Target Tracking Based on Mean Shift and KALMAN Filter with Kernel Histogram Filtering

Sara Qazvini Abhari (Corresponding author*)*

Faculty of Electrical, Computer and IT Engineering Islamic Azad University – Qazvin Branch

Qazvin-Iran

Tel: 98-91-2380-7173   E-mail: s.qazvini@qiau.ac.ir

Towhid Zargar Ershadi

Faculty of Electrical, Computer and IT Engineering Islamic Azad University – Qazvin Branch

E-mail: zargar@qiau.ac.ir

Abstract

Visual object tracking is required in many tasks such as video compression, surveillance, automated video analysis, etc. mean shift algorithm is one of popular methods to this task and has some advantages comparing to other tracking methods. This method would not be appropriate in the case of large target appearance changes and occlusion; therefore target model update could actually improve this method. KALMAN filter is a suitable approach to handle model update. We performed mean shift algorithm with model update ability for tracking in this paper and achieve good results.

Keywords: KALMAN filter, Mean shift, Tracking, Hypothesis testing

1. Introduction

Visual object tracking is required in many tasks such as video compression, surveillance, automated video analysis, etc. Object tracking should deal with some challenges like occlusion, target appearance changes, cluttering. Mean shift algorithm is an efficient tracking method which has been employed in the literature (Wang and Yagi Wang,2009; Comaniciu and Meer, 2003; Collins, 2004; Xin and Song, 2006; Sen, Wei, and Sen, 2009). This algorithm needs less computational complexity compared with methods performed in the whole image. In the traditional mean shift algorithm, the target model is unchanged during the tracking and belongs to the first frame, so this algorithm is not robust in the case of large changes of appearance, size and direction of the target. In order to compensate the target size changes, Comaniciu (Comaniciu and Meer, 2003) changes the Kernel size about 10% and selected the kernel with maximum similarity value and (Collins, 2003) used scale space. Peng et al. (Peng N.,Yang J.& Liu Z. ,2005) integrated adaptive KALMAN filter into mean shift algorithm to update the target model to cope with target appearance changes. Target model update in each frame makes the tracker sensitive to occlusion and noise. So model update should be performed with some cautions. Nguyen (Nguyen, Worring and van den Boomagaard, 2001) used filter residual to test model update and (Peng, Yang and Liu, 2005) used filter residual to test the update model hypothesis in the form of t distribution.

This paper is organized as follows. The mean shift algorithm is reviewed in section. KALMAN filter is considered in section. Model update is coming in section. Results are in section.

2. Mean shift algorithm

Mean shift algorithm is a nonparametric mode seeking method introduced by Fukanaga (Fukunaga and Hostetler, 1975) and considered by Cheng (Cheng, 1995). Comaniciu (Comaniciu and Meer, 2003) incorporated mean shift algorithm into tracking. This method optimizes the similarity measure between target model and candidate model in the consecutive frames.

To characterize the target, first a feature space should be defined. This feature space could be selected as the color space of the target. Among the color spaces, RGB has widely been used.

2.1 Object representation

The object model is represented by its probability density function (pdf) in the feature space. In order to decrease computational cost imposed by real-time processing, m-bin histograms are so desirable. The object model is called target model and candidate model in reference frame and subsequent frames respectively. Fig 1 illustrates the histogram construction of a 3×3 gray scale image.

Target model is defined as:

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

Where x_i is the normalized coordinates of pixels, δ is the dirac function. C is normalization constant to keep $\sum_{u=1}^{m} \hat{q}_u = 1$ condition. k is the kernel which Gaussian and Epanechnikov ones are more popular and defined as follows:

Epanechnikov kernel:

$$k(x) = \begin{cases} \frac{1}{2} c_d^{-1}(d+2)(1-x) & x \le 1 \\ 0 & else \end{cases} \quad (2)$$

Gaussian kernel:

$$k(x) = exp\left(-0.5 * \left\|\frac{y - x_i}{h}\right\|^2\right) \quad (3)$$

Candidate model is defined as:

$$(4)$$

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

As it could be seen, the only difference between the target and candidate models is h, the scale of kernel.

