where $\text{Time}(\varphi_H(x))$ is the time taken to evaluate the kernel function. This time is simply much too prohibitive for real-time video surveillance systems at current CPU speeds.

The second limitation of the above nonparametric dynamic background model is its reliance on a foreground model, which is usually not realistic for most real-world systems. In a real-world system, there are usually insufficient foreground samples available for offline training.

The third limitation of typical non-parametric background models is that when a foreground object passes an area which has similar color as the foreground object, the foreground object cannot be detected. This is termed as the *camouflage problem* in [10]. Taking the pixel location (x, y) into consideration in the background model only serves to worsen the problem, because if a foreground object is in the spatial proximity of a background area which has a similar color as the foreground object, the foreground object will most likely go undetected.

In this paper we primarily address the problem of high computational complexity of conventional nonparametric background models [2, 8]. A simple solution would be to reduce the number of background samples. However, a simple reduction in the size of the background sample set would result in a loss of modeling accuracy. We propose to use an importance sampling technique to generate a small subset of representative background samples from the very large set of original background samples. The much smaller subset of representative background samples could then be used to represent the background distribution without serious loss of accuracy. The CONDENSATION algorithm [4] is a simple but efficient sampling method, which could be exploited for this purpose.

To address the lack of readily available foreground data to formulate a foreground model, we use a uniform distribution for the foreground model as is done in [2, 8, 11]. To address the camouflage problem, we limit the bandwidth of the pixel samples in the spatial dimensions (x, y) such that the probability computation at a given pixel location is unaffected by spatially distant pixels. The spatial bandwidth is controlled by the kernel function. For a diagonal Gaussian kernel function $G(\sigma) = G(\sigma_x, \sigma_y, \sigma_r, \sigma_g, \sigma_b)$, the σ_x and σ_y values control the bandwidth of the kernel function limiting its influence within a certain area [3].

2 Review of the CONDENSATION Algorithm

We use an importance sampling technique based on the CONDENSATION algorithm [4] to maintain a smaller representative sample set for the background model. The CONDENSATION algorithm has three basic components: population sampling, sample prediction and sample measurement. The sample set is represented as $\{(z_i^{t-1}, w_i^{t-1}), i = 1, ..., N\}$, where N is the total number of samples and w_i^{t-1} is the weight of sample z_i^{t-1} at time $t-1$. A new variable $c_i^{t-1} = \sum_{k=1}^{i} w_k^{t-1}$ is computed to facilitate the sampling process. During the sampling process, N random numbers $r \in [0, 1]$ are generated. If $r \in [c_{i-1}^{t-1}, c_i^{t-1})$, then sample z_i^{t-1} is chosen. A binary search algorithm is used to determine the value of i, so that the computational complexity of the sampling procedure is $O(N \log N)$. In the prediction process, the sample z_i^t is predicted using a predefined proposal distribution $p(z_i^t | z_i^{t-1})$. In the measurement process, the probability $w_i^t \sim p(I | z_i^t)$ is computed based on the actual model where I is the observed image data. The CONDENSATION algorithm iteratively performs the above three processes to continuously update the samples.

The CONDENSATION algorithm provides a simple yet flexible means to perform importance sampling that can be extended to many problems. In our case, the samples are simply the background samples that combine both the color and spatial location information. The background samples are used to compute the background probability distribution via the kernel density function. The observed images are used to measure the probability distribution $w_i^t \sim p(I | z_i^t)$. The prediction model (proposal distribution) can also be computed using the measured accumulated errors.

Proceedings of the IEEE International Conference
on Video and Signal Based Surveillance (AVSS'06)
0-7695-2688-8/06 $20.00 © 2006 IEEE

3 The Proposed Nonparametric Background Model

The proposed background model is illustrated in Figure 1. A set of background samples $\{(s_i^{t-1}, w_i^{t-1}), i = 1, ..., M\}$ is maintained to compute the background probability distribution using a kernel function. The value of s_i^{t-1} encodes the color and location information of sample i whereas w_i^{t-1} is the weight of sample i at time $t-1$. The weights of the samples are adapted based on the observation(s) I, that is, $w_i^{t-1} \sim p(I|s_i^{t-1})$. The computed background probability, in conjunction with the MRF model, is used to compute the foreground mask. The pixels classified as the foreground are also used as candidates for the background samples. The CONDENSATION algorithm is used to resample the background samples such that the background samples are adapted to changes in the underlying dynamic scene. Note that the probability of a background sample to be selected by the importance sampling procedure is proportional to the weight associated with the sample. Therefore, the speed at which the model adapts to changes in the background is tunable with the proposed weight updating scheme. The proposed system is a negative feedback system since the output is fed back to decrease the intensity of the input signal, thus resulting in a stable system.

