An Improvement on ArUco Marker for Pose Tracking Using Kalman Filter

Ho Chuen Kam*, Ying Kin Yu, Kin Hong Wong
Department of Computer Science and Engineering
The Chinese University of Hong Kong
Shatin, Hong Kong
**E-mail: hckam@cse.cuhk.edu.hk*

***Abstract*—This paper presents a robust but simple object pose tracking algorithm based on Kalman filtering. Compared to markerless pose tracking, a fiducial marker called ArUco provides a fast and accurate solution to the problem. With these advantages, this marker-based technique is ready to be used in virtual reality and operate in low-cost wearable devices. However, it still suffers from the problem of occlusion and noise. If a large part of the marker is occluded, no pose information can be acquired for the moment. Noises due to hand shaking also affect the quality of the resulting pose. This is not desirable in a real-time environment. We tackle the problems by employing a linear Kalman filter. The pose information can be estimated even if the camera view is blocked temporarily. We have performed real experiments to demonstrate the effects of the application of Kalman filter. The results are satisfactory.**

***Keywords*-Object Tracking, Pose estimation, Virtual Reality, Kalman Filter**

I. Introduction

Object tracking is an important topic in computer vision. The pose of an object is found using a camera. With the pose information, a number of interesting applications can be developed, such as augmented reality [1], [2], [3] and robot navigation [4], [5], [6]. To locate and track the moving object in the environment, there are two classes of vision-based techniques, namely the markerless and the marker-based approaches. For markerless tracking, colours, patterns and other features of the target are extracted and compared among frames. This requires no a-priori knowledge such as the size or arrangement of patterns. The drawback is mainly the high computation requirement due to feature extraction and matching. In order to simplify the tracking process, the marker-based tracking can be used in various occasions, especially in handheld devices equipped with cameras. With fiducial markers, the pose information can be acquired accurately at a high speed. In this work, we develop our method based on the ArUco marker [7]. This is one of the most popular fiducial markers and available as an open source software in OpenCV.

Common problems in object pose tracking are occlusion and noises. Although the pose of the fiducial marker can be recovered instantly when it appears again in the image, pose information is lost when the object is occluded. In addition, the marker-based tracking approaches are sensitive to noises and the output may have some undesirable jerk. In this situation, Kalman filter can be utilized to provide an estimate of the marker positions during occlusion and stabilize the resulting pose.

In the paper, we propose a simple and robust tracking method based on Kalman filter. The system can track the object with an ArUco marker attached. The motion information can be estimated even if the marker is occluded. The resulting pose is smoothed by the Kalman filter. Our method can be run in smartphones because of its low computation requirement.

The rest of the paper is organized as follows. Section II presents the background of our work. Section III describes the theory and methodology of our approach. Then experimental results are presented and illustrated in section IV. Finally, section V concludes our work.

II. Background

A. Marker-based Tracking

Tracking requires correspondences in the target objects. Traditionally, features such as corners and edges, are compared and matched among the image frames. Another kind of tracking methods uses fiducial markers [8] to solve the problem. These approaches require the knowledge of the existence of the marker in advance, but in turns it gives a higher accuracy.

There are literatures discussing the fiducial marker systems. Some popular markers are in circular shape, such as [9] and [10]. However, the design of the pattern inside a circular marker is difficult and such a kind of markers is hard to be represented in a rectangular array [11].

A more common choice is square markers. Examples include ARToolkit [12], Matrix [13], ArUco [8] and etc. This is more preferable than the circular markers, as the four corners can be identified more easily and the 6 degree-of-freedom (DOF) pose information can be acquired.

Thanks to its high identification rate, accuracy and computation speed, it is ready to be used in virtual reality applications [12] [14] [15]. Some other real-life applications [5] [16] [17] use fiducial markers to track vessels and patients.

