RobotEye Technology for Thermal Target Tracking Using Predictive Control

Ansu Man Singh[1], Quang Ha[1], David K. Wood[2], Mark Bishop[2], Quang Nguyen[3] and Albert Wong[3]

[1.] School of Electrical, Mechanical and Mechatronic Systems, University of Technology Sydney, Australia
[2.] Ocular Robotics Pty Ltd, NSW, Australia
[3.] Defence Science and Technology Group (DSTG), Department of Defence, Australia

Abstract –
Thermal cameras are widely used in the fatigue analysis of mechanical structures using the thermoelastic effect. Nevertheless, such analysis is hampered due to blurry images resulting from the motion of structure-under-test. To address the issue this paper presents a system that utilizes robotic vision and predictive control. The system comprises of a thermal camera, a vision camera, a RobotEye, and a fiducial detection system. A marker is attached to a thermal target in order to estimate its position and orientation using the proposed detection system. To predict the future position of the thermal moving object, a Kalman filter is used. Finally, the Model Predictive Control (MPC) approach is applied to generate commands for the robot to follow the target. Results of the tracking by MPC are included in this paper along with the performance evaluation of the whole system. The evaluation clearly shows the improvement in the tracking performance of the development for thermal structural analysis.

Keywords –
Automation and control, Computer Vision

1 Introduction

Kelvin's Law has laid the scientific foundation on which the mechanical stress measurements can be derived from the temperature measurements [1]; and the formation of the Thermoelastic Stress Analysis (TSA) system has further taken shape alongside with the emergence of the suitable modern thermal camera where it can be used to remotely measure the stress distribution of a structure under load [2].

The effective use of TSA for full-field stress measurement, structural fatigue detection and crack propagation have since extended its prognostic capabilities to the construction industry, where TSA has been used to inspect and measure the usage stress load of aging infrastructures such as steel bridges, railway tracks, power poles, pipelines, heavy industrial equipment and vehicles [3], [4]. Even though full-field TSA measurement can be remotely conducted for large civil engineering structures, it would often demand for expensive optical setup and the resources required.

Since the structure is continuously under loading motion, with the thermal camera fixed looking at a region of interest, the resultant motion blur would stop TSA from being useful. To account for the resultant shift of pixels in the thermal images, the thermal camera can directly be mounted onto the structure under test; unfortunately, this approach has not always been practically or physically feasible. Thus specific attempts have often been made to compensate for this motion blur under each individual experimental setup; Wong & Ryall first presented a single axial loading, video-processing motion compensation technique [5], yet the search for an elegant and effective motion compensation method is still ongoing.

Attempts to visually track a moving structure with camera mounted on motion platforms such as a servo controlled pan-tilt mechanism, active gimbal devices and 6-DoF robotic systems etc. are some of the tactics that have been widely explored and developed [6]. Nevertheless, due to the demand on complex multidimensional loading dynamics, and due to nonlinear dynamics such as backlash and friction, tracking performance of such system degrades when it comes to random motion.

In order to tackle this problem, researchers from Defence Science and Technology Group have put forward a unique solution by having a stationary camera with a pointing mechanism to track moving structure. A major advantage of this technique is in having the stationary thermal camera staring at a low-mass pointing mirror, which undergoes a small angular motion to track an external structure's motion. This is where the complementary Ocular Robotics' RobotEye innovative optical pointing technology has made this solution practically possible [7].

This paper presents the solution proposed by Defence Science and Technology Group, where the high speed visual feedback is used to enable the optical pointing sensor to track the structural motion via a defined target, whilst at the same time it allows thermal camera to take long exposure, stable, un-blurred images of the region in

the vicinity of this defined target.

This solution comprises of a RobotEye, the state-of-the-art pointing technology which has been developed and commercialized by Ocular Robotics Limited [7], a high-speed vision camera, a marker detection system [8], and a thermal camera embedded within the optical pointing system to capture the view where the eye is pointing. The pointing technology uses mirror-based technology [9], which significantly reduces the inertia of the sensor in comparison to the pan-tilt system. A fiducial marker is attached to the target, where its position is calculated using the marker detection system, as explained in [8]. Included also is the design of a predictive tracking controller based on the Model Predictive Control (MPC) for the optical pointing sensor's head motion. Since the robot is connected to a computer network, its tracking performance is affected by delays. Nevertheless, owing to the MPC, the control system for the actuator can generate control signals even in the presence of the unexpected delays.

