

# Kernel Density Estimation Method Basing on Color and Motion Features Frame for Moving Object Detection

Yu Guo<sup>a</sup>, Ziqiang Shen<sup>b</sup>

School of Computer Science and technology, Nantong University, Nantong, China

<sup>a</sup>13120008@yjs.ntu.edu.cn, <sup>b</sup>305137530@qq.com

**Abstract**—Due to large calculations and complex background updating problems of kernel density estimation, this paper proposes a feature frame building method based on the combination of color feature and motion information, using this method to extract the number of samples N, it can not only reflect the global information of image but also reflect local information variations of image, besides it can effectively solve the inaccurate problem of the sample numbers, thereby enhancing the instantaneity of kernel density estimation algorithm.

**Keywords**—kernel density; feature frame; color feature; motion feature

Kernel density estimation algorithm [1] is of good robustness, it can solve parameter setting problems in many algorithms, no need in parameter setting is a big advantage of this algorithm [2]. However, due to large calculations and complex background updating problems of kernel density estimation, its instantaneity is poor. The key to improve its real-time performance is to choose a more reasonable sample number N. One method [3] is directly using fewer samples, but this method is bound to lose more information. The second [4] is using key frames extraction, although this method greatly improve the instantaneity of target detection, but it lose the sample feature information, the instantaneity is still not satisfactory. The third [5] is through extracting the color feature information of the whole sample, and building feature frames, which to some extent overcome the problem of information loss and improve the instantaneity of target detection, but only using color features can only reflect the global information of image, but can't reflect local information variations of image. Video is composed of a sequence of time-varying frame image and motion is its main feature, so the use of the motion information is a video processing trend. This paper proposes a feature frame building method based on the combination of color feature and motion information, which will be further discussed and studied.

## I. INTRODUCTION OF PRINCIPLES AND FRAMEWORK

### A. Principle of Feature Frame

Video is composed of a sequence of time-varying frame image. In the situation of appearing no moving targets, the gray value of image pixel changes relatively stable, but in the case of appearing moving targets, the gray value of image pixel changes very sharply. By this phenomenon, we can use the difference between the sample and the average

template to build feature frame. But when moving targets moves small may regards pixels as background pixels mistakenly, to solve this problem, you can take certain steps to divide its video data.

Step one: according to the step, divide the input N frame images into groups  $K_1, K_2, \dots, K_m$ , each group containing M frame samples.

$P_i$  represents pixel values of each position in the mean template

$$P_i = \frac{1}{M} \sum_{\lambda=1}^M K_i(x_\lambda) \quad (1)$$

Where  $K_i(x_\lambda)$  represents the pixel value  $x_\lambda$  of  $\lambda$ -th frame in i-th set of data

Step two: data partitioning for pixel values in each set of image. We use maximum differential distance sharing method:

$$th_{ij} = j * (\max |x_\lambda - P_i| - \min |x_\lambda - P_i|) / l \quad (2)$$

Where  $l$  is determined by the number of feature frame samples  $(S_1, S_2, \dots, S_k)$  k and the grouping number m

$$l = k / m$$

The value of j-th pixel in i-th feature frame:

$$S_{ij} = \frac{1}{n_{ij}} \sum_{\lambda=1}^{n_{ij}} x_\lambda \quad (3)$$

Where  $th_{ij} < |x_\lambda - P_i| < th_{i(j+1)}$ ,  $i = 1, 2, \dots, m$ ,  $j = 0, 1, 2, \dots, l-1$ .

Where  $n_{ij}$  represents the number of pixels in i-th group that distant from average pixel value  $P_i$  between  $th_{ij}$  and  $th_{i(j+1)}$ .

### B. Principle of Kernel Density Estimation

In probability theory Kernel Density Estimation is to estimate the unknown density function, is one of the non-parametric test method, proposed by Rosenblatt (1955) and Emanuel Parzen (1962), also called Parzen window. Kernel density estimation method can cope with the small-scale amount of data and can quickly generate density

estimation approaching to true data with superior probability properties.