978-1-5386-5889-5/18/$31.00 ©2018 IEEE
SNPD 2018, June 27-29, 2018, Busan, Korea

B. Kalman Filter

Kalman filter is first developed in 1960 by R.E.Kalman [18]. It is a recursive algorithm to estimate the state of a process in discrete time, for example, ballistic missile in the early days [19]. With the assumption of linear systems and Gaussian noises, it can provide optimal estimates based on the prediction models and measurements. Kalman filter has been applied to pose tracking in computer vision [20] [21] [22] [23] [24]. Traditional Kalman filter can only deal with linear systems. Recently, there are different variations of filters, such as Extended Kalman Filter (EKF) [25], Unscented Kalman Filter (UKF) [26] and Cubature Kalman Filter (CKF) [27], for non-linear and high dimensional systems.

III. Theory and Design

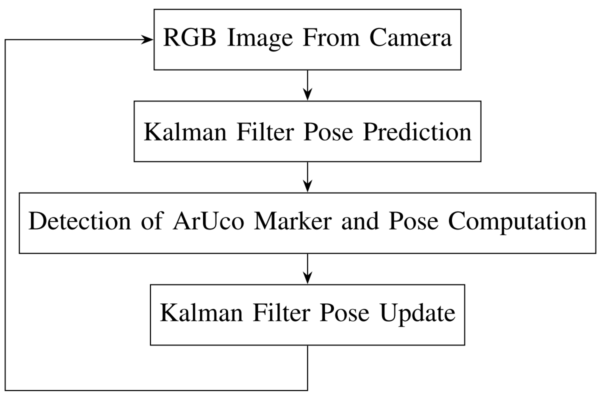

Figure 1. Overview of the proposed system

The overview of the system is shown in figure 1. We use an ordinary RGB webcam for the input of the system. The camera is already calibrated beforehand. Then the system constantly searches for the AruCO marker in the scene. If the marker is detected, pose estimation is carried out to find out the pose of the marker. Kalman filter predicts the *a priori* poses based on the dynamic model. After getting the latest measurement, the filter uses it to update the prediction. The state is now *a posteriori*.

A. ArUco Marker Tracking

We build our method based on the ArUco marker fiducial system. It has a few advantages over its previous counterparts in terms of low false negative rate and robustness to occlusion. Firstly, image segmentation is done on the input image to extract the marker region. Then contour extraction is performed to extract the polygons in the interior of the marker images. Further identification of marker codes by the pre-defined dictionary is performed. Finally, pose information is computed using four square corners with Levenberg-Marguardt algorithm [28].

B. Tracking with Kalman Filter

1) Input to Kalman Filter: The ArUco marker algorithm outputs the marker poses directly upon the successful detection. A linear Kalman filter is applied to post-process the pose information. In our approach, the pose is represented by the rotation R_t and the translation T_t at time t. In particular, quaternion is used to describe the rotation as this can avoid the gimbal lock problem [29]. Then R_t is represented by the four quaternion parameters q_x, q_y, q_z, q_w and the translation becomes $T_t = [T_x, T_y, T_z]^T$ where T_x, T_y, T_z is the translation along the x,y,z axes, respectively. In addition to the displacement, velocities of the above parameters are also considered in the system. They are assumed constant. State x_t is defined as:

$$x_t = [T_x, T_y, T_z, q_x, q_y, q_z, q_w, \dot{T}_x, \dot{T}_y, \dot{T}_z, \dot{q_x}, \dot{q_y}, \dot{q_z}, \dot{q_w}]_t^T \tag{1}$$

2) The Dynamic System and Measurement Model: The states of a linear Kalman filter are related by the following equations:

$$x_t = Ax_{t-1} + w_t \tag{2}$$
$$z_t = Hx_t + v_t \tag{3}$$

Equation 2 defines the state transition by a discrete-time dynamic system. Equation 3 defines the measurement model. x_t is the state vector, t is the time index and z is the measurement vector. Matrix A in equation 2 relates the state at the previous time frame $t-1$ and the current frame t, which is shown in equation 4 below.

$$A = \begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & dT & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & dT & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & dT & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & dT & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & dT & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & dT & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & dT \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix} \tag{4}$$

Matrix H in equation 3 maps state x to the measurement z.

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{5}$$

w and v are random variables with covariance Q and R, respectively.

IV. Experiments And Results

We have built a testing platform to demonstrate the effects of the application Kalman filter in our proposed system. The platform is shown in figure 2. It consists of a stepping-motor controlling the ArUco marker position. A precise movement can be made. The webcam at top captured the motion of the marker and the images were fed to our system for processing. In order to test the robustness of our algorithm, the marker is occluded in some parts of the image sequence.

