Target Tracking using Kalman Filter

Bingbing Hu[1]
[1]SIST of ShanghaiTech University, Shanghai, China

Target tracking is often complicated by the measurement noise. The noise must be filtered out in order to predict the true path of a moving target. In this study of linear filtering, the Kalman filter, a recursive linear filtering model, was used to estimate tracks. Kalman filter was successful in smoothing random deviations from the true path of the targets, improving in its ability to predict the path of each target as more measurements from the tracker were processed.

***Index Terms*—Kalman filtering, Target tracking, Image segmentation.**

I. Introduction

Random process or signal is ubiquitous in engineering. On the one hand, any deterministic signal, after being measured, will often introduce a random error so that it will be randomized. On the other hand, any signal itself has random interference. A signal is called noise if it interfered signal or system functions. Power spectral density (PSD) divides noise into white noise and color noise, and we call a zero mean white noise pure random signal. Therefore, any random signal can be regarded as a mixture of pure random signal and deterministic signal, or just called a random signal.

We are concerned about extract useful signal from a the mixture signal with noise. Wiener and Kalman filter are some of the methods to solve the problem. Kalman filter uses a series of measurements observed over time, containing statistical noise and other inaccuracies, and produces estimates of unknown variables that tend to be more accurate than those based on a single measurement alone, by estimating a joint probability distribution over the variables for each timeframe.

Therefore, we can use Kalman filter to track moving target. The detection and tracking of moving objects in sequence images means to find and extract moving objects in real time in a continuous sequence of images, and the moving objects are continuously tracked according to the changes of the edge, local motion and gray level of the target to obtain the track of the moving objects. This provides the data of target recognition, motion analysis and other processing for the next stage. It is an important subject in computer vision. It has a wide range of applications in military visual guidance, video surveillance, traffic flow observation, robot navigation and video image compression and transmission. The commonly used moving target detection methods in video surveillance mainly include the algorithm based on the adjacent frame difference or based on the difference between background image and the current frame. Most of the methods based on background difference mainly involve the estimation of background image and the real-time updating. When the moving objects changes it state form moving to static in a long time, due to the real-time updating of the background, it is possible to mistake the moving object into the background image. Moreover, when the target moves suddenly from a standstill, subtracting the current frame from the background may obtain a false detection results. In order to overcome these shortcomings, we use Kalman filter to predict the area of the moving target may be in next time slot. This can narrow the search space so as to achieve the fast track of moving target.

II. Kalman Filter

One can use a Kalman filter in any place where one have uncertain information about some dynamic system, and one can make an educated guess about what the system is going to do next. Even if messy reality comes along and interferes with the clean motion one guessed about, the Kalman filter will often do a very good job of figuring out what actually happened. And it can take advantage of correlations between crazy phenomena that one maybe wouldnt have thought to exploit! Kalman filters are ideal for systems which are continuously changing. They have the advantage that they are light on memory (they dont need to keep any history other than the previous state), and they are very fast, making them well suited for real time problems and embedded systems.

Now we will give a detail introduction to the Kalman filter. For convenience, we assume the data model is Gauss-Markov, which has the form

$$s[n] = as[n-1] + u[n] \quad n \geq 0 \tag{1}$$

Now we want to estimate $s[n]$ based on the data $\{x[0], x[1], \ldots, x[n]\}$ as n increases. This is called filtering. The approach computes the estimator $\hat{s}[n]$ based on the estimator for the previous time sample $\hat{s}[n-1]$ and so is recursive in nature. This is so-called Kalman filter. Consider the scalar state equation and the scalar observation equation

$$\begin{aligned} s[n] &= as[n-1] + u[n] \\ x[n] &= s[n] + w[n] \end{aligned} \tag{2}$$

where $u[n]$ is zero mean Gaussian noise with independent samples and $\mathbb{E}(u^2[n]) = \sigma_u^2$, $w[n]$ is zero mean Gaussian noise with independent samples and $\mathbb{E}(w^2[n]) = \sigma_n^2$. Furthermore, we make two assumptions:

(i) $s[-1], u[n]$ and $w[n]$ are all independent,
(ii) $s[-1] \sim \mathscr{N}(\mu_s, \sigma_s^2)$.

