Proceedings of the American Control Conference
San Diego, California • June 1999

5D Target Trajectory Detection via Intelligent Monocular Visual Tracking System in Real-Time with Air-Target Orientation Recognition

Chih-Yu Chen [1], and Li-Chen Fu [1,2]
(12th Dragon Thesis Award)

Department of Electrical Engineering[1]
Department of Computer Science and Information Engineering[2]
National Taiwan University
Taipei, Taiwan, Republic of China
E-mail: karlchen@ms14.hinet.net, lichen@ccms.ntu.edu.tw

Keywords: Visual Tracking, Real-time Target Detection, 3D Object Orientation Recognition, Adaptive Trajectory Filter

Abstract

In this paper, a night-and-day operational monocular visual tracking system (VTS) for a gun turret system is proposed. Its functions include detection and tracking an air-target in realistic noisy environment as well as automatic aiming of the turret gun at the target. For detection purpose, a real-time (about 100ms) integrated detector is proposed to obtain the 2-D target location. In addition, for helping the gun fighter to make a better firing decision, we also try to retrieve the other 3-D orientational information of the target via the model based pattern recognition. Thus, the task of 5-D target trajectory detection can be achieved in our monocular VTS. However, in this paper only the first 2-D target location tracking is implemented.

Next, a two-layer structure fuzzy controller, named recurrent fuzzy visual gain controller (RFVGC), based on the usual perspective visual model but with additional consideration of the target velocity is proposed. Inside this controller, a mixed Kalman filter with IMM algorithm is also developed to estimate the 2-D target position when the target is temporarily missing and its 2-D velocity i.e., both 3-D information are projected onto the image plane. This mixed Kalman filter will be mainly used to tune the first-layer de-fuzzy membership function. Finally, some experimental results show that this proposed monocular VTS can achieve the satisfactory performance for tracking the maneuverable air-target.

1. Introduction

Automated weapon systems have been developed devotedly by many countries for decades. Toward the automation of the weapon system, target surveillance or tracking is one of the most important tasks. For traditional tracking system, radar plays an important role. However, there are several disadvantages about the radar system. First, it is too expensive for the small-scaled close-in weapon systems. Second, some interference signals such as the clutter, RFI (friendly R.F. interference), or multi-path will affect its abilities seriously. Third, because of the uncertainty ellipse generated from the ambiguity function of the radar signal, it will not provide enough accuracy for such close-in weapons. Last, since radar is an active tracking system, it may be easily expressed to hostile and dangerous environment due to its traceable signal. This is fatal disadvantageous for a fixed defense weapon system.

With the advances of the computer, real-time image processes become possible. Therefore, here, we propose a computer aided visual tracking system (VTS) that consists of the real-time target detection, visual control, target trajectory filter, and camera platform control. The features of this tracking system include smaller volume, lower cost, and higher accuracy. Notably, this visual system tracks the target position in a passive manner, which takes to the fixed defense weapon systems with higher survival rate. In the future, this monocular VTS will be physically associated with a gun turret and play the rule of automatic target aiming auxiliary system.

Here, we will propose this monocular VTS which is able to provide the functionalities mentioned above under the following assumptions and specifications.

Environment descriptions:

- There is only one genuine target in the ROI (region of interest) image at any time.
- The plane size of the target airplane is always assumed within 12x12-meter square area, such as the F-16 fighter.

Image processing:

- The real-time processing adopted takes 100ms.
- The motion of the target will be assumed rather continuous from an image frame to its next frame.

Camera platform:

- Throughout the course of tracking, the target never leaves the entire 3-D detectable workspace of the camera platform.

The amount of researches in visual tracking control in the past is far beyond our descriptions, but the techniques for real-time implementation have not been well discussed. Here, we would just point out a few.

First, about the target detection, there are several distinct methods which have been used. Aggarwal and Yalamanchili discussed the motion-energy method [1][2][3] (differencing approach) based on the MTI (moving target indicator) of the radar system. However, unfortunately, it is too sensitive to the noise so that it can not be used in real environment. To resolve such difficulty, several pattern matching methods [3][5] have been proposed and been tried to detect the target in the noisy environment.

Second, about the visual tracking, Koivo and Houshangi [4] introduced an adaptive control technique in conjunction with the information obtained from a stationary camera in order to grasp a moving object. Papanikolopoulos [5][6] applied several well-known control algorithms and experimentally examined the performance of the resulting controllers. Observer-based control approach and optimal control approach have also been proposed by Hashimoto et al. in 1996 [7].