Mean shift algorithm optimizes the similarity between target and candidate models. Bhattacharyya coefficient is a common similarity measurement and is defined as:

$$\hat{\rho}(y) = \rho(\hat{p}(y), \hat{q}) = \sum_{u=1}^{m} \sqrt{\hat{p}_u(y) \hat{q}_u} \quad (5)$$

This coefficient could be considered as the cosine of the angle between the unit vectors $\sqrt{\hat{p}_1, \ldots, \hat{p}_m}$ and $\sqrt{\hat{q}_1, \ldots, \hat{q}_m}$. Bhattacharyya coefficient optimization leads to mean shit algorithm as the following equation (see Comaniciu D., 2003 for more details).

$$d = \frac{\sum_{i=1}^{n_h} x_i w_i g\left(\left\|\frac{y_0 - x_i}{h}\right\|^2\right)}{\sum_{i=1}^{n_h} w_i g\left(\left\|\frac{y_0 - x_i}{h}\right\|^2\right)} \quad (6)$$

Where d is the displacement vector of kernel center, $g(x) = -k'(x)$ and the weight w_i is defined as:

$$w_i = \sum_{u=1}^{m} \sqrt{\frac{\hat{q}_u}{\hat{p}_u(\hat{y}_0)}} \, \delta(b(x_i) - u) \quad (7)$$

If the Epanechnikov kernel is used, mean shift equation is simplified to:

$$d = \frac{\sum_{i=1}^{n_h} x_i w_i}{\sum_{i=1}^{n_h} w_i} \quad (8)$$

KALMAN filter

KALMAN filter was first introduced by KALMAN (Kalman, 1960) to estimate the state of a system. This filter has been used to estimate the position of target in many literature (Comaniciu., 2003; Kalman, 1960; Bai, 2010; Zhu, 2002; Zhao Qiao and Men, 2009). KALMAN filter consists of two stages called prediction step and correction step. The state of system is predicted using state equation in the prediction step and is corrected using measurement equation in the correction step.

The State and measurement equations of a linear KALMAN filter are respectively as follows:

$$x_k = A_k x_{k-1} + B_k u_k + w_{k-1} \qquad (9)$$

$$z_k = H_k x_k + v_k$$

Where x_k and x_{k-1} are the state vector of current and previous time respectively. $w_{k-1} \sim N(0, Q)$ and $v_k \sim N(0, R)$ are the system and measurement noise. Performance of KALMAN filter depends on two key parts: the reasonable linear model and the initial parameters including the noise statistical properties and the initial state variance (Peng ,Yang and   Liu , 2005).

Adaptive KALMAN filter is used in (Peng.,Yang, and Liu, 2005) to update the target model. State and error variance prediction equations are as follows:

$$\hat{q}_b^{i|i-1} = \hat{q}_b^{i-1} \qquad (10)$$

$$P_b^{i|i-1} = P_b^{i-1} + Q^i$$

Where $\{\hat{q}_b^i\}_{b=1,...,m}$ is the target model and $\{P_b^i\}_{b=1,...,m}$ is the variance of prediction error. $b = 1, ..., m$ and i are the histogram bins and frame index respectively. KALMAN gain is calculated as:

$$K_b^i = \frac{P_b^{i|i-1}}{P_b^{i|i-1} + R} \qquad (11)$$

And finally state and error variance correction equations are:

$$\hat{q}_b^i = \hat{q}_b^{i|i-1} + K_b^i \left(p_b^i - \hat{q}_b^{i|i-1} \right) \qquad (12)$$

$$P_b^i = \left(1 - K_b^i \right) \hat{q}_b^{i|i-1}$$

It is important to tune a KALMAN filter with appropriate initial value of prediction error variance and state and measurement noises.   used KALMAN residual to do this.