Furthermore, this study canvasses a transformational idea that aptly fits to the advancement of modern day robotics and inspection technology, where our proposed solution here can be retrofitted onto a suitable autonomous vehicle platform to deliver a low, cost-effective in-situ structural health prognostic monitoring and management system. We envisage the potential benefits of having large infrastructures such as bridges, high speed rail tracks, pipelines and wind turbines inspected using such TSA autonomous systems, where the collected structural health information can be used to facilitate sound engineering analyses and decisions, and thus avoiding any surprise failures.

This paper is arranged in following sections. Details about the RobotEye technology and the proposed system are presented in Section 2, followed by the MPC for the robot's head motion in Section 3. Results are presented in Section 4. Finally, this paper concludes with a conclusion and recommendation for future works.

2 RobotEye Technology

The sensing technology used in this paper has a specific advantage in cutting-edge precision and responsiveness: it can be applied to enhance the dynamic control of LiDAR sensors for 3D mapping, navigation and automation, through to simultaneously tracking multiple dynamically moving objects even when mounted on a moving platform itself. In addition, the technology can be integrated with various types of sensors including vision, thermal and hyper-spectral, adding an extra dimension to sensing with these technologies [8].

2.1 Variables representing gazing direction

Orientation of the RobotEye pointing device or its the gazing direction can be described by two variables namely azimuth (ψ) and elevation (θ) as shown in Figure 1. The azimuth angle represents the rotation of the head about the vertical axis as depicted in the figure. Similarly, the elevation angle represents the angle made by the viewing direction with respect to the horizontal plane.

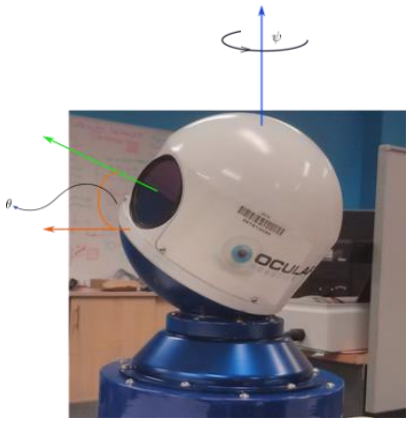

Figure 1: Variables representing robot eye head motion

2.2 Thermal target tracking

The advantage of light weight design makes the RobotEye technology a prefered choice in tracking applications. The laboratory set-up for this application consists of a high-speed vision camera, a RobotEye, a circular patches marker, and a thermal camera.

In the thermal target tracking system, the target is placed at the center of a fiducial marker which is within the field-of-view of the vision camera. The marker consists of circular patches. Details on the marker design and the detection algorithm are provided in [9]. The thermal camera is embedded inside the pointing sensor, thus, it captures the view where the RobotEye is pointing. This system has a huge implication in the fatigue analysis of mechanical structures which are under constant motion. As mentioned previously, fatigue analyses of such moving structures are adversly affected due to the blurry images captured by static cameras.

2.3 System Configuration

Coordinate systems assigned to the RobotEye optical sensor pointing device and the vision camera for tracking are shown in Figure 2. The reference frames for the vision camera and the RobotEye are represented by O_V and O_R, respectively. Similarly, the R_V^R is the rotational matrix that represents the orientation of the pointing device with respect to the vision camera reference frame. Consider an

observation of the target position with respect to the vision camera be represented as p_V^t. Then, the observation can be represented with respect to the RobotEye reference frame as

$$p_R^t = R_V^R p_R^t + p_R^V, \quad (1)$$

where p_R^t is the position of target and p_V^t is the location of vision camera with respect to the pointing device.

