Applied Mathematics and Computation, Vol. 205, No. 2, 2008, pp 899-907

A New Image Thresholding Method Based on Gaussian Mixture Model*

Zhi-Kai Huang[1,3], Kwok-Wing Chau [2]

[1.] *Intelligent Computing Lab, Hefei Institute of Intelligent Machines, Chinese Academy of Sciences, P.O.Box 1130, Hefei, Anhui 230031, China*

[2.] *Department of Civil & Structural Engineering, The Hong Kong Polytechnic University，Hong Kong，China*

[3.] *Department of Machinery and Dynamic Engineering Nanchang Institute of Technology, Nanchang, Jiangxi 330099, China*

E-mail address: *huangzk@iim.ac.cn*

ABSTRACT

Abstract: In this paper, an efficient approach to search for the global threshold of image using Gaussian mixture model is proposed. Firstly, a gray-level histogram of an image is represented as a function of the frequencies of gray-level. Then，to fit the Gaussian mixtures to the histogram of image, the Expectation Maximization (EM) algorithm is developed to estimate the number of Gaussian mixture of such histograms and their corresponding parameterization. Finally, the optimal threshold which is the average of these Gaussian mixture means is chosen. And the experimental results show that the new algorithm performs better.

Keywords: Histogram, Optimization, Thresholding.

1. Introduction

Segmentation of images into homogeneous regions is an important ongoing research of computer vision. Image thresholding， one of the most important techniques for image segmentation, which is defined as partitioning an image into homogeneous regions, is regarded as an analytic image representation method [1] and plays a very important role in many tasks of pattern recognition, computer vision, and image and video retrieval [2]. Previous work on automatic threshold selection has been described in a number of literatures. For example, Otsu et al [3] described a method that maximizes the between-class variance. The method is a global thresholding method, but it gave poor results when the background of image appears to be darker. Prewitt and Mendelson [4] suggested selecting the threshold at the valleys of the histogram, while Doyle [5] suggested the choice of the median of the histogram. Ridler and Calvard [6] proposed an iterative method of thresholding (also see [7]). Tsai [8] suggested selecting a threshold at which the resulting binary images have the same first three moments. All these traditional algorithms considered a fixed threshold value according to gray-level histogram and cannot process the images whose histograms are nearly unimodal, especially when the target region is much smaller and low-contrasted to the background area. Thresholding methods based on entropy functions can be found in literature [9-10], but entropy function methods do not always give a good solution. Sometimes, results obtained by the entropic thresholding methods

* This work was supported by the grants of the National Science Foundation of China, Nos. 60772130 and 60705007, the grant from the National Basic Research Program of China (973 Program), No.2007CB311002, the grants from the National High Technology Research and Development Program of China (863 Program), No. 2007AA01Z167, the grant of the Guide Project of Innovative Base of Chinese Academy of Sciences (CAS), No.KSCX1-YW-R-30.

are found to be biased. There are some unsupervised training neural systems that can be used for the segmentation of images. Given some ground-truth exemplars they could learn the general features used in segmentation. They also could learn how to classify image attributes according to their similarity without predefined categories. In addition, some neural network models or neural computation methods can be adopted for addressing this problem [2，11-14].

An effective and adaptive approach to background subtraction is to construct a statistical model which represents the probabilistic distribution of the pixel's intensity or color. Wren et al. adopted a single Gaussian to represent the background model [15]. However, this system is sensitive to the initialization, and is improper to process multi-modal and clutter scenes. Another statistical model for background subtraction is the finite Gaussian mixture model (GMM) [16-18]. Friedman and Russell [15] used a mixture of three Gaussian distributions to model the pixel value for traffic surveillance applications. Stauffer et al [17, 18] proposed a similar algorithm, which used a mixture of Gaussian distribution to model a multi-modal background. However, all these methods have a drawback that the number of the mixture components is a pre-set and fixed value. Because the number of the mixture components mostly determines the number of the need-estimating parameters, this drawback may make foreground segmentation time-consuming.

This paper presents an efficient method for image segmentation based on GMM. A gray-level histogram of an image can be represented as a function of the frequencies of gray-level value. When a given image contains objects/regions of different gray-level values, different modes will appear on the histogram of the image. This type of histogram is called "multi-modal" histogram that we will describe. This method accurately models data distributions using GMM. However, when objects/regions in the image have close gray-level averages, they may overlap to give a single mode. Our hypothesis is that each mode corresponds to a normal distribution. This is acceptable in a large number of practical applications [10]. Then the Expectation Maximization (EM) algorithm can be developed to estimate the number of modes of such histograms. We use the EM algorithm to fit GMM to data, determining the number of components and the parameters by means of an iterative procedure.