KALMAN filter residual is the difference between predicted state and measurement as:

$$r = p_b^i - \hat{q}_b^{i|i-1} \qquad (13)$$

The variance of filtering residuals over last L frame is calculated as:

$$(\bar{r}^i)^2 = \frac{1}{mL} \sum_{l=i-L+1}^{i} \sum_{b=1}^{m} \left(p_b^l - q_b^{l|l-1} \right)^2 \qquad (14)$$

Then measurement noise variance is assumed constant but variance of state noise is estimated as:

$$Q^i = (\bar{r}^i)^2 - R - \frac{1}{m}\sum_{b=1}^{m} q_b^{i-1} \qquad (15)$$

Which is not constant, therefore the used KALMAN filter is adaptive. The initial value for variances of state noise, measurement noise and prediction error are calculated (Peng ,Yang and Liu, 2005) as:

$$Q^1 = 0 \qquad\qquad (16)$$

$$R = 0.5(r^1)^2 \qquad\qquad (17)$$

$$P^1 = 0.5(\bar{r}^1)^2 \qquad\qquad (18)$$

4. Hypothesis testing

Over model update could be as harmful as under model update and some considerations should be taken to avoid that. Nguyen (Nguyen ,Worring and van den Boomagaard, 2001) used filter residual to avoid over update model. (Peng ,Yang and Liu, 2005) used t distribution of residual population to test the hypothesis of model update. Let suppose filter residuals $\{r_t\}_{t=1,...,m}$ be the samples from totality T with normal distribution $T \sim N(\mu, \sigma^2)$, since the σ^2 is unknown t distribution is used as follows:

$$t = \frac{\bar{r} - \mu_0}{s/\sqrt{N}} \qquad\qquad (19)$$

Where N is the number of samples and s is the unbiased estimation of standard deviation.

If the null hypothesis H_0 is accepted, model update is performed; else the hypothesis H_1 would be accepted and previous target model would be held. The following conditional probability assigns the critical region.

$$P(refuse\ H_0|H_0\ is\ true) = P\left(\frac{\bar{r} - \mu_0}{s/\sqrt{N}} \geq cr\right) = \alpha \quad (20)$$

The critical region is defined as:

$$t = \frac{\bar{r} - \mu_0}{s/\sqrt{N}} \geq t_\alpha \qquad\qquad (21)$$

Where μ_0 is set to $\bar{r}^i$ and the critical value t_α is determined from t distribution table.

We have made some modifications to mean shift-KALMAN filter method of (Peng ,Yang and Liu, 2005). If the large appearance changes of target are not due to the occlusion, the target model should be updated. We supposed at most 6 frames for occlusion duration, and if the number of model update hypothesis rejections (H_1) exceeds this number, model update procedure is performed.

5. Results

We have performed mean shift algorithm with model update capability to two image sequence here. In the first experiment the head of a tennis playerhas been tracked (Figure.2). Since there is no occlusion in this sequence model update procedure could be omitted. To construct the target and candidate models 16 quantization levels is considered; therefore the histogram size is be 16×16×16. We have used mean shift algorithm with variable kernel size in the first experiment. As could be seen, the result is satisfying. Bhattacharyya coefficient is demonstrated in Fig. 3 for more considerations.

In the next experiment we want to investigate the operation of model update in the mean shift algorithm. Model update comes in useful especially in target occlusion cases. To prove this claim, tracking a person who severely occluded by another person is demonstrated in the third and forth figures.

Figure 4 shows the tracking result of a person using mean shift algorithm without target model update. As it could be seen, the tracking is completely challenging because the target appearance severely changes and occlusion has occurred in some frames. Therefore as it is expected, target is totally lost in the middle frames of

Fig. 4. In the next experiment, mean shift algorithm with model update capability is performed to track the person and the result is shown in Fig. 5. As it could be seen, tracking is satisfying and the man has been fully tracked even in the case of occlusion. The tracked path of person is shown in Fig. 6.