As presented in Section 2.1, gazing direction for the RobotEye optical pointing sensor is represented by azimuth ψ and elevation θ angles which are in the spherical coordinate system. Therefore, for the representation of the thermal target as the gazing direction from the RobotEye, conversion from Cartesian to spherical coordinates should be applied as

$$\begin{aligned} \psi &= \tan^{-1}\frac{y}{x} \\ \theta &= \tan^{-1}\left(\frac{\sqrt{y^2+x^2}}{z}\right), \end{aligned} \quad (2)$$

where $p_R^t = [x\ y\ z]$.

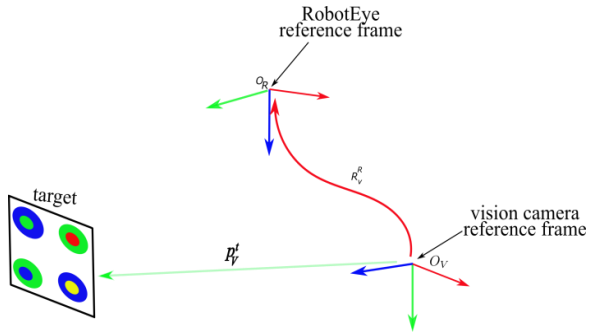

Figure 2. Arrangement of RobotEye and vision camera

2.4 Control architecture for thermal tracking

The system architecture for the proposed system comprising of a RobotEye, vision camera, and thermal camera is shown in Figure 3. In the system, the position and orientation of the target is estimated from the frames captured by vision camera which utilizes the circular marker detection algorithm [9].

After estimating the position of the thermal target with respect to the vision camera, it is calculated with respect to the pointing device in the azimuth and elevation space using equations (1) and (2). The conversion requires the extrinsic parameters such as R_V^T and p_R^V. After that, a prediction algorithm based on Kalman filter is utilised to predict its future trajectory. The filter uses a constant velocity model to represent the target motion. The model is represented as:

$$\begin{aligned} \hat{\boldsymbol{x}}[n+1] &= A_0\hat{\boldsymbol{x}}[n] \\ \hat{\boldsymbol{y}}[n] &= C_0\hat{\boldsymbol{x}}[n], \end{aligned} \quad (3)$$

where $\hat{\boldsymbol{x}} = [\psi\ \theta\ \dot{\psi}\ \dot{\theta}\ ]^{\mathrm{T}} \in \mathbb{R}^4, \hat{\boldsymbol{y}} = [\psi\ \theta\ ]^{\mathrm{T}} \in \mathbb{R}^2$,

$$A_0 = \begin{bmatrix} 1 & 0 & T & 0 \\ 0 & 1 & 0 & T \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

$$C_0 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix},$$

and T is the sample period.

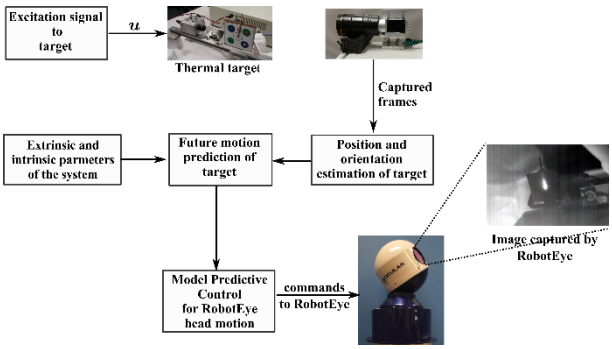

Figure 3. System architecture.

The predicted trajectory is then fed to a model predictive controller (MPC) which generates commands for the RobotEye drive controller. Here, the objective of MPC is to compensate for the lags in the system, which are introduced through various sources such as sensing and network delays, in order to improve the overall performance of the system. Modelling and Control System Design

2.5 Modelling

For the representation of the motion dynamics of the RobotEye sensor head, we have considered the state-space model in discrete-time as

$$\begin{aligned} \boldsymbol{x}[n+1] &= A\boldsymbol{x}[n] + B\boldsymbol{u}[n] \\ \boldsymbol{y}[n] &= C\boldsymbol{x}[n], \end{aligned} \quad (4)$$

where $\boldsymbol{x} \in \mathbb{R}^n$, $\boldsymbol{u} \in \mathbb{R}^m$, $\boldsymbol{y} \in \mathbb{R}^p$ are, respectively, the states, inputs and outputs of the system. In the case of the pointing actuator $\boldsymbol{y} = [\psi\ \theta]^{\mathrm{T}} \in \mathbb{R}^2$ and $\boldsymbol{u} = [\psi_{cmd}\ \theta_{cmd}]^{\mathrm{T}}$, where ψ_{cmd} and θ_{cmd} represent the command inputs to the system.