2. Method

2.1 Thresholding Technique

Thresholding is a technique for segmentation of colored or grey scaled images based on the color or grayscale value, which transforms an image into a binary image by transforming each pixel according to whether it is inside or outside a specified range. The user chooses lower and upper threshold values to process the histogram. If a pixel is inside of this range, it is assigned an "inside" value. Otherwise it is assigned an "outside" value. So, thresholding may be viewed as an operation that involves tests against a function T,

$$T = T[x, y, p(x, y), f(x, y)] \tag{1}$$

where $f(x, y)$ is the grey level of point (x, y) and $p(x, y)$ denotes some local properties of this point, e.g, the average grey level of a neighborhood. The actual part of thresholding consists of setting background values for pixels below a threshold value T and a different set of values for the foreground. A thresholded image, $g(x, y)$ is then defined as:

$$g(x) = \begin{cases} 0 & if \quad f(x) < T \\ 1 & otherwise \end{cases}, \tag{2}$$

Notes:

(1). Global $-T$ depends on $f(x,y)$ only.

(2). Local $-T$ depends on $f(x,y)$ and $p(x,y)$.

(3). Dynamic $-T$ depends on (x,y) as well.

The input to a thresholding operation is typically a grayscale or color image. In the simplest implementation, the output is a binary image representing the segmentation. Black pixels correspond to background and white pixels correspond to foreground (or *vice versa*). In simple implementations, the segmentation is determined by a single parameter known as the *intensity threshold*. In a single pass, each pixel in the image is compared with this threshold. In our study, if the pixel's intensity is higher than the threshold, the pixel is set to white in the output. If it is less than the threshold, it is set to black.

Figure 1 shows two probability density functions. Optimal thresholds can usually be extracted from bimodal histograms. If the histogram is not bimodal, then threshold determination will become difficult (see Figures l and 2)

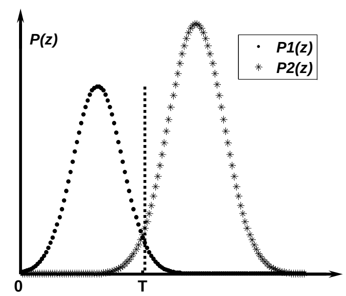


Fig.1. Gray-level probability density functions of two regions in an image

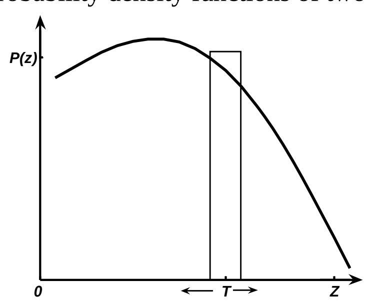


Fig. 2. Non-optimal histogram for threshold selection.

2.2 Gaussian Mixture Model (GMM) and EM Algorithm

It is assumed that there are a finite number of Gray-level probability density functions in the image, say k, and each pixel distribution can be modeled by one Gaussian. With this assumption, the whole image can be modeled by a mixture of g component Gaussian distributions in some unknown proportions $\pi_i, i=1,2,...,k$. The probability density function (PDF) of a data point x will be

$$f(x|\psi)=\sum_{i=1}^{k}\pi_i f(x,\mu_i,\Sigma_i) \tag{3}$$

where, $0\le\pi_i\le 1$, $\sum_{i=1}^{g}\pi_i=1$, ψ is a vector containing parameters π_i,μ_i,Σ_i for $i=1,2,...,k$.

Hence,

$$f(x,\mu_i,\Sigma_i)=\frac{1}{2\pi}\left|\Sigma_i\right|^{-\frac{1}{2}}e^{\frac{1}{2}(X-\mu_i)^t\sum_i^{-1}(X-\mu_i)} \tag{4}$$

Equition 4 describes the $i-th$ component Gaussian distribution with mean μ_i, and covariance Σ_i.