Only the histogram bins with none zero values are participated in the model update process and the last three frames is used for hypothesis testing i.e. *L=3*. We used significance level α equal to 0.1 and t_α is assumed equal to 1.3 in the hypothesis testing stage because 21 nonzero value of target model is used (t distribution table).

Bhattacharyya coefficient of tracking of the figures 4 and 5 is demonstrated in Fig.7. Histogram of KALMAN filter residuals of a frame is brought in Fig.8. As it is representing, the normal assumption which was made in the last section is met. Fig.9 shows the t value of each frame and t_α. As it is expected, the *t* value exceeds t_α in some frames in which large appearance changes and occlusion occurrs and model update is not performed in these frames. To consider the KALMAN filter operation, variance of state noise and averaged variance of prediction error are shown in Fig. 10. Variance of state noise suddenly raises in the frames with large appearance change and occlusion to cope the KALMAN filter with them.

References

Bai K. (2010). A new object tracking algorithm based on Mean Shift in 4-D state space and On-line feature selection,*Third International Conference on Information and Computing, IEEE* .

Cheng Y. (1995). Mean shift, mode seeking, and clustering, *IEEE Transactios on Pattern Analysis and Machine Intelligence.*, 17(8), 790–799.

Comaniciu D. and Meer p. (2003). Kernel-based object tracking, *IEEE Transactios on Pattern Analysis and Machine Intelligence.*, 25(5), 564–577.

Collins R. T. (2003). Mean shift blob tracking through scale space, *IEEE International Proceeding on Computer Vision Pattern Recognition,* 2, 234–240.

Collins R. T. (2004). Mean-shift blob tracking through scale space, *IEEE Computer Society Conference on Computer Vision and Pattern*.

Fukunaga K., Hostetler L. D. (1975). The Estimation of the gradient of a density function, with applications in pattern recognition, *IEEE Transaction on Information Theory*, 21, 32-40.

Kalman R. E. (1960). A new approach to linear filtering and prediction problems, *Transaction of the ASME Journal of Basic Engineering*, 35-45.

Noceti N., Destrero A., Lovato A., & Odone F. (2009). Combined motion and appearance models for robust object tracking in real-time*, Advanced Video and Signal Based Surveillance.*

Nguyen H. T., Worring M. & van den Boomagaard R. (2001). Occlusion robust adaptive template tracking. *IEEE International Conference Computer Vision,* 1, 678–683.

Peng N.,Yang J. & Liu Z. (2005). Mean shift blob tracking with kernel histogram filtering and hypothesis testing, *Pattern Recognition Letters,* 26, 605–614.

Sen G., Wei L., Xin L. & Sen L. Y. (2009). An adaptive weight values updating mean shift tracking algorithm, *Proceedings of the IEEE International Conference on Information and Automation*.

Wang J. & Yagi Y. (2009). Adaptive mean-Shift tracking with auxiliary particles, *IEEE Transaction on Systems, Man, Cybernetics.*, 39(6).

Xin P. J. and Song P. N. (2006). Adaptive kernel based tracking using mean-shift,„*Springer-Verlag Berlin Heidelberg* .

Zhu Z., Ji Q., Fujimura K. (2002). Combining KALMAN filtering and mean shift for real time eye tracking under active IR illumination, *IEEE International Conference on Pattern* Recognition, 4, 318-321.