The noise process $w[n]$ differs from WGN only in that its variance is allowed to change with time. To simplify the derivation we assume that $\mu_s = 0$, so that $\mathbb{E}(s[n]) = 0$ for $n \geq 0$. We wish to estimate $s[n]$ based on the observations $\{x[0], \ldots, x[n]\}$ or to filter $x[n]$ to produce $\hat{s}[n]$. More generally, the estimator of $s[n]$ based on the observations $\{x[0], \ldots, x[m]\}$ will be denoted by $\hat{s}[n|m]$. Our criterion of optimality will be the minimum Bayesian MSE or

$$\mathbb{E}[(s[n] - \hat{s}[n|n])^2] \tag{3}$$

where the expectation is with respect to $p(x[0], x[1], \ldots, x[n])$. But the MMSE estimator is just the mean of the posterior PDF or

$$\hat{s}[n|n] = \mathbb{E}(s[n]|x[0], \ldots, x[n]) \tag{4}$$

which gives

$$\hat{s}[n|n] = \mathbf{C}_{\theta x}\mathbf{C}_{xx}^{-1}\mathbf{x} \tag{5}$$

since $\theta = s[n]$ and $\mathbf{x} = [x[0], \ldots, x[n]]^T$ are jointly Gaussian. Because we are assuming Gaussian statistics for the signal and noise, the MMSE estimator is linear and is identical in algebraic form to the LMMSE estimator. The algebraic properties allow us to utilize the vector space approach to find the estimator. The implicit linear constraint does not detract from the generality since we already know that the optimal estimator is linear. Furthermore, if the Gaussian assumption is not valid, then the resulting estimator is still valid but can only be said to be the optimal LMMSE estimator. Returning to the sequential computation of (5), we note that if $x[n]$ is uncorrelated with $\{x[0], \ldots, x[n-1]\}$, then from (4) and the orthogonality principle we will have

$$\begin{aligned} \hat{s}[n|n] &= \mathbb{E}(s[n]|x[0], \ldots, x[n-1]) + \mathbb{E}(s[n]|x[n]) \\ &= \hat{s}[n|n-1] + \mathbb{E}(s[n]|x[n]) \end{aligned} \tag{6}$$

which has the desired sequential form. Unfortunately, the $x[n]$ is correlated due to their dependence on $s[n]$, which is correlated from sample to sample.

There are some useful properties of MMSE estimator:

(i) The MMSE estimator of θ based on two uncorrelated data vectors, assuming jointly Gaussian statistics, is

$$\begin{aligned} \hat{\theta} &= \mathbb{E}(\theta|\mathbf{x_1}, \mathbf{x_2}) \\ &= \mathbb{E}(\theta|\mathbf{x_1}) + \mathbb{E}(\theta|\mathbf{x_2}) \end{aligned}$$

if θ is zero mean.

(ii) The MMSE estimator is additive in that if $\theta = \theta_1 + \theta_2$, then

$$\begin{aligned} \hat{\theta} &= \mathbb{E}(\theta|\mathbf{x}) \\ &= \mathbb{E}(\theta_1 + \theta_2|\mathbf{x}) \\ &= \mathbb{E}(\theta_1|\mathbf{x}) + \mathbb{E}(\theta_2|\mathbf{x}) \end{aligned}$$

Now let $\mathbf{X}[n] = [x[0], \ldots, x[n]]^T$ and $\tilde{x}[n]$ denote the innovation, noting that the innovation is jointlythe part of $x[n]$ that is uncorrelated with the previous samples $\{x[0], \ldots, x[n-1]\}$ or

$$\tilde{x}[n] = x[n] - \hat{x}[n|n-1] \tag{7}$$

This is because by the orthogonality principle $\hat{x}[n|n-1]$ is the MMSE estimator of $x[n]$ based on data $\{x[0], \ldots, x[n-1]\}$, the error or $\tilde{x}[n]$ being orthogonal with the data. The data $\mathbf{X}[n-1]$, $\tilde{x}[n]$ are equivalent to the original data set since $x[n]$ may be recovered from

$$\begin{aligned} x[n] &= \tilde{x}[n] + \hat{x}[n|n-1] \\ jointly &= \tilde{x}[n] + \sum_{k=0}^{n-1} a_k x[k] \end{aligned}$$

where a_k are the optimal weighting coefficients of the MMSE estimator of $x[n]$ based on $\{x[0], \ldots, x[n-1]\}$. Now we can rewirte (4) as

$$\hat{s}[n|n] = \mathbb{E}(s[n]|\mathbf{X}[n-1], \tilde{x}[n]) jointly \tag{8}$$

and because $\mathbf{X}[n-1]$ and $\tilde{x}[n]$ are uncorrelated, we have

$$\hat{s}[n|n] = \mathbb{E}(s[n]|\mathbf{X}[n-1]) + \mathbb{E}(s[n]|\tilde{x}[n]).$$

