Proc. VIIth Digital Image Computing: Techniques and Applications, Sun C., Talbot H., Ourselin S. and Adriaansen T. (Eds.), 10-12 Dec. 2003, Sydney

Background Modeling and Subtraction Using a Local-Linear-Dependence-Based Cauchy Statistical Model

Ying Ming[1] Jingjue Jiang[2] Jun Ming[3]

[1]The research center for spatial information and digital engineering, Wuhan University, 129 Luoyu Road, China, 430079
whumike@sina.com
[2]School of Computer Science, Wuhan University, 129 Luoyu Road, Wuhan, China, 430079
bigeyes@public.wh.hb.cn
[3]Yangtse University, 88 Jingbi Road, Jingzhou, China, 434025

Abstract: Many motion object detection algorithms rely on the process of background subtraction, an important technique which is used for detecting changes from a model of the background scene. The background model affects object detecting algorithm tolerating changes in background scene and the granularity of the detected foreground objects. An algorithm using a subtracted background modeling based on Cauchy statistical distribution the purpose of object detecting is presented. The paper concludes that the ratios of the intensity values between background image and current image are fitted to a Cauchy distribution. The Cauchy has much heavier tails and better represents the tails of the histogram than the Gaussian. The Cauchy based method without exponential operation is more cost-efficient than the Gaussian. The proposed approach takes the advantages of the statistic distribution characteristic of pixels and spatial correlativity of the region around a pixel to subtract background. The paper also discusses the changes in background scene in detail. At last, a robust object detecting approach being invariant or adapting to the changes in background scene is acquired by hypothesis test. Experimental results demonstrate the proposed algorithms can tolerate the whole or local sudden or slow change in illumination, filter clutter noises caused by small motion in background scene, and adapt to rain

Key Words: image processing; change detection; background extraction; Local Linear Dependence;Cauchy distribution

1. Introduction

Automated video surveillance has emerged as an important research topic in computer vision community. The growth in this area is being driven by the increased availability of inexpensive computing power and image sensors, as well as the inefficiency of manual surveillance and monitoring system. Applications, such as event detection, human action recognition, and semantic indexing of video, etc, are based on real time motion detecting and tracking algorithms which provide low-level functionality upon which higher level recognition capabilities are built.

The detection of motion in many current visual surveillance systems relies on the technique of background subtraction. Scene background modeling is an important task of background subtraction. In [1], Elgammal et.al provides a brief review of the relevant work. Many researchers have proposed methods to address some of the issues regarding the background modeling and background subtraction. All of the previous background models based on statistical modeling of single pixel intensity with the ability to adapt the model. However, a sudden change in illumination presents a challenge to those models. Most of such models only make use of the information on temporal property in successive images, not regarding the property of pixels spatial distribution in each image.

This paper proposed a novel method of background modeling and subtraction based on local-linear-dependence-based Cauchy statistical model. The new idea is triggered by [2] and [1]. A ratio of pixels intensity or color between two images or two difference images is used as the feature for subtracted background modeling. We find that the ratio distribution in spatial distribution obeys Cauchy distribution, and a pixel is local-linear-dependence to all or part of its nearby pixels if there hasn't been a change or they are belong to a same object. A robust background modeling and subtraction approach based on this is acquired. At last, Experimental results demonstrate the proposed algorithm can tolerate the whole or local sudden or slow change in illumination, filter clutter noise caused by small motion in scene background, and adapt to rain.

2. Changes in scene background

In real indoor or outdoor scene, there are changes that occur over time and may be classified as changes to the scene background. These changes can be local, affecting only part of the background, or global, affecting the entire background. It is important that the background model tolerates these kinds of changes, either by being invariant to them or by adapting to them. The study of these changes is essential to understand the motivations behind different background subtraction techniques.

2.1 Illumination changes

In practice environment, illumination changes often occur over time. We classify these changes in terms of their source.

– Gradual change in illumination

This change might occur in outdoor or indoor due to the change in the location of sun over time, and be a kind of whole change. It results that all of the image intensity lighted by sun are smoothly changed simultaneously.

– Sudden change in illumination

It might happen in an indoor environment by switching lights or off, or in an outdoor environment by a quick change between cloudy and sunny conditions. All the affected image intensity is suddenly changed synchronously.