Zhao J., Qiao W., Men G. (2009). An Approach based on mean shift and KALMAN filter for target tracking under occlusion, *Proceedings of the Eighth International Conference on Machine Learning and Cybernetics.*
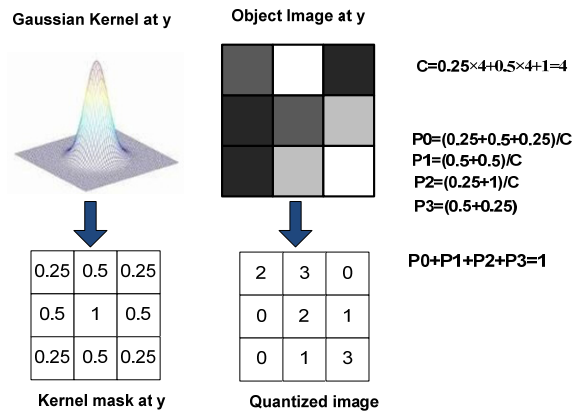
Figure 1. histogram construction of a gray scale image (Peng N., 2005)
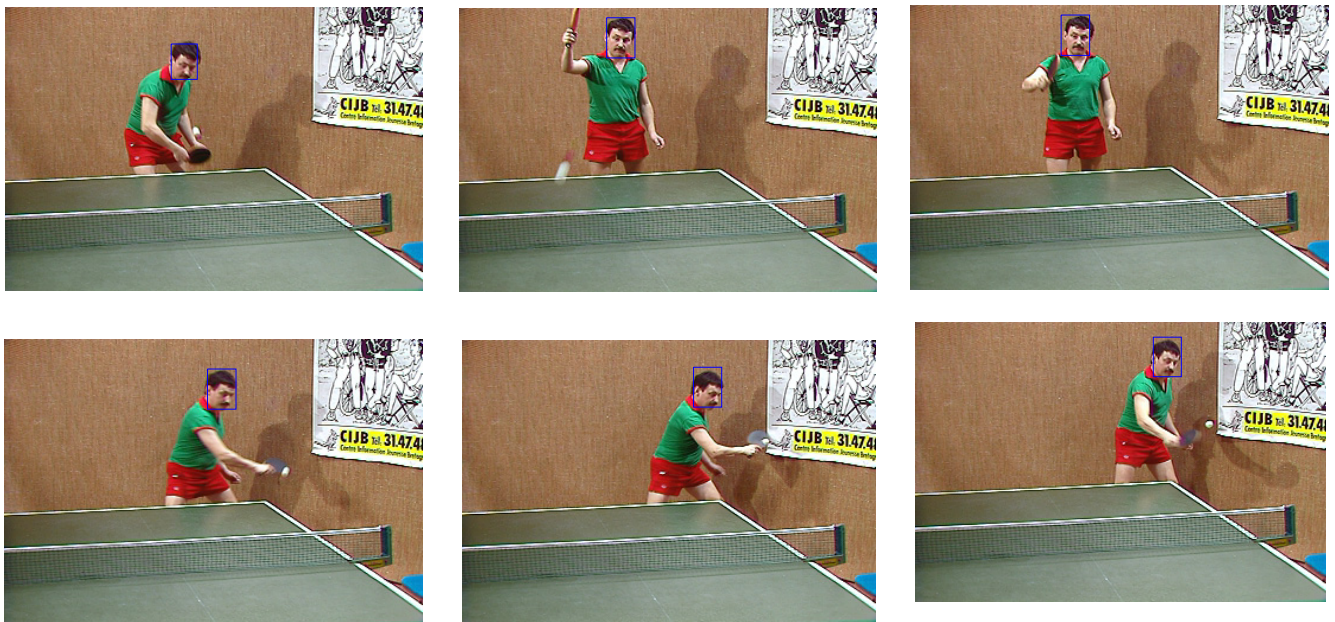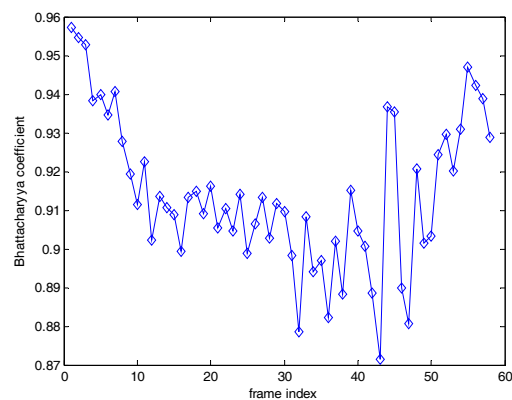

