# One Approach to the Analysis Influence of Change Background Statistical Parameters on the Capability of Tracking Objects using "Mean Shift" Procedure

DIMITRIJE BUJAKOVIC, MILENKO ANDRIC
Military Academy
University of Belgrade
General Pavla Jurisica Sturma 33, 11000 Belgrade
SERBIA
dbujakovic@verat.net, asmilenko@beotel.net

*Abstract:* - A quantitative analysis of change background statistics on the capability of tracking objects using "mean shift" procedure is present in this paper. Change of background statistics assumed changing of mean of brightness and changing noise variance in the scene. Quantitative analysis implies detection error and number of iteration needed for position determination using "mean shift" procedure.

*Key-Words:* - Object detection, object tracking, background statistical parameters, "mean shift" procedure, quantitative analysis, detection error, number of iterations,

## 1 Introduction

In real scenes, probability density function of brightness could be often assumed as a Gaussian mixture probability density function. „Mean shift" procedure is a method for determinating statistics of modes that probability density function. This procedure could be used in various applications: image filtration, image segmentation [1], detecting of countour [2] and tracking objects [3,4]. Influence of change mean gray and variance on the capability of tracking object is analysed in this paper. Analysis is done using MATLAB.

## 2 Detection object position in the scene using "mean shift" procedure

Algorithm for detection object using "mean shift" procedure is presented in [3, 4]. A part of picture where is object model, is described by modified histogram. Modified histogram defines as:

$$\hat{q}_u = C\sum_{i=1}^{n} k\left(\left\|x_i^*\right\|^2\right)\delta\left[b(x_i^* - u)\right] \quad (1)$$

$u$ is probability of appearance gray-level, and $b(x_i^*)$ is a bin of $i$-th pixel. $C$ is normalization constant. Modified histogram ascribes bigger values to pixels nearer to the center of tracking object. This can be achieved due using of monotonic decreasing kernel.

Let we assumed that a part of picture where is candidate for object in current frame, is moved for $y$. Modified histogram of object candidate can be defined as

$$\hat{p}_u(y) = C_h\sum_{i=1}^{n_h} k\left(\left\|\frac{y-x_i}{h}\right\|^2\right)\delta\left[b(x_i^* - u)\right] \quad (2)$$

$h$ is kernel height and $C_h$ is normalization constant.

An idea of object detection using "mean shift" procedure is based on the position determination that candidate for object which modified histogram and modified histogram of object model has a maximum likelihood. As a measure of likelihood, it has been used a function of likelihood, defined as

$$\hat{\rho}(y) \equiv \rho\left[\hat{p}(y), \hat{q}\right] = \sum_{u=1}^{m}\sqrt{\hat{p}_u(y)\hat{q}_u} \quad (3)$$

The distance between modified histograms defines as

$$d(y) = \sqrt{1-\rho\left[\hat{p}(y), \hat{q}\right]} \quad (4)$$

Using Taylor expansion around $\hat{p}(y_0)$, function of likelihood is:

$$\rho[\hat{p}(y), \hat{q}] \approx 0.5\sum_{u=1}^{m}\sqrt{\hat{p}_u(\hat{y}_0)\hat{q}_u} + \frac{C_h}{2}\sum_{i=1}^{n_h} w_i k\left(\left\|\frac{y-x_i}{h}\right\|^2\right) \quad (5)$$

Weighting coefficients can be calculate as

$$w_i = \sum_{u=1}^{m}\delta\left[b(x_i) - u\right]\sqrt{\frac{\hat{q}_u}{\hat{p}_u(\hat{y}_0)}} \quad (6)$$

A minimum distance between modified histograms is obtained by maximization second part in (5). If it is used an Epanechnikov kernel, object position for which this distance is minimal, can be defined as

$$y_{j+1} = \frac{\sum_{i=1}^{n} x_i w_i}{\sum_{i=1}^{n} w_i} \quad (7)$$

Algorithm for detection object using "mean shift" procedure execute through few steps, where are a priori known a modified histogram of object model $\boldsymbol{q_0}$ and object position in the previous frame $\boldsymbol{y_0}$

Step 1: calculation of modified histogram of object candidates in the position where object was in the previous frame

Step 2: calculation of weighting coefficients using (6)

Step 3: object position in the current frame is calculated using an Epanechnikov kernel using (7)

Step 4: if it is a diference of succesive object position smaller than previosly defined ε, the object position is defined in the last position. If this diference is grater than ε, step 1 is repeated, , where modified histogram calculating is done on the position determinate in step 3.