– Shadow

Shadows cast on the background by objects in the background itself or by foreground objects. This changes cause local image affected. All the image intensity in the shadow are changed simultaneously and have a smooth change alone the projection direction of the shadow.

2.2 Dynamic changes

– Motion in parts of scene background

It might be caused by tree branches moving with wind, or rippling water, and snow or rain, which results that parts of image intensity are changed randomly.

– Global motion

These changes are occurred due to small camera displacements, they are common in outdoor situations due to wind or other sources of motion which causes global motion in the images.

– Physical changes

These include any changes in the geometry or the appearance of the scene background introduced by objects.

3. Background modeling and updating

3.1 Learning initial background parameter

A simple and common background modeling method involves subtracting each image from a background model and thresholding the results of difference image to determine moving object pixels. The features to describe background can be divided into two classes. One is the level of pixel characteristic, such as pixel intensity, edge and intensity difference, etc. The other is the level of region identity, such as connection, similarity and so on. The pixel intensity of a completely stationary background can be can be modeled with a normal distribution or Gaussian, and it can adapt to slow change in the scene by recursively updating the model [9]. In our work, the pixel intensity is used as the feature of background modeling.

In order to make background model can be used to describe both color images and intensity images, the YCbCr color space is selected as the image format. The YCbCr color space is widely used for digital video. In this format, luminance information is stored as a single component (Y), and chrominance information is stored as two color-difference components (Cb and Cr). Cb represents the difference between the blue component and a reference value. Cr represents the difference between the red component and a reference value.

We obtain the initial background parameters from a short video sequence without any moving objects. The median value of intensity at each pixel location in all images is determined as the value of intensity at the corresponding pixel location in the initial background model.

3.2 Background modeling and updating

The surveillance area does not stay the same for a long period of time. There could be illumination changes, motion changes and physical changes. As our background model still can be classified into models based on intensity feature of background pixel, some changes in illumination can cause false positives. Additionally, any foreground object detected for a long time without motion, such as a parked car, can cause false negatives, e.g., a person cannot be detected while he is getting out of the parked car. [8] is a good reference that explain the problems of background scene maintenance for surveillance systems. In this paper, we use a pixel-based update method as

$$X_{t+1}(i) = \begin{cases} \alpha \cdot X_t(i) + (1-\alpha) \cdot Y_t(i), 0 < \alpha < 1, if \quad i \quad is \quad background\, pixel \\ X_t(i), if \quad i \quad is \quad foreground\, pixel \end{cases}$$

where α controls the updating speed of background image, which is a experiential value estimated mainly by the frame rate of video image captured and the expectant interval of background update.

4. Background subtraction and object detection

In this paper, the ratio of image pixel intensity or color between current image and background image is used as basic feature to subtract scene background. In the scene background model, a sample value of each pixel in background image is kept so that the pdf of the ratio between current image and background image can be acquired. Therefore, the probability of any new measure intensity can be estimated by the background model.

4.1 The statistic characteristic of a ratio image

4.1.1 the distribution of the ratio between two image noise

According to [2], in a sequence of an image with size $M \times N$, the intensity of an arbitrary pixel $x_{i,j} \in \{x_{i,j}, i = 1,2,\cdots,M; j = 1,2,\cdots,N\}$ can be defined as

$$\hat{X} = X_b + \Theta \qquad \textbf{(1)}$$

Where $\hat{X}$ is the measure value of the pixel intensity, X_b is the real value without nose, and Θ is a stochastic noise variable. If X_b is regarded as the intensity of the pixel in the background image, we can get noise variable of the pixel in the current image

$$\Theta = \hat{X} - X_b \qquad \textbf{(2)}$$

This is the common formula of different image. Under the hypothesis that no change occurs at location $x_{i,j}$, the corresponding difference Θ obeys a Gaussian distribution and the camera noise is uncorrelated between different frames.