The definition of Kernel Density Estimation is as follows:

Sample  $\{x_1, x_2, \dots, x_n\}$  is a group of data of independent and identically distributed, its probability density is  $f(x)$ . Assuming region  $R$  ( $x$  is center,  $h$  is the side length,  $D$  is a hypercube of space dimension),  $k$  samples in region  $R$ . Process of estimating sample data in region  $R$  is as follows, the kernel function  $K(\lambda)$  is defined as:

$$K(\lambda) = \begin{cases} 1, & |\lambda_j| < 0.5, \quad j = 1, \dots, D \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

So the sample number in region  $R$  is:

$$m = \sum_{i=1}^n K\left(\frac{x - x_i}{h}\right) \quad (5)$$

Where  $x$  is measured sample points

From the non-parametric probability density estimation formula (6) can get kernel density estimation expression, shown in equation (2.3):

$$\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x - x_i}{h}\right) \quad (6)$$

Where  $x$  is the measured sample points

Common kernel density estimation function in figure1:

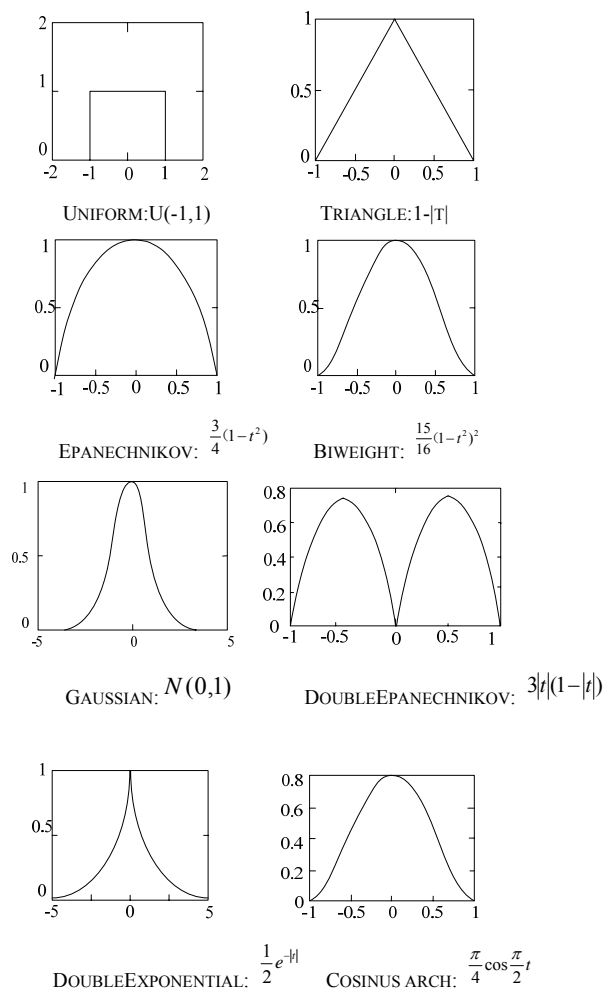


Figure 1. Common kernel function in one-dimensional space

### C. System Block Diagram

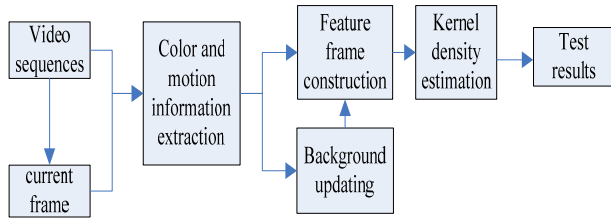


Figure 2. Kernel function is commonly used in one-dimensional spatial

## II. COLOR AND MOTION INFORMATION EXTRACTION AND INTEGRATION MECHANISM

### A. Color Feature Extraction

Color feature is the most widely used in image retrieval, mainly due to color often relate to object or scene in the image. Color feature representation are: color histogram [6], color moment, color sets, color polymerization vector and color related maps. Image color space that can fully express global information image. This paper adapt color histogram method, and then maps the image from RGB space to YCbCr space [7], the mapping formula is as follows:

$$\begin{aligned} Y &= 0.299 \times R + 0.587 \times G + 0.114 \times B; \\ Cb &= -0.169 \times R - 0.331 \times G + 0.500 \times B; \\ Cr &= 0.500 \times R - 0.419 \times G - 0.081 \times B; \end{aligned} \quad (7)$$

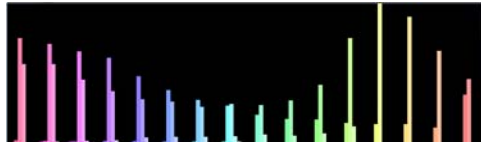


Figure 3. YcbCr histogram

Next, extract color component from each frame image YCbCr space, calculating the average of the respective colors, and then carry on DCT transform. Finally, do scan quantization processing for each component of DCT coefficients, finally obtaining color feature information of the frame image.