2.6 System Identification

For the identification of the system's parameter of equation (4), we measured the angles and inputs through the application programming interface (API) provided by

the RobotEye. Then, we applied a standard identification algorithm such as N4SID in the framework provided by Matlab. Figure 4 shows the validation results of the identified method, which clearly indicate that the responses from the identified model is similar to that of the actual system. For instance, the similarity of responses in elevation and azimuth angles are, respectively, 95.4% and 94.39%. System matrices of the identified model are obtained as

$$\mathrm{A} = \begin{bmatrix} -0.4141 & 288.6 & -848.2 & 54.97 \\ -75.28 & -40.89 & -495.1 & -127 \\ 757.5 & 43.73 & -546.4 & -80.56 \\ 75.16 & 48.69 & -210.6 & -13.13 \end{bmatrix},$$

$$\mathrm{B} = \begin{bmatrix} -96.69 & -82.96 \\ -42.41 & -45.76 \\ -108.7 & -92.47 \\ -28.4 & -24.58 \end{bmatrix},$$

and

$$C = \begin{bmatrix} -9.633 & -9.392 & -8.936 & 72.69 \\ 19.64 & -11.09 & 7.066 & -98.29 \end{bmatrix}.$$

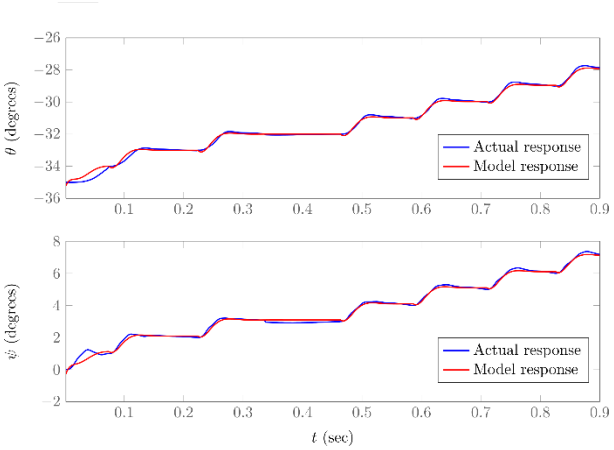

Figure 4. Validation of the identified model.

3 Model Predictive Control (MPC)

In order to apply the predictive control, we utilized the MPC in tracking the output reference signal. The problem is formulated as

$$\min \begin{Bmatrix} \frac{1}{2}\sum_{i=1}^{N}\{(\boldsymbol{y}_i - \boldsymbol{r}_i)^{\mathrm{T}}\boldsymbol{Q}_y(\boldsymbol{y}_i - \boldsymbol{r}_i) + \boldsymbol{u}_i{}^{\mathrm{T}}\boldsymbol{Q}_u\boldsymbol{u}_i\} \\ +(\boldsymbol{y}_N - \boldsymbol{r}_N)^{\mathrm{T}}\boldsymbol{Q}_N(\boldsymbol{y}_N - \boldsymbol{r}_N) \end{Bmatrix} \quad (5)$$

$$\text{subject to: } \begin{cases} \boldsymbol{u}_{min} \le \boldsymbol{u} \le \boldsymbol{u}_{max} \\ \boldsymbol{x}_{i+1} = A\boldsymbol{x}_i + B\boldsymbol{u}_i \\ \boldsymbol{y}_i = C\boldsymbol{x}_i, \end{cases}$$

where $\boldsymbol{y}_i = \boldsymbol{y}[n+i]$ is the i^{th} future output of the system, $\boldsymbol{u}_i = \boldsymbol{u}[n+i]$ is the i^{th} future input to the system, $\boldsymbol{r}_i$ is the i^{th} future reference, N is the prediction horizon, A is the system matrix, and B is the input matrix.