The ratio of two stochastic image differences or two image noise is expressed as

$$\frac{\Theta^m}{\Theta^n} = \frac{\hat{X}^m - X_b}{\hat{X}^n - X_b} \qquad \textbf{(3)}$$

Where, *m* and *n* denotes two different image frames. Θ^m and Θ^n obey Gaussian distribution, e.g., $\Theta^m(x) \sim N(\mu_m, \sigma_m)$, $\Theta^n(x) \sim N(\mu_n, \sigma_n)$.

Let $R(x) = (\hat{X}^m - X_b)/(\hat{X}^n - X_b) = \Theta^m/\Theta^n$, according to [4], we can draw a conclusion that the distribution of $R(x)$ obeys a Cauchy distribution with parameters μ and λ, that is,

$$R(x) = (\hat{X}^m - X_b)/(\hat{X}^n - X_b) = \Theta^m/\Theta^n \sim C(\mu, \lambda) \qquad \textbf{(4)}$$

and its probability density function (pdf) is

$$P_r(x) = \frac{1}{\pi} \cdot \frac{\lambda}{\lambda^2 + (x - \mu)^2} \qquad \textbf{(5)}$$

Figure 1 shows that the histogram for ratio of difference pixel intensity value and its fitted Cauchy distribution. The Figure.1.a is the background image or reference image. The Figure.1.b and Figure.1.c are two current images without moving object at different moment. The Figure.1.d is the histogram for the ratio of pixel intensity value between a pairs of different images, e.g., a-b and a-c. The distribution of the ratio of two image noise follows a sharp Cauchy distribution, see Figure.1.e.

a. b c d e

Fig. 1. Histogram for ratio of pixel intensity value between a pairs of different images and its fitted Cauchy distribution

4.1.2 The distribution of the ratio of image intensity

From [5], In a sequence of images with size $M \times N$, each pixel $x_{i,j} \in \{x_{i,j}, i = 1,2,\cdots,M; j = 1,2,\cdots,N\}$ is modeled as an independent statistical process. The distribution of the intensity value is fitted with a Gaussian or a mixture Gaussian. Each Gaussian may correspond to the distribution of background or individual moving objects covering this pixel over time. If the intensity distribution at the location of a pixel *x* in two arbitrary frames of image with same scene are respectively $\hat{X}^m(x) \sim N(\mu_m^{'}, \sigma_m^{'2})$, $\hat{X}^n(x) \sim N(\mu_n^{'}, \sigma_n^{'2})$, where, *m* and *n* denote different frame image.

Let $R^{'}(x_m/x_n) = \hat{X}^m(x) / \hat{X}^n(x)$, we can acquire similarly that $R^{'}(x_m/x_n)$ obeys a Cauchy distribution, e.g.,

$$R^{'}(x_m / x_n) = \frac{\hat{X}^m(x)}{\hat{X}^n(x)} \sim C(\mu^{'}, \lambda^{'}) \quad \textbf{(6)}$$

and its pdf is

$$P^{'}_r(x) = \frac{1}{\pi} \cdot \frac{\lambda^{'}}{\lambda^{'2} + (x - \mu^{'})^2} \quad \textbf{(7)}$$

Figure 2 shows that histogram for ratio of pixel intensity value between two images and its fitted Cauchy distribution. The two original images in Figure.2.a have a same scene without evident background changes. Therefore, the histogram for the ratio of pixel intensity (color) value between them is sharp and narrow and smooth, and is coincident with a same figuration Cauchy distribution. In Figure.2.b, the histogram is not so sharp and smooth as Figure.2.a, there are some dynamic background changes duo to rain. So its background distribution is a sharp Cauchy and the distribution of dynamic background changes is a flat Cauchy.

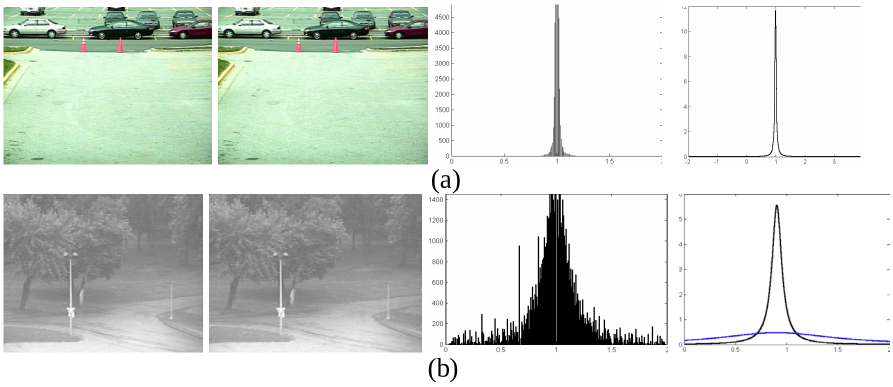

Fig. 2. Histogram for ratio of Pixel intensity value and its fitted Cauchy distribution