### B. Motion Information Extraction

Motion information of image is a prominent feature of video sequence, the video can be more accurately processed further by this feature. Generally, we adapt block motion compensation and motion estimation methods to extract video motion information [8]. Motion compensation is describing how does each small of previous frame move to a location in the current frame. A major disadvantage is the discontinuities between blocks, often referred as block effects. This paper will use the motion estimation method, basic idea of motion estimation is to divide each frame of image sequence into many non-overlapping macro blocks, and consider the displacement of all pixels in macro block are the same, then according to certain matching criteria, to

find the most similar to the current block (matching block) for each macro block within a given range of specific search, relative displacement between current block and matching block is the motion vector [9].

(1) Divide each frame image into many 8\*8blocks;

(2) For the  $u$ -th block  $b_u(i)$  in  $i$ -th frame, to find the best match  $v$ -th block  $b_v(i+1)$  in the  $(i+1)$ -th, get motion vector of this block

$$D_{i,i+1}(u) = D(b_u(i), b_v(i+1)) \quad (8)$$

(3) Summing of all blocks I-frame motion vectors  $D_{i,i+1}(u)$  is image motion-information distance.

$$d_2 = \sum D_{i,i+1}(u) \quad (9)$$

### C. Integration Mechanism

In order to get color features and motion information together, we use normalization method, then weighted calculating the distance between two adjacent frames

$$d = w_1 d_1 + w_2 d_2 \quad (10)$$

$w_1$  and  $w_2$  are the weight of color features and motion information, satisfies

$$w_1 + w_2 = 1 \quad (11)$$

Weight values represents the proportion of color features and motion information in a video sequence,  $w_1$  is larger when color feature is obvious,  $w_2$  is larger when motion information is obvious. In this experiment, consider the color features and motion information is equally important, the weight values of the two feature are all 0.5.

## III. MOVING TARGET DETECTION AND BACKGROUND UPDATE

From the last chapter we get a set of feature frame based on color features and motion information, here we regard this group of feature set as sample for kernel density estimation.

When do moving targets detection [10] for real-time image, based on kernel density formula (6) we can obtain

kernel density estimation value  $\hat{f}(x_t)$  of current pixel, that is, the pixel value is the probability of the background points.

Next, we will compare  $\hat{f}(x_t)$  with threshold  $th$  to detect moving targets, the threshold  $th$  may be obtained through

several experiments, when  $\hat{f}(x_t)$  less than the threshold  $th$ , then consider this pixel as a moving target, assign the current image pixel value to  $M_t$ , otherwise is background point,  $M_t$  is 0.

$$M_t = \begin{cases} x_t, & \hat{f}(x_t) < th \\ 0, & \hat{f}(x_t) > th \end{cases} \quad (12)$$

Where  $x_t$  is pixel values of current image,  $th$  To see the formula (2).

#### IV. EXPERIMENT RESULTS AND ANALYSIS

Based on these ideas, this paper take Visual Studio 2013 + OpenCV for the experimental environment to achieve the frame extraction of features based on color and motion, in order to detect the performance of this algorithm ,we select two sets of video for experiment, and then compare this algorithm to other algorithms.

We select two sets of video for experiment and finish it on PC of 4G memory, Intel(R) Core(TM)i3-41604GHz CPU:

	Image format	Resolution	Frame rate
<b>Video 1</b>	bmp	720×576	25
<b>Video 2</b>	bmp	320×240	25

Original image: A1, A2,A3,A4,A5,A6  
 GMM results: B1, B2, B3,B4,B5,B6  
 KDE results: C1,C2,C3,C4,C5,C6

TABLE I. EXPERIMENT RESULTS OF VIDEO 1

	5-th frame	23-thframe	45-thframe
<b>GMM(ms)</b>	180	194	213
<b>KDE(ms)</b>	882	892	924

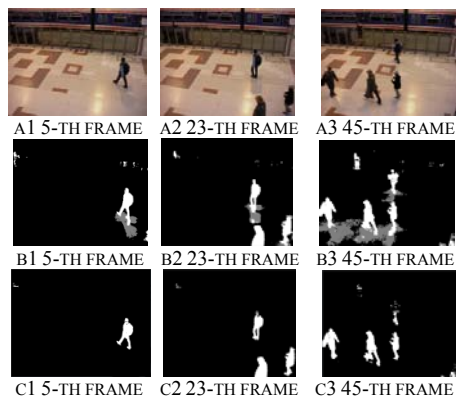


Figure 4. Experiment results of video 1

TABLE II. EXPERIMENT RESULTS OF VIDEO 2

	11-th frame	22-th frame	40-th frame	53-th frame	72-th frame	90-th frame
<b>GMM(ms)</b>	38	82	75	71	60	66
<b>KDE(ms)</b>	259	346	352	277	296	376