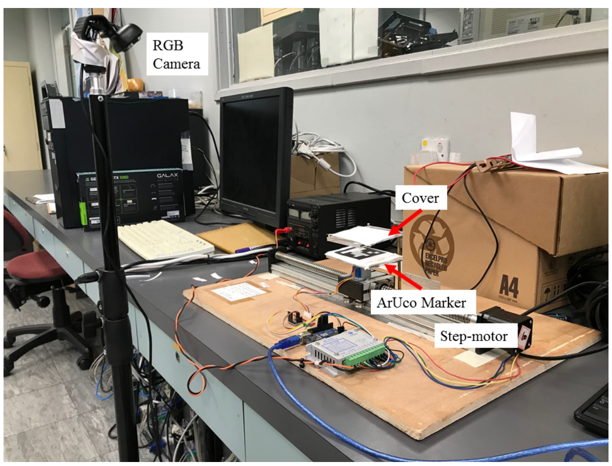

Figure 2. The setup of our experiments

We performed two experiments. The first one was to compare the resulting pose with and without the application of Kalman filtering in the tracking process. The second one simulated hand shaking by vibrating the fiducial marker.

A. With and Without Kalman Filter

In this test, we mounted the ArUco marker on a platform. It was moved along the path with a stepping-motor. A cover was placed in the middle of the path to occlude the marker. Figure 3 shows the results of this experiment. No pose information were available from the ArUco marker when the marker was occluded. With the Kalman filter, our method could still estimate the state and predict the locations of the markers in this period of time. The result shows that the algorithm can handle the occlusion problem and give an optimal estimate.

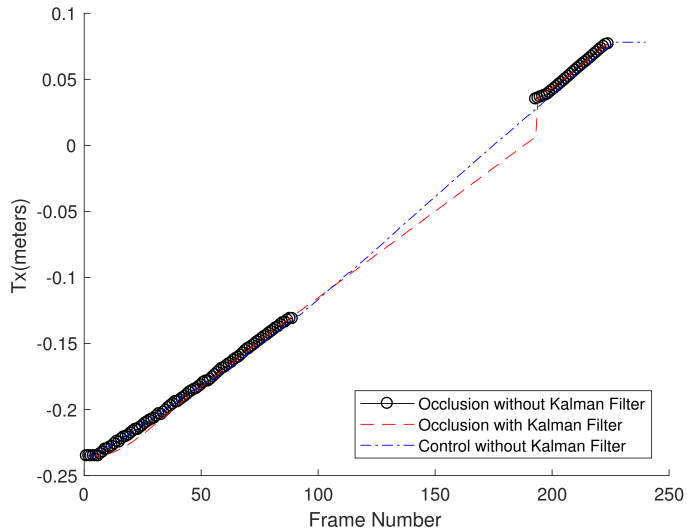

Figure 3. Experiment 1: Tracking with and without Kalman filter

B. Resistant to Noise

If there are noises in the images or vibration of markers due to handshaking, the accuracy of the pose obtained by the ArUco marker is affected. In order to improve the resulting pose, Kalman filter can be applied. This experiment imposed random noises on the marker motion by vibrating it as shown in figure 4.

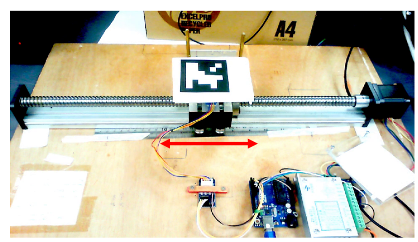

Figure 4. Experiment 2: Producing marker vibration

Figure 5 shows the results of experiment 2. Without Kalman filtering, the pose follows the unwanted vibration noise closely. This is not desirable in some applications such as virtual reality. If the Kalman filter is applied, the pose remains relatively stable.