4.2 Background pixel model in a ratio image

Let x_i be a sample of the intensity value of a pixel in a background image, $x_i \in \vec{X}_i, \vec{X}_i = [x_1, \cdots, x_i, \cdots, x_n]$. For the given samples, the distribution probability for the measure intensity of each pixel in current image at moment t can be estimated by the pdf of the ratio and the samples, where, $y_i^t \in \vec{Y}_i, \vec{Y}_i = [y_1, \cdots, y_i, \cdots, y_n]$. $\vec{Y}_i$ is the image vector corresponding to $\vec{X}_i$, whose probability estimate is

$$P_r(x_t) = \frac{1}{n} \sum_{i=1}^{n} R_{\mu,\lambda}(x_t / x_i) \quad \textbf{(8)}$$

where, $R_{\mu,\lambda}$ is the pdf of a Cauchy distribution following parameters μ and λ. This estimate can be expanded to use the ratio of color feature as

$$P_r(x_t) = \frac{1}{N} \sum_{i=1}^{N} \prod_{j=1}^{d} R_{\mu_j,\lambda_j}(x_{tj} / x_{ij}) \quad \textbf{(9)}$$

where x_{tj} is a 3-dimensional color feature, and R_{μ_j,λ_j} is a Cauchy pdf with parameters μ and λ in the jth color space dimension which use (5) or (7). How to estimate the parameters μ and λ can see [6].

From [7], we can know that a image space is a linear space, and each image can be regarded as a linear combination of moving objects and background. Hence, except for the parts of the current images occluded by moving objects or occurring motion changes of scene background, the ratio of the rest of current image to the corresponding parts of background still satisfies the condition of local linear dependent, and obeys a Cauchy distribution. Using this probability estimate, the pixel x_t is considered to be a background pixel if $P_r(x_t) \geq Threshold$, otherwise, if $P_r(x_t) < Threshold$, the pixel x_t is a change pixel, where the threshold *Threshold* is a global threshold over all the images that can be adjusted to achieve a desired percentage of false positives.

The Cauchy based background subtracting method has several attractive features relative to those based on the Gaussian. First, the Cauchy pdf has much heavier tails and better represents the tails of the histogram than the Gaussian. Second, the Cauchy method without exponential operation is more cost-efficient than the Gaussian.

4.3 Static background subtraction

Let x_t be intensity value of a arbitrary pixel $\hat{X}_c(i,j)$ in a current image, and then its probability $P_r(x_t)$ can be calculated by the formula (8) with R_{μ_j,λ_j} using the expressions (6). According to its probability estimate and the following criterion, each pixel is classified into background pixel and change pixel. The criterion is that

- If $P_r(x_t) \geq Threshold$, the pixel $\hat{X}_c(i,j)$ is labeled as a background pixel.
- If $P_r(x_t) < Threshold$, the pixel $\hat{X}_c(i,j)$ is labeled as a change pixel.

The intensity value of the entire labeled background pixel is set as 0, so all the labeled background pixels is abstracted, and only those labeled change pixels are held. Whereas, some pixels among the labeled change pixels are caused by motion change of scene background, so all the labeled change pixels must continue to be sorted by dynamic background subtraction.

4.4 Dynamic background subtraction

In outdoor environments, those previously mentioned background changes are the main source of false detection. This can occur locally, for example, if a tree branch moves further than it did during model generation. This can also occur globally in the image as a result of small camera displacements caused by wind load, or as results of rain or snow, which is common in outdoor surveillance and causes much false detection. These kinds of false detections are usually spatially clustered in the image, and they are not easy to eliminate using morphological techniques or noise filtering because these operations might also affect detection of small and/or occlude targets.