The fit of a model to the data can be measured by the total log likelihood of the data

$$L(\psi)=\sum_{j=1}^{N}\log f(x_j\mid\psi) \tag{5}$$

Now, the GMM [8] addresses the problem in which the joint conditional density of the observations is $f_x(\vec{x}\mid\vec{\pi},\vec{\theta}^1,\vec{\theta}^2,\ldots,\vec{\theta}^L)$ formed as:

$$f_x(\vec{x}\mid\vec{\pi},\vec{\theta}^1,\vec{\theta}^2,\ldots,\vec{\theta}^L)=\prod_{i=1}^{N}\sum_{j=1}^{L}\vec{\pi}_j f_j(\vec{x}_j\mid\vec{\theta}^j) \tag{6}$$

The $\vec{\pi}_j$ are termed the mixing weights and the $f_j(\cdot)$'s are termed the component densities.

The Expectation Maximization (EM) algorithm can be used to find such an estimation of the parameter $\hat{\psi}$. However, it has been noticed that the results of the EM algorithm are generally very sensitive to the initial values of the parameters because of local maxima for the total likelihood in the parameter space.

In this study, a simple method to learn the parameters of the mixture model is to use the EM algorithm with a predefined number of Gaussians and some initial means and covariance. We derived an EM algorithm for ML estimation of the parameters of the component densities, for the case where we assumed all are univariate Gaussian and measurements $\vec{x}$ are i.i.d. Let $L\times 1$ vectors $\vec{\mu}$ and $\vec{\sigma}$ contain, respectively, the unknown means and standard deviations so that

$$f(\vec{x}_i\mid\vec{\psi}_j)=f(\vec{x}_i\mid\vec{\mu}_i,\vec{\sigma}_i)=\frac{1}{\sqrt{2\pi}\vec{\sigma}_j}\exp[-\frac{(\vec{x}_i-\vec{\mu}_i)^2}{2(\vec{\sigma}_j)^2}] \tag{7}$$

Then, the iterative EM algorithm for estimating the parameters of the component densities is given by:

$$\vec{\omega}_j^{i(k)}=\frac{\vec{\pi}_j^{(k)}f_i(\vec{x}_i\mid\vec{\mu}_j^{(k)},\vec{\sigma}_j^{(k)})}{\sum_{l=1}^{L}\vec{\pi}_l^{(k)}f_l(\vec{x}_i\mid\vec{\mu}_l^{(k)},\vec{\sigma}_l^{(k)})} \tag{8}$$

$$\vec{\pi}_j^{(k+1)}=\frac{1}{N}\sum_{i=1}^{N}\vec{\omega}_j^{i(k)} \tag{9}$$

$$\vec{\mu}_j^{(k+1)}=\frac{1}{N\vec{\pi}_j^{(k+1)}}\sum_{i=1}^{N}\vec{\omega}_j^{i(k)}\vec{x}_i \tag{10}$$

$$[(\vec{\sigma}_j)^2]^{(k+1)}=\frac{1}{N\vec{\pi}_j^{(k+1)}}\sum_{i=1}^{N}\vec{\omega}_j^{i(k)}[\vec{x}_i-\vec{\mu}_j^{(k+1)}]^2 \tag{11}$$

Parameters $\vec{\pi},\vec{\mu},\vec{\sigma}$ all describe the global properties of the measurements.

Our global threshold algorithm takes the following steps to determine thresholding of image:

(1) Given an image with gray-level range, compute the gray-level frequency distribution and histogram.

(2) Initial Gaussian components.

(3) The EM is used to determine the Gaussian parameters $(\omega_j,\pi_j,\mu_j,\sigma_j)$ by maximizing the equation 6;

(4) If the estimation error for step (3) is less than a predefined threshold or the number of iteration is greater than T, stop and output results. Otherwise, go to step (3).

(4) Add a new Gaussian component and perform step (2) again.

(5)In the end, we choose the optimal threshold which is the average of these means μ_m.