V. Conclusion

We have proposed a marker-based pose tracking method in this paper. The pose of the object can be tracked effectively by attaching the ArUco marker to it. Although the original result is fast and accurate, it is still vulnerable to occlusion and hand shaking. To tackle these problems, we apply a linear Kalman filter to improve the robustness of the ArUco marker. It can estimate the pose of the object

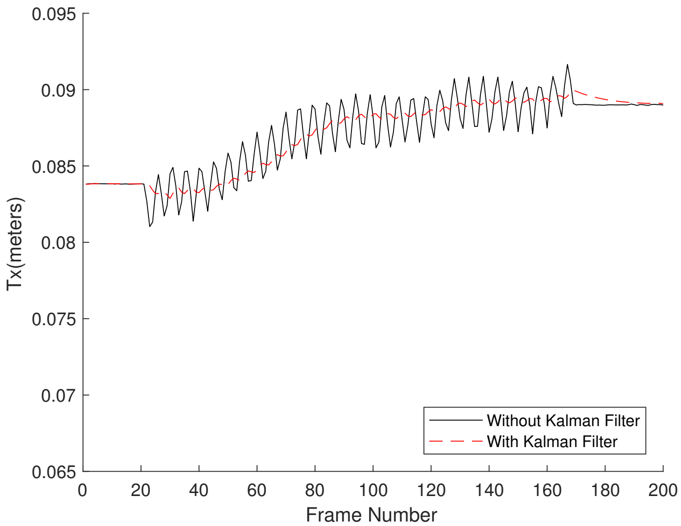

Figure 5. Experiment 2: The recovered marker motion

even if the target object is blocked temporarily. It can also stabilize the resulting pose. Experimental results show that the application of Kalman filtering can improve the results of ArUco marker effectively.

REFERENCES

[1] R. Azuma, Y. Baillot, R. Behringer, S. Feiner, S. Julier, and B. MacIntyre, "Recent advances in augmented reality," *IEEE Computer Graphics and Applications*, vol. 21, no. 6, pp. 34–47, 2001.

[2] D. Van Krevelen and R. Poelman, "A survey of augmented reality technologies, applications and limitations," *International Journal of Virtual Reality*, vol. 9, no. 2, p. 1, 2010.

[3] M. Billinghurst, A. Clark, G. Lee *et al.*, "A survey of augmented reality," *Foundations and Trends in Human–Computer Interaction*, vol. 8, no. 2-3, pp. 73–272, 2015.

[4] R. Acuna, Z. Li, and V. Willert, "Moma: Visual mobile marker odometry," *arXiv preprint arXiv:1704.02222*, 2017.

[5] J. Bacik, F. Durovsky, P. Fedor, and D. Perdukova, "Autonomous flying with quadrocopter using fuzzy control and aruco markers," *Intelligent Service Robotics*, vol. 10, no. 3, pp. 185–194, 2017.

[6] J. L. Sanchez-Lopez, J. Pestana, P. de la Puente, A. Carrio, and P. Campoy, "Visual quadrotor swarm for the imav 2013 indoor competition," in *First Iberian Robotics Conference (ROBOT2013)*. Springer, 2014, pp. 55–63.

[7] R. Munoz-Salinas, "Aruco: a minimal library for augmented reality applications based on opencv," *Universidad de Córdoba*, 2012.

[8] S. Garrido-Jurado, R. Muñoz-Salinas, F. J. Madrid-Cuevas, and M. J. Marín-Jiménez, "Automatic generation and detection of highly reliable fiducial markers under occlusion," *Pattern Recognition*, vol. 47, no. 6, pp. 2280–2292, 2014.

[9] V. A. Knyaz, "The development of new coded targets for automated point identification and non-contact 3d surface measurements," *IAPRS*, vol. 5, pp. 80–85, 1998.

[10] L. Naimark and E. Foxlin, "Circular data matrix fiducial system and robust image processing for a wearable vision-inertial self-tracker," in *Proceedings of the 1st International Symposium on Mixed and Augmented Reality*. IEEE Computer Society, 2002, p. 27.

[11] C. B. Owen, F. Xiao, and P. Middlin, "What is the best fiducial?" in *The First IEEE International Workshop on Augmented Reality Toolkit*. IEEE, 2002, p. 8.