Figure 2. Tracking of tennis player, using mean shift algorithm with variable kernel size in frames

1,13,21,47,53,58

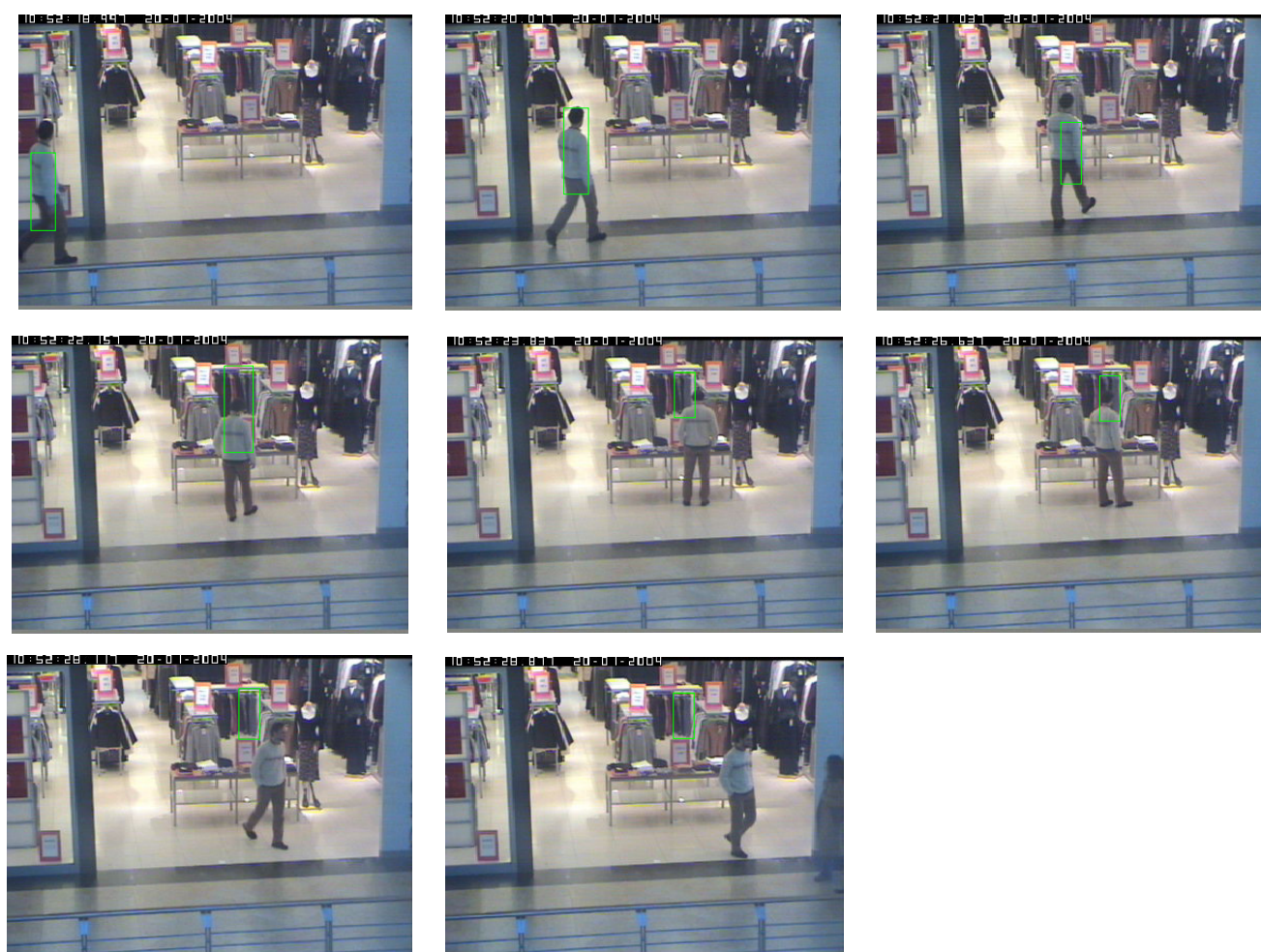Figure 3. Bhattacharyya coefficient of tennist tracking

