Comparison between Kalman and Unscented Kalman Filters in Tracking Applications of Computational Vision

Raquel R. Pinho & João Manuel R. S. Tavares
Instituto de Engenharia Mecânica e Gestão Industrial, Lab. de Óptica e Mecânica Experimental
Universidade do Porto, Faculdade de Engenharia

ABSTRACT: In this paper, the problem of tracking feature points along image sequences is addressed. The establishment of correspondences between points and their tracking along image sequences is a complex problem in Computational Vision; especially, when intricate motions, erroneously detections or cases of occlusion or appearance/disappearing of features are involved. To overcome some of those difficulties, a statistical approach is frequently used in a multi-object data association and state estimation framework. Additionally, the correspondence between each measurement and predicted feature can be performed by minimizing the overall Mahalanobis distance. Under these circumstances, the estimation of the system can be accomplished using different stochastic filters. Hereby, a comparison is made between the results obtained, with the described framework, either by the Kalman Filter or the Unscented Kalman Filter, in the tracking of linear and nonlinear motions of feature points along image sequences.

1 INTRODUCTION

Tracking features along image sequences has been a long-standing problem in Computational Vision, as no known framework has been able to demonstrate its robustness to all kinds of changeling problems that may occur (for instance, image sequences with complex motions, cluttered scenes, noisy data, occlusions, topology variations or high number of mage frames to analyze). To accomplish that goal, many attempts have been made. However, none is known to be able to automatically detect, track and interrupt the procedure when appropriate, independently of the scenes' conditions or features' motion.

The tracking results are not independent of the used methodology, and their outcome may vary according to the used approaches. In this paper, the tracking of feature points is done on a stochastic strategy. Thus, to contribute to the purpose of robustness, we compare the results obtained with two stochastic filters, which can be used in a tracking framework.

The Kalman Filter (KF) is a well known efficient recursive approach that optimally estimates the state of a linear dynamic system from a series of noisy measurements. It is composed by a set of mathematical equations that provide an efficient computational mean to estimate the state of a system, in a way that minimizes the mean of the squared error. The KF is very powerful in several aspects: it supports estimations of past, present and even future states, even when the precise nature of the modeled system is unknown, [Welch & Bishop (2006)]. However, its major drawback is it linear assumption; that is, the KF tries to estimate the state of a discrete-time controlled process that is governed by a linear stochastic difference equation; consequently, it does not reliably overcome nonlinear occurrences.

Therefore, to overcome KF's drawbacks, other filters have been proposed, as the Unscented Kalman Filter (UKF). These filters endeavor to propagate mean and covariance information through nonlinear transformations. Regarding the UKF, the authors present it as more accurate, easier to implement, and indicate that it uses the same order of computation as the linearization, [Julier & Uhlmann (2004)]. In this work, we will present and analyze the results obtained by these filters in a tracking framework.

1.1 *Related Work*

The Kalman Filter has been widely used, but to surpass its limitation several approaches have been suggested. For instance, the Extended Kalman Filter (EKF) is probably the most commonly used estimation algorithm for nonlinear systems. It linearizes about the current mean and covariance. However, more than 35 years of experience in the estimation community have shown that is difficult to implement, complicated to tune, and only reliable for systems that are almost linear on the time scale of the updates, [Julier & Uhlmann (2004)].

A central and crucial operation performed in the KF is the propagation of a Gaussian Random Variable (GRV) through the system's dynamics. In the EKF, the state distribution is approximated by a GRV, which is then propagated analytically through the first-order linearization of the nonlinear system. This can introduce significant errors in the true posterior mean and covariance of the transformed GRV, which may lead to sub-optimal performance and sometimes divergence of the filter.

Other filters have been proposed, such as the Iterated Extended Kalman Filter (IEKF), the Central Difference Filter (CDF) and the first order Divided Difference Filter. The conventional EKF and IEKF algorithms do not take the linearization errors into account, which leads to inconsistent state estimates when these errors cannot be neglected, [Lefebvre et al. (2003)].

The UKF addresses that problem by using a deterministic sampling approach. The state distribution is again approximated by a GRV, which is represented using a set of carefully chosen sample points. These sample points completely capture the true mean and covariance of the GRV, and when propagated through the *true* nonlinear system, also approach the posterior mean and covariance accurately to the third order (Taylor series expansion) for any nonlinearity. The EKF, in contrast, only achieves first order accuracy. Remarkably, the computational complexity of the UKF is of the same order as the EKF's, [Wan & Van Der Merwe (2000)].