In order to make possible a changing of object dimension, algorithm must be done three times for diferent kernel heights. Kernel height for which likelihood function is maximalized is kernel height in the next frame

## 3 Results

In this paper is done quantitative analysis of changing background statistical feature influence through two experiments. Analysis is implemented by a detection error and a number needed iteration for determination position using "mean shift" procedure. Detection error defines as:

$$d = \sqrt{(x_d - x_r)^2 + (y_d - y_r)^2} \quad (8)$$

$(x_d, y_d)$ is determinate object position and $(x_r, y_r)$ is real object position in current frame.

In first experiment object is a trapeze with basis 40 and 20 pixels and height 20 pixels. This object is on the position (200,100) in the first frame. Background is modelled by a Gaussian probability density function mean value 97.7 and shifty standard deviation, which is constant during one sequence. During the experiment there are analyzed 50 sequences, where the change of standard deviation is 1 from sequence to sequence. Initial value of standard deviation is one. Illumination of object can be defined as:

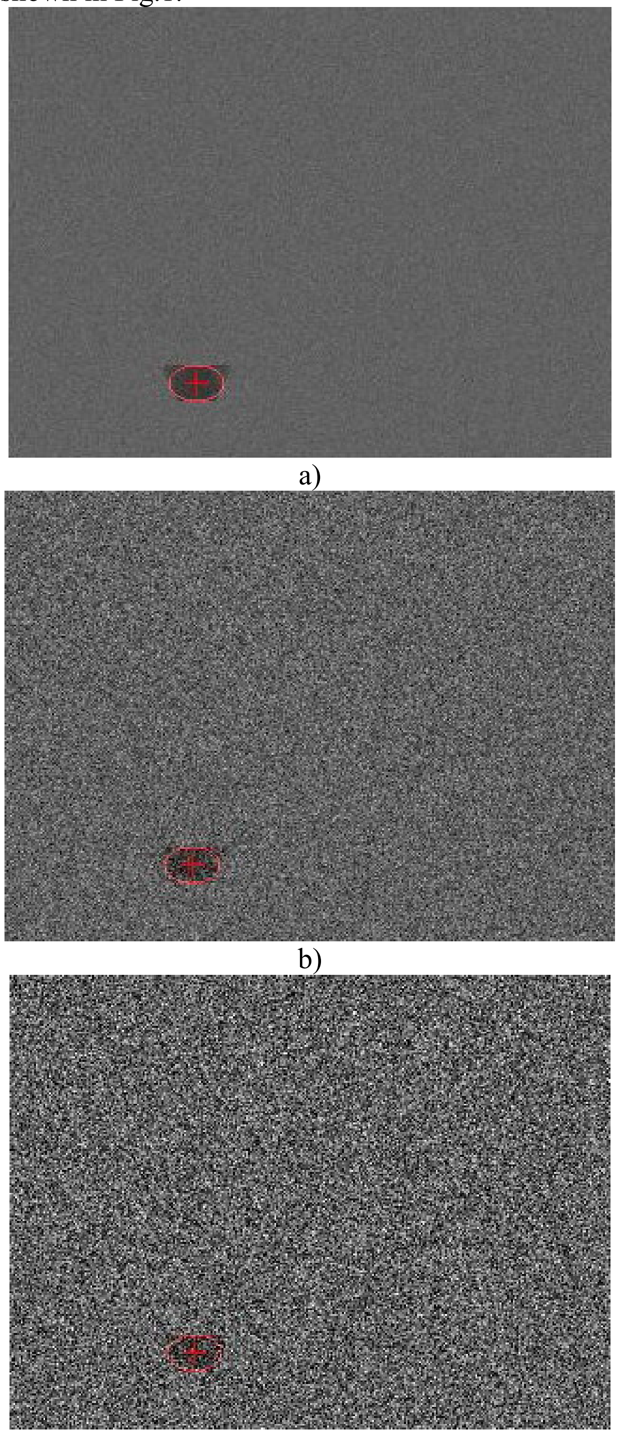
$$I(x,y) = I_{\max} e^{\frac{(x^2+y^2)}{2}} \quad x,y \in R_d \quad (9)$$

$I_{max}$ is a maximum illumination and its value is 55.3, and Rd is an object definition region. Object of interest move with velocity of 1 pixel/frame from left to right. Each sequence has 100 frames.

Modified histogram calculates in 256 bins. If the distance of successive position is smaller than 0.25 pixels, this procedure is stopped.

In this paper, it is used an Epanechnikov kernel height (15, 10) pixels on the initial position (200,100) pixels.