Last, about the adaptive trajectory filtering, Bar-Shalom [8] proposed the IMM (interacting multiple model) algorithm, and then several types of the trajectory filters with this algorithm for distinct conditions were presented such as Bradshaw et al. [9] used this method in extended Kalman filters and obtained satisfactory results.

The paper is organized as follows: Chapter 1 is an overall introduction. In Chapter 2, we give an overall description of a gun turret carrying our proposed monocular visual tracking subsystem. In Chapter 3, a real-time integrated method is presented to achieve the real-time target detection. Moreover, for helping the gun fighter to make a better firing decision, the other 3-D orientational information of the target via model-based pattern recognition is also retrieved.

Chapter 4, a new structure of the visual servo controller, uses the so-called recurrent fuzzy approach to tune the best visual gain. Inside this controller, a IMM Kalman filter is proposed which not only provides the target velocity to tune the de-fuzzy membership function of this RFVGC, but also filters the trajectory and predicts the next target position when the temporary target missing occurs.

The hardware configuration of our system and some experimental result are described in Chapter 5. Finally, the paper is concluded in Chapter 6.

2. System Overview

In the lab-scale experiment with a focus on this visual tracking task, a stereo vision VTS has been built in Advanced Control Lab. in Dep. of Electrical Engineering of National Taiwan University [10] as shown in Fig. 1. In this paper, we only adopt one of the a-pair-of this stereo vision VTS to conduct the monocular visual tracking research. Like the gun turret system, the monocular VTS can also make rotational movements both in the vertical and horizontal planes via two DC servomotors.

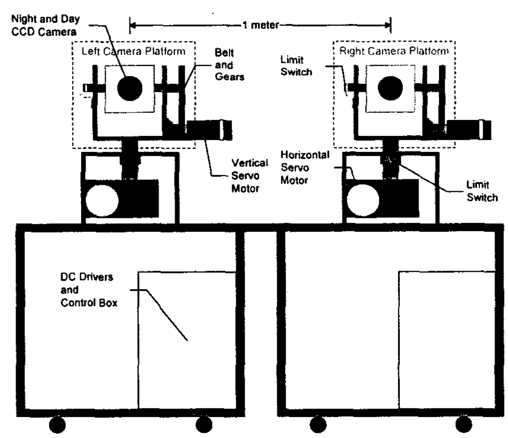

Fig. 1: Stereo vision visual tracking system.

In Fig. 2, a reasonable architecture of monocular visual tracking research can be clearly expressed in a block diagram. At first, an image frame is grabbed, in which we detect the target location in real-time. Of course, these image processes should work for all kinds of noisy sky environment night and day, and their correctness will be very critical for the following steps.

Next, the relations between the task of gradual centering the target centroid at the image center and the task of moving the camera to achieve the former task have to be identified. In this paper, we derive a visual model for that based on the traditional perspective projection principle with the considerations of the non-zero velocity target. The underlying idea of this model is roughly that: the trajectories to move the camera platform at the next samples should be well dependent on where the target image appears in the image plane and on its image velocity. Furthermore, based on these information, the size of the next ROI images are also determined to reduce the image processing time and the false detection probability. Later, we will show that this visual model is realized in terms of — two-layer fuzzy structure.

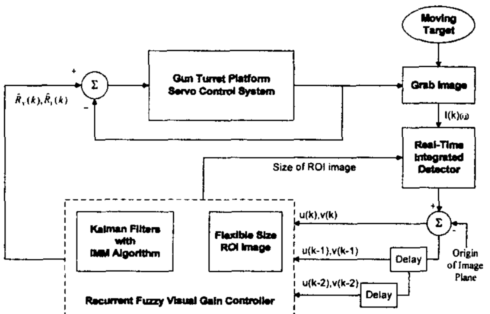

Fig. 2: The architecture of the monocular VTS.

Finally, after the moving displacement of the camera platform is determined, the PD control with gain scheduling is exercised to move the camera to traverse the desired trajectory. About the controllers of this stereo camera platform and the calibrations of their coordinate frames were proposed in our previous paper [10].

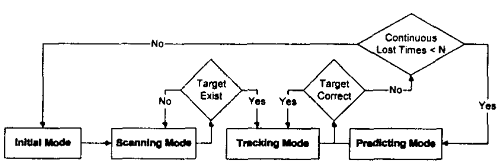

Fig. 3: The logic switching strategies.