A variation of the original UKF algorithm (called the Reduced Sigma Point Filters [Julier & Uhlmann (2002)]) chooses only $n+1$ sampling points. This means that the linear regression through those points is exact; therefore, it does not take the linearization errors into account. Hence, as was the case for the (I)EKF, the state estimates are generally inconsistent, [Lefebvre, Bruyninckx and De Schutter (2003)].

An arising estimation model is the Particle Filter (PF): the key idea is to represent the posterior density function by a set of random samples with associated weights and to compute estimates based on these estimates and weights. Comparing PF to KF, the former has a more robust performance in the case of non-Gaussianity and non-linearity due to the simulated posterior distribution. In a PF, a large number of particles are desirable to represent the posterior distribution, especially in situations where new measurements appear in the tail of the prior, or if the likelihood is strongly peaked. To solve the computational problem raised by the large particle numbers that can be generated, for example, in [MacCormick & Isard (2000)] a partitioned sampling is proposed, which requires that the state-space can be sliced. With the same purpose, Sullivan et al., in [Sullivan et al. (1999)], proposed layered sampling using multiscale processing of images. It turns out that these solutions significantly reduce the computational costs, but in-depth efforts are desirable for better efficiency, [Zhou et al. (2009)].

In [Merwe et al. (2000)] an Unscented Particle Filter (UPF) is introduced. The UKF allows the PF to incorporate the latest observations into a prior updating routine. In addition, the UKF generates proposal distributions that match the true posterior more closely and also has the capability of generating heavier tailed distributions than the well known EKF. Consequently, the convergence results predict that UPF should outperform standard PF, EKF and UKF. However, its computational cost can also be questioned.

Other comparison studies between predictive filters have been published; for example, in [LaViola (2003)] the performance of the UKF and EKF is compared towards improving human head and hand tracking using quaternions, and in [Lefebvre, Bruyninckx and De Schutter (2003)] the process and measurement update performances are studied separately, instead of analyzing their overall behavior of some KF variants.

Along the last years, numerous applications of state estimation have been reported, such as in sensor fusion and navigation, [Metzger et al. (2005)], state estimation for chemical processes, [Vachhani et al. (2005)], weather forecasting, [Carme et al. (2001)], and training of neural networks [Rios Neto (1997)].

1.2 *Adopted Tracking Framework*

In the applied tracking framework, the measured data is integrated using optimization upon the correspondence step. The proposed criterion minimizes the global matching cost based on the Mahalanobis distance, under the assumption that the matching cost is superior to a threshold value. To deal with the appearance and disappearance of features during the tracking process a management model is considered, [Pinho et al. (2007)].

1.3 *Paper Overview*

This paper is organized as follows. In the next sections, a brief introduction is made to KF and UKF. In section 4, some experimental results are presented with synthetic and real image sequences. In the last section, the main conclusions are held as well as perspectives of future work.

2 THE KALMAN FILTER

The equations of the KF fall into main two groups: time update (or prediction) and measurement update (or correction) equations. The time update stage is responsible for projecting forward (in time) the current state and error covariance to obtain the a priori estimates for the next time step. The measurement update phase is responsible for the feedback (i.e. for incorporating a new measurement into the a priori estimation to obtain an improved a posteriori approximation), [Welch & Bishop (2006)].

The prediction step is based on the Chapman-Kolmogorov equation for a first order Markov process:

$$x_t^- = \Phi x_{t-1}^+ , \tag{1}$$

where Φ relates the system state x_{t-1}^+ at the previous time step $t-1$ to the state x_t^- at the current step t. The superscripts $^+$ and $^-$ indicate if the measurement data have been or not considered, respectively. The related uncertainty is given by:

$$P_t^- = \Phi P_{t-1}^+ \Phi^T + Q , \tag{2}$$

where P is the covariance matrix and Q models the process noise.

The correction equations that update the predicted estimations upon the incorporation of new u_t measurements are given by:

$$K_t = P_t^- H^T \left[H P_t^- H^T + R_t \right]^{-1} , \tag{3}$$

$$x_t^+ = x_t^- + K_t \left[u_t - H x_t^- \right] , \tag{4}$$

$$P_t^+ = \left[I - K_t H \right] P_t^- , \tag{5}$$

where K is chosen to be the gain that minimizes the a posteriori error covariance equation, H processes the coordinates transformation between the predicted space and the measurement space, R_t is the measurement noise involved, and I is the identity matrix.