If a part of the background (a tree branch, for example) moves to occupy a new pixel, but it was not part of the model for that pixel, then it will be detected as a foreground object. However, this object will have a high probability of being a part of the background distribution corresponding to its original pixel. Assuming that only a small displacement can occur between consecutive frames, we decide if a detected pixel is caused by a background object that has moved by considering the background distributions of a small neighborhood of the detection location.

Let x_t be the observed value of a pixel x detected as a change pixel at time t. We define the pixel displacement probability $P_N(x_t)$ to be the maximum probability that the observed value, x_t, belong to the background distribution of some points in the neighborhood N of x

$$P_N(x_t) = \max_{y \in N(x)} P_r(x_t \mid B_y) \tag{10}$$

where B_y is the background sample for pixel y, and the probability estimation $P_r(x_t \mid B_y)$ is calculated using the formula (8) with R_{μ_j,λ_j} using the expressions (4). The size of the neighborhood N of x can be 3×3, 5×5 and 7×7, according to image resolution and size of object image. By thresholding P_N for detected pixels, much false detection due to small motions in the background scene can be eliminated. Meantime, to avoid losing true detections that might accidentally be similar to the background of some nearby pixel (e.g., camouflaged targets), a constraint is added that the whole detected foreground object must have moved from a nearby location, and not only some of its pixels. The area displacement probability P_c is defined to be the probability that a detected connected area C has been displaced from a nearby location. This probability is estimated by

$$P_c = \prod_{x \in C} P_N(x) \tag{11}$$

For a connected area corresponding to a real target, the probability that this area has displaced from the background will be very small. So, the criterion to abstract dynamic background pixel can be acquired as following:

If $P_N(x) > Threshold1$ and $P_C(x) > Threshold2$

Then x is a background pixel; else x is a foreground pixel.

5. Results

In order to verify the proposed approach, two kinds of video sequence data are used in our experimental. One is color image, shown in Figure.2.a; another is grey image, shown in Figure.2.b. All test data get from [1]. The data of Figure.2.a are sampled at a resolution of 240×320 and a rate of 30 frames per second, and saved in YC_bC_r color format. The data of Figure.2.b are sampled at a resolution of 240×320 and a rate of 30 frames per second , and saved as grey image.

In Figure.3, the first row (a) are original image, the second row (b) are the results of change detecting. The detecting results can be acquired by static background subtraction. Reflections from the windows of the three cars and the shadow of the pedestrian are eliminated.

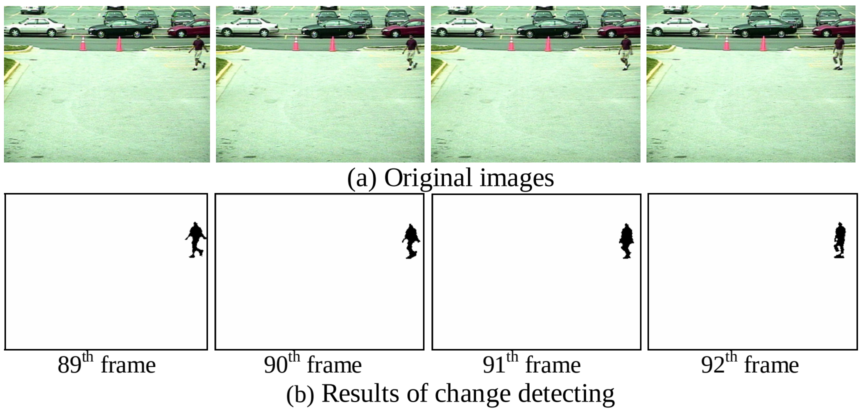

Fig. 3. Original images and results of change detecting in static background by the proposed approach

Figure.4 shows a rainy day. There are some small motions in the background scene duo to rain. In Figure.4.a, there is a transiting van. Figure.4.b shows detecting results corresponding to their original image, in which the moving object is detected correctly without any filtering process. In Figure.4.d, a walking man and a car are detected accurately.