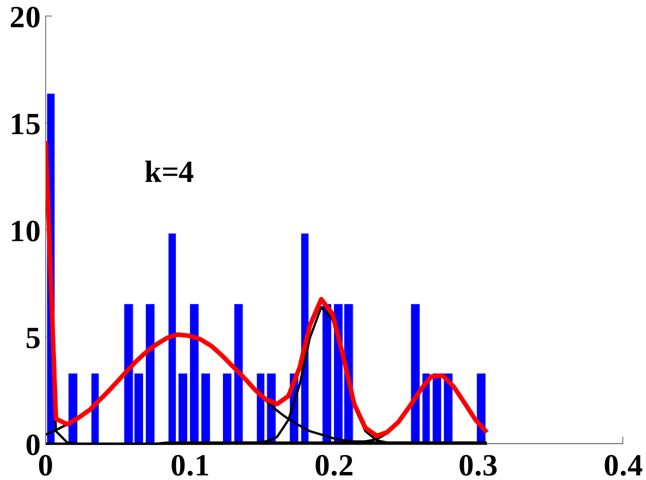
$$T_{opt} = \frac{1}{m}\sum_{i=1}^{m}\mu_m \,, \qquad (12)$$

where m is the number of Gaussian mixtures.

3. Experimental Results and Discussion

To verify the performance of our method, a set of various images was tested by our methods. For all the tested images, the images labeled (a) are original images. Figure 3b is a comparison between the probability density function (PDF) of image vs. its Gaussian mixture model. Figure 3c is thresholding images of our method. We examine the performance of the EM algorithms with respect to the mixing weights, the mean value of each class, the variances in each class, and the number of classes in the image [16].

(a)

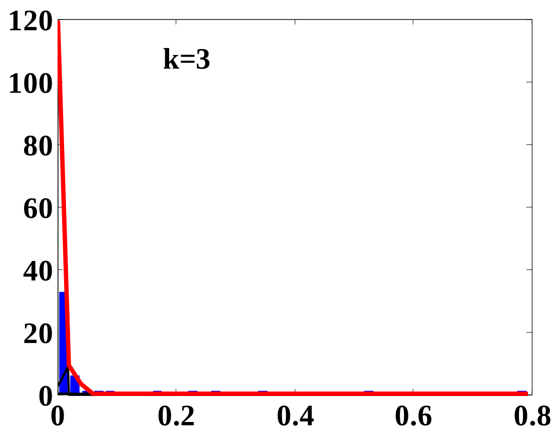


(b)

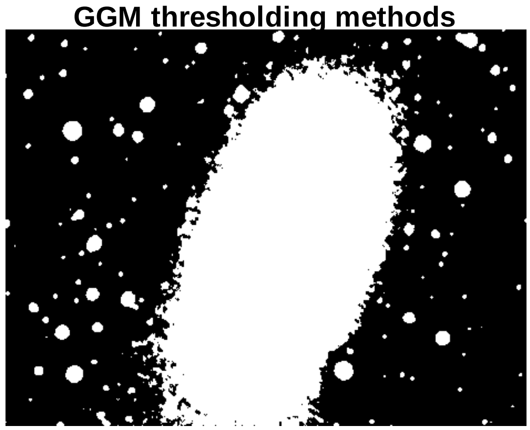
(c)

Fig.3. (a) Original image; (b) A comparison between the PDF of image vs. its Gaussian mixture model. The solid line is image histogram and the '-' line is mixture model. (c) Result of an image conversion to a binary image [15]

(a)



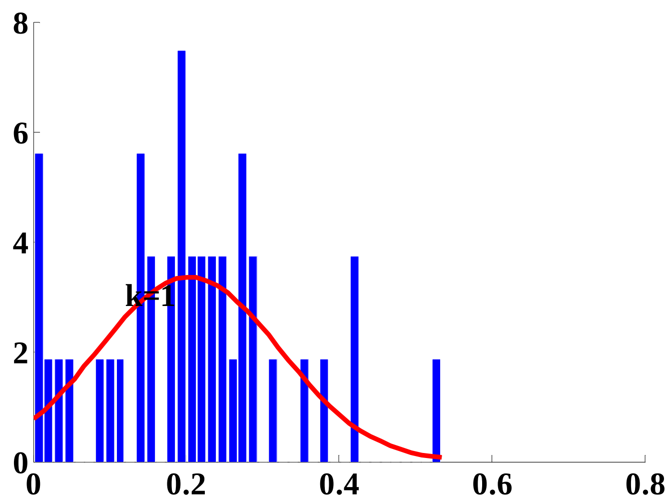
(b)



(c)

Fig.4. (a) Original image; (b) A comparison between the PDF of image vs. its Gaussian mixture model. The solid line is image histogram and the '-' line is mixture model. (c) Result of an image conversion to a binary image.