3 THE UNSCENTED KALMAN FILTER

As formerly mentioned the UKF addresses the major shortcomings of the EKF and is more appropriate for nonlinear movement. It considers a set of sigma-points from the distribution of the state vector x at the beginning of the time step. Those points are all propagated through true nonlinearity, and the parameters of the Gaussian approximation are then re-estimated. If these sigma-points are appropriately chosen, it can be shown that the mean and covariance of x at the end of the time step are p^{th}-order accurate for any nonlinearity, [Mariani & Ghisi (2007)]. The order p of accuracy depends primarily on the number of sigma-points. In general, $2L+1$ points are considered, being L the dimension of the state vector, although other proposals have been made with fewer points (see, for instance, [Julier & Uhlmann (2002)]).

In the prediction phase, the sigma-points $S_i = \{w_i, \chi_i\}$, $i = 0, \ldots, 2L$ are obtained using the unscented transformation:

$$\begin{cases} \chi_0 = \bar{x} \\ \chi_i = \bar{x} + \left(\sqrt{(L+\lambda) P_x} \right)_i & \text{if} \quad i = 1, \ldots, L \\ \chi_i = \bar{x} - \left(\sqrt{(L+\lambda) P_x} \right)_i & \text{if} \quad i = L+1, \ldots, 2L \end{cases} , \tag{6}$$

with:

$$w_0^m = \frac{\lambda}{L+\lambda} , \; w_0^c = \frac{\lambda}{(L+\lambda)} + 1 - \alpha^2 + \beta , \tag{7}$$

and:

$$w_i^m = w_i^c = \frac{\lambda}{2(L+\lambda)} , \text{ if } i = 1, \ldots, 2L , \tag{8}$$

where w_i is the weight associated with the i^{th} sigma-point such that:

$$\Sigma_{i=0}^{2L} w_i = 1 , \tag{9}$$

and λ is a scaling parameter in order that $\left(\sqrt{(L+\lambda) P_x} \right)_i$ is the i^{th} column (or row) of the square root of the weighted covariance matrix, $(L+\lambda) P_x$. The values of α and β parameters should be determined experimentally.

Each sigma point is now propagated through the nonlinear function:

$$Y_i = g(\chi_i) , \tag{10}$$

with $i = 0, \ldots, 2L$, and the approximated mean and covariance of y are as follows:

$$\bar{y}_i \approx \sum_{i=0}^{2L} w_i Y_i , \tag{11}$$

$$P_y \approx \sum_{i=0}^{2L} w_i (Y_i - \bar{y})(Y_i - \bar{y})^T , \tag{12}$$

$$Z_i = h(\chi_i) , \tag{13}$$

$$\bar{z}_i \approx \sum_{i=0}^{2L} w_i Z_i . \tag{14}$$

The measurement update equations are given by:

$$P_{zz} \approx \sum_{i=0}^{2L} w_i (Z_i - \bar{z})(Z_i - \bar{z})^T , \tag{15}$$

$$P_{yz} \approx \sum_{i=0}^{2L} w_i (Y_i - \bar{y})(Z_i - \bar{z})^T , \tag{16}$$

$$K = P_{yz} P_{zz}^{-1} , \tag{17}$$

$$\bar{y} = \bar{y} + K(Z_i - \bar{z}), \quad (18)$$

$$P = P - KP_{zz}K^T. \quad (19)$$

The computational cost of the above algorithm is $\vartheta(n^3)$, being the most expensive operations the computation of the matrix square root and the outer products required to obtain the covariance of the projected sigma points.

4 EXPERIMENTAL RESULTS

The first example considers a synthetic sequence of eight images of 300x300 pixels, in which the centre of a square is moving according to a linear model given by (in pixels):

$$\begin{cases} x_0=5 \\ y_0=250 \\ x_i=x_{i-1}+30 \\ y_i=y_{i-1}+250 \end{cases}. \quad (20)$$

The tracking error, that is, the differences between the estimated and measured features' positions, along the sequence is represented in Figure 1. In that figure, one can notice that the UKF gets better estimations for the undergoing motion, but as time goes by KF tends to draw near UKF results.

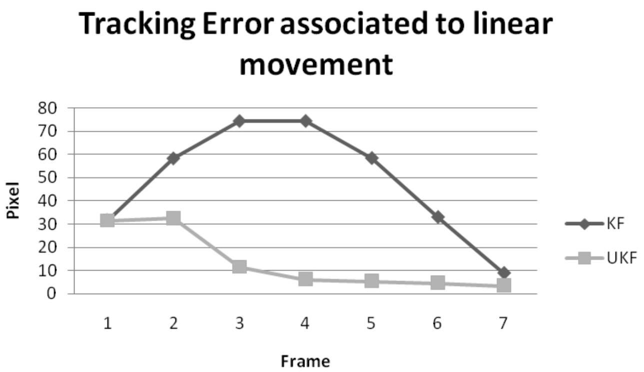

Figure 1. Tracking error associated to a point that undergoes the linear motion according to (20).