In this paper, the monocular VTS is designed to assist the operator of the gun turret. To play a successful role, the control strategies adopted by the VTS and control flow governing the cooperation between the VTS and the human are very crucial. Here, the proposed control flow can be further divided into four basic modes: initial mode, scanning mode, tracking mode, and predicting mode as shown in Fig. 3.

3. Real-time Integrated Detector

In this chapter, we present a real-time integrated detector in order to achieve the goal of real-time target detection in realistic noisy sky environment. In Fig. 4, the block diagram of this detector is illustrated, and the demonstration of it applied to deal with a real airplane image in the morning is illustrated in Fig. 5.

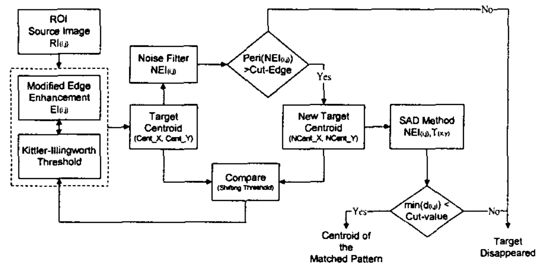

Fig. 4: Block diagram of the real-time integrated detector

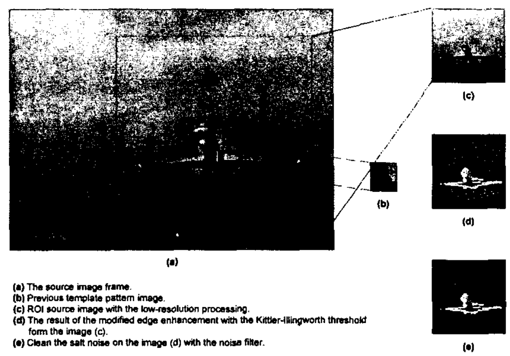

Fig. 5: Real-time integrated detector applied to real airplane in the morning.

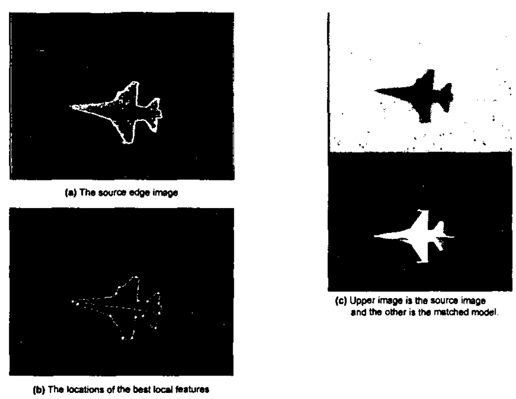

Fig. 6: Orientation recognition to the model airplane in the night time.

Model-based pattern recognition

Then, we also try to retrieve the 3 DOF rotational information with the model-based pattern recognition. Here, the K-curvature method with different distance K to extract the corner features from every sections of the binary target edge image that is developed by the former steps. Then, these extracted features are re-expressed against one reference frame and achieve our goal to find the features invariant to translation, rotation, and scaling. Then, a hash table is built based on the quantized coordinates of the invariant features for all sampled orientation models.

By voting and matching the orientation models, which one has the minimum Euclidean distance is the most similar model for this recognized target image. As a result, the orientation of the target in terms of row, pitch and yaw angles can be detected. Fig. 6 shows the sequential results of these recognition processes.

4. Recurrent Fuzzy Visual Gain Controller (RFVGC)

In practice, our VTS is a dynamic look-and-move system. In other words, the camera platform should move in tune with the target. This is a standard MCMO (moving camera, moving object) case. So, how to derive its visual model is a very complicated task, especially for the unknown moving trajectories of target. Here, we try to use a two-layer fuzzy structure to solve this sophisticated task.

First, the MCSO (moving camera, stationary object) case is considered to derive the visual model based on the perspective projection principle. In practice, our monocular VTS only has 2 DOF, ω_X and ω_Y, and the origin of the platform coordinate frame is calibrated closed to the origin of the camera coordinate frame [10]. Hence, this invertiable image Jacobian J_{image} can almost be rewritten as:

$$J_{image} \approx \begin{bmatrix} -\frac{uvs_v}{\lambda} & \frac{\lambda}{s_u}+\frac{u^2 s_u}{\lambda} \\ -(\frac{\lambda}{s_v}+\frac{v^2 s_v}{\lambda}) & \frac{uvs_u}{\lambda} \end{bmatrix}, \quad (1)$$

and the control law is discretized into the following:

$$\begin{bmatrix} \Delta R_X \\ \Delta R_Y \end{bmatrix} = J_{image}^{-1} \begin{bmatrix} \Delta u \\ \Delta v \end{bmatrix}, \quad (2)$$

where ΔR_X and ΔR_Y are the rotational angles of the horizontal and the vertical servomotors. Furthermore, the off-diagonal terms of J_{image}, $\frac{\lambda}{s_u}+\frac{u^2 s_u}{\lambda}$ and $\frac{\lambda}{s_v}+\frac{v^2 s_v}{\lambda}$, are much larger than the terms, $\frac{uvs_v}{\lambda}$ and $\frac{uvs_u}{\lambda}$, the u-direction and v-direction of the image plane can be separately fuzzified.

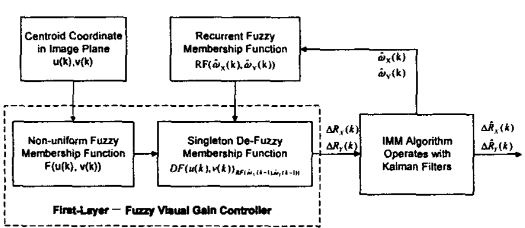

Fig. 7: Block diagram of RFVGC.

As shown in Fig. 7, the block diagram of the RFVGC is described. Since the motions of the camera platform are in tune with the motions of the target, the IMM Kalman filter can be used to describe the trajectory of the camera platform and hence that of the target motions. It can also provide the information about the estimated 2-D projected target velocity, i.e.: the velocity of camera platform, to tune the singleton de-fuzzy membership function of the first-layer.

5. Experiments

In this chapter, some experiments are illustrated with this monocular VTS in the indoor environment. A target, a model airplane, moves along a rail, and its velocity can be controlled by tuning the input voltage of the DC servomotor. In our research, the target motions are assumed free since we can arbitrarily increase or decrease the input voltage as desired. In the following, the main equipment of this monocular VTS is listed.

- AMD K6-MMX200 CPU.
- DIPIX XPG-1000 image grabber card (T.I. TMS-320C40 DSP on board).
- PULNiX TM-7CN/EX night and day CCD camera (RS-170 analog CCD camera).

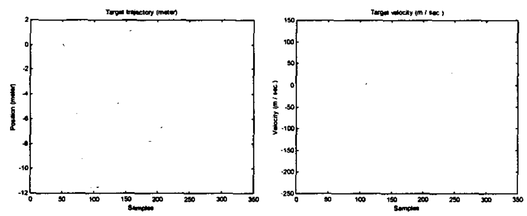
Fig. 8: The measured trajectory and velocity of the target in Condition 1.

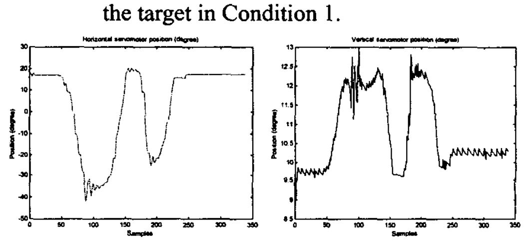
Fig. 9: Servomotor position trajectories in Condition 1.

Moving Target (Condition 1)

In Condition 1, a standard task of this monocular VTS is executed for tracking a moving free target. Here, the target moves in a way that it repeats its break-and-forth motion twice with varying velocities as shown in Fig. 8. In the figure of the target velocity, the distance between the CCD camera to the target is assumed 1 km. Then, the tracking results of the servomotor positions measured from the encoders are illustrated in Fig. 9. Here, special attention should be paid for the situation where the target just changes its moving direction. In these occasions, due to the rapid reverse of the motion directions of the target, the camera platform has to spend more time to get back to follow these abrupt maneuvers closely and hence larger errors will be inevitably generated. The distribution of the tracking errors in the image frame is shown in Fig. 10, and the means of them in u-direction and v-direction of the image coordinate frame are 11.433 and 9.668 pixels separately.

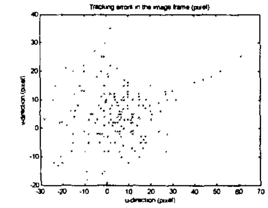

Fig. 10: Tracking errors in the image frame (pixels).