Initial frames of 5th, 25th and 48th sequence is shown in Fig.1.



a)

b)

c)

Fig.1: Initial frames of: a) 5th, b) 25th and c) 48th sequence

It has been calculated the detection error, (8). Detection error graphic illustration in 5th, 25th and 48th is shown in Fig.2.

It is noticeable that with increasing a variation, detection errors increase, too. Detection error mean values are 1.849, 4.588 and 12.7 pixels, respectably, and standard deviations of these errors are 0.5942, 1.489 and 5.216, respectably.

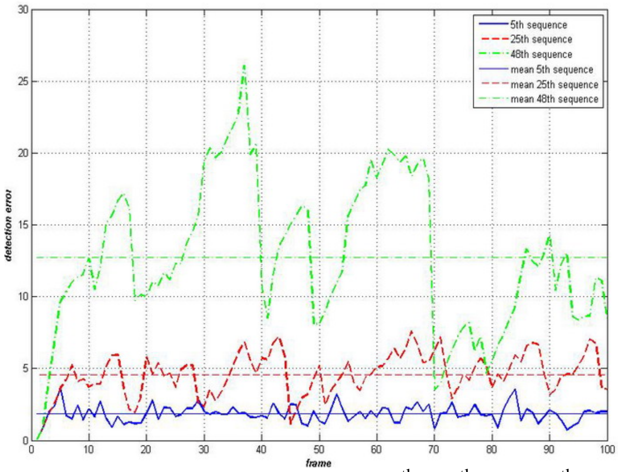



Fig.2: Detection error in 5th, 25th and 48th sequence

It is noticeable that with increasing a variation, detection errors increase, too. Detection error mean values are 1.849, 4.588 and 12.7 pixels, respectably, and standard deviations of these errors are 0.5942, 1.489 and 5.216, respectably.

On the ground of calculated detection error, it has been determine the detection error mean values in each sequence, what is shown in Fig.3.

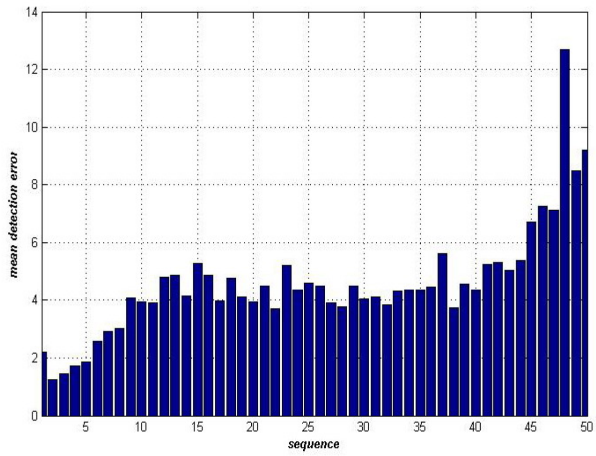



Fig.3: Detection error mean value in analyzed sequences

It is noticeable that after initial grows of detection error, there is not major variations detection error mean value, for standard deviations from 10 to 40. For greater values of standard deviations, detection error mean value grow. It is granted by experiments that tracking objects using "mean shift" procedure cannot be used for standard deviations greater than 70.

The graphic illustrations of needed number iterations for detecting object are shown in Fig.4.

According to Fig.4 it is noticeable that when standard deviations increase, needed number of iterations increase, too. Numbers of iterations mean values are 2.26, 2.73 and 4.47, respectable, and standard deviations of needed number of iterations are 1.661, 2.785 and 6.762, respectable.

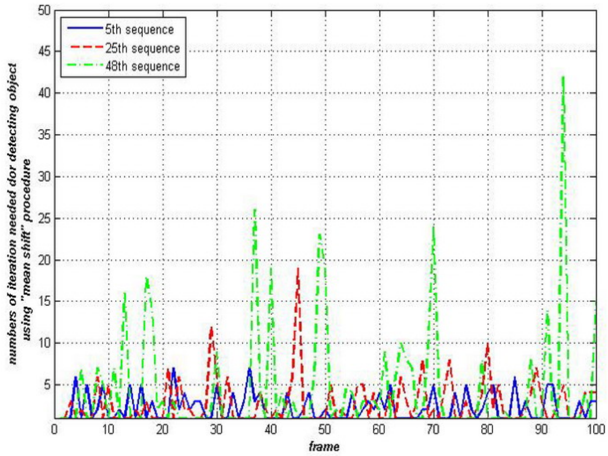



Fig.4: Needed number iterations for detecting object in 5th, 25th and 48th sequence

Graphic illustration numbers of iterations mean value used for determines object position using "mean shift" procedure is shown in Fig.5.