The second example is an image sequence composed by 6 frames with the same dimensions of the previous example, also synthetically built (Figures 2 and 3). In those images, the + crosses represent the filters' prediction, the x crosses in the centre of the tracked squares represent the accurate measurements and the x crosses between the other two represent the corrected results after the measurement incorporation. But, in this case, the motion of the tracked feature is defined by a nonlinear model (in pixels):

$$\begin{cases} x_0=6 \\ y_0=10 \\ x_i=x_{i-1}+2(i-1)^2+12.5 \\ y_i=x_{i-1}+12.5 \end{cases}. \quad (21)$$

In this second example, the tracking error, represented in Figure 4 shows that as increases the features' distance between frames, the UKF gets more accurate results than the KF, which can also be noticed in the obtained images, some of which represented in Figure 3.

In the third example, a real sequence of 30 images is considered, where three mice are tracked in images of 320x240 pixels. Several difficulties arise in tracking the center of the mice's bodies in the acquired images. One of which comprehends the fast movement of the mice, as they may go back and forth changing drastically their movement direction, or may move very quickly along one direction. Figure 5 represents the tracking results obtained in 3 frames, at which the upper mouse changes severely its movement direction. However, as by visual observation it is difficult to distinguish accurately the shown images results, in Figure 6 the tracking position error of the mice is represented. Hence, from Figure 6, one can notice that in general UKF (results represented by dashed lines) achieves better approaches to the undergoing motion; nevertheless, in the case of the third mouse the results obtained by each filter are very similar. Moreover, as tracking was successfully accomplished by both filters, mainly due to the used association approach, taking into account the computational cost related to the unscented transformation, one can think that in this case KF could be more advantageous.

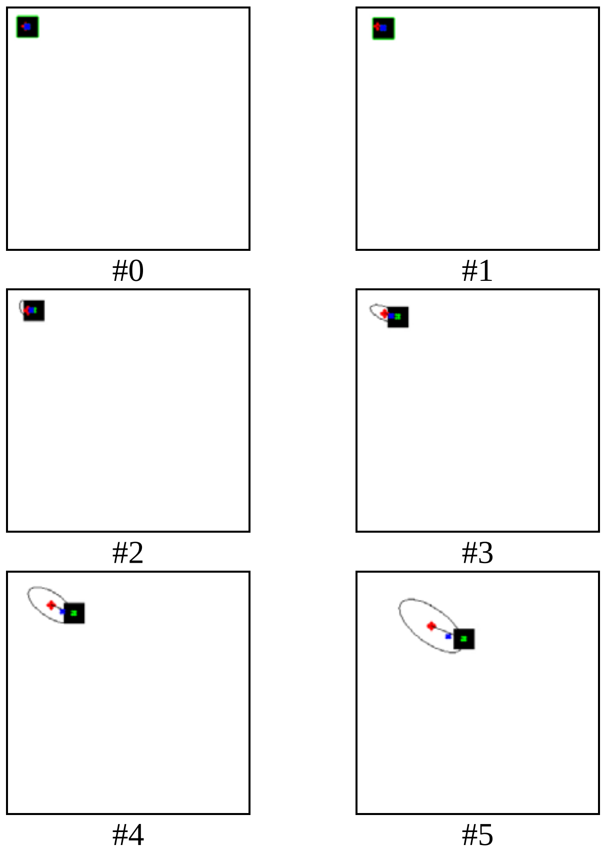

Figure 2. Tracking a point (the centre of the square) in a nonlinear sequence using KF.

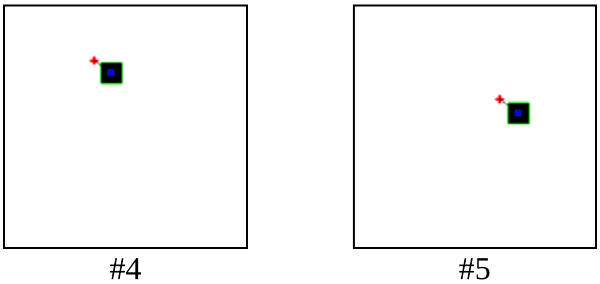

Figure 3. Tracking results of the last frames of sequence in Figure 2 using UKF.