The background probability for a given observation z^t is given by $p(z^t|B) = \frac{1}{M}\sum_{s_i^{t-1} \in o(z^t)} \varphi_H(z^t - s_i^{t-1})$, which is identical to the expression for the background probability in the nonparametric background model described in [8]. A diagonal Gaussian kernel function is used in our model where $o(z)$ is the set of background samples located in the neighbourhood of the observation z. In order to efficiently compute $o(z)$, all the samples are indexed using their locations (x, y). Thus, it is not necessary to enumerate all the background samples in order to compute $o(z)$. In our current scheme, $o(z) = \{(x, y, r, g, b);\ |x - z_x| < w,\ |y - z_y| < w,\ |r - z_r| < d_c,\ |g - z_g| < d_c$ and $|b - z_b| < d_c\}$ where w is a predefined window size and d_c is a preset color threshold. Note that w is directly related to the kernel function bandwidth along the dimensions x and y. In our case, we use $w = 3\sigma$, where $\sigma = \sigma_x = \sigma_y$. By indexing the samples using their locations (x, y), the computation of $o(z)$ is limited to the samples located in a window $\{(x, y); |x - z_x| < w, |y - z_y| < w\}$ within the index table.

For an adaptive background model, the learning rate is a critical performance parameter. Too large a learning rate will result in the background model adapting rapidly to a temporarily static foreground object causing the foreground object detection algorithm to generate false negatives. Too small a learning rate will make the background model insensitive to gradual changes in the scene background over time, causing a foreground object detection algorithm to generate false positives. In a nonparametric background the learning rate is proportional to $1/K$ where K is the number of most recent image frames that are used to generate the background samples [8]. An empirically determined reasonable learning rate is $0.01 \sim 0.001$ for most dynamic scenes, depending on the speeds of the moving foreground objects and rate of change of the background. The faster the (average) speed of the foreground objects, the higher the learning rate needs to be and vice versa. In our case, the learning rate is 0.001, that is, the background samples are obtained from the 1000 most recent image frames when computing the background probability.

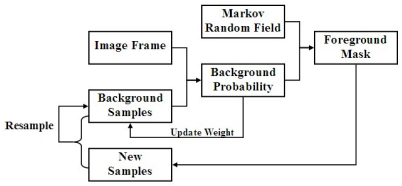

Figure 1. Background Model

The number of background samples is M. The background samples are directly initialized using the first $K = \frac{M}{W \cdot H}$ frames. Each sample is assigned an initial weight $w_i^0 = 1; i = 1, \ldots, M$. It is preferable to use clean background images (without moving foreground objects) to initialize the background samples. However, since the model is adaptive, it works correctly even if moving foreground objects are present in the initial frames. The weights w_i^{t-1}'s are computed using the observed images, and reflect the probability of the samples $w_n^t \sim p(I|z_n^t)$.

For each pixel in a new frame, its weight is chosen to be αK. Thus, the total weight of a new image is $\alpha K \cdot W \cdot H$ and the total weight of the sample set is $K * W * H$. The learning rate is approximately α which, as stated previously, is chosen to be 0.001. The proposal distribution is computed by accumulating the error between the observations and the samples. Given the average accumulated error ϵ and the learning rate α, the proposal distribution is given by $p(s^t|s^{t-1}) = \mathcal{N}((1-\alpha)s^{t-1} + \alpha\epsilon, R)$, where $\mathcal{N}(\mu, R)$ represents Gaussian white noise with mean μ and covariance R. The algorithm used to update the background samples and compute the foreground mask is given below:

(1) Initially, all the pixels of the first K frames are

Proceedings of the IEEE International Conference
on Video and Signal Based Surveillance (AVSS'06)
0-7695-2688-8/06 $20.00 © 2006 IEEE

directly used to constitute the sample set $S = \{s_i^0, w_i^0\}$ *where* $K = 5 \sim 10$ *in our experiments.*

(2) Create a 2-D index table for all the samples based on their locations (x, y) *in the image frame. For each sample, the weight* w_i^0 *is set to* 1 *and the accumulated error* ϵ_i *is set to* 0*;*