But $\mathbb{E}(s[n]|\mathbf{X}[n-1])$ is the prediction of $s[n]$ based on the previous data, and we denote it by $\hat{s}[n|n-1]$, thus have

$$\begin{aligned}\hat{s}[n|n-1] &= \mathbb{E}(s[n]|\mathbf{X}[n-1]) \\ &= \mathbb{E}(as[n-1]+u[n]|\mathbf{X}[n-1]) \\ &= a\mathbb{E}(s[n-1]|\mathbf{X}[n-1]) \\ &= a\hat{s}[n-1|n-1].\end{aligned}$$

Therefore we now have

$$\hat{s}[n|n] = \hat{s}[n|n-1] + \mathbb{E}(s[n]|\tilde{x}[n]) \tag{9}$$

where

$$\hat{s}[n|n-1] = a\hat{s}[n-1|n-1].$$

To determine $\mathbb{E}(s[n]|\tilde{x}[n])$ we note that it is MMSE of $s[n]$ based on $\tilde{x}[n]$. As such, it is linear, and because of the zero mean assumption of $s[n]$, it takes the form

$$\begin{aligned}\mathbb{E}(s[n]|\tilde{x}[n]) &= K[n]\tilde{x}[n] \\ &= K[n](x[n] - \tilde{x}[n|n-1])\end{aligned}$$

where

$$K[n] = \frac{\mathbb{E}(s[n]\tilde{x}[n])}{\mathbb{E}(x^2)}. \tag{10}$$

This follows from the general MMSE estimator for jointly Gaussian θ and x

$$\hat{\theta} = C_{\theta x}C_{xx}^{-1}x = \frac{\mathbb{E}(\theta x)}{\mathbb{E}(x^2)}x.$$

But $x[n] = s[n] + w[n]$, so that

$$\begin{aligned}\hat{x}[n|n-1] &= \hat{s}[n|n-1] + \hat{w}[n|n-1] \\ &= \hat{s}[n|n-1]\end{aligned}$$

since $\hat{w}[n|n-1] = 0$ due to $w[n]$ being independent of $\{x[0], \ldots, x[n-1]\}$. Thus,

$$\mathbb{E}(s[n]|\tilde{x}[n]) = K[n](x[n] - \hat{s}[n|n-1])$$

and from (9) we now have

$$\hat{s}[n|n] = \hat{s}[n|n-1] + K[n](x[n] - \hat{s}[n|n-1]) \tag{11}$$

where

$$\hat{s}[n|n-1] = a\hat{s}[n-1|n-1]. \tag{12}$$

It remains only to determine the gain factor $K[n]$. From (10) the gain factor is

$$K[n] = \frac{\mathbb{E}[s[n](x[n] - \hat{s}[n|n-1])]}{\mathbb{E}[(x[n] - \hat{s}[n|n-1])^2]}.$$

Moreover, we can derivate that

$$\begin{aligned}K[n] &= \frac{\mathbb{E}[(s[n] - \hat{s}[n|n-1])(x[n] - \hat{s}[n|n-1])]}{\mathbb{E}[(s[n] - \hat{s}[n|n-1] + w[n])^2]} \\ &= \frac{\mathbb{E}[(s[n] - \hat{s}[n|n-1])^2}{\mathbb{E}[(s[n] - \hat{s}[n|n-1])^2] + \sigma_n^2}\end{aligned} \tag{13}$$

But the numerator is just the minimum MSE incurred when $s[n]$ is estimated based on the previous data or the minimum one-step prediction error. We will denote this by $M[n|n-1]$, so that

$$K[n] = \frac{M[n|n-1]}{M[n|n-1] + \sigma_n^2} \tag{14}$$

Moreover, the minimum prediction error is

$$\begin{aligned}M[n|n-1] &= \mathbb{E}[(s[n] - \hat{s}[n|n-1])^2] \\ &= \mathbb{E}[(as[n-1] + u[n] - \hat{s}[n|n-1])^2] \\ &= \mathbb{E}[(a(s[n-1] - \hat{s}[n-1|n-1]) + u[n])^2]\end{aligned}$$