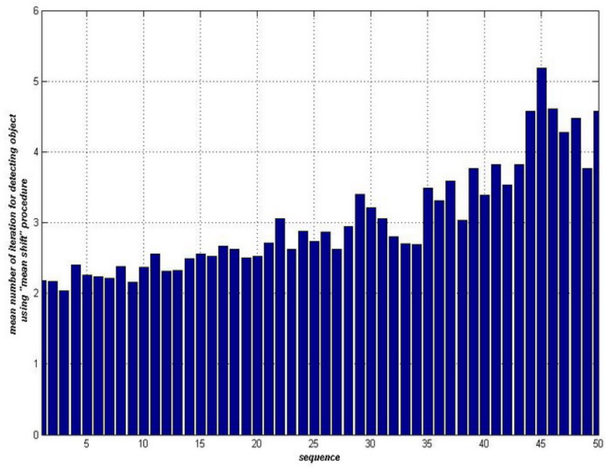



Fig.5: Mean value of needed number iterations for determinate object position in analyzed sequences

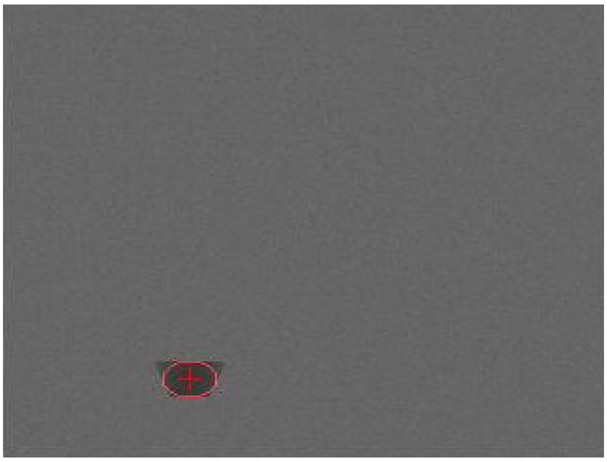
It is noticeable that with increasing standard deviations, mean value of needed number iterations increase.

Second experiment is similar to the first one. In second experiment background standard deviation is constant and it is 10. Changes of background illumination mean value goes from 20 with the step two due sequences. Initial frames of 5th, 14th and 36th sequences are shown in Fig.6.

Graphic illustration of detection error in 5th, 14th and 36th sequence is shown in Fig.7.

a)

b)



c)

Fig.6**:** Initial frames of: a) 5th, b) 14th and c) 36th sequence

According to Figure 7 it is noticeable that tracking object using "mean shift" procedure is impossible for 14th sequence. Detection error mean values in 5th, 14th and 36th sequence are 1.374, 7.372 and 1.43 pixels, respectable, and standard deviations of these errors are 0.521, 7.633 and 0.527, respectable.

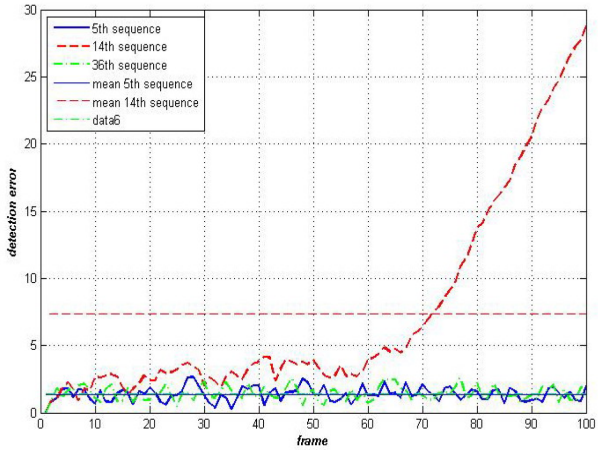



Fig.7**:** Detection error in 5th, 14th and 36th sequence

On the ground of calculated detection error, it has been determine the detection error mean values in each sequence, what is shown in Fig.8.