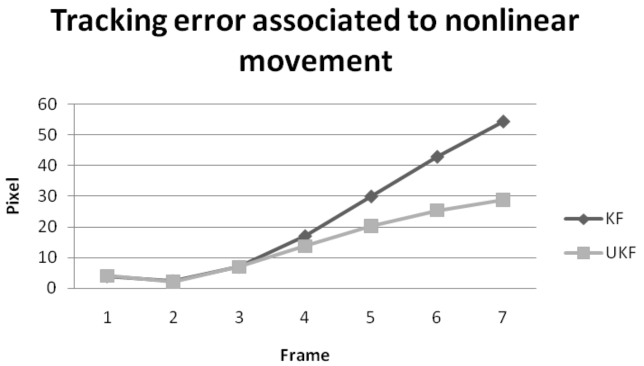

Figure 4. Error involved in the nonlinear tracking of the square shown in Figure 2.

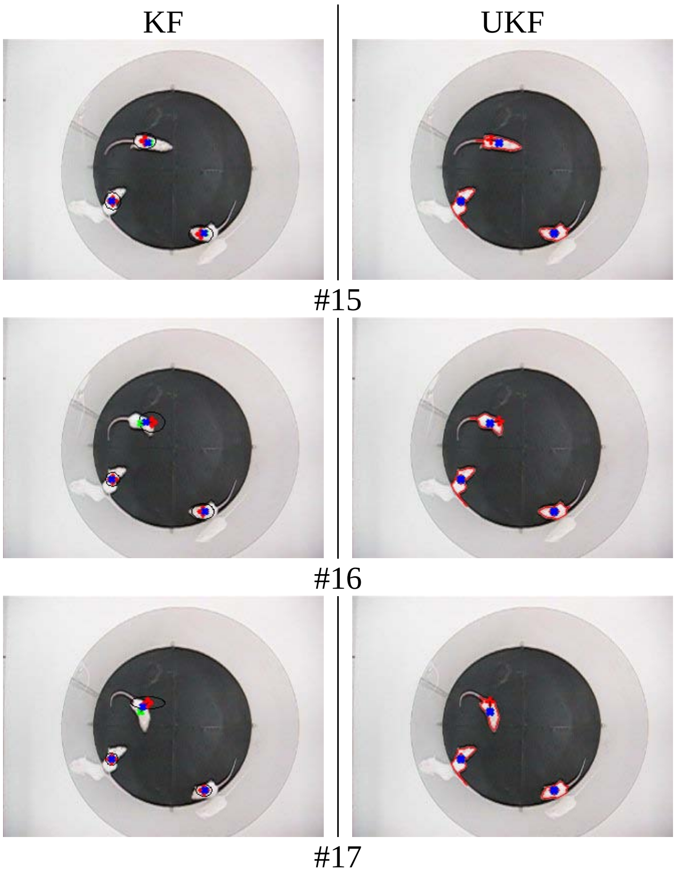

Figure 5. Tracking results for frames 15, 16 and 17 of a 30 image sequence obtained using KF and UKF.

5 CONCLUSIONS AND FUTURE WORK

In this paper, we have analyzed the influence of the used filter (KF and UKF) in the tracking of points undergoing linear and non-linear motions. In this comparison one noticed that if the movement is highly nonlinear, then the UKF is worth its computational load; otherwise, the KF with the undertaken matching methodology suits reasonably the tracking process with lower computational requirements.

Hence, the decision to use KF or UKF in this tracking framework is application dependent: if the computational requirements are strongly constrained, then KF can be considered with a good matching strategy; when such constrains are not presented, then UKF should be considered as it generally obtains more robust tracking results.

ACKNOWLEDGMENTS

We would like to thank Prof. Hemerson Pistori and Prof. João Bosco Monteiro (Dom Bosco Catholic University, Brazil) for gently sharing the mice image sequences used in the last experimental example.

This work was partially done in the scope of the research project *Segmentation, Tracking and Motion Analysis of Deformable (2D/3D) Objects using Physical Principles*, with reference POSC/EEA-SRI/55386/2004, financially supported by the Portuguese FCT - Fundação para a Ciência e Tecnologia.

The first author would like to thank the support of her PhD grant SFRH/BD/12834/2003 from FCT.