(3) Each observed pixel $z^t = (x, y, r, g, b)$ *in a new image, is assigned a weight* $w^0 = \alpha K$*. Its background probability is computed as* $p(z^t|B) = \frac{1}{M}\sum_{s_i^{t-1} \in o(z^t)} \varphi_H(z^t - s_i^{t-1})$*. If* $o(z^t)$ *is empty, then* $p(z^t|B) = p^0$*, where* p^0 *is a preset small probability value.*

(4) If $o(z^t)$ *is not empty, then* s_i^{t-1} *is determined such that* $s_i^{t-1} \in o(z^t)$ *and* $\varphi_H(s_i^{t-1} - z^t)$ *is a minimum. The weight of* s_i^{t-1} *is updated by addition of* $w^0 = \alpha K$ *i.e.,* $w_i^{t-1} = w_i^{t-1} + w^0$*. The accumulated error* ϵ_i *for* s_i^{t-1} *is updated as* $\epsilon_i = \epsilon_i + (z^t - s_i^{t-1})$*.*

(5) The computed background probability $p(z^t|B)$ *for every pixel in the new image, the simple foreground model* $p(z^t|F) = c^0$ *and the MRF model are used to obtain the foreground mask as the output.*

(6) If a pixel is classified as a foreground pixel, it is assigned a weight $w^0 = \alpha K$ *and inserted in the set of new background samples.*

(7) Repeat steps (3)–(6) for T *frames.*

(8) For once every T *frames, each sample is updated as* $s_i^t = \frac{s_i^{t-1} + (w_i^{t-1} - 1)\frac{\epsilon_i}{T}}{w_i^{t-1}} + \mathcal{N}(R)$*. where* $\mathcal{N}(R)$ *is zero mean Gaussian white noise with variance* R*.* M *new background samples are generated from the updated* M *samples and the new background samples based on their updated weights.*

(9) Repeat steps (2)–(8).

In step 3, a very small probability value p^0 is used instead of 0, because a value of 0 causes problems in the computation of the likelihood. The accumulated error is used to track gradual changes in the background samples. The accumulated error is set to 0 in step 2. In step 4, the error between the samples and the observations is accumulated. Step 4 thus defines the measurement process, in which the weight w_i^{t-1} of the ith sample is continuously updated as $w_i^{t-1} = w_i^{t-1} + w^0$ if the observed color is close to this sample. Therefore $w_i^{t-1} \sim p(I|s_i^{t-1})$. In step 8, the accumulated error is used to update the samples. Since the weight of the original sample is 1; the accumulated error is $w_i - 1$. The average accumulated error is ϵ_i/T, hence the sample value is adjusted as $s_i^t = (s_i^{t-1} + (w_i^{t-1} - 1)\epsilon_i/T)/w_i^{t-1}$. Thus step 8 also constitutes the prediction process in the CONDENSATION algorithm.

The sampling process is performed for every $T = 30$ frames instead of for every frame since, at the given learning rate of 0.001, the contribution of each successive frame to the sample set is not significant. In step 8, the sampling process is specially designed to perform faster sampling. In the weight-based sampling scheme, a new variable is computed as $c_i = \sum_{j=1}^{i} w_j$. In order to generate M new samples, M random numbers $r_i \in [0, c_G)$ are generated, where $G = M+$ the number of samples in the new sample set. For each value of r_i, a value of k is determined such that if $c_{k-1} < r_i \leq c_k$, then s_k is copied into the new background sample set. In order to determine k, many sampling algorithms use a binary search method which takes $O(M \log G) \approx O(M \log M)$ time since $G \approx M$. However, in our case, computation time that close to $O(M)$ can be achieved. In the proposed algorithm, the weights of most of samples are close to 1. Therefore, the difference between c_k and c_{k-1} is approximately 1. An index table I is created to facilitate faster sampling such that $I[n] = k$, where k is the smallest integer which satisfies the condition $s_k > n$. Given a random number r_i, a sequential search can start from the $I[\lfloor r_i \rfloor]^{th}$ sample. It only takes one or two steps to find the value of k, such that $c_{k-1} < r_i \leq c_k$. Therefore, the proposed algorithm is computationally much more efficient. The proposal distribution in step 8 is computed as $p(s^t|s^{t-1}) = \mathcal{N}(\frac{s_i^{t-1} + (w_i^{t-1} - 1)\frac{\epsilon_i}{T}}{w_i^{t-1}}, R)$.