Target Temporarily Missing (Condition 2)

In Condition 2, this monocular VTS is still tracking a moving free target, but a small cloud will block the target so that it will temporarily disappear from the image frame. Then, the abilities of our monocular VTS to predict the target trajectories which keeping detecting the missing target are demonstrated. In this case, the target penetrates into a small cloud and waits there for a moment as shown in Fig. 11. Thus, this small cloud should occlude the target, and the tracking flows should switch to the predicting mode immediately as shown in Fig. 12. In this figure, the standard tracking mode is defined as the low-level of the dash line, whereas the predicting mode is defined as the high-level.

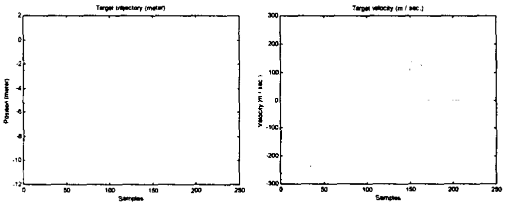

Fig. 11: The measured trajectory and velocity of the target in Condition 2.

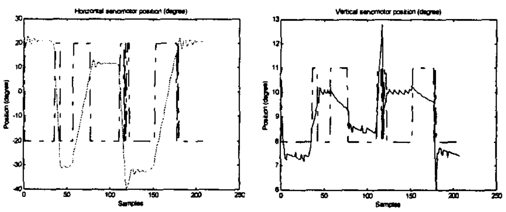

Fig. 12: Servomotor position trajectories in Condition 2.

6. Conclusions

In this paper, a real-time (about 100ms) integrated detector incorporating some modified versions of the proposed detection methods is presented to achieve following objectives: real-time detection, locking one target, and running in day and night. Then, for helping the gun fighter to make a better firing decision, the target orientation recognition is also proposed.

Next, the visual model with the consideration of the target velocity is fuzzified as a proposed two-layer structure, RFVGC, to obtain a superior performance for the maneuverable target. Inside this controller, the IMM Kalman filter is used to filter and predict the trajectory and velocity of the target.

Finally, in this paper, the tracking strategies of this real-time monocular VTS are also proposed. We believe that this developed system can be regarded as a target information provider and an effective assistant of the gun fighter to enhance the precision of the automated weapon systems.

Reference:

1. Aggarwal and N. Badler, “Special Issue on Motion and Time-Varying Imagery”, IEEE Trans. PAMI, PAMI-2, 1980.
2. S. Yalamanchili, W. Martin, and J. Aggarwal, “Extraction of Moving Object Descriptions via Differencing”, Computer Graphics and Image Processing, Vol. 18, 1982, pp. 188-201.
3. M. Sezgin, S. Birecik, D. Demir, I.O. Bucak, S. Cetin, and F. Kurugollu, “Comparison of Visual Target Tracking Methods in Noisy Environment”, IECON Proceedings (Industrial Electronics Conference), Part 2, 1995, pp. 1360~1365.
4. A. Koivo and N. Houshangi, “Real-time Vision Feedback for Servoing robotic Manipulator with Self-Tuning Controller”, IEEE Trans. Syst., Man, Cyber., Vol. 21, No. 1, 1991, pp. 134-142.
5. N. Papanikolopoulos, P. Khosla, and T. Kanade, “Visual Tracking of a Moving target by A Camera Mounted on A Robot”, IEEE Trans. on Robotics and Automation, Vol. 9, No. 1, 1993, pp. 14-35.
6. N. Papanikolopoulos, B. Nelson, and P. Khosla, “Six Degree-of-Freedom Hand/Eye Visual Tracking with Uncertain Parameters”, IEEE Trans. on Robotics and Automation, Vol. 11, No. 5, 1995, pp. 725-732.
7. K. Hashimoto, T. Ebine, and H. Kimura, “Visual Servoing with Hand-Eye Manipulator-Optimal Control Approach”, IEEE Trans. on Robotics and Automation, Vol. 12, No. 5, 1996, pp. 766-774.
8. H. A. P. Blom and Y. Bar-Shalom, “The Interacting Multiple Model Algorithm for Systems with Markovian Switching Coefficients”, IEEE Trans. on Automatic Control, Vol. 33, No. 8, 1988, pp. 780-783.
9. Kevin J. Bradshaw, Ian D. Reid, and David W. Murray, “The Active Recovery of 3D Motion Trajectories and Their Use in Prediction”, IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 19, No. 3, 1997, pp. 219-234.
10. Huah Tu, Chien-Hsiang Chen, Chih-Yu Chen, and Li-Chen Fu, “A System for 3-D Target Trajectory Detection via Stereo Visual Tracking”, IEEE Trans. on Automatic Control, 1998. (revised)