The cost function $V(\boldsymbol{U})$ can also be represented as

$$\begin{aligned} V(\boldsymbol{U}) \quad = \quad & (\boldsymbol{y}_0 - \boldsymbol{r}_0)^{\mathrm{T}}\boldsymbol{Q}_0(\boldsymbol{y}_0 - \boldsymbol{r}_0) \\ & +(\boldsymbol{Y} - \boldsymbol{R})^{\mathrm{T}}\bar{\boldsymbol{Q}}_y(\boldsymbol{Y} - \boldsymbol{R}) \\ & \boldsymbol{U}^{\mathrm{T}}\bar{\boldsymbol{Q}}_u\boldsymbol{U}, \end{aligned} \quad (6)$$

where

$$\boldsymbol{Y} = \begin{bmatrix} \boldsymbol{y}_1 \\ \boldsymbol{y}_2 \\ \vdots \\ \boldsymbol{y}_N \end{bmatrix}, \ \boldsymbol{X} = \begin{bmatrix} \boldsymbol{u}_0 \\ \boldsymbol{u}_1 \\ \vdots \\ \boldsymbol{u}_{N-1} \end{bmatrix},$$

and

$$\boldsymbol{R} = \begin{bmatrix} \boldsymbol{r}_1 \\ \boldsymbol{r}_2 \\ \vdots \\ \boldsymbol{r}_N \end{bmatrix}.$$

With the system dynamics obtained and by using forward substitution, one can formulate the optimization problem as:

$$\begin{aligned} & \text{minimize } V(\boldsymbol{U}) = \boldsymbol{U}^{\mathrm{T}}\bar{H}\boldsymbol{U} + \bar{F}^{\mathrm{T}}\boldsymbol{U} \\ & \text{subject to: } E\boldsymbol{U} \le \boldsymbol{\delta} \end{aligned}, \quad (7)$$

where

$$\begin{aligned} \bar{H} &= \bar{S}^{\mathrm{T}}\bar{Q}\bar{S} + \boldsymbol{R}, \\ \bar{F}^{\mathrm{T}} &= 2\boldsymbol{x}_0\bar{T}\bar{Q}\bar{S} - \boldsymbol{R}^{\mathrm{T}}\bar{Q}\bar{S}, \end{aligned}$$

$$\bar{E} = \begin{bmatrix} I \\ -I \\ \bar{S} \\ -\bar{S} \end{bmatrix},$$

and,

$$\boldsymbol{\delta} = \begin{bmatrix} \Delta\,\boldsymbol{U}_{max} \\ -\Delta\,\boldsymbol{U}_{min} \\ \boldsymbol{Y}_{max} - \bar{T}\boldsymbol{x}_0 \\ -\boldsymbol{Y}_{min} + \bar{T}\boldsymbol{x}_0 \end{bmatrix}.$$

Now, this problem can be solved using the available QP solver [10].

4 Tracking results

4.1 Simulation

Figure 5 shows the results of the MPC controller for the sinusoidal reference signal. From the figure it is clear that the controller is able to track the reference signal. Similarly, from the plots of error signals, as presented in Figure 6, it is clear that the error in the Azimuth space is less than 0.03 degrees, compared to 0.01 degrees in case of the elevation angle.

The tracking performance can also be evaluated in terms of the performance index, e.g. Integral Square Error (ISE) which is defined as:

$$ISE = \sum_{i=1}^{M} e^2\,[n]. \quad (7)$$

The ISE errors for Azimuth and Elevation angles are 0.34937 and 0.017448 degree^2, respectively.