When the background probability $p(z|B)$ and foreground probability $p(z|F)$ are known, the logarithm (log) of the posterior of the foreground labels is given by:

$$\log(p(L|Z)) \propto \sum_{i=1}^{p} l_i \log \frac{p((x, y, r, g, b)|F)}{p((x, y, r, g, b)|B)} + \lambda \sum_{i=1}^{p} \sum_{j \in o(i)} (l_i l_j + (1 - l_i)(1 - l_j)) \quad (1)$$

where Z is the observation, L is the set of foreground/background labels for the entire image, p is the total number of pixels in the image, and l_i is the label of the i^{th} pixel such that $l_i = 0$ if pixel i is classified as background and $l_i = 1$ if pixel i is classified as foreground. The first term in the right hand side of equation (1) indicates that if the log-likelihood is larger than 0, then $l_i = 1$ will make the posterior probability larger just as in the common threshold-based classification method. The second term in the right hand side of equation (1) imposes spatial constraints on the labels. When the label at a pixel location is the same as the labels of its neighboring pixels, the log of the posterior probability is larger. As mentioned in section 1, $p(z|F)$ is a uniform distribution. Equation (1) is derived in a

Proceedings of the IEEE International Conference
on Video and Signal Based Surveillance (AVSS'06)
0-7695-2688-8/06 $20.00 © 2006 IEEE

manner similar to the expression for the log-posterior probability presented in [8].

The optimization problem defined by equation (1) is solved using the maximum cut/minimum flow graph algorithm described in [6] which guarantees a globally optimum solution. The source code for maximum cut/minimum flow graph algorithm is available at http://www.cs.cornell.edu/People/vnk/software.html. Note that iterative techniques to solve the above optimization problem, such as Gibbs sampling and simulated annealing used in [11], are computationally more intensive and do not guarantee a global optimum.

(a) (b) (c)

Figure 2. Trees swaying in the breeze. (a) The original image. (b) Foreground object detection using pixel-based multiple Gaussian mixture model. (c) Foreground object detection using the proposed nonparametric background model.

4 Experimental Results

Since the original nonparametric background model in [8] is too slow to run on our system (20 seconds per frame with only 100 frames of background samples), no results are available for comparison. Instead, the well known pixel-based MGM background model described in [9] is used for the purpose of comparison with our proposed method. Three prototypical video clips are used to compare the proposed nonparametric dynamic background modeling algorithm with the pixel-based MGM background modeling algorithm. The first video contains trees swaying in the breeze. The second video shows undulating waves on the surface of a lake. The third video contains strong camera motion. Videos 2 and 3 are provided by a commercial unit. Video 1 has been shot by us. The results of the proposed nonparametric dynamic background modeling algorithm on Video 1 are available at http://www.cs.uga.edu/~xingzhi/avss.zip.

The proposed nonparametric dynamic background modeling program runs on a Pentium IV desktop with a 3.0 GHz CPU and 1 GB of memory (RAM). The video frames are 320×240 pixels or 360×240 pixels in size. The learning rate of the pixel-based MGM background model is set to 1/1000. When the MRF model is not used, the pixel-based MGM background model takes 95 milliseconds to process a single frame, whereas the proposed nonparametric background model takes 100 milliseconds to process a single frame. The introduction of the MRF model adds a processing overhead of 132 milliseconds per frame which is the same for both, the MGM-based background model and the proposed nonparametric background model. A reason to consider the time taken by the MRF model separately from that taken by the dynamic background model is that the current MRF model is implemented using source code provided to us by other researchers [6]. Also, the current implementation of the proposed nonparametric dynamic background model uses a precomputed lookup table instead of computing the Gaussian kernel function directly in the interest of improving the run time performance. There is also substantial room to speed up the current MRF model implementation since most of computation therein can be performed offline. Thus, the proposed nonparametric background model is comparable to the pixel-based MGM background model in terms of computational efficiency. The computational efficiency of the proposed nonparametric background model stems from the incorporation of the CONDENSATION algorithm in the modeling scheme.

The current evaluation of the accuracy of the two background modeling schemes is done by visual examination of their outputs. Video 1 (shot by us) contains trees swaying in the breeze, a fountain of water and undulating waves on the surface of a lake. Results of foreground object detection using the proposed nonparametric background model show that the proposed nonparametric background model is able to suppress the effect of the motion of the trees to a much greater extent than the pixel-based MGM background model (Figure 2). In the provided demo videos, it can be observed that the proposed nonparametric background model can also detect small moving foreground objects (such as ducks swimming in the lake) and filter out most of the random noise. However when the breeze is too strong and not constant over time at some pixel location, not all the random noise can be filtered out. This is a common problem for most statistical background models that rely on the stationarity assumption about the random noise.