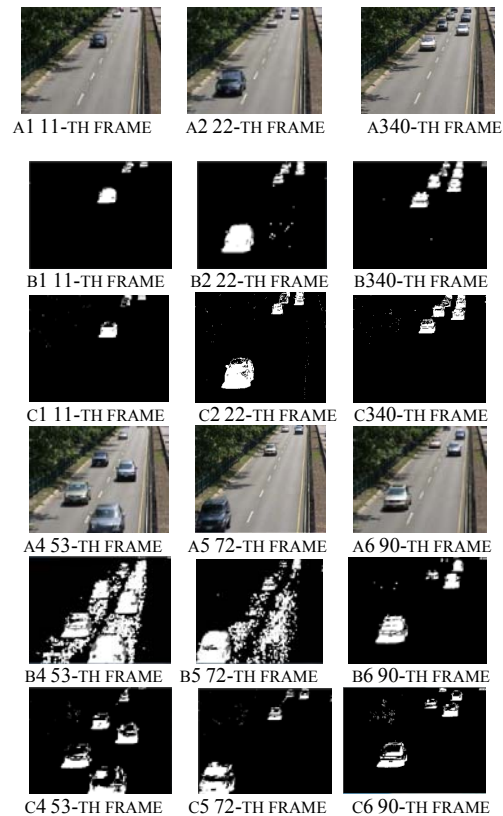


Figure 5. Experiment results of video 2

Utilizing kernel density estimation method basing on color and motion features frame for moving object detection, respectively, with indoor surveillance video and outdoor traffic video for experiment. From the experimental results we find this method has a good detection effect, and strong anti-interference ability.

#### V. CONCLUSION

When detecting the video target, we select the color features and motion information for feature space, use the differences between background and target to extract feature frame set as samples for kernel density estimation. Because the sample number N has a large influence on kernel density estimation, when N is too small, poor resistance to noise, when N is too large, will reduce the computing speed of kernel density estimation. In the premise of not affect the computing speed, as far as possible to overcome the serious problem of information loss, to some extent, we improve the ability of anti-noise interference.

#### ACKNOWLEDGMENT

This research was financially supported by the National Science Foundation (61340037, 61171132).

#### REFERENCES

- [1] Kai Zhang, Ming Tang, James T Kwok. Applying neighborhood consistency for fast clustering and kernel density estimation[C].

- Proceedings of the International Conference on Computer Vision and Pattern Recognition, 2005, 2(2): 1001-1007
- [2] D.Comaniciu,V.Ramesh, P. Meer. Kernel-Based Object Tracking[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, 25(5): 564-575
  - [3] WANG Jin-song, WEI Fa jie, YAN Yi-an. Moving object detection method using background Gaussian kernel density estimation.[J] Infrared and Laser Engineering, 2009, 38(2): 373-376
  - [4] Wang Jinsong moving target detection method, Yan Yian, Wei Fajie. By way of background Gaussian kernel density estimation [J]. Infrared and Laser Engineering, 2009,38 (2): 373-376
  - [5] LiuMinghua.Based on the characteristics of constructed frame moving target detection algorithm [D] Harbin: Harbin Engineering University, 2012
  - [6] P.Perez, C.Hue, J.Vermaak, M.Gangnet. Color-Based Probabilistic Tracking[C]. Proceedings of the 7th European Conference on Computer Vision 2002: 661-675
  - [7] BI Houjie. A new generation of video compression coding standard-H.264/AVC [M].The People's Posts and Telecommunications Press,2005.
  - [8] OU Yang he, HAN Jun .H. 264 and MPEG-4 video compression: a new. generation multimedia technology of video encoding[M]. National Defense University Press 2004.
  - [9] Zeeshan Rasheed, Mubarak Shah. Detetion and representation of scenes in video [J]. IEEE,2005
  - [10] M.K Singh N.Ahuja. Mean-shift segmentation with wavelet-based bandwidth selection[C]. Proceedings of Sixth IEEE Workshop on Applications of Computer Vision, 2002: 43-47