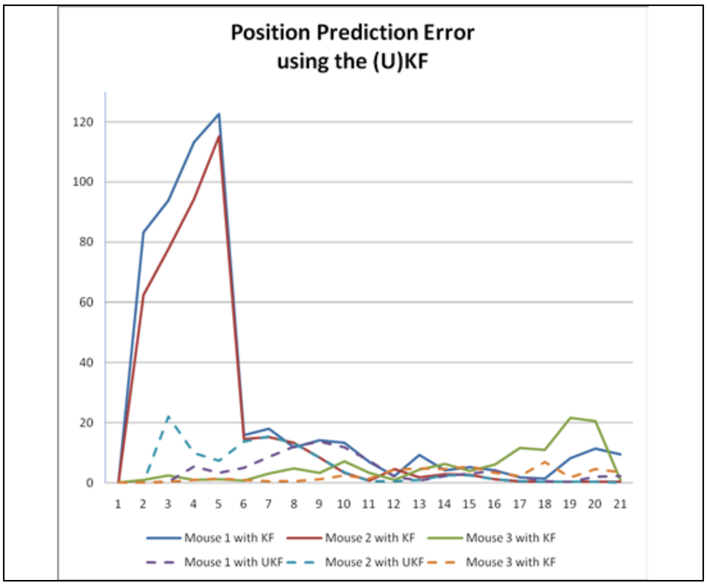

Figure 6. Errors of the predicted position of the mice's centre point motion in a real image sequence.

REFERENCES

Carme, S.; Pham, D. T.; Verron, J. 2001. Improving the singular evolutive extended Kalman filter for strongly nonlinear models for use in ocean data assimilation. *Inverse Problems* 17(5): 1535/1559.

Julier, S.; Uhlmann, J. 2002. Reduced Sigma Point Filters for the Propagation of Means and Covariances Through Nonlinear

Transformations. *American Control Conference* Alaska, USA:887/892.
Julier, S. J.; Uhlmann, J. K. 2004. Unscented Filtering and Nonlinear Estimation. *Proceedings of the IEEE* 92(3): 401-422.
LaViola, J. J. 2003. A comparison of unscented and extended Kalman filtering for estimating quaternion motion. *Proceedings of the 2003 American Control Conference* 3:2435/2440.
Lefebvre, T.; Bruyninckx, H.; De Schutter, J. 2003. Kalman filters for nonlinear systems: a comparison of performance. *International Journal of Control* 77(7): 639-653.
MacCormick, J.; Isard, M. 2000. Partitioned sampling, articulated objects, and interface-quality hand tracking. *Proceedings of the European Conference on Computer Vision*:390-395.
Mariani, S.; Ghisi, A. 2007. Unscented Kalman filtering for nonlinear structural dynamics *Nonlinear Dynamics Journal* 49 (1-2): 131/150.
Merwe, R.; Doucety, A.; Freitas, N.; Wan, E. 2000. The Unscented Kalman Filter *Advances in Neural Information Processing Systems* 2000 Vancouver, Canada
Metzger, J.; Wisotzky, K.; Wendel, J.; Trommer, G. 2005. Sigma-point filter for terrain referenced navigation. *AIAA Guidance, Navigation, and Control Conference* San Francisco, California (U.S.A.)
Pinho, R.; Tavares, J.; Correia, M. 2007. An Improved Management Model for Tracking Missing Features in Computer Vision Long Image Sequences *WSEAS Transactions on Information Science and Applications* 4(1): 196-203.
Rios Neto, A. 1997. Stochastic optimal linear parameter estimation and neural nets training in systems modelling. *Journal of the Brazilian Society of Mechanical Sciences* 19(2): 138/146.
Sullivan, J.; Blake, A.; Isard, M.; MacCormick, J. 1999. Object localization by bayesian correlation. *Proceedings of the Seventh International Conference on Computer Vision*:1068-1075.
Vachhani, P.; Rengaswamy, R.; Gangwal, V.; Narasimhan, S. 2005. Recursive estimation in constrained nonlinear dynamical systems. *AIChE Journal* 51 (3): 946/959.
Wan, E. A.; Van Der Merwe, R. 2000. The unscented Kalman filter for nonlinear estimation. IEEE Adaptive Systems for Signal Processing, Communications, and Control Symposium 2000 Lake Louise, Alta., Canada:153-158.
Welch, G.; Bishop, G. 2006. An Introduction to Kalman Filter. University of North Carolina at Chapel Hill
Zhou, H.; Yuan, Y.; Shi, C. 2009. Object tracking using SIFT features and Mean Shift. *Computer Vision and Image Understanding* 113(3): 345-352.