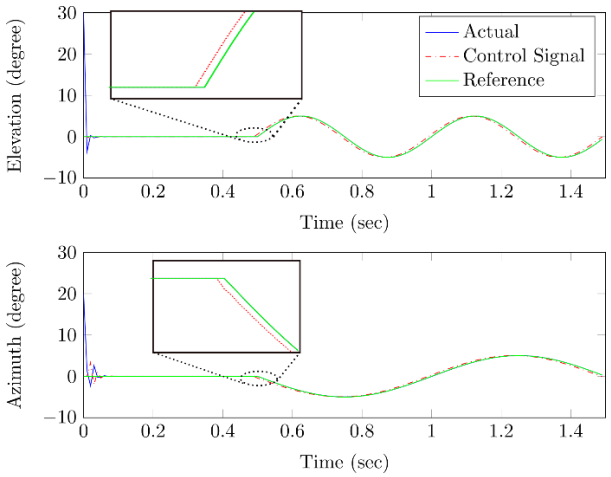

Figure 5. Plot of the angles along with command and actual values

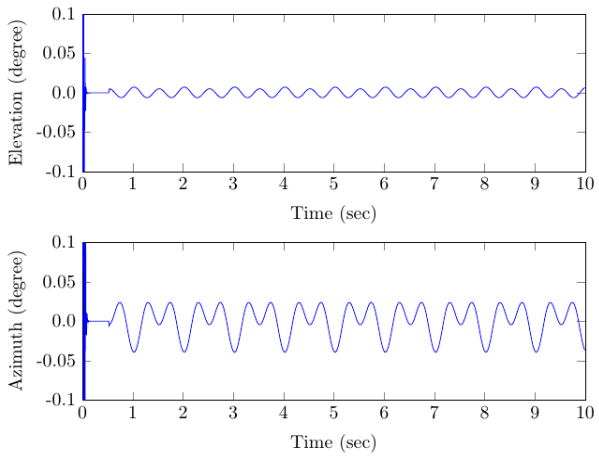

Figure 6. Plot of errors in Azimuth and Elevation angles

4.2 Real time tracking results

For the real-time testing, the MPC based predictive control algorithm was applied to a RobotEye with the specifications provided in Table 1. The block diagram for the experiment is provided in the Figure 7.

In the experiment, sinusoidal signals were provided as the reference signals. The frequencies of the reference signals were 1 Hz and 2 Hz for azimuth and elevation angles, respectively. To send the angular commands and receive the measurements we used the proprietary application programming interface.

The tracking errors during the experiment are shown in Figure 8. The error is less than 0.2 degrees for elevation and 0.1 degrees for the azimuth angle. Similarly, ISE errors for the azimuth and elevation angles are 3.947 and 33.73 degree^2, respectively.

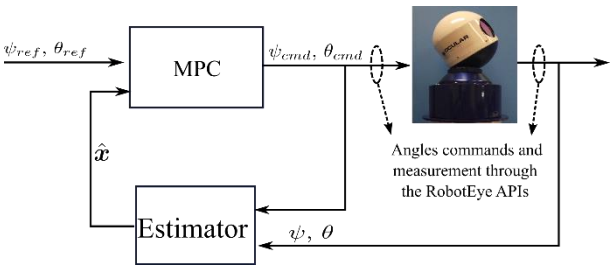

Figure 7: Real time experiment for MPC.

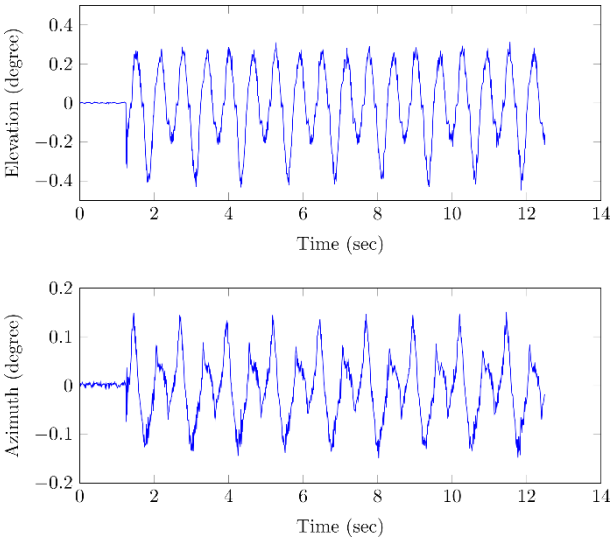

Figure 8. Tracking errors in azimuth and elevation angles in real experiment.