Note that

$$\mathbb{E}[(s[n-1] - \hat{s}[n-1|n-1])u[n]] = 0$$

since $s[n-1]$ depends on $\{u[0], \ldots, u[n-1], s[-1]\}$, which are independent of $u[n]$, and $\hat{s}[n-1|n-1]$ depends on past data samples, which also are independent of $u[n]$. Thus,

$$M[n|n-1] = a^2 M[n-1|n-1] + \sigma_u^2.$$

Summarize all above, we have, for $n \geq 0$:

- Prediction:

$$\hat{s}[n|n-1] = a\hat{s}[n-1|n-1]$$

- Minimum Prediction MSE:

$$M[n|n-1] = a^2 M[n-1|n-1] + \sigma_u^2$$

- Kalman Gain:

$$K[n] = \frac{M[n|n-1]}{M[n|n-1] + \sigma_n^2}$$

- Correction:

$$\hat{s}[n|n] = \hat{s}[n|n-1] + K[n](x[n] - \hat{s}[n|n-1])$$

- Minimum MSE:

$$M[n|n] = (1 - K[n])M[n|n-1]$$

III. Target Tracking

A. Basic target tracking methods based on Kalman filter

There are several steps to follow using Kalman filter in target tracking:

(i) Calculate the feature of the moving object. In order to track the moving target, firstly, calculate the center of mass and the width and height of the bounding rectangle.
(ii) Initialize Kalman filter with the information obtained. Due to the speed and the bounding rectangle change rate is unknown when initializing, set these variables to 0.

(iii) Using Kalman filter to predict the corresponding target area in next frame. Then run a target matching in target area after the arrival of next arrival.

(iv) If it matches, update Kalman filter, and record the target information in the current frame.

B. Problem modeling

In order to reduce the search space of the target and improve the speed of the target tracking, it is usually necessary to estimate the target's motion parameters at the next moment, so as to reduce the search complexity. The Kalman filtering is an optimal recursive algorithm for data processing. It does not need to store all the previous data in the memory and only needs to process the new measured value taken at each moment to estimate the optimal state.

Assume we have a stochastic dynamic system, with statistical properties as follows:

$$x(k+1) = \Phi(k+1,k)x(k) + \xi(k) \quad (15)$$

$$y(k) = \Theta(k)x(k) + \eta(k) \quad (16)$$

where $x(k)$ is a $n \times 1$ random state vector, $y(k)$ is a $p \times 1$ measurement vector, $\Phi(k+1,k)$ is a one-step transition matrix, $\Theta(k)$ is a $p \times n$ measurement vector, $\xi(k)$ is a $n \times 1$ noise vector while $\eta(k)$ is a $p \times 1$ measured noise vector.

In general, the statistical noise in video detection can be regarded as a zero mean Gaussian r.v., then it can be assumed that the system noise and the measurement noise satisfy the independent and identically-distributed WGN with zero mean. That is to say, for all k we have

$$\mathbb{E}(\xi(k)) = 0, \quad \mathbb{E}(\xi(k)\xi^T(k)) = Q(k) \quad (17)$$

$$\mathbb{E}(\eta(k)) = 0, \quad \mathbb{E}(\eta(k)\eta^T(k)) = R(k) \quad (18)$$

where the noise covariance matrix Q and measurement noise covariance matrix R are all positive definite.

According to the measurement information, the Kalman filtering equation should be:

- Prediction:

$$\hat{x}(k+1|k) = \Phi(k+1,k)\hat{x}(k)$$

- Filtering:

$$\hat{x}(k+1) = \hat{x}(k+1|k) + K(k+1)[y(k+1) - \theta(k+1)\hat{x}(k+1|k)]$$

- Kalman Gain:

$$K(k+1) = P(k+1|k)\Theta^T(k+1)[\Theta(k+1)P(K+1|k)\Theta^T(k+1) + R(k+1)]^{-1}$$

- Correction:

$$P(k+1|k) = \Phi(k+1,k)P(k)\Phi^T(k+1,k) + Q(k)$$

- Initial value:

$$\hat{x}(0) = \mathbb{E}(x(0))$$
$$P(0) = \mathrm{Var}(x(0))$$

Kalman gain $K(k+1)$ expresses the effect of innovation, with larger value obtaining much more correction on prediction, and small value denoting the prediction is much more accurate, the innovation making little effect. Therefore, the error, of time $k+1$, between the optimal estimation and real state is decided by two factors: the difference between prediction $\hat{x}(k+1|k)$ and real state $x(k+1)$; how much the Kalman gain $K(k+1)$ will correct the prediction $\hat{x}(k+1|k)$.