(a)

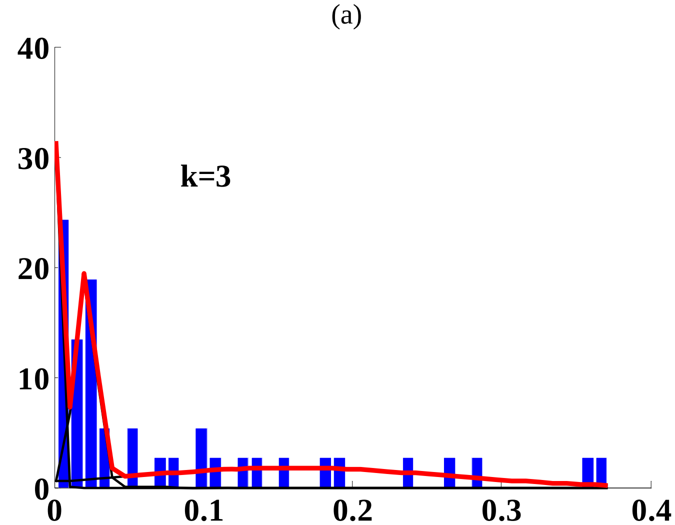


(b)

(c)

Fig.5. (a) Original image; (b) A comparison between the PDF of image vs. its Gaussian mixture model. The solid line is image histogram and the '-' line is mixture model. (c) Result of an image conversion to a binary image

(a)

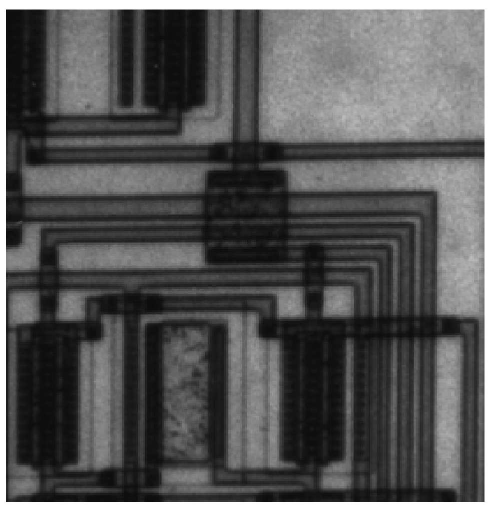


(b)

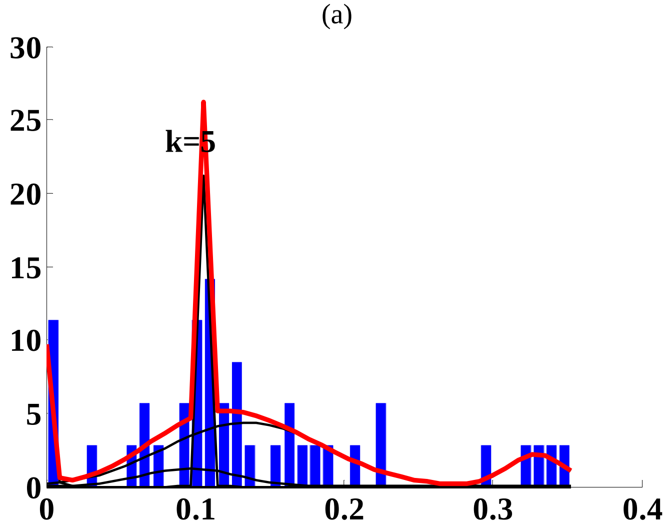
GGM thresholding methods

(c)

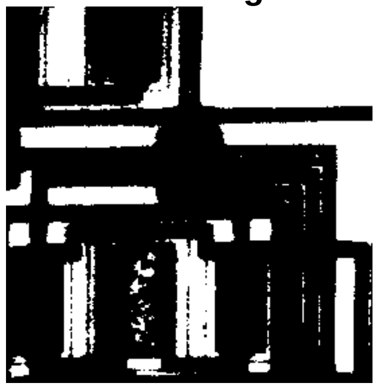
Fig.6. (a) Original image; (b) A comparison between the PDF of image vs. its Gaussian mixture model. The solid line is image histogram and the '-' line is mixture model. (c) Result of an image conversion to a binary image [16].



(a)



(b)



(c)

Fig.7. (a) Original image; (b) A comparison between the PDF of image vs. its Gaussian mixture model. The solid line is image histogram and the '-' line is mixture model. (c) Result of an image conversion to a binary image [16].