Figure 4. Tracking of a person using mean shift without model update in frames 1,14,26,40,61,96,115,124
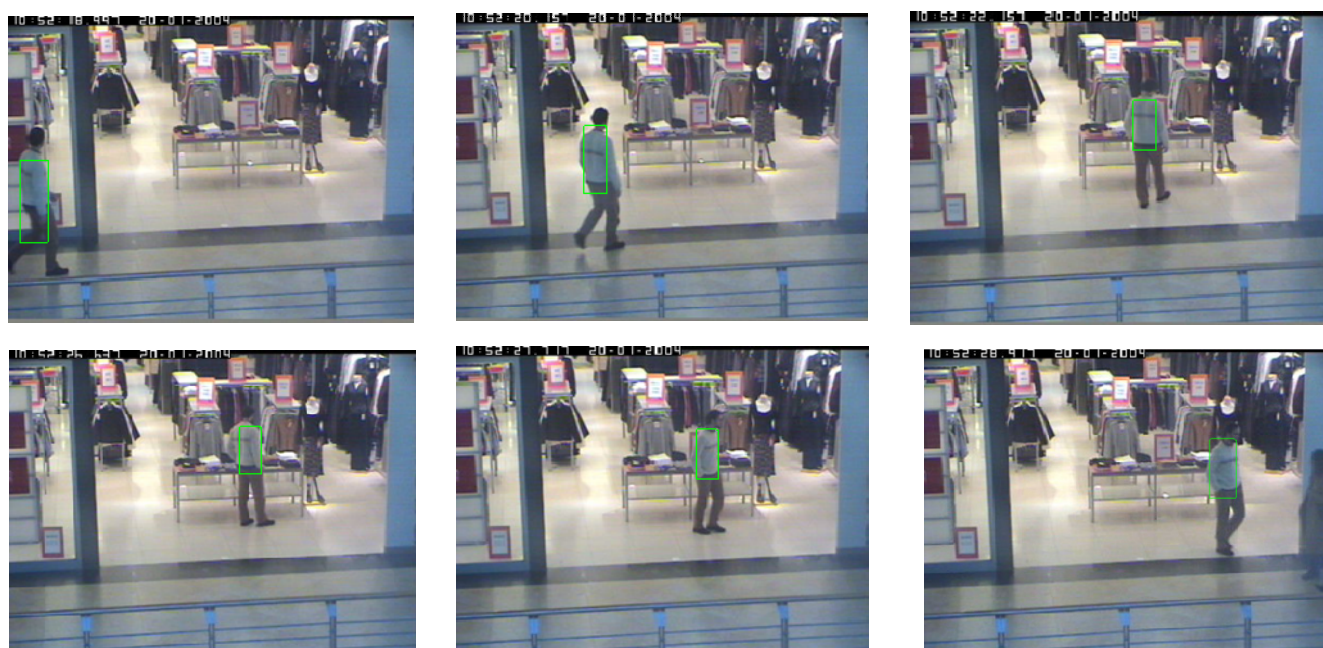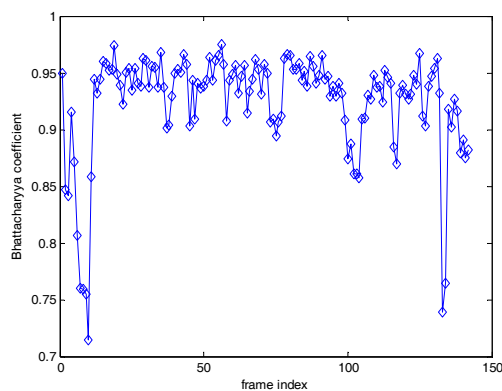
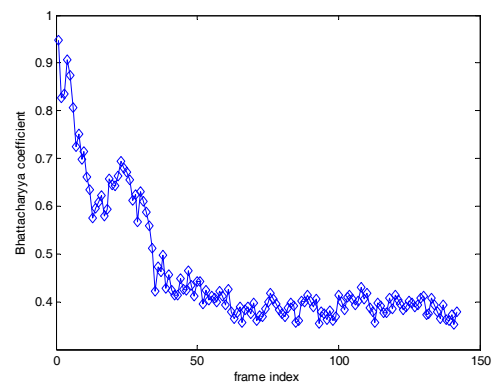Figure 5. Tracking of a person using mean shift with model update in frames