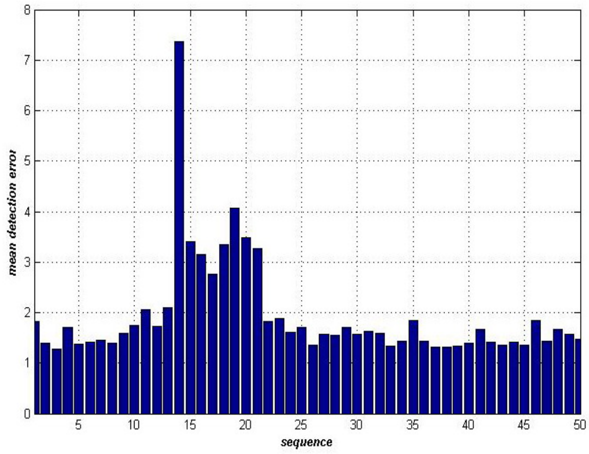



Fig.8**:** Detection error means value in analyzed sequences

It is noticeable that there is a greater detection error from 14th to 21st sequences. In other sequences detection error mean values are no significant. Greater detection error mean value can be explained as likelihood of background and object statistical features.

Graphic illustration of needed number iteration for detecting object in 5th, 14th and 36th sequence is shown in Fig.9.

According to Fig.9 it is noticeable that there are no greater variances of needed number iterations for detecting objects using "mean shift" procedure. Mean values of needed number iterations for 5th, 14th and 36th sequence are 2.14, 1.98 and 2.12, respectable, and standard deviations for those sequences are 1.365, 1.101 and 1.274, respectable.

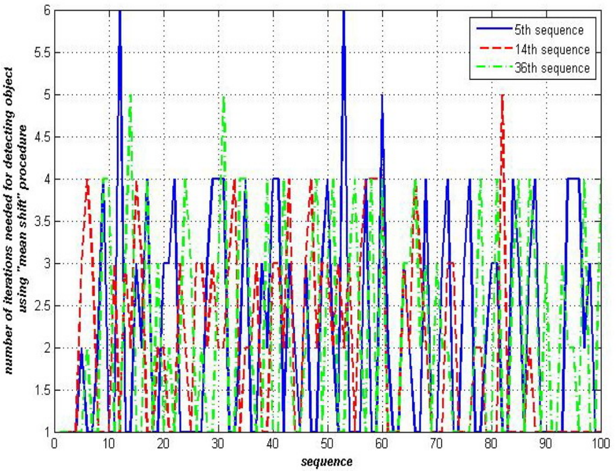



Fig.9**:** Needed number iterations for detecting object in 5th, 25th and 48th sequence

Graphic illustration of needed number of iterations mean value is shown in Fig.10.

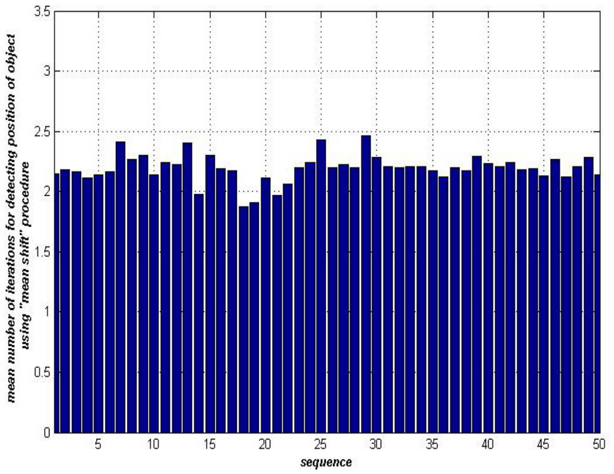



Fig.10**:** Mean value of needed number iterations for determinate object position in analyzed sequences

According to Fig.10 it is noticeable that needed number of iterations for detecting object using "mean shift" procedure is independent from changing background illumination mean value.

## 4 Conclusion

On the presented analysis, it can conclude that increasing illumination variance in the scene, causes increasing detection error and needed number of iteration for detecting object using "mean shift" procedure. On the other hand, if object illumination mean value is similar to the background illumination mean value, detection error is quickly increased and needed numbers of iterations are approximately same.

Further work will be direct to the problem of initialization detection algorithm and to the statistical classification objects that are tracked using "mean shift" procedure.

*References:*

[1] D. Comaniciu, P. Meer, Mean Shift: A Robust Approach Toward Feature Space Analysis, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 24, No.5, May 2002.

[2] H. Tek, D. Commaniciu, J. Williams, Vessel Detection by Mean Shift Ray Propagation, *Work, and Math, Models in Biomedical Image Analysis*, 2001.

[3] D. Comaniciu, V. Ramesh, P. Meer, Kernel-Based Object Tracking, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 25, No.5, May 2003.

[4] D. Comaniciu, V. Ramesh, P. Meer, Real-time tracking of non-rigid objects using mean shift, *IEEE Conf. on Computer Vision and Pattern Recognition*, Hilton Head, SC, pp. 142-149, 2000