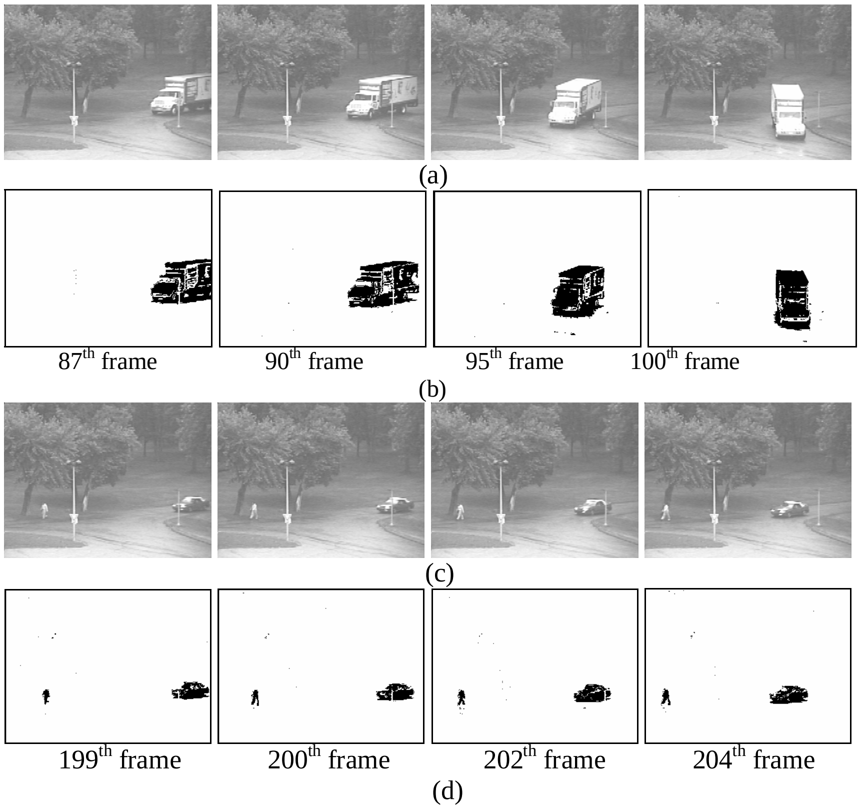

Fig. 4. Original images and results of change detected in dynamic background by the proposed approach. (a) and (c) are original images; (b) and (d) are detecting results.

6. Conclusions

The experimental results shows that the proposed approach of background modeling and subtraction can accurately detect semantic objects, eliminate noise caused by illumination change and motion change in background. The Cauchy based approach without exponential operation is more cost-efficient than traditional statistic algorithms based on Gaussian distribution model.

Reference

1. Ahmed Elgammal, Ramani Duraiswami, and David Harwood et al. Background and foreground modeling using nonparametric kernel density estimation for visual surveillance. Proceedings of the IEEE, 2002, Vol. 90(7): 1151~1164.

2. Emrullah Durucan, Touradj Dbrahimi. Change Detection and background Extraction by linear algebra. Proceedings of the IEEE, Vol.89, No.10, October 2001: 1368~1381.
3. K.Skifstad and R.Jain. Illumination independent change detection for real world image sequence. CVIP,1989,Vol.46(3):387~399
4. Li Yuqi, He Pin.The theory of probability and statistics. National defense industry publishing company, Beijing, 2001: 130～131
5. Ying Ren, Chinseng Chua et al. Motion detection from time-varied background. International Journal of Image and Graphics, Vol. 2, No.2 (2002): 163~178
6. Tsihrintzis, G.A. Nikias, C.L. Fast estimation of the parameters of alpha-stable impulsive interference using asymptotic extreme value theory. ICASSP-95，Vol. 3 , 9-12 May 1995: 1840 –1843
7. Li Yun. Basic Engineering mathematics. Wuhan traffic science and technology university publish，1996：30～31,28~29
8. Ismail Haritaoglu, David Harwood, and et al.. W4: Real-time Surveillance of People and Their Activites. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 22, No. 8 August 2000:809~829
9. Ismail Haritaoglu, David Harwood and Larry S. Davis, 2000, A fast Background scene modeling and maintenance for outdoor surveillance, in Proceedings of 15th International conference on pattern recognition, Vol.4:179~183.