[12] H. Kato and M. Billinghurst, "Marker tracking and hmd calibration for a video-based augmented reality conferencing system," in *Proceedings of the 2nd IEEE and ACM International Workshop on Augmented Reality 1999 (IWAR99)*. IEEE, 1999, pp. 85–94.

[13] J. Rekimoto, "Matrix: A realtime object identification and registration method for augmented reality," in *Proceedings of the 3rd Asia Pacific Computer Human Interaction 1998*. IEEE, 1998, pp. 63–68.

[14] H. Kato, M. Billinghurst, I. Poupyrev, K. Imamoto, and K. Tachibana, "Virtual object manipulation on a table-top ar environment," in *Proceedings of IEEE and ACM International Symposium onAugmented Reality 2000.(ISAR2000)*. Ieee, 2000, pp. 111–119.

[15] F.-e. Ababsa and M. Mallem, "Robust camera pose estimation using 2d fiducials tracking for real-time augmented reality systems," in *Proceedings of the 2004 ACM SIGGRAPH International Conference on Virtual Reality Continuum and its Applications in Industry*. ACM, 2004, pp. 431–435.

[16] T. Shchory, D. Schifter, R. Lichtman, D. Neustadter, and B. W. Corn, "Tracking accuracy of a real-time fiducial tracking system for patient positioning and monitoring in radiation therapy," *International Journal of Radiation Oncology, Biology, Physics*, vol. 78, no. 4, pp. 1227–1234, 2010.

[17] S. S. Tørdal and G. Hovland, "Relative vessel motion tracking using sensor fusion, aruco markers, and mru sensors," 2017.

[18] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Journal of Basic Engineering*, vol. 82, no. 1, pp. 35–45, 1960.

[19] G. M. Siouris, G. Chen, and J. Wang, "Tracking an incoming ballistic missile using an extended interval kalman filter," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 33, no. 1, pp. 232–240, 1997.

[20] Y. K. Yu, K. H. Wong, M. M.-Y. Chang, and S. H. Or, "Recursive camera-motion estimation with the trifocal tensor," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 36, no. 5, pp. 1081–1090, 2006.

[21] Y. K. Yu, K. H. Wong, S. H. Or, and C. Junzhou, "Controlling virtual cameras based on a robust model-free pose acquisition technique," *IEEE Transactions on Multimedia*, vol. 11, no. 1, pp. 184–190, 2009.

[22] Y. K. Yu, K. H. Wong, and M. M. Y. Chang, "A fast and robust simultaneous pose tracking and structure recovery algorithm for augmented reality applications," in *Proceedings of the 2004 IEEE International Conference on Image Processing (ICIP-2004)*. IEEE, 2004, pp. 1029–1032.

[23] K. K. Lee, Y. K. Yu, K. H. Wong, and M. M. Y. Chang, "Tracking 3-d motion from straight lines with trifocal tensors," *Multimedia Systems*, vol. 22, no. 2, pp. 181–195, 2016.

[24] Y. K. Yu, K. H. Wong, S. H. Or, and M. M.-Y. Chang, "Robust 3-d motion tracking from stereo images: A model-less method," *IEEE Transactions on Instrumentation and Measurement*, vol. 57, no. 3, pp. 622–630, 2008.

[25] L. Ljung, "Asymptotic behavior of the extended kalman filter as a parameter estimator for linear systems," *IEEE Transactions on Automatic Control*, vol. 24, no. 1, pp. 36–50, 1979.

[26] E. A. Wan and R. Van Der Merwe, "The unscented kalman filter for nonlinear estimation," in *The IEEE 2000 Symposium on Adaptive Systems for Signal Processing, Communications and Control 2000*. Ieee, 2000, pp. 153–158.

[27] I. Arasaratnam and S. Haykin, "Cubature kalman filters," *IEEE Transactions on Automatic Control*, vol. 54, no. 6, pp. 1254–1269, 2009.

[28] D. W. Marquardt, "An algorithm for least-squares estimation of nonlinear parameters," *Journal of the Society for Industrial and Applied Mathematics*, vol. 11, no. 2, pp. 431–441, 1963.

[29] A. J. Hanson, "Visualizing quaternions," in *ACM SIGGRAPH 2005 Courses*. ACM, 2005, p. 1.