4.3 Performance evaluation for thermal target tracking

In this section, we evaluate performance of the proposed system described above and compare the performance with and without predictive control. The experimental setup for the evaluation of the system is presented in the Figure 9 with the specifications of all the components in Table 1.

The figure also shows the target consisting of the circular patched marker with a heat source at the centre of the marker. The heat source was used to detect the target in the thermal camera.

In this real time application, there is no direct way of evaluating the tracking performance of the system. However, as mentioned previously, the thermal camera is embedded inside the optical pointing device, to capture

the view where it is directed to. Therefore, the thermal targets are detected via images taken by thermal and vision cameras, and then compared for evaluation.

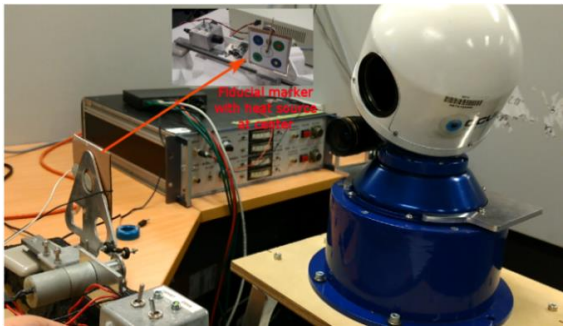

Figure 9: Experimental setup for tracking system.

The calculation of target pixels in the vision camera is obvious because the circular fiducial marker is attached with the target. However, the calculation of the target pixels in the thermal images is not straight forward. Therefore, a thresholding process is applied in every image, followed by a blob detection algorithm. The purpose of the blob detection algorithm is to detect hot spot on the image which is the location of the target.

Table 1: Parameters of the system.

	Parameter	Values
Vision camera	Focal length	16mm
	Resolution	658×492
	Frame Rate	100 fps
Thermal camera	Focal length	19mm
	Resolution	640×512
	Frame Rate	30fps
RobotEye	Model no	RELW 50
	Aperture diameter	50mm

After the calculation of target points in both vision and thermal camera, one can evaluate the target tracking performance of the RobotEye by comparing standard deviations of the target pixels in x and y-axis. Nevertheless, this approach does not give the clear indication of the performance. For example, if the variance of the target pixels in the vision camera is along y-axis but the variance of the pixels in the thermal camera is along x-axis then by comparing the above mentioned parameters does not represent the performance. In such situation, to better evaluate the tracking performance, the principal component analysis (PCA) of the target points can be applied. An example of the PCA of the target points is presented in Figure 10, wherein it is clear that the major axis and the minor axis of the PCA represents the direction of maximum and minimum variations of the target points. By using the principal component analysis, the following parameters:

- Maximum eigenvalues of the PCA of target pixels
- Minimum eigenvalues of the PCA of target pixels

are calculated for evaluation. It should be noted that in this system small variations of the target points in thermal cameras signifies better performance.

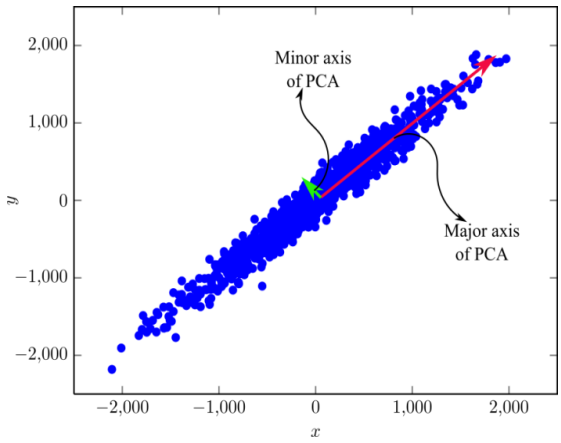

Figure 10. An example of the PCA of target points showing major and minor axis.

From the calculation of the above mentioned parameters, ratios of the parameters are obtained. Table 2 shows the summaries of the performance evaluations. From the table, it can be observed that there are improvements in all parameters. For instance, in terms of ratio of maximum eigenvalues there is the improvement by 74 percent. Similarly, in the case of the ratio of standard deviation in x-axis the improvement is about 42 percent.