The above results show that our methods are able to select the number of components in an unsupervised way. Our proposal obtains near-to-optimal data models, requiring far less computational time than other proposals in the bibliography. This makes it appealing to be used in practical problems. Our algorithm refines iteratively an initial GMM constituted by only one Gaussian, augmenting its order incrementally until a good model is obtained.

4. Conclusions

This paper proposed a thresholding method based on Gaussian mixture model. According to the fact that the histogram of image can be used to represent the statistical character of probability density function, the Gaussian mixture is used to estimate the image's PDF of image's grey level. The optimal number of mixtures (Gaussian function) is searched for the candidate by EM algorithm. The optimal threshold has been determined as the average of these means. Experimental results show that our method can achieve better threshold result and it is more robust. The algorithm is unsupervised and fully automatic. Nevertheless this method is still dependent on the initial parameter estimation. In our future work, the combination strategy such as multiple random starts will be adopted to choose highest likelihood estimation. For some poor contrast image, some imprecision is present in the region. Therefore, improvement of the method based on a priori knowledge is needed and is actually in progress in our work.

5. References

[1] Rafael C.gonzalez, Richard E.Woods, "Digital Image Processing," Publishing House of Electronics Industry, Beijing, 2002
[2] D.S.Huang, Systematic Theory of Neural Networks for Pattern Recognition. Publishing House of Electronic Industry of China, Beijing, 1996.
[3] N. Otsu, "A Threshold Selection Method from Gray-Level Histogram," IEEE Trans. Systems, Man, and Cybernetics, vol. 8, pp. 62-66, 1978.
[4] J.M.S. Prewitt and M.L. Mendelsohn, "The Analysis of Cell Images," Ann. New York Academy Science, vol.128, pp. 1035-1053, 1966.
[5] W. Doyle, "Operations Useful for Similarity-Invariant Pattern Recognition," J. ACM, Vol.9, pp. 259-267, 1962.
[6] T. Rider and S. Calvard, "Picture Thresholding Using an Iterative Selection Method," IEEE Trans. Systems, Man, and Cybernetics, vol. 8, pp. 630-632, 1978.
[7] H.J. Trussell, "Comments on `Picture Thresholding Using an Iterative Selection Method," IEEE Trans. Systems, Man, and Cybernetics, vol. 9, p. 311, 1979.
[8] W. Tsai, "Moment-Preserving Thresholding: A New Approach," Computer Vision, Graphics, and Image Processing, vol. 29, pp. 377-393, 1985.
[9] T. Pun, "A New Method for Gray-Level Picture Thresholding Using the Entropy of the Histogram," Signal Processing, vol. 2,

pp. 223-237, 1980.
[10] A.K.C. Wong and P.K. Sahoo, “A Gray-Level Threshold Selection Method Based on Maximum Entropy Principle,” IEEE Trans. Systems, Man, and Cybernetics, vol.19, pp. 866-871, 1989.
[11] D.S.Huang, Wen-Bo Zhao, “Determining the centers of radial basis probabilities neural networks by recursive orthogonal least square algorithms, ” Applied Mathematics and Computation, vol 162, no.1 pp 461-473, 2005.
[12]D.S.Huang and S.D.Ma,“Linear and nonlinear feedforward neural network classifiers: A comprehensive understanding,”Journal of Intelligent Systems, Vol.9, No.1, 1-38,1999.
[13] D.S.Huang, Horace H.S.Ip and Zheru Chi, “A neural root finder of polynomials based on root moments”, Neural Computation, Vol.16, No.8, pp.1721-1762, 2004.
[14] D.S.Huang,“Radial basis probabilistic neural networks: Model and application,”International Journal of Pattern Recognition and Artificial Intelligence, Vol.13(7), 1083-1101,1999.
[15] N. Friedman, S. Russell, “Image Segmentation in Video Sequences: A Probabilistic Approach”. in Proc. of the Thirteenth Conference on Uncertainty in Artificial Intelligence, Aug. 1997.
[16] MathWorks Inc: http:\\www.mathworks.com
[17] W.E.L. Grimson, C. Stauffer, R. Romano, and L. Lee, “Using Adaptive Tracking to Classify and Monitor Activities in a Site”. Computer Vision and Pattern Recognition, 1998.
[18] C. Stauffer and W.E.L. Grimson. “Adaptive Background Mixture Models for Real-Time Tracking”. Computer Vision and Pattern Recognition, 1999.