1,15,40,96,110,125,131,135,137,139,142

Figure 6. Tracked path of a person using mean shift with model update in frames

(a)

(b)

Figure 7. Bhattacharyya coefficient of tracking a) with model update b) without model update
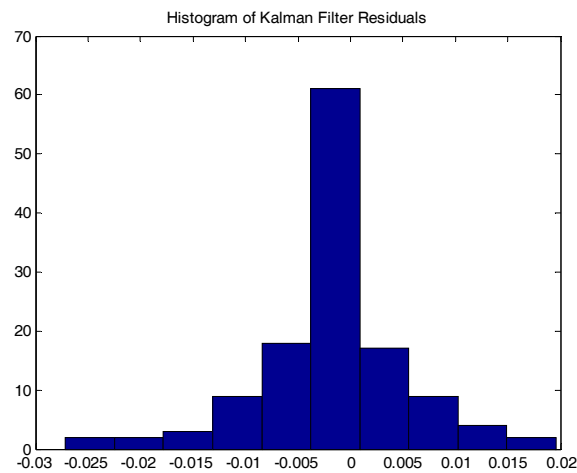
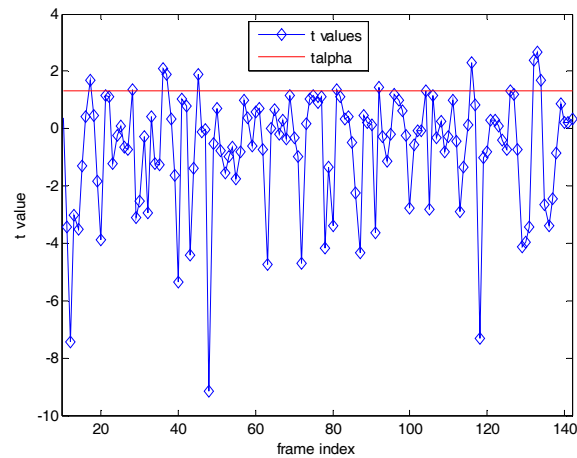Figure 8. histogram of residual of a frame
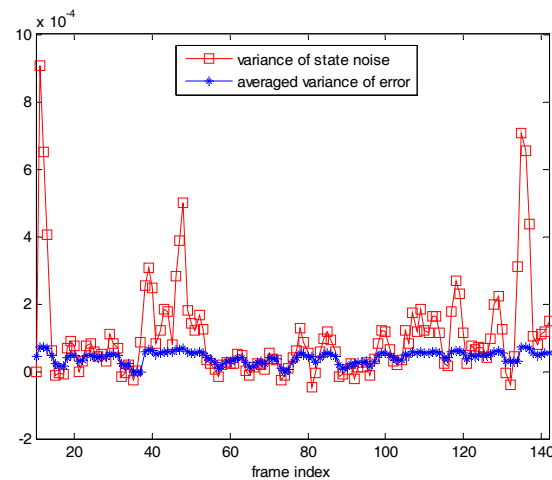

Figure 9. t distribution values of each frame

Figure 10. variance of state noise and averaged variance of prediction error