We establish a Kalman filter for each moving object. Taking the center of the bounding rectangle of the target as the tracking point, $x(k)$ consists of 4 components:

$$x(k) = [p_x(k), p_y(k), v_x(k), v_y(k)]^T$$

where $p_x(k)$ is the x component of the position of the tracking point at time k, $p_y(k)$ has the similar meaning, $v_x(k)$ is the x component of the speed of tracking point at time k and $v_y(k)$ has the similar meaning.

Since the time interval between each frame is relatively small, considering that each frame can be approximated as uniform motion, the system gain matrix is considered to be invariant. Supposing the time interval per frame is T, linearize and discretize the system to obtain the state transition matrix and the measurement matrix as follow:

$$\Phi(k+1|k) = \begin{bmatrix} 1 & 0 & T & 0 \\ 0 & 1 & 0 & T \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (19)$$

$$\Theta(k) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \quad (20)$$

IV. Experiment Results

We use the the video sequence with image resolution of 384×288 and frame rate 25 per second. Figure 1 shows the change, in position, of the error of a Kalman filter that stabilizes from initial state to stationary state.

We then use Kalman filtering to track the moving object. Because Kalman filter estimation can obtain the continuous motion state of the moving target, it is easy to maintain the correspondence between the targets of different frames, the

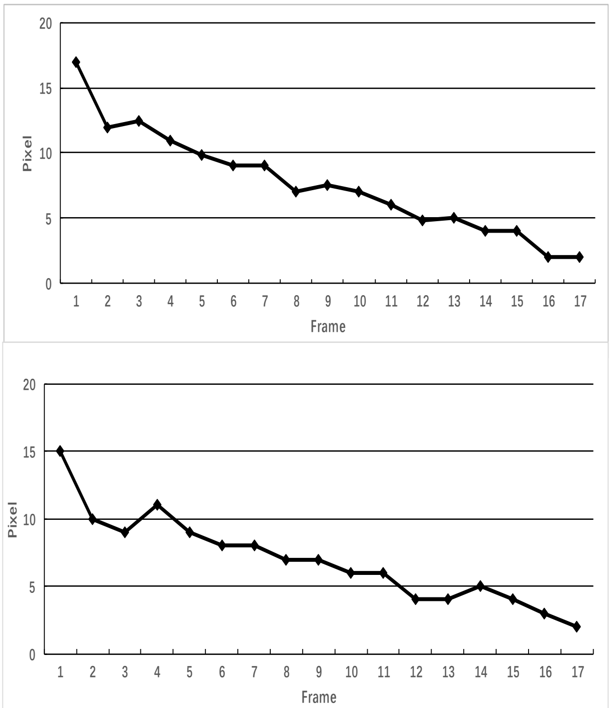

Fig. 1. Changes of error of position estimation, x and y component respectively

only thing needed to do is to confirm the existence of the target within the estimated range so as to reduce search complexity.

V. Conclusion

We elaborate, in detail, the application of kalman filter in predicting the target information, track the different motion states separately by establishing the inter-frame relationship matrix. From the above we can conclude that Kalman filter performs well in target tracking, it can well predict the target area next time slot based on present information.

References

[1] J. Pan, S. Li, and Y.-Q. Zhang, Automatic extraction of moving objects using multiple features and multiple frames, in Circuits and Systems, 2000. Proceedings. ISCAS 2000 Geneva. The 2000 IEEE International Symposium on, 2000, vol. 1, pp. 36-39.

[2] L. Rong, W.-C. Wang, M. Logan, and T. Donohue, Multiplatform multi-sensor fusion with adaptive-rate data communication, IEEE Transactions on Aerospace and Electronic Systems, vol. 33, no. 1, pp. 274-281, 1997.

[3] B. D. Anderson and J. B. Moore, Optimal filtering, Englewood Cliffs, vol. 21, pp. 22-95, 1979.

[4] C.-T. Chen, Linear system theory and design. Oxford University Press, Inc., 1995.

[5] A. Gelb, Applied optimal estimation. MIT press, 1974.

[6] A. H. Jazwinski, Stochastic processes and filtering theory. Courier Corporation, 2007.

[7] D. J. DePriest, Lessons in Digital Estimation Theory. Taylor & Francis Group, 1988.