Video 2 (Figure 3) shows undulating waves on the surface of a lake and two small boats moving on the lake surface. Both the pixel-based MGM background model and the proposed nonparametric background model were able to detect the small moving boats. This indicates that both background models are sensitive to small foreground objects. Video 3 (Figure 4)

Proceedings of the IEEE International Conference
on Video and Signal Based Surveillance (AVSS'06)
0-7695-2688-8/06 $20.00 © 2006 IEEE

(a) (b) (c)

Figure 3. Waves on a lake. (a) The original image. (b) Foreground object detection using a pixel-based multiple Gaussian mixture model. (c) Foreground object detection using the proposed nonparametric background model.

(a) (b) (c)

Figure 4. Shaking camera. (a) The original image. (b) Foreground object detection using pixel-based multiple Gaussian mixture model. (c) Foreground object detection using proposed nonparametric background model.

taken with a shaking camera shows that the proposed nonparametric background model exhibits better performance than the parametric pixel-based MGM background model in the presence of camera jitter.

5 Summary and Future Work

In this paper, a novel nonparametric dynamic background model using the CONDENSATION algorithm is presented. To the best of our knowledge, this is the first nonparametric dynamic background model that incorporates the importance sampling technique commonly employed in particle filters. The proposed background model maintains very few background samples to represent the background probability using kernel-based density functions. The computational complexity of the proposed nonparametric dynamic background model is comparable to that of the popular parametric pixel-based MGM background model but considerably lower than that of existing nonparametric dynamic background models described in the literature [8]. The learning rate of our proposed background model is also adaptable to suit different applications. The proposal distribution in the proposed model enables the samples to adapt to gradual illumination changes. Experiments on real videos show that the proposed nonparametric background model yields much better results than the pixel-based MGM background model with little increase in computational overhead. The proposed nonparametric background model is well suited for use in real-time video surveillance systems.

References

[1] S.M. Bhandarkar and X. Luo, A Robust and Fast Background Updating for Real-time Surveillance Systems, *Proc. IEEE Intl. Workshop on Machine Vision for Intelligent Vehicles*, San Diego, CA, June 2005.

[2] A. Elgammal, R. Duraiswami, D. Harwood and L.S. Davis, Background and Foreground Modeling Using Nonparametric Kernel Density Estimation for Visual Surveillance, *Proc. IEEE*, Vol. 90, No. 7, July 2002.

[3] D. Butler, S. Sridharan and V.M. Bove, Jr., Real-time Adaptive Background Segmentation, *Proc. IEEE ICME*, Baltimore, MD, July 2003.

[4] M. Isard and A. Blake, CONDENSATION – Conditional density propagation for visual tracking, *Intl. Jour. Computer Vision*, 1998, Vol. 29, No. 1, pp. 5-28.

[5] P. KaewTraKulPong and R. Bowden, An improved adaptive background mixture model for real-time tracking with shadow detection, *Proc. Wkshp. Adv. Vision-based Surveillance Sys.*, Kingston, UK, Sept. 2001.

[6] V. Kolmogorov and R. Zabih, What Energy Functions Can Be Minimized via Graph Cuts, *IEEE. Trans. PAMI*, Vol. 26, No. 2, 2004, pp. 147-159.

[7] S.J. McKenna, S. Jabri, Z. Duric, A. Rosenfield and H. Wechsler, Tracking Groups of People, *CVIU*, Vol. 80, 2000, pp. 42-56.

[8] Y. Sheikh and M. Shah, Bayesian Object Detection in Dynamic Scenes, *Proc. IEEE Conf. CVPR*, San Diego, CA, 2005.

[9] C. Stauffer and W.E.L. Grimson, Adaptive Background Mixture Models for Real-time Tracking. *Proc. IEEE Conf. CVPR*, Ft. Collins, CO, June 1999, pp. 246-252.

[10] K. Tooyama, J. Krumm, B. Brumit and B. Meyers, Wallflower: Principles and Practice of Background Maintenance. *Proc. ICCV*, Corfu, Greece, Sept. 1999, pp. 255-261.

[11] Y. Zhou, W. Xu, H. Tao and Y. Gong, Background Segmentation using spatio-temporal multi-resolution MRF, *Proc. MOTION05*, Colorado, USA, 2005.

Proceedings of the IEEE International Conference
on Video and Signal Based Surveillance (AVSS'06)
0-7695-2688-8/06 $20.00 © 2006 IEEE