Table 2: Summaries of the performance of the thermal tracking system with and without MPC.

	Without MPC	With MPC	% improvement
Ratio of maximum eigenvalues of PCA (thermal/vision)	0.0001	0.00002	74
Ratio of minimum eigenvalues of PCA (thermal/vision)	1.2213	0.8455	30
Ratio of standard deviation in X-axis (thermal/vision)	0.0101	0.005	42
Ratio of standard deviation (thermal/vision)	1.00659	0.8153	19

5 Conclusion

In this paper, we have presented a thermal target tracking system using RobotEye technology. The hardware comprises of the RobotEye optical pointing device, a thermal camera and a vision camera. Furthermore, the proposed system consists of an estimator for the target orientation and position with a Kalman filter using the application program interface for modeling and a Model Predictive Controller for command generation. In addition, a fiducial marker consisting of circular patches is attached to the target. The objective of the fiducial marker is to estimate the position of marker from the images taken from the vision camera. After the estimation of the target position, its future position is predicted using a Kalman filter, and a predictive control is then applied to generate command for the RobotEye. Experiments have been conducted and the tracking performance was evaluated for the system. The improvements of up to 74 percent is observed with the predictive control compared to the non MPC control case.

Acknowledgement

This work was supported by Ocular Robotics Limited, Australian Mathematical Science Institute (AMSI), and Defence Science and Technology Group (DSTG), Department of Defence of the Australian Government. The authors would also like to thank Mr Len Meadows (DSTG) for his persistent support.

References

[1] Wong, A. K., Dunn S. A., and Sparrow J. G, Residual Stress Measurement by Means of the Thermoelastic Effect. *Nature* 332, 6165, 1988, doi: https://doi.org/10.1038/332613a0

[2] Wong, A. K., Rajic, N., & Nguyen, Q. Seeing stresses through the thermoelastic lens-a retrospective and prospective from an Australian viewpoint. *An International Journal of Experimental Mechanics (Strain)*, 51(1), 2014, doi: https://doi.org/10.1111/str.12116

[3] Sakagami T., Izumi Y., and Kubo S. Application of infrared thermography to structural integrity evaluation of steel bridges. *Journal of Modern Optics*, 57(18), 2010, doi: https://doi.org/10.1080/09500340.2010.511289

[4] Haldorsen L. M., Thermoelastic Stress Analysis System Developed for Industrial Applications, *PhD Thesis, Alborg University,1998,* ISBN 82-7644-057-6

[5] Wong, A. K., and T. G. Ryall. The use of FAST on structures undergoing large motions. *Proceedings of the First Australasian Congress on Applied Mechanics: ACAM-96*. 1996

[6] Helgesen H. H., Leira F. S., Fossen T. I., and Johansen T. A., Tracking of Ocean Surface Objects from Unmanned Aerial Vehicles with a Pan/Tilt Unit using a Thermal Camera, *Journal of Intelligent & Robotic Systems*, 1–19, 2017, doi: https://doi.org/10.1007/s10846-017-0722-3.

[7] Wood D. and Bishop M., A novel approach to 3D laser scanning, in *Proceedings of Australian Conference on Robotics and Automation (ACRA)*, 2012,

[8] Singh A. M., Ha Q. P., Wood D. K., and Bishop M., Low-latency Vision-based Fiducial Detection and Localisation for Object Tracking, in *34th International Symposium on Automation and Robotics in Construction* pages. 706–711, Taiwan, 2017, doi: https://doi.org/10.22260/ISARC2017/0098.

[9] Okumura K., Oku H., and Ishikawa M., High-speed gaze controller for millisecond-order pan/tilt camera, in *IEEE International Conference on Robotics and Automation (ICRA)*, pages 6186–6191, 2011, doi: https://doi.org/10.1109/ICRA.2011.5980080.

[10] Tran, T. and Ha, Q.P. Quadratic Constraint Approach to Model Predictive Control of Interconnected Systems, *Springer*, ISBN 978-981-